# macroeconomics

Errol D'Souza

## A highly readable and relevant text:

"The text is easy to read and understand. The chapters are technical in nature, but the language is quite lucid…this book is better than my current textbooks."

"The writing style of the chapters is simple and very approachable."

"It is written in a lucid manner and is useful for teaching. The chapters contain some standard material and also cover some specific approaches by specific authors."

"The language is lucid, the examples make the message clearer, and the emphasis on Indian examples makes it easier to relate to our own surroundings and apply the knowledge readily."

## Uniquely detailed and provides in-depth explanation of topics:

"The text narrows down the explanation from a larger unit to individual units. This approach is not followed in other texts and that makes this book unique and special. One really comes to know the basics."

"This book is much more detailed than most books I've read so far! After reading through all the 'extra' portions (detailed reasons for a phenomenon, for instance, while most other books would just define the phenomenon and mention a few reasons), students can understand the topic much better than they would have from other books."

"Another thing that impresses me is the importance given to details and explanation. The book scores a point with its detailed structure and ability to strengthen the basics of a student."

## Replete with a host of innovative and value-adding features:

"The marginalia and macrofocus boxes are illuminating and make the reading quite interesting. The online application problems are quite innovative."

"The definitions and the extra information are highly value-adding to the readers. The end-of-chapter readings and appendices are useful to the students who would want to learn more than what is discussed in the chapter."

"Where this book really shines is in the clarity and presentation of the mathematical relationships, and its analytical content, which is clearly ahead of others. This book is far superior to any other macroeconomics textbook written in India."

## *Praise for* **Macroeconomics**

"Written in a lucid manner ..."

—Sanjaya Baru, *Media Adviser to the Prime Minister of India and former editor-in-chief of* The Financial Express

"It is a comprehensive and complete book with excellent features."

—Suma Damodaran, *XLRI, Jamshedpur*

"The book will be considered as the bible for students of economics."

—Mihir Mahapatra, *Goa Institute of Management*

"The novelty of the book is its excellent balance between theory and practice and the numerous examples and data."

—Sunil Mani, *Centre for Development Studies, Thiruvananthapuram*

"Owning this book would be tantamount to holding a lifelong treasure."

—K. S. S. Uduman Mohideen, *University of Madras*

"A well-written book."

—A. G. Moss, *OU College for Women, Hyderabad*

"*Macroeconomics* by Errol D'Souza is probably the first book written by an Indian author that caters significantly to the needs of the management students."

—Jaydeep Mukherjee, *Xavier Institute of Management, Bhubaneswar*

"The book is sure to get the attention of students for its lucidity."

—Paramita Mukherjee, *The ICFAI Business School, Kolkata*

"The pedagogical features like macrofocus boxes, summary, test yourself, and online-application questions are really interesting."

—K. C. Padhy, *Institute of Management and Information Sciences, Bhubaneswar*

"Concepts are explained using Indian data, which is really good."

—C. Bhujanga Rao, *Madras School of Economics*

"The book will provide a knowledge base to the students, researchers, and teachers in the area of macroeconomics."

—S. N. Sahdeo, *Birla Institute of Technology, Ranchi*

"The depth and 'correctness' of theory coupled with a very clear focus on learning from the Indian economy make it a much-needed book on macroeconomics."

—Runa Sarkar, *Indian Institute of Technology Kanpur*

"The organization of the book is worth mentioning …[The book] seeks to present the chapters in such a manner as to facilitate comprehension of the subject on a graduated and staggered basis."

—Prabal Kr. Sen, *XLRI, Jamshedpur*

"The USP of the book is the analysis of the variables in relation to real world, that is, the practical aspect of macroeconomic variables."

—Kumudani Sinha, *Patna University*

"Errol D'Souza has written a text on macroeconomics in a way that allows students to understand the basic macroeconomic issues."

—T. H. Thangiah, *International School of Business and Media, Pune*

"The book covers a wide range of macroeconomic issues and theories."

—I. Thiagarajan, *Loyola Institute of Business Administration, Chennai*

**Errol D'Souza**
INDIAN INSTITUTE OF MANAGEMENT AHMEDABAD

PEARSON Education

Delhi • Chennai • Chandigarh
Upper Saddle River • Boston • London
Sydney • Singapore • Hong Kong • Toronto • Tokyo

Library of Congress Cataloging-in-Publication Data

D'Souza, Errol.
Macroeconomics / Errol D'souza.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-8131708187 (pbk.)
1. Macroeconomics. I. Title.
HB172.5.D76 2008
339—dc22

2008010113



ISBN 978-81-317-0818-7

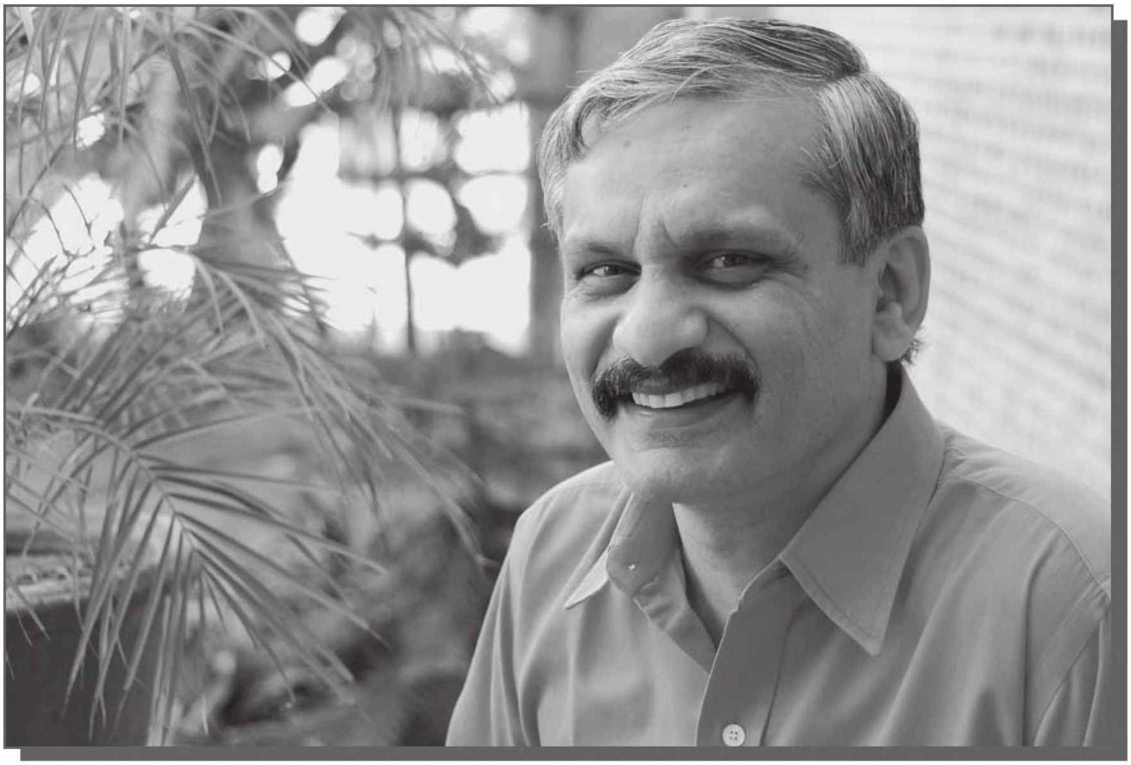
Head Office: 482 FIE, Patparganj, Delhi 110 092, India
Registered Office: 14 Local Shopping Centre, Panchsheel Park, New Delhi 110 017, India

Layout designed by Satwinder Singh Channey
Typeset in 10/12 Georgia by Astral Pre Media, NOIDA
Printed in India at India Binding House

Pearson Education Inc., Upper Saddle River, NJ
Pearson Education Ltd., London
Pearson Education Australia Pty, Limited, Sydney
Pearson Education Singapore, Pte. Ltd
Pearson Education North Asia Ltd, Hong Kong
Pearson Education Canada, Ltd., Toronto
Pearson Educacion de Mexico, S.A. de C.V.
Pearson Education-Japan, Tokyo
Pearson Education Malaysia, Pte. Ltd.

*For*
*Saachi and Hriya*

# About the Author



Errol D'Souza studied economics and statistics at the University of Mumbai where he was awarded the Kashinath Trimbak Telang gold medal for obtaining the first rank in his MA degree. He went on to receive his Ph.D. as a University Grants Commission National Research Fellow from Jawaharlal Nehru University, New Delhi, with a study of the fiscal limitations on state intervention in post-Independence India. He joined Goa University as an assistant professor in the Department of Economics and, for some time, was a consultant to the Finance Department, Government of Goa. In 1994, he moved to the University of Mumbai as an associate professor in industrial economics and public finance. In 1995, he was awarded a Ford Foundation fellowship and was a visiting scholar at Columbia University, New York, where he did theoretical work on the effectiveness of fiscal policy and tax evasion. Subsequently, in January 1999, he was appointed Industrial Finance Corporation of India Chair Professor at the Department of Economics, University of Mumbai.

Currently, Professor D'Souza is a professor of economics at the Indian Institute of Management Ahmedabad. He is a consultant to the Indian industry and the banking sector, the Reserve Bank of India, the United Nations Development Programme, the International Labour Organization, the Planning Commission, as well as the central and various state governments; he has worked as a member of the academic committees of the University Grants Commission and the Indian Council of Social Science Research; and he was an associate of the Indian Institute of Advanced Study, Shimla. He is also a director on the board of the National Housing Bank and an honorary senior fellow at the National University of Singapore.

Professor D'Souza's academic interests and areas of expertise include tax reforms, fiscal and monetary policy, the structure of corporate finance, social security and livelihood issues in the informal sector, and personnel economics.

Professor D'Souza teaches and lives at Ahmedabad.

# Brief Contents


*MacroFocus Boxes: Tying Theory to the Real World* xx
*Preface* xxi
*The Development of This Book* xxvii
*A Word of Thanks* xxx

1 Macroeconomics: What Is It About? 2
2 Savings, the Balance of Payments, and the Money Supply 30
3 Consumption 58
4 Investment 90
5 The Trade Balance and Exchange Rates 130
6 The Demand for Money 160
7 The Labour Market 202
8 The *IS–LM* Model 234
9 Aggregate Demand and Aggregate Supply 266
10 More on Unemployment 292
11 Open Economy I: The Mundell–Fleming Model 304
12 Stabilization and Government Deficits 326
13 Open Economy II: The Current Account in an Intertemporal Framework 346
14 Financial Sector Reforms 372
15 Monetary Policy Objectives and Targets 400
16 Economic Growth 434

*Bibliography* 469
*Glossary* 473
*Index* 477

# Contents


*MacroFocus Boxes: Tying Theory to the Real World* xx
*Preface* xxi
*The Development of This Book* xxvii
*A Word of Thanks* xxx

Macroeconomics: What Is It About? 2
Critical Questions
1.1 Aggregate Income and Its Dimensions 3
1.1.1 Income and Welfare 4
1.1.2 Omissions in the Measurement of Aggregate Income 5
1.2 Measuring Output 6
⌘ The Underground Economy: Beneficial or Subversive? 7
1.2.1 Connecting Output with Income 9
1.2.2 Aggregate Income Categories 11
1.2.3 Components of Aggregate Expenditure 15
⌘ Chinese and Indian GDP: Re-emerging Economies? 19
1.3 Real and Nominal Incomes 21
1.3.1 Comparing Income Across Time 21
1.3.2 Measures of Inflation 23
Summary 26
Notes 27
Test Yourself 27
Online Application 27
Appendix 1.1 The GDP Deflator 28

Savings, the Balance of Payments, and the Money Supply 30
Critical Questions
2.1 The Sources and Use of Savings 31
2.2 The Balance of Payments 34
⌘ Forex Reserves: Index of Economic Strength? 37
2.3 The Money Supply 39
2.3.1 Bank Balance Sheets and the Money Supply 39
⌘ Does Money Have a Future? 43
2.3.2 The Money Multiplier 45
**2.3.3 Open-Market Operations and Reserves 51
Summary 55
Notes 56
Test Yourself 56
Online Application 56
Appendix 2.1 The Formula for the Money Multiplier 57

Consumption 58
Critical Questions
3.1 Keynes on Consumption 59
3.2 Consumption Smoothing 63
3.2.1 Intertemporal Constraints and Preferences 63
3.2.2 Time Preference and the Permanent Income 67
⌘ Emotions and Intertemporal Decisions 67

3.3 Temporary and Permanent Shocks 71
**3.4 Stochastic Income Expectations 75
⌘ Joint Families and Borrowing Constraints 75
3.5 Effect of Interest Rates 77
3.6 Aggregating Across Individuals 79
**3.7 Savings and Portfolio Choice 81
⌘ House Prices and Consumption 81
Summary 86
Notes 86
Test Yourself 87
Online Application 87
Appendix 3.1 Borrowing Constraints 88

Investment 90
Critical Questions
4.1 Profit Maximization and the Optimal Capital Stock 91
4.2 Adjustment Costs and Investment Decisions 95
**4.3 Financial Structure and Investment 100
⌘ Overconfidence and Investment 100
4.3.1 The Equity and the Value of the Firm 101
4.3.2 Debt and Investment Financing 104
4.3.3 Mature Firms and External Financing 105
4.3.4 Collateral and Liquidity Constraints 105
4.4 Residential and Inventory Investment 107
4.5 Irreversibility and Investment 111
⌘ India: An Inverted Pecking Order? 111
4.6 Investment in Developing Countries 114
4.7 Investment in India 115
Summary 116
Notes 117
Test Yourself 119
Online Application 119
Appendix 4.1 The Optimal Capital Stock When the Production Function Is Cobb–Douglas 120
Appendix 4.2 The Firm's Optimal Decision for Hiring Labour and Accumulating Capital (Investment) 121
Appendix 4.3 The Dynamic Optimization Approach to the Firm's Factor Accumulation Decision 124
Appendix 4.4 Value of the Firm's Equity 128
Appendix 4.5 Tobin's *Q* and the Rate of Interest 129

The Trade Balance and Exchange Rates 130
Critical Questions
5.1 The Real Exchange Rate 131
5.1.1 Real Exchange Rates and Net Exports 132
5.1.2 The Marshall–Lerner Condition 134
5.1.3 The J-Curve 137
5.2 Other Approaches to the Trade Balance 139
5.2.1 Devaluation in Developing Countries 140
⌘ Does Devaluation Raise the Price of Imported Goods? 140
5.2.2 Evidence on Devaluation 141
5.3 Exchange Rates and Assets 142
5.3.1 Uncovered and Covered Interest Parity 142

5.3.2 Exchange Rate Determination 144
5.4 Purchasing Power Parity 145
⌘ Big Macs, iPods, and Professors' Salaries 147
5.5 Choice of Exchange-Rate Regimes 148
⌘ The Euro 151
Summary 153
Notes 154
Test Yourself 155
Online Application 156
Appendix 5.1 The Trade Balance and the Real Exchange Rate 157

The Demand for Money 160
Critical Questions
6.1 Money, Bonds, and Private Wealth 161
6.1.1 Bond Pricing 161
6.1.2 Private Sector Net Wealth 162
6.2 Nominal and Real Interest Rates 164
6.3 Financial Assets and the Budget Constraint 165
6.3.1 The Determinants of the Real Money Demand 168
6.3.2 Parametric Changes on the Money Demand 169
⌘ Stock Markets, Exchange Rates, and the Money Demand 169
6.4 Money as a Store of Value 171
6.5 Seigniorage 175
**6.5.1 The Optimal Level of Seigniorage 176
⌘ Saddam Dinars and Swiss Dinars 178
6.5.2 Alternative Means of Obtaining Seigniorage 179
Summary 181
Notes 182
Test Yourself 183
Online Application 183
Appendix 6.1 Preference Maximizing Choice of Consumption and Real Money Balances 184
Appendix 6.2 The Two-Period Budget Constraint 186
Appendix 6.3 Seigniorage and the Real Money Demand 187
Appendix 6.4 Currency Crises and Speculative Attacks 194
Appendix 6.5 The Inventory Demand for Money (Baumol–Tobin) 198

The Labour Market 202
Critical Questions
7.1 Profit Maximization and Labour Demand 203
7.2 Utility and Labour Supply 204
7.3 Aggregate Supply with/without Money Illusion 207
7.4 Introducing Unemployment 210
⌘ Why Has Wage Inequality Increased? 210
7.4.1 Keynesian Aggregate Supply 211
7.4.2 Friedman on Mistaken Expectations 212
7.5 Cyclical Unemployment and the Output Gap 214
⌘ Unemployment and Wages in India 217
7.6 The Static Phillips Curve 219
**7.7 The Dynamic Phillips Curve 221
Summary 227
Notes 228
Test Yourself 229

Online Application 229
Appendix 7.1 The Mistaken Expectations Aggregate Supply Curve 230
Appendix 7.2 Separability Between Aggregate Demand and Supply 233

The *IS–LM* Model 234
Critical Questions
8.1 Walras' Law 235
8.2 Nominal Versus Real Rate of Interest 236
8.3 The *IS* Curve 237
8.3.1 Parametric Shifts of the *IS* Curve 240
8.3.2 Slope of the *IS* Curve 242
8.4 The *LM* Curve 243
8.4.1 Impact of Varying the Money Supply 245
8.4.2 Slope of the *LM* Curve 245
8.5 *IS* and *LM*—Fiscal and Monetary Policy 247
⌘ Interest Rate Targets and Monetary Policy 247
8.5.1 Crowding Out 249
8.5.2 Fiscalist and Monetarist Intervention 253
⌘ Zero Interest Rates and Macro Policy in Japan 254
8.6 *IS–LM* in India 255
8.7 Ricardian Equivalence 256
Summary 260
Notes 260
Test Yourself 261
Online Application 262
Appendix 8.1 *IS–LM* in Japan 263

Aggregate Demand and Aggregate Supply 266
Critical Questions
9.1 The Aggregate Demand Curve 267
9.2 *AD–AS* Equilibrium in the Keynesian Case 271
⌘ Fiscal Policy and Working Capital 273
9.3 *AD* and the Mistaken Expectations Aggregate Supply 274
9.4 Policy Lags 276
9.5 Real Business Cycles 279
⌘ Nominal Anchors and Intertemporal Policy Constraints 279
Summary 286
Notes 287
Test Yourself 288
Online Application 288
Appendix 9.1 The Aggregate Demand Curve 290

More on Unemployment 292
Critical Questions
10.1 Shirking and Efficiency Wages 293
10.2 The Long-Run and Staggered Wage Setting 296
10.2.1 Staggered Wage Contracts 297
10.2.2 Wage Adjustment 298
⌘ Henry Ford and Efficiency Wages 298
10.3 The Insider–Outsider Approach 299
⌘ Macroeconomic Policy with Insiders 301

Summary 302
Notes 302
Test Yourself 303
Online Application 303

Open Economy I: The Mundell–Fleming Model 304
Critical Questions
11.1 Accounting for an Open Economy 305
11.2 Capital Mobility and the Balance of Payments 307
⌘ Capital Account Convertibility 311
11.3 Monetary Policy under Flexible Exchange Rates 312
11.4 Fiscal Policy under Flexible Exchange Rates 314
11.5 Fixed Exchange Rates and the Monetary Policy 315
11.6 Fixed Exchange Rates and the Fiscal Policy 316
11.7 Capital Flows and Macro Policy 317
⌘ The Pressure on China to Float 318
Summary 321
Notes 322
Test Yourself 323
Online Application 323
Appendix 11.1 Slopes of the *BP* and *LM* Curves 324

Stabilization and Government Deficits 326
Critical Questions
12.1 Limits to Discretionary Policy 327
12.2 The Government Budget Constraint and Debt Dynamics 329
12.2.1 Primary Deficits and Stability 331
⌘ Fiscal Surpluses 332
12.2.2 Primary Surpluses 333
12.3 Debt and Deficits in India 333
12.4 Sustainable Primary Deficits 337
12.5 Fiscal Rules for Adjustment 338
⌘ Are Fiscal Rules a Good Way to Bind Policymakers? 339
12.6 Fiscal Restructuring 340
Summary 344
Notes 344
Test Yourself 345
Online Application 345

Open Economy II: The Current Account in an Intertemporal Framework 346
Critical Questions
13.1 Intertemporal Accounting and the Current Account Determination 347
13.1.1 Intertemporal Accounts of a Nation 347
13.1.2 Intertemporal Budget Constraint in a Two-Period Model 349
13.2 Current Account Determination 352
⌘ Global Imbalances 352
13.2.1 Three Determinants of Borrowing/Lending 354
13.2.2 Current Account and Shocks to the Output 357
13.3 The Current Account in the Medium Run 358
13.3.1 Determination of Expenditure–Output 361

13.3.2 Fiscal Adjustment and the Aggregate Income 363
⌘ Expansionary Fiscal Contractions 364
13.4 The External Debt Burden and Sustainability 365
Summary 369
Notes 370
Test Yourself 370
Online Application 371

Financial Sector Reforms 372
Critical Questions
14.1 Features of Financial Markets 373
14.1.1 Adverse Selection 373
14.1.2 Moral Hazard 374
14.1.3 *Ex post* State Verification 375
14.1.4 Banks and Debt Markets 375
14.1.5 Banks and Capital Markets 376
14.1.6 Functions of Financial Markets 377
14.1.7 Types of Banking Reforms 379
14.2 Financial Repression and Major Financial Reforms in India 379
⌘ Financial Repression—Revenues and Liberalization 381
14.3 The Supply of Loans and the Demand for Deposits 385
14.3.1 Risk and the Supply of Credit 385
14.3.2 The Determination of Banks' Asset Portfolio 388
14.4 Banks' Response to Financial Liberalization 390
⌘ Liberalization, Growth, and Collateral Benefits 393
14.5 Human Capital: A Keystone of Financial Reforms 394
Summary 396
Notes 397
Test Yourself 398
Online Application 398

Monetary Policy Objectives and Targets 400
Critical Questions
15.1 Goals of Monetary Policy and Intermediate Targets 401
⌘ Evolution of the Operating Instruments of the Monetary Policy in India 403
15.2 Choosing Intermediate Targets in the Case of Demand Shocks 404
15.2.1 *IS* Curve Shocks 404
15.2.2 *LM* Curve Shocks 407
15.3 Choosing Intermediate Targets in the Case of Supply Shocks 409
15.4 Targeting the Nominal GDP 411
15.5 Rules Versus Discretion in the Monetary Policy 416
15.5.1 The Social Welfare Function of Policymakers 417
15.5.2 The Private Sector and Central Bank Policy 418
15.5.3 Discretion and Surprise (Unanticipated) Inflation 420
15.5.4 Commitment to a Policy Rule 422
15.5.5 Achieving Credibility 424
⌘ Monetary Policy Rules 424
Summary 426
Notes 426
Test Yourself 427
Online Application 427

Appendix 15.1 Nominal Spending, the Short-Run and the Long-Run Full-Information Supply Curve 429
Appendix 15.2 Derivations of the Equations in Section 15.5 on Rules Versus Discretion 432

Economic Growth 434
Critical Questions
16.1 Sources of Growth 435
⌘ Growth Accounting—China and India 437
**16.2 Potential and Feasible Output 438
16.3 The Determinants of Long-Run Growth 441
16.4 The Solow Diagram 443
16.5 Comparative Statics 445
16.6 The Golden Rule 449
⌘ Calculating the Golden Rule Saving Rate 450
16.7 Convergence 451
16.8 Poverty Traps 453
⌘ Lawyers—Do They Reduce Economic Growth? 455
16.9 Endogenous Growth 456
Summary 460
Notes 461
Test Yourself 462
Online Application 463
Appendix 16.1 Derivations of Important Equations 464

*Bibliography* *469*
*Glossary* *473*
*Index* *477*

# MacroFocus Boxes: Tying Theory to the Real World


| | | |
|---|---|---|
| 1.1 | The Underground Economy: Beneficial or Subversive? | 7 |
| 1.2 | Chinese and Indian GDP: Re-emerging Economies? | 19 |
| 2.1 | Forex Reserves: Index of Economic Strength? | 37 |
| 2.2 | Does Money Have a Future? | 43 |
| 3.1 | Emotions and Intertemporal Decisions | 67 |
| 3.2 | Joint Families and Borrowing Constraints | 75 |
| 3.3 | House Prices and Consumption | 81 |
| 4.1 | Overconfidence and Investment | 100 |
| 4.2 | India: An Inverted Pecking Order? | 111 |
| 5.1 | Does Devaluation Raise the Price of Imported Goods? | 140 |
| 5.2 | Big Macs, iPods, and Professors' Salaries | 147 |
| 5.3 | The Euro | 151 |
| 6.1 | Stock Markets, Exchange Rates, and the Money Demand | 169 |
| 6.2 | Saddam Dinars and Swiss Dinars | 178 |
| 7.1 | Why Has Wage Inequality Increased? | 210 |
| 7.2 | Unemployment and Wages in India | 217 |
| 8.1 | Interest Rate Targets and Monetary Policy | 247 |
| 8.2 | Zero Interest Rates and Macro Policy in Japan | 254 |
| 9.1 | Fiscal Policy and Working Capital | 273 |
| 9.2 | Nominal Anchors and Intertemporal Policy Constraints | 279 |
| 10.1 | Henry Ford and Efficiency Wages | 298 |
| 10.2 | Macroeconomic Policy with Insiders | 301 |
| 11.1 | Capital Account Convertibility | 311 |
| 11.2 | The Pressure on China to Float | 318 |
| 12.1 | Fiscal Surpluses | 332 |
| 12.2 | Are Fiscal Rules a Good Way to Bind Policymakers? | 339 |
| 13.1 | Global Imbalances | 352 |
| 13.2 | Expansionary Fiscal Contractions | 364 |
| 14.1 | Financial Repression—Revenues and Liberalization | 381 |
| 14.2 | Liberalization, Growth, and Collateral Benefits | 393 |
| 15.1 | Evolution of the Operating Instruments of the Monetary Policy in India | 403 |
| 15.2 | Monetary Policy Rules | 424 |
| 16.1 | Growth Accounting—China and India | 437 |
| 16.2 | Calculating the Golden Rule Saving Rate | 450 |
| 16.3 | Lawyers—Do They Reduce Economic Growth? | 455 |

# Preface

It would be naïve to assume that most students taking a course in economics want to go on to specialize in the discipline. Instead, *Macroeconomics* understands that what they will definitely do is participate in the economy and want to understand how policy influences their economic lives.

Thus, the focus of this textbook rests, first and foremost, on the analysis of macroeconomic thought in terms of the intuition and underlying logic that forms its basis.

This book has been designed to help students think independently about real-world situations, by helping them master the basic technical tools that enable them to do this. Therefore, macroeconomics is discussed in the form in which researchers, policy watchers, and policymakers think about the economy. This is just a part of the unique approach of this book, which sets it apart.

## Introducing Students to Foundational Issues

Most macroeconomics texts are based on the presumption that students will feel intimidated if they are required to grasp the methodologies essential for developing an understanding of the basic concepts. *Macroeconomics*, on the other hand, bases itself on the belief that students who are first introduced to the economic intuition behind a fundamental concept are motivated to think more deeply about the rigour that underlies the intuition.

## Providing Insight into a Real-World Economy

The techniques and tools employed in this book are not discussed as ends in themselves. Books that have the sole aim of being "accessible" are forced to be merely descriptive, and gloss over many real-world complications. This book, on the other hand, understands that students are prepared to "dirty their hands," with technique, provided it is not overwhelming, which helps them relate to the practical world.

## Focusing on Current Issues

Over the years, macroeconomics has progressed from one controversy and school of thought to another, and attached itself to school after school of thought. Over time, however, some consensus has begun to emerge about the methods of the discipline. This emphasizes thinking about macroeconomics as involving the interactions between aggregate markets. These are where choices are made in the knowledge that they result in current as well as future consequences—what economists refer to as intertemporal models. *Macroeconomics* maintains focus on the most current and relevant issues, while also understanding the fluidity of the subject.

## Organization

The unique strengths of *Macroeconomics* are highlighted through a conceptual framework, which organizes the chapters into five major modules. The following flowchart presents an outline of the topics within the structure of the book. The topics addressed in each of the chapters are as follows:

### Part I: Definition and Measurement of Aggregate Variables

Macroeconomics studies a broad aggregation of markets in an attempt to explain what happens in the economy as a whole rather than what goes on in an individual market. The first building block in macroeconomics, thus, involves the definition and measurement of aggregate variables.

Chapters 1 and 2 define and measure aggregate measures of income, output, inflation, savings, the balance of payments, and money supply in the economy. These chapters also examine what the data reveal about these variables and familiarize readers with the substance that macroeconomics deals in.

Macroeconomics, at its core, is concerned with the interaction between choices made in four aggregate markets—the market for goods, or the demand for output; the money market; the bonds market; and the market for labour services, which enables the production of the output. Insights into these foundations of macroeconomics are presented in a staggered manner in Parts II and III.

**1. DEFINITION & MEASUREMENT OF AGGREGATE VARIABLES**

**CHAPTER 1.**
Aggregate Income, Output and Inflation

**CHAPTER 2.**
Aggregate Savings, Balance of Payments and Money Supply

**2. AGGREGATE MARKETS IN MACROECONOMICS**

**GOODS MARKET**

**CHAPTER 3.**
Consumption

**CHAPTER 4.**
Investment

**CHAPTER 5.**
Foreign Trade

**FINANCIAL MARKET**

**CHAPTER 6.**
Money and Bond Markets

**LABOUR MARKET**

**CHAPTER 7.**
Labour Market

**3. SHORT-RUN MACROECONOMICS**

**CHAPTER 8.**
Goods & Financial Market Equilibrium

**CHAPTER 9.**
Aggregate Demand & Aggregate Supply

**CHAPTER 11.**
Open Economy Extensions

**4. MACROECONOMIC POLICY & ISSUES**

**CHAPTER 10.**
Unemployment

**CHAPTER 12.**
Deficits & Adjustments

**CHAPTER 13.**
The Current Account

**CHAPTER 14.**
Financial Sector Liberalization

**CHAPTER 15.**
Monetary Policy Targets & Objectives

**5. LONG-RUN MACROECONOMICS**

**CHAPTER 16.**
Economic Growth

### Part II: Aggregate Markets in Macroeconomics

Part II examines how choices are made, given the opportunities available in each of the aggregate markets. Having introduced readers to some of the major terms in macroeconomics in Part I, Part II develops a conceptual framework that provides them with insights into the determinants of the aggregate measures central to macroeconomics. Chapters 3 to 7 deal with the foundations of macroeconomics, and are, thus, the building blocks of the subject.

In macroeconomics, the expenditure on goods is partitioned into expenditure by individuals on consumption and expenditure by businesses on investment. (There is also public expenditure by a government which macroeconomists, in a first cut, take as given or exogenous.) Chapters 3 and 4 focus on the determinants of consumption and investment expenditure. Some expenditure is incurred by domestic residents on goods produced abroad (imports). Also, some goods produced domestically are sold abroad (exports). Chapter 5 examines the determinants of these expenditures. Subsequently, Chapter 6 discusses the financial market—the markets for money and bonds—and Chapter 7 deals with the market for labour.

### Part III: Short-Run Macroeconomics

Part III discusses how the interactions of the markets discussed in Part II result in aggregate outcomes. It combines the goods and financial markets to generate a framework for considering how the demand for output and the interest rate are determined in the short run—a period within which the economy's capacity is underutilized and where the price level does not vary. This is the famous workhorse *IS–LM* model of macroeconomics, which is introduced in Chapter 8. The immediate short-run impacts of macroeconomic policy, such as fiscal and monetary policy, are analysed in this chapter.

The *IS–LM* model ignores the supply of output. Therefore, we need to integrate the aggregate supply relationship with the aggregate demand relationship implicit in the *IS–LM* model. This is the basic aggregate demand–aggregate supply model or *AD–AS* model, which determines the price level and aggregate output in the economy and, by implication, the level of employment, interest rate, and (for some versions of macroeconomics) the nominal wage rate as well. This *AD–AS* model, discussed in Chapter 9, is very useful when confronted with a macroeconomic question.

Given that economies are increasingly integrating themselves with the global economy, Chapter 11 extends the basic approach of Chapter 8 to an open economy. The openness of an economy is not just a matter of trading in goods and services—the extent to which an economy that is closed to capital flows affects macroeconomic outcomes is discussed in Chapter 11. Moreover, the way different exchange rate regimes—fixed versus flexible—mediate policy changes in terms of macro outcomes is also shown to be an important insight that the closed-economy approach does not offer. Therefore, a basic macroeconomic course would require a student to master Chapters 1 to 9 and Chapter 11.

### Part IV: Macroeconomic Policy and Issues

Part IV deals with additional issues that fall within the realm of macroeconomic policy. Chapters 10, and 12 to 15, deal with contemporary macroeconomic-policy issues. These involve questions such as: should governments be put on a leash as far as deficits and the incurring of debt go? What are the impacts of financial liberalization? Is unemployment the result of employer policy, or can workers collude to exclude others from the labour market? Is a reining-in of government expenditure necessarily contractionary? What should the operating instruments and targets of monetary policy be?

### Part V: Long-Run Macroeconomics

One of the central issues in macroeconomics concerns how policy can impact the cyclical fluctuations of the economy that result in variations in output, employment, and prices. Another central issue is concerned with the questions of why some economies, over long periods, grow faster than others. This is the subject matter of economic growth, dealt with in Chapter 16.

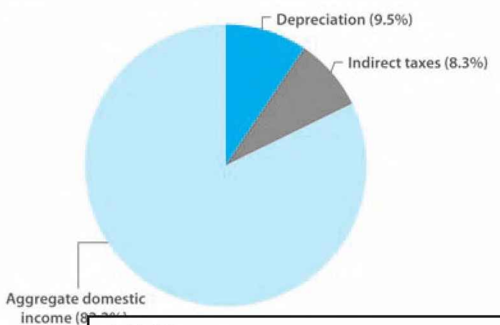
## Features

Each chapter includes some key features aimed at helping readers gain an in-depth understanding of the concepts discussed in the chapter.

*Learning Goals:* The learning goals of each chapter are stated upfront in terms of Critical Questions.

CRITICAL QUESTIONS

- » *What is the cost of employing factor inputs such as physical capital and labour for a firm?*
- » *How does a firm select the optimal stock of capital?*
- » *What is Tobin's marginal Q?*
- » *How does the cost of external finance affect the investment decision?*
- » *What are the determinants of investment in housing and in inventory?*
- » *How does the irreversibility of investment affect the decision to invest now rather than to wait and invest later?*
- » *What are the other circumstances applicable to developing countries that affect investment*




› **Figure 1.2**
Components of GDP at Market Prices. The aggregate domestic income earned from the production of output within the nation is the major component of measured GDP.
This figure is based on Table 1.1.
*Source:* Central Statistical Organisation, Government of India, National Accounts Statistics, 2006.

*Real-world Validation:* Each chapter highlights the salience of the results of the conceptual framework developed in terms of whether the economy's behaviour matches the way theory depicts it. This makes macroeconomics about theory in the real world. Graphs and tables are used throughout the book to explain each concept and its application as well.

› **Table 1.7**
India's Foreign Trade (INR Billion)
*Source:* Reserve Bank of India, *Handbook of Statistics on Indian Economy.*

| | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|---|
| **Exports** | 2078.52 | 2133.45 | 2600.79 | 3039.15 | 3817.85 | 4657.05 |
| Agriculture | 272.88 | 281.44 | 324.73 | 346.16 | 380.78 | 452.20 |
| Gems and Jewellery | 337.33 | 348.45 | 437.01 | 485.86 | 618.34 | 687.53 |
| Chemicals and Related Products | 268.89 | 288.62 | 360.80 | 434.05 | 559.11 | 653.90 |
| Engineering Goods | 311.50 | 331.83 | 437.15 | 570.05 | 779.49 | 961.57 |
| Textiles | 515.55 | 486.77 | 562.21 | 587.79 | 609.06 | 726.18 |
| **Imports** | 2645.89 | 2683.00 | 3117.76 | 3673.01 | 5335.50 | 6951.31 |
| Petrol, Oil, and Lubricants | 714.97 | 667.70 | 853.67 | 945.20 | 1340.94 | 1946.40 |
| Capital Goods | 408.47 | 471.30 | 653.25 | 839.94 | 1129.36 | 1667.61 |
| Gold and Silver | 211.89 | 218.54 | 207.53 | 315.06 | 500.99 | 501.08 |
| Chemicals—Organic and Inorganic | 111.65 | 133.52 | 146.40 | 185.27 | 256.10 | 309.21 |
| Pearls, Precious and Semi-precious Stones | 219.64 | 220.46 | 293.41 | 327.57 | 423.38 | 404.41 |
| **Trade Balance** | −567.37 | −549.55 | −516.97 | −633.86 | −1517.65 | −2294.26 |
| Invisibles (Net) – of which | 451.39 | 713.81 | 823.57 | 1273.69 | 1395.91 | 1887.04 |
| Foreign Travel | 79.05 | 158.89 | 176.44 | 65.20 | 62.87 | 61.98 |
| Transport | −68.08 | −61.60 | −35.67 | 40.26 | 6.58 | −68.72 |
| Insurance | 2.30 | 0.35 | 0.96 | 2.50 | 6.64 | 0.69 |
| Software Exports (Net) | 263.08 | 328.36 | 428.59 | 566.06 | 758.25 | 986.78 |
| Net Factor Incomes from Abroad | −227.33 | −200.68 | −166.90 | −182.50 | −223.75 | −249.69 |
| Transfers (Net) | 599.67 | 755.60 | 814.03 | 1016.96 | 931.35 | 1076.73 |
| **Current Account Balance** | −115.98 | 164.26 | 306.60 | 639.83 | −121.74 | −407.22 |

*Notes:* (1) Only major components of exports, imports, and invisibles have been depicted. The list is not exhaustive.
(2) Chemicals and pearls, precious and semi-precious stones are mainly for further processing and re-export.

*Recent Research and Data:* Data from government organizations like the Central Statistical Organisation of India and the Reserve Bank of India are used to illustrate and clarify concepts.

› Money's most important function is as a medium of exchange, when it is used to purchase goods and services.

› As a unit of accounting, money is the yardstick for measuring economic value. It measures the value of goods, services and assets, and thereby enables a comparison of these based on their values.

› As a store of value, money is one way of holding wealth.

operating as a MEDIUM OF EXCHANGE—a device for making transactions—money allows people to trade with less time and effort.

Besides, money is also a unit of accounting—the basic unit that is used to assess economic value. In India, for instance, all prices, wages, asset values, debts, and financial contracts are expressed in INR. A single and uniform measure of value makes it straightforward to compare different goods and assets. As goods and services are exchanged for money—the medium of exchange function—it is also natural to express economic value in terms of money—as a UNIT OF ACCOUNTING FUNCTION.[1]

The third function of money is as a STORE OF VALUE. Money can be used to buy goods not only today but also tomorrow and so it embodies and stands for future purchasing power. Money can, thus, be used as a store of value or a way of holding wealth. However, there are other assets—stocks, bonds, gold, real estate—that can be used as a store of value and they typically outperform money in this role as they normally yield a higher rate of return than money. The distinguishing feature of money is, therefore, its role as the medium of

*Margin Notes:* The marginalia, or margin notes, along the text recreate some of the student–teacher interactions that take place in the classroom and provide material that is complementary to the matter contained in the text. These notes are also used to summarize crucial points and highlight definitions.

*Summary:* Each chapter ends with a summary of the main topics covered in the chapter.

SUMMARY

» Firms pay higher than the market-clearing wages to elicit effort, reduce turnover, recruit better workers, and improve worker morale.
» When firms have difficulty monitoring employees, they have an incentive to offer a wage–employment package that deters workers from shirking on the job.
» Efficiency wages are the wages that maximize firm profits when it is difficult to monitor whether workers work or shirk.
» Efficiency wages are higher than the competitive market wage and generate involuntary unemployment.
» The New Keynesian natural rate of unemployment is an equilibrium labour market outcome where wages deter shirking behaviour, firms maximize profits, and there is an excess supply of labour.
» The New Keynesians argue that nominal wage setting is staggered so that wages are sticky in the short run.
» Insiders are incumbent employees who have market power in the determination of wages. This market power arises from insiders' ability to exploit turnover costs of hiring and firing in the wage bargain compared to outsiders.
» Insiders can also reduce labour productivity by refusing to cooperate with outsiders who attempt to underbid prevailing wages. Firms then find outsiders unattractive to hire. Harassment by insiders also raises the opportunity cost of employment and deters outsiders from offering their labour services.
» Established firms setting wages on an insider–outsider basis are susceptible to competition from new firms that have no insiders and can employ workers at lower reservation wages.
» The rate of unemployment depends on the barriers to entry faced by new firms in the economy when insider effects prevail in the labour market.

MACROFOCUS 1.2

**Chinese and Indian GDP: Re-emerging Economies?**

Definitions of emerging economies vary. The United Nations counts Hong Kong, Singapore, South Korea, and Taiwan as emerging economies. The International Monetary Fund, however, counts all four as "developing economies" in its International Financial Statistics but as "advanced economies" in its World Economic Outlook. In policy discussions, these are labelled as emerging economies. By 2005, the emerging economies accounted for about 30 per cent of the world GDP and have begun to dominate the global economy in terms of exports, energy consumption, and foreign-exchange reserve holdings. Currently, two fast growing emerging market economies are China and India which together contributed to about 25 per cent of the increase in emerging market economies' GDP in 2005.

In a study of the world economy the economic historian, Angus Maddison* found that until the late 19th century, China and India were the world's two biggest economies. These two economies dominated the world output before the steam engine and the industrial revolution placed Britain up in front. Maddison estimates that in the eighteen centuries up to 1820, the current emerging market economies produced on average a staggering 80 per cent of the world's GDP. After a gap of about 180 years they are now back with amongst the fastest growth rates of GDP.

Given this long perspective it may be more appropriate to label these economies as re-emerging economies. Two important features of economies such as China, Hong Kong, Indonesia, Malaysia, Oman, Malta, Singapore, and Thailand that have seen sustained high increases in the GDP in the past couple of decades have been (1) the high levels of investment—attempting high growth without investment in infrastructure and skills is a sure recipe to ignite inflation rather than increase the output of goods and services; and (2) the large fraction of GDP associated with exports and imports—an indicator of a country's ability to leverage the demand and resources of the global economy. In 2004 for instance, China's total investment expenditure was 43.2 per cent of its GDP compared with India's 27.2 per cent. Similarly, international trade (exports plus imports) as a percentage of GDP was 58.2 per cent in China whereas it was 27.4 per cent in India.

*Angus Maddison, *The World Economies: A Millennial Perspective*, OECD, 2001.

*MacroFocus Boxes:* Each chapter has at least a couple of boxed items that discuss applications, present curious and interesting macroeconomic details, or indicate the areas of current research in the discipline.

*Levels of Difficulty:* Starred sections indicating a relatively difficult or advanced topic are provided in chapters for the eager student. These may be skipped during the first read without losing the sense of continuity in the text.

*Test-Yourself Questions:* Test-yourself questions at the end of each chapter are designed to help students gauge their understanding of the concepts discussed in the chapter.

TEST YOURSELF

1. Distinguish between automatic and discretionary fiscal policy.
2. Define the intertemporal budget constraint of the government. How is it related to the sustainability of the debt?
3. Why is the level of debt not as suitable a measure as the debt/GDP ratio of a country's debt obligations?
4. What is the difference between the fiscal deficit and the primary deficit? Which is a more appropriate measure of the deficit of the government and why?
5. Define a steady state. When is a steady state stable and when is it unstable? What condition must be satisfied to ensure that the debt/GDP ratio results in a stable steady state?
6. India is said to be experiencing a "demographic dividend" in that a large section of the population is in the labour force. What is the likely impact of this on future deficits?
7. In the case where the debt/GDP ratio is on an unstable steady-state path what should the government do to reduce the debt?
8. Discuss the deficit and debt scenario in India between 1980–1981 and 2002–2003. In what way would FRBMA alter the debt and deficit position of the government?
9. What are fiscal rules and how effective are they in restraining the growth of government debt?
10. What correlated macroeconomic policy is required to sustain a fiscal adjustment?

*Online Application Questions:* Students are exposed to the available databases on the economy and online applications of the tools and techniques are indicated.

**ONLINE APPLICATION**

1. a. Go to the home page of the Reserve Bank of India URL: www.rbi.org.in/home.aspx.
   b. Click on the "DataBase" icon.
2. Then under the classification of annual data, click on "Handbook of Statistics on Indian Economy"
3. Click on "Table 185: Components of the Money Stock" and create a spreadsheet of monthly data on narrow money (M1) for the period April 2003 to June 2006.
4. Then click on "Table 182: Wholesale Price Index—Monthly Average" and in your spreadsheet insert monthly data on the wholesale price index from April 2003 to June 2006.
5. The wholesale prices are in the form of indices. To create an indexed data series of real money balances, index the data you inserted in Step 3. For example, let the M1 data for April 2003, INR 4933.33 billion, equal the index value of 1100. In that case the index for May 2003 for INR 4955.35 billion is 1105.
6. Divide the index for M1 by the price index of the Wholesale Price Index in that month. For instance, in April 2003, the WPI index is 173.1. Thus, the index for real money balances in April 2003 is (1100/173.1) = 6.35.
7. Next go back to the RBI Handbook of Statistics on Indian Economy. Click on "Table 192: Weighted Average Call Money Rates". Take the data for the corresponding period April 2003 to June 2006 on the average call money interest rate each month.
8. Create a graph depicting the real money balance index on the horizontal axis and the interest rate on the vertical axis. Does the real money balance vary inversely with the interest rate?
9. You would have noticed that there is no clear-cut relationship between the money balances and the interest rate. In fact, if you fit a trend to the series of dots in your graph obtained in Step 9 you would obtain a graph like that in Figure 8.21, where initially there is a negative relationship followed by a positive and then again negative relationship indicating a tilted *S* curve. Also, if you create a graph between the change in the interest rate in a year and the change in the real money balances in that year and fit a linear trend to that graph, you would observe a mildly negative relationship that seems to indicate a negative relationship between changes in money balances and changes in the interest rate.
10. Give reasons for why the relationship between real money balances and the interest rate may not be strictly negative.

## Appendix 5.1 The Trade Balance and the Real Exchange Rate

**Section A: Imports and the Real Exchange Rate**

Suppose that the demand for imported goods is inversely related to the real exchange rate and this relationship is linear, or

$$Q_M = \mu_0 - \mu_2 q \qquad \text{(A5.1.1)}$$

where $\mu_0$ and $\mu_2$ are positive constants. Then, the demand for imported goods measured in terms of domestic goods or imports is equal to

$$qQ_M = \mu_0 q - \mu_2 q^2 \qquad \text{(A5.1.2)}$$

This is a quadratic equation in the real exchange rate which results in an inverted U-shaped curve as depicted in Figure 5.2.

The maximum value of $qQ_M$ is attained where

$$\frac{\partial(qQ_M)}{\partial q} = \mu_0 - 2\mu_2 q = 0,$$

or

$$q = \frac{\mu_0}{2\mu_2}$$

Substituting this into Eq. (A5.1.1),

$$Q_M = \mu_0 - \mu_2\left(\frac{\mu_0}{2\mu_2}\right) = \frac{\mu_0}{2} \qquad \text{(A5.1.3)}$$

*Appendices:* Mathematical derivations for each chapter are provided in the appendices at the end of the chapter so that the student can become well-versed in the theory behind the concepts.

## The Teaching and Learning Package

A full range of resources that support teaching and learning is available with this book. These include:

- An *Instructors' Resource Manual* that provides an introduction to the major concepts of each chapter, definitions and background for each concept, application of data, analytical problems to be offered to students, and related issues that can be discussed in the class.
- A *companion Web site* that contains chapter-wise summary and convenient links where instructors can access and download the instructors' manual and lecture slides.
- *Media resources* including detailed step-by-step explanations of key topics through effective animation.
- *PowerPoint lecture slides* that provide an animated set of outlines, summaries, graphs, and equations for each chapter.

## A Note on the Language and Terms Used in the Book

- We have used the ISO 4217 code for the Indian currency—*INR*—instead of *Rs* throughout the book.
- The terms *million* and *billion* have been used to represent large numbers, and the use of *lakh* and *crore*, although common in Indian English, have been avoided.
- In the interest of gender equality, we have avoided using terms like *he* and *she* when the pronoun is indefinite. Instead, we have used *they*, *their*, and *them* in the singular sense.

# The Development of This Book

*Macroeconomics* has benefited from an extensive development process. Over 150 faculty reviewers, students, and industry professionals provided feedback about the accuracy and relevance of the content as well as suggestions for its improvements. While we could not incorporate every suggestion from everyone, we do acknowledge that their feedback was invaluable in our attempt at creating the best possible macroeconomics textbook.

## Consultant Board

The consultant board provided us with a detailed and critical analysis of each chapter and worked with us throughout the development of the book. We would like to thank the following for their time and commitment:

Sujoy Chakravarty
*Indian Institute of Technology Delhi*

Ajitava Raychaudhuri
*Jadavpur University, Kolkata*

V. K. Seth
*Faculty of Management Studies, Delhi*

S. N. V. Sivakumar
*K. J. Somaiya Institute of Management Studies and Research, Mumbai*

## Supplement Authors

T. J. Joseph of the ICFAI Institute of Management Teachers and C. S. Shylajan of the ICFAI Business School prepared the Instructor's Manual for *Macroeconomics*. This supplement includes introduction to the major concepts of each chapter, definitions and background for each concept, application of data, and analytical problems that can be discussed in the class.

## Student Reviewers

We took the help of many students who class-tested the manuscript, evaluated it for clarity, and assessed each feature. Their comments helped us expand the book's content, improved the pedagogical features, and strengthened the assessment features. We are thankful to the following:

Sandip Agarwal
*University of Delhi*

Richa Bhardwaj
*University of Delhi*

Yajnaseni Das
*Jadavpur University*

Yasodhara Das
*University of Calcutta*

Kunal Singh Dhaliwal
*Rabindra Bharati University*

Chinmayee Gopal
*Delhi School of Economics*

Sheikh Azeem Hussain
*Jadavpur University*

Ishita Mohanty
*London School of Economics*

Amiya Sharma
*Delhi School of Economics*

Ashish Tyagi
*Jawaharlal Nehru University*

## Reviewers

The guidance and thoughtful recommendations of many helped us improve this book. We are grateful for the comments and helpful suggestions received from the following reviewers:

Debashis Acharya
*Indian Institute of Technology Chennai*

Sunil Ashra
*Management Development Institute, Gurgaon*

Benni B. S.
*Sinhgad Institute of Management, Pune*

Sanjaya Baru
*Prime Minister's Office, New Delhi*

Bimal Beri
*Narsee Monjee Institute of Management Studies, Mumbai*

Laveesh Bhandari
*Indicus Analytics*

Manisha Chakrabarty
*Indian Institute of Management Calcutta*

Harnita Choudhary
*Symbiosis Institute of Management Studies, Pune*

Suma Damodaran
*XLRI, Jamshedpur*

Mrutyunjay Dash
*Asian School of Business Management, Bhubaneswar*

V. Renuka Devi
*Ethiraj College for Women, Chennai*

A. C. Fernando
*Loyola Institute of Business Administration, Chennai*

Mridula Goel
*Jaipuria Institute of Management, Lucknow*

C. Kanagaraj
*Happy Valley Business School, Coimbatore*

Raj Kumar
*Maharaja Agrasen Institute of Management and Technology, Jagadhri*

Mihir Mahapatra
*Goa Institute of Management*

Sunil Mani
*Centre for Development Studies, Thiruvananthapuram*

B. Deevena Margaret
*Yogi Vemana University, Kadapa*

Meeta K. Mehra
*Jawaharlal Nehru University, New Delhi*

K. S. S. Uduman Mohideen
*University of Madras*

A. G. Moss
*OU College for Women, Hyderabad*

Jaydeep Mukherjee
*Xavier Institute of Management, Bhubaneswar*

Paramita Mukherjee
*The ICFAI Business School, Kolkata*

Ram Ji Narayan
*PGDAV College, New Delhi*

Amar Nath
*National Institute of Public Finance and Policy, New Delhi*

P. C. Padhan
*XLRI, Jamshedpur*

K. C. Padhy
*Institute of Management and Information Sciences, Bhubaneswar*

Pranab. K. Pani
*Indian Institute of Management Indore*

G. A. Prabhakar
*GVP College for PG Studies, Visakhapatnam*

C. Bhujanga Rao
*Madras School of Economics*

D. Tripati Rao
*Indian Institute of Management Lucknow*

S. N. Sahdeo
*Birla Institute of Technology, Ranchi*

Runa Sarkar
*Indian Institute of Technology Kanpur*

Prabal Kr. Sen
*XLRI, Jamshedpur*

Rudra Sensharma
*Indian Institute of Management Lucknow*

Anup K. Sinha
*Indian Institute of Management Calcutta*

Kumudani Sinha
*Patna University*

T. H. Thangiah
*International School of Business and Media, Pune*

I. Thiagarajan
*Loyola Institute of Business Administration, Chennai*

N. S. Viswanath
*Bharatiya Vidya Bhavan, Bangalore*

# A Word of Thanks

This book began as lectures delivered at the University of Mumbai and the Indian Institute of Management Ahmedabad. I have always maintained that the quality of the output of institutions of higher learning is directly attributable to the quality of the students (that determines the extent and the depth of the disciplinary ground that can be covered), the knowledge embodied in the teacher (that accumulates as the teacher engages in research and not just teaching), and the time and effort devoted to teaching by the teacher. Over the years, many students have made critical observations whenever they spotted an error, an inconsistency, or an ambiguity in my presentation of the subject matter. I have been fortunate to have this intervention from their side. I would like to thank in particular my doctoral students Vanditaa Dar, Anjani Sarma, and Vineet Virmani.

Over the years, I have also had the good fortune to obtain the support of the following academic associates in my teaching—Indrani Manna, Swapan Chakraborty, Poonam Arora, Dipti Lalit, Suparna Banerjee, Vivek Pandey, and Renu Pothen.

The writing of this book was made easier by the able secretarial assistance and word-processing prowess of Reena Narendran and Sarala Nair. I am grateful to the Research and Publications Committee at the Indian Institute of Management Ahmedabad for providing me with a grant. The credit for the photograph goes to Sheetal Bhalla.

The editorial and development support that I received from Amarjyoti Dutta and Jonaki Ray at Pearson Education made the task of completing this manuscript a pleasurable one.

A part of this book was completed in the summer of 2007, which I spent at the Institute of South Asian Studies, National University of Singapore. The hospitality of the Institute faculty and staff, the Director, Tan Tai Yong, and the Head of Research, S. Narayan, provided a conducive environment in which I could work on parts of this book.

Over the years, I have benefited from seminars on macroeconomics and discussions with numerous colleagues, which have influenced my thinking on the discipline. I must state the disclaimer that some of them may not recognize and may indeed disagree with some of the writing they see in this text. I would like to acknowledge my debt to Amit Bhaduri, the late Krishna Bhardwaj, Shubham Chaudhuri, Vikas Chitre, Romar Correa, Devesh Kapur, Satish Jain, Shubhashis Gangopadhyay, Subir Gokarn, Arun Kumar, Sugata Marjit, Anjan Mukherji, Dilip Nachane, Avadhoot Nadkarni, R. Nagaraj, Mihir Rakshit, M. Ramachandran, Ajit Ranade, the late M. J. Manohar Rao, Abhijit Sen, Pushpa Trivedi, and John Williamson. I am especially grateful to Partha Sen for his remarks on the exposition in parts of the book.

Errol D'Souza

# 1

# Macroeconomics: What Is It About?

CRITICAL QUESTIONS

- » *What is macroeconomics and what are its distinguishing features?*
- » *What is gross domestic product (GDP) and how is it measured?*
- » *How is the aggregate output related to the aggregate income and expenditure?*
- » *How do international transactions affect the measurement of the GDP?*
- » *What is the difference between the nominal GDP and real GDP?*
- » *What is the difference between price indices such as the producer price index, the consumer price index, and the GDP deflator?*

## 1.1 Aggregate Income and Its Dimensions

Why do some countries have higher incomes than others? Why are the incomes of some countries growing faster than others? How is income related to employment, wages, prices, and interest rates? How do government policies affect incomes? Are higher incomes sure to make us better off?

These are the questions that macroeconomists try to answer, and one of the ways they do it is by studying income. The income that macroeconomists focus on is aggregate income—the income of an entire country. Macroeconomists study aggregate income using the same methods by which economists study everything else—through an examination of how people make choices. In what follows, the emphasis will be on INTERTEMPORAL CHOICES—choices that have consequences across more than one time period. This stress on intertemporal choices is because major decisions, such as, how much to work, how much to save, and how much to consume are heavily influenced by our mental images of the future. The choice to work less or consume more today necessarily involves a choice to work more or consume less in the future. The choice about how much to work or consume as the result of a decision of today has both present and future consequences. As a result, today's decisions will be made with an eye to the future.

› Intertemporal choices are choices that take into consideration the consequences for the current time period as well as those for future periods.

Choices in an intertemporal setting are, of course, from an economic point of view, made in the context of markets. In MACROECONOMICS, we are concerned with the choices related to four markets.

› **Macroeconomics** is concerned with choices in an intertemporal setting within four markets—the markets for goods, for bonds, for labour, and for money.

The first market is the goods or output market—the market for goods available for immediate use, or, current goods. There are, of course, separate markets for textbooks, iron and steel, shoes, cars and clothes. The interactions between these separate markets are the subject matter of microeconomics. Macroeconomics, on the other hand, aggregates all these markets into a single market for goods, and uses a single unit in which all of these goods can be measured.

The second market studied by macroeconomics is the market for future goods or equivalently the market for bonds. A bond is a promise to deliver commodities at some specified time in the future. The bond market allows goods that will be produced in the future to be traded for goods currently produced. If we purchase a bond for one Indian rupee or INR 1, and get paid back a sum, inclusive of interest, amounting to INR 1.17 when we redeem the bond after a year, then the interest on the bond is 17 per cent per annum. At the same time, if the price of a good like an apple changed from INR 1 to 1.05 between the time of the bond's purchase and its redemption, then the bond can equivalently be thought of as exchanging an apple today for 1.17/1.05 or 1.11 apples tomorrow.

For goods to be exchanged, they must be produced and that requires factors of production. Macroeconomics, therefore, denotes the third market to be the market for a factor of production—the labour market. In this market, the availability of goods changes as the amount of labour utilized in production is varied.

Finally, in order to exchange goods, individuals need an asset that can be used to purchase and sell things—money. In economics, assets that are actually used to purchase goods are referred to as money. Such assets include coins or pieces of paper that bear pictures of national leaders.[1] The final market that macroeconomics deals with is, accordingly, the market for money or the market for the holding and provision of money balances.

Macroeconomics, thus, attempts to understand the markets for goods, bonds, labour, and money, and the way in which they influence each other. Agents making intertemporal choices in these four interconnected markets then become the natural subject domain of macroeconomics.

Before examining these four markets, we need to get familiar with some of the jargon used by macroeconomists. Let us begin with income.

› The aggregate income is the sum of all incomes accruing to the citizens of a country within a given period of time—a year or a quarter.

› The per capita income is aggregate income divided by the population of the country.

In a sense, we all know what this is. Individuals work or own small or big businesses that fetch them an earning, which is deployed for expenses on food, shelter, clothing, and education among other items of expenditure. Included in their income would also be the earning on interest or dividends, or rents from savings, or holdings of equity stock, or housing stock. The sources of income are thus varied—wages and salaries, rental income, interest income, and profits from business activity. The AGGREGATE INCOME of a country is the sum of all the incomes that accrue to all its citizens within a given period of time—a year or a quarter. The sum total of all incomes or aggregate income divided by the population of the country is the PER CAPITA INCOME.

### *1.1.1* Income and Welfare

Another reason economists are interested in income is because it is an indicator of the welfare of the citizens of a country. The notion is that with greater incomes people are better off—they can live more comfortably and afford more goods and also feel more secure as they can save more for the future. Economists are actually more interested in welfare, or rather happiness as measured in terms of welfare, and would like to increase it. Unfortunately, it is not easy to measure, and most economists measure income as a rough approximation of welfare. But is income really a true measure of welfare?

While Robert Redford seemed to think so in the film *Indecent Proposal*, the fact is that money can't buy love, as declared by the Beatles. Still, greater income can get people a better education and memberships to clubs, and access to many other things that improve their welfare. Many economists also argue that improvements in the quality of life—lower crime rates, cleaner air, less congestion, more parks, less violence against women—matter more than incomes. Macroeconomists would only point out that higher aggregate incomes are surely of some relevance when we tackle each of these issues.

However, the correlation between income and happiness is not certain. As economists like Lane[2] have shown recently, giving money to people below the poverty line makes them happy. But above the poverty line, this relationship between the level of income and subjective well-being is weak. The rich are no more satisfied with their lives than the merely comfortable. They, in turn, are only slightly, if at all, more satisfied with their lives than the lower middle classes. Lane found that it is people—family solidarity, social inclusion, and warm friendships—and not money that makes people happy.

It is also important to recognize that an increase in aggregate income is not the only way to solve many social and economic problems. It is not true that all increases in income make individuals better off because there is no perfect correspondence between income and welfare. For instance, the government could cause aggregate income to increase by mandating that everyone work for longer hours and also work harder. The retirement age could also be relaxed and the increased work life may contribute to increased incomes but it will not make people better off.

Again, aggregate income is a deceptive measure of welfare when inappropriate goods are produced. The Soviet Union in the 1980s, for instance, had a measured income that was relatively high because of the production of all kinds of machinery and military equipment that were not of much good for average citizens. For the average citizen, another warplane or machine was not of much value when bread was scarce—the only way to buy it was lining up at the bakery every morning. Similarly, greater income was meaningless when married couples and their children lived with their grandparents in two-room apartments and were not allowed to move elsewhere.

Finally, macroeconomists do not pass judgments on important ethical problems involving aggregate income such as whether a country where everyone has exactly the same income is better off than a country where some people have extremely high incomes and everyone else has an extremely meagre income.

These omissions, however, do not imply that there is no merit in studying the determinants of aggregate income. A caricature of macroeconomists is that they believe that the worth of a person is revealed when he dies and goes to heaven where Saint Peter stops him at the gates and asks what his contribution to aggregate income has been so that he may decide whether or not to let him in. Of course, income is not identical to welfare, but it is worth studying because it is a gauge, however imperfect, by which macroeconomists can evaluate whether the welfare or well-being of citizens has been enhanced.

### *1.1.2* Omissions in the Measurement of Aggregate Income

A country can find out the aggregate income of its citizens by requiring them to fill out forms detailing their income. Since this is time consuming and people would surely vote out a government that required them to spend valuable time reporting their income to a bureaucrat each day, most countries measure income on an annual basis. Although, when individuals report their income received during the previous year to the tax authorities, the resulting aggregate misses significant sources of a country's income.

An important omission in measures of aggregate income includes underreported income. Business people often overstate their expenses and understate reported incomes; teachers often do not report earnings from tuition classes; and secretaries often take in an extra income by typing evenings on the side. All these individuals understate their incomes so as not to pay taxes and work with an accounting system that leaves no paper trails so that the government finds it difficult to track their true incomes.

In addition, there are all sorts of underground activities that people are engaged in, such as drug dealing and gambling where they do not report their incomes. Telgi allegedly selling fake stamp paper or Dawood reportedly indulging in extortion from celebrities and film stars, or arms dealers paying off politicians are instances of people who do not report their incomes to the government.

It may be just as well that the incomes from these activities are not included in aggregate incomes as they are illegal, but if they promote welfare it is not easy to conclude in this manner. If an individual engages in betting for a cricket game of his own free will without directly decreasing someone else's welfare as would be the case with extortion, and if it makes him happy, even though the activity may be illegal, it contributes to his welfare. To that extent, it merits being included in a measure of aggregate income.

Finally, there are many activities performed that are not marketed through formal markets. These non-market activities do not get included in aggregate income. In India, many farmers live in subsistence and produce their own food—a form of income that many farmers do not reckon as income and which may not show up in official statistics as the goods (food) produced are not traded in markets. In developing countries, many households do much of their own production. Many households produce goods within families that do not get traded in markets. This makes comparisons of per capita incomes between countries problematic.

Another problem arises when, say, instead of washing your own dishes, cooking your own meals, and scrubbing your own floors, you contracted out these activities for a market wage. Just as much work gets done and probably you have as well-run a house but the measured income has grown since other people are getting paid. So why do goods and services that we provide ourselves instead of contracting out not enter into official statistics? They do not because we do not pay ourselves for such activities and so do not report such income. Ignoring these activities in measures of aggregate income distorts the correspondence between income and welfare. For example, if more women start going to work as happened in the 1990s, and start paying for services like housekeeping and child care, which they earlier used to provide themselves without a payment, measured income can go up. But to what extent welfare rises is a subject of many unending household debates.

But as long as the level of non-market activities and under-reported incomes stays constant, we could still compare incomes from one year to another and make inferences about changes in welfare that correspond to the change in income. A change in the level of these activities should, however, induce cautions in data handling. For instance, in the mid-1990s tax rates were cut, which reduced the disincentive to report income. This resulted in a situation where the reported income might have increased while the actual income did not change at all. It is a problem worthy of consideration as to how much of the increase in the reported aggregate income of India in the 1990s is due to increased compliance with the tax system and as such represents no change in actual incomes.

## 1.2 Measuring Output

Before measuring income, however, we must consider how we are earning income. When an income gets earned, an economic unit must produce something like shoes, haircuts, and entertainment. Goods and services that are produced are called production or output. The addition of the values of all the goods produced in a particular period in a country is that country's aggregate output or the gross domestic product (GDP).

Let us for a moment suppose that there are only two goods in an economy—*lungis* and shoes. If 20 *lungis* were produced in a year and sold for INR 12 each, and 8 pairs of shoes that were produced were sold for INR 32.50 each, then the aggregate output of the economy or the GDP in that year is INR $(12 \times 20) + (32.5 \times 8) =$ INR 500.

GDP is the total value of the current production of final goods and services within the national territory during a given period of time—a quarter or a year. This definition implies that measuring output is an exercise fraught with difficulties. For instance, the fact that GDP includes only current production means that we do not count the resale of items. Similarly, since only final goods are included means that we do not count raw materials and the intermediate goods used as inputs. Let us take an illustration.

› GDP is the market value of the final goods and services produced in a country during a given time period.

Suppose a farmer produces INR 50 worth of wheat which is then sold to a baker. The baker, in turn, puts in effort worth INR 50 to produce bread from the wheat, which is then sold for INR 100 (see Figure 1.1). If we ask the farmer and the baker to report their output and simply add their outputs we would falsely conclude that INR 150 of output has been produced in the economy. What has caused the error in this operation of counting is that we have counted the wheat, which is not a final good but rather an intermediate good that disappears once it is converted into bread.

We could define an INTERMEDIATE GOOD (or service) as one that is used up in the production of other goods (and services) during the same period in which it was produced. Intermediate goods like wheat, oil and advertising should not be double-counted when output is computed.

› An intermediate good is used up in the production of other goods during the same period in which it was produced. Such goods are not counted as part of GDP.

To avoid errors due to inclusion of intermediate goods, economic agents would be asked either to report their sales of final goods to consumers (the baker reports INR 100 and the farmer INR 0), or to report the contribution each makes to the total output (the baker reports INR 50 of effort and the farmer INR 50 worth of wheat, which totals to an aggregate output of INR 100). An economic agent's contribution to output is called the agent's VALUE ADDED. In the baker's case the value added is calculated by subtracting costs (INR 50 of wheat) from his revenue (INR 100). The baker's value added is INR 50. The farmer's value added is INR 50 as in the example we did not allow him to pay for any costs. Adding the value added by all the economic entities gives us a measure of GDP.

› The value added is the market value of the product or service of an agent minus the cost of intermediate inputs purchased from other agents or firms.

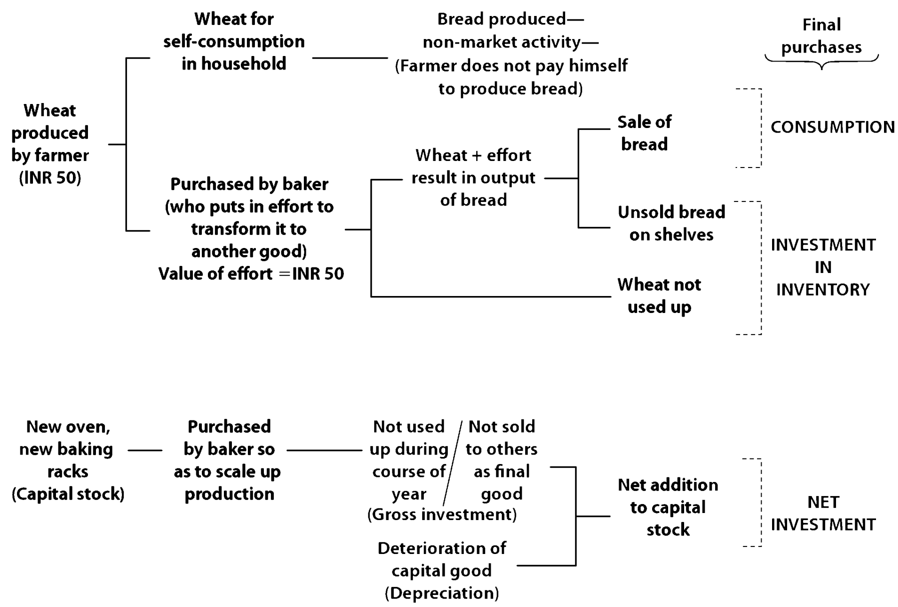
MACROFOCUS 1.1

### The Underground Economy: Beneficial or Subversive?

The underground economy constitutes economic activities that are not recorded in government statistics. These activities have also been referred to as the shadow economy, unofficial economy, black economy, and parallel economy. Scholars who have attempted to indirectly estimate the share of the underground economy in GDP have found a significant size of this activity. Schneider and Enste* found the underground economy in Nigeria, Egypt, and Thailand to be nearly three-quarters of the officially recorded GDP as averaged over 1990–1993. The OECD countries underground economies ranged from 8 per cent to 30 per cent of their GDP. In India, the estimates vary from 22 to 50 per cent, in Taiwan from 25 to 45 per cent, and in Pakistan from 20 to 50 per cent. In Brazil, the estimates started at 7 per cent in the early 1980s and shot up to more than 100 per cent by the early 1990s. In the 1980s, Mexico's underground economy is reported to have trebled while the legal economy registered virtually no growth.

While there is a natural tendency to think of the underground economy as undesirable, this is a contentious issue. One of the reasons for the existence of the underground economy, for instance, is that it is a refuge for entrepreneurs from a government that resorts to excessive taxes, regulation, and even graft. In response to a demand by corrupt government officials (say for a license), an entrepreneur can make one of three decisions—(1) to make the bribery payment and enter the business, (2) to stay out of business, or (3) to produce without the required license in the underground economy. The last saves the cost of bribes but bears the cost of possible detection and the subsequent punishment and evasion of law-enforcement costs. In such a situation where entrepreneurs take the outside option and move to the underground sector, corrupt officials are constrained in their ability to demand graft. As a result, there is a lower handover required to be paid to corrupt officials from business than would be the case if this outside option of migrating to the underground sector did not exist. This lowers the cost of doing business in the official sector and induces more entrepreneurs to enter it, which, in turn, results in additional economic activity. The official economy then is complementary to the underground economy, which is beneficial to its growth. The underground economy in such a situation improves social welfare and efforts to eradicate it are not socially valuable unless efforts are made to curtail corruption in the first place.

*F. Schneider and D. H. Enste, "Shadow Economies: Size, Causes, and Consequences," *Journal of Economic Literature*, no. 38 (2000): 77–114.




› **Figure 1.1**
Production Processes and the National Income. The production of goods includes intermediate goods and goods for self-consumption that are not counted as final goods. Final goods include the purchases for consumption, additions to inventory or investment in inventory, and net additions to the capital stock or net investment.

We now need to think about the estimation of aggregate output when the baker does not sell all the bread during the accounting period, or he does not use up all the wheat purchased from the farmer in that period. The unused wheat and unsold bread is not used up and so cannot be deemed to be intermediate goods. Rather, they are both deemed to be inventories that are classified as final goods and counted as part of output (see Figure 1.1). The accounting is done by classifying the change in the value of an economic unit's inventory from one year to the next as a final good called the INVENTORY INVESTMENT. If the baker had an inventory worth INR 200 at the end of one year and INR 300 at the end of the next year, the baker is considered to have undertaken INR 100 worth of inventory investment during that year. This INR 100 is reported as output in the form of inventory.

› A firm's unused raw materials or unsold output is its inventory. The change in the stock of inventory in an accounting year is treated as inventory investment and is classified as a final good.

Now, suppose the baker decides that he wants to scale up his business and buys new baking racks and a new oven. These objects are not used up during the accounting period and so cannot be intermediate goods. Neither does the baker sell them as final goods. Rather these objects are called capital goods and the economic entity (here the baker) who purchases them is considered to be the final user of the CAPITAL GOOD (see Figure 1.1). As the capital good is not used up during the period in which it was bought it is thought of as a final good. A capital good is a long-lived good that is itself produced and is used to produce other goods. It is similar to an intermediate good in that it is used to produce other goods, but it is unlike an intermediate good in that it is not used up right away. The total quantity of a country's capital good is called its capital stock. The change in the country's capital stock from the beginning to the end of the year is denoted as the country's INVESTMENT for that year. If the baker began the year with a stock of INR 500 of ovens and baking racks and he ended the year with a stock of INR 800 of ovens and racks, he is considered to have invested INR 300 during the year—he bought INR 300 worth of capital goods. The change in a stock is called a FLOW. Thus, the flow of capital goods is the investment that is a portion of the output for the year.

› A capital good is a long-lived good that is used in the production of other goods and services.

› The expenditure resulting in the increase in the stock of capital goods is investment expenditure.

› A measure that is defined at a point in time is a stock. A measure that is defined per unit of time is a flow.

A capital good is not used up right away but it diminishes in its material respect and is used up eventually. In short, it undergoes DEPRECIATION. If in a given year INR 500 worth of new capital is created while INR 150 worth of old capital wears out, then the total capital stock would have increased by INR 350. This situation is where the gross investment is INR 500 and the net investment is INR 350. Gross investment is the amount of new capital created while NET INVESTMENT is gross investment minus depreciation.

› The decrease in the value of a capital good as it is used and becomes old or becomes obsolete is depreciation.

› Net investment is the gross investment less depreciation.

The contribution of private economic units to aggregate output is, therefore, given by the sum of the value added by all the units to final goods, including the value of capital goods and inventories and then subtracting out the value of depreciation.

Some output in a country is produced by the government and, therefore, the total output in the country is the sum of the government output and the private sector output. Measuring the contribution of government to the output, however, is not an easy task. Many things that the government produces, such as defence and police services, are not income generators. Also, the vehicle taxes and tolls the government charges for roads are never enough to cover the costs incurred in constructing and maintaining the roads. In practice, therefore, the output of the government is measured by its costs, which, in turn, are the expenses it incurs on its employees and what it pays for goods and services. The actual valuation of the output of the government could be more or less than what it costs to produce that output but due to the lack of a proper metric valuing government output at cost is the best that we can do.

The aggregate output of a nation is, thus, computed by its GDP and is the sum of the private-sector output and the government costs.

### *1.2.1* Connecting Output with Income

GDP is a measure of the aggregate output of the economy and this output must be either sold or added to stocks. Moreover, the proceeds of sales of goods must cover the costs of production and allow for profits, which could be negative. Hence, the total value of the output produced must equal the total value of incomes generated in producing the output. The equivalence of output and income can be seen by recollecting that output is either sold or not sold and the firm may not have used up all its intermediate goods. The unsold output becomes an increase in the stock of finished goods and the unused intermediate good constitutes work in progress. The change in stock of finished goods and work in progress constitutes inventory or final purchases by the business.

$$\text{Output} = \text{Output Sold} + \underbrace{\text{Unsold Output} + \text{Work in Progress}}$$

$$= \text{Output Sold} + \text{Change in Stocks (Inventory)}$$

The output sold in turn comprises the following elements:

$$\begin{aligned}\text{Output Sold} = {} & \text{Final Purchases by Consumers (Consumer Goods)} \\ & + \text{Final Purchases by Producers (Capital Goods)} \\ & + \text{Intermediate Input Purchases by Producers} \\ & \quad \text{(Interfirm Purchases)}\end{aligned}$$

› The final demand is the goods and services consumed by the ultimate user who is what the end product of the production process serves. This equals final purchases by consumers and producers and the changes in stocks.

The change in stocks is an inventory investment and is considered to be the FINAL DEMAND by producers. Accordingly, we may write,

$$\text{Output} = \text{Final Purchases} + \text{Interfirm Purchases}$$

A firm produces output so as to generate revenues for the business. A firm has three claims on its revenues. It must pay for intermediate goods it has purchased from suppliers; it must pay for workers it has hired; and it must pay interest on the funds it has borrowed. After these claims on its revenues have been discharged, the residual amount left is the profit of the firm. Thus,

$$\begin{aligned}\text{Revenues} &= \begin{array}{c}\text{Payments to Inputs}\\ \text{from Other Firms}\\ \text{(Suppliers)}\end{array} + \begin{array}{c}\text{Payments to}\\ \text{Workers}\\ \text{(Employees)}\end{array} + \begin{array}{c}\text{Interest}\\ \text{Payments}\end{array} + \text{Profits}\\ &= \begin{array}{c}\text{Interfirm}\\ \text{Purchases}\end{array} + \text{Wages} + \text{Capital Incomes}\end{aligned}$$

Interest payments and profits may be considered as returns to capital assets delegated to the management of the firm by their original owners, or, capital incomes. With revenues equal to the value of output,

$$\begin{array}{c}\text{Interfirm}\\ \text{Purchases}\end{array} + \text{Wages} + \text{Capital Incomes} = \begin{array}{c}\text{Final Purchases}\\ \text{by Consumers}\\ \text{and Producers}\end{array} + \begin{array}{c}\text{Interfirm}\\ \text{Purchases}\end{array}$$

Or, cancelling out interfirm purchases from both sides,

$$\text{Wages} + \text{Capital Incomes} = \text{Final Purchases}$$

Recall that value added is the revenue received from the output minus the cost of inputs purchased from other firms.

$$\begin{aligned}\text{Value Added} &= \text{Revenues} - \text{Interfirm Purchases}\\ &= \text{Wages} + \text{Capital Incomes}\end{aligned}$$

Thus, we may write

$$\text{Wages} + \text{Capital Incomes} = \text{Final Purchases}$$

as

$$\begin{array}{c}\text{Value Added}\\ \text{(Income)}\end{array} = \begin{array}{c}\text{Final Demand}\\ \text{(Output)}\end{array}$$

However, when we get down to the actual measurement we must take care to relate measured output to income. A part of the revenues received by the firm is in the form of indirect taxes levied by the government, over which the firm has no claims. Then again, a part of the revenues must be set aside to replace the equipment worn out during the production process—a process called depreciation. The revenues a firm has after selling its output constitutes its net output. In the process of producing an output, say, of INR 10 million if the equipment and plant depreciated by INR 1 million, then even though the gross output of the firm is INR 10 million, its net output is INR 9 million. It is the revenue from the net output that covers the claims on the firms by its suppliers, its employees, and those who have delegated their funds to the firm.

First, then, a part of current output must be reinvested in any given period simply to make up for depreciation. The value of the output of the firms left after replacing the depreciated capital is called the NET DOMESTIC PRODUCT or NDP.

› Net domestic product at market prices is GDP less depreciation.

$$\text{GDP} - \text{Depreciation} = \text{NDP}$$

It is the NDP that constitutes the output sold to generate revenues, which the business then uses to pay out to suppliers, employees, and owners of capital or assets. Second, before that income is paid out, firms have to pay indirect taxes. Sales taxes are imposed on the sales of goods to final consumers so that the sum of the consumer's final expenditures on output will exceed the proceeds of sales by the amount of the tax imposed. Only after paying out that indirect tax to the government will firms be able to pay out to suppliers, employees, and asset owners. Thus,

$$\text{NDP} - \text{Indirect Taxes} = \text{Aggregate Domestic Income}$$

Combining the two previous equations gives us

$$\text{GDP} = \text{Aggregate Domestic Income} + \text{Indirect Taxes} + \text{Depreciation}$$

Or,

$$\text{Output} = \text{Income}$$

Thus, macroeconomists measure income as the sum of aggregate domestic income and indirect taxes and depreciation. And income, as measured this way, always equals output.

For macroeconomists, GDP is a measure of both output and income as the two are identical. Aggregate domestic income is *not* the measure a macroeconomist would use to measure the aggregate income or the output of a country.

### *1.2.2* Aggregate Income Categories

When aggregate output is measured at the prices at which it is actually sold, it is referred to as GDP at market prices. When the net output—GDP less depreciation—is measured at market prices, it is referred to as NDP at market prices. Subtracting out indirect taxes less subsides paid by the businesses to the government from the net output sold to final consumers gives the NDP AT FACTOR COST. NDP at factor cost is the aggregate of domestic incomes earned from the output produced within the national territory.

› NDP at factor cost is GDP at market prices less depreciation less indirect taxes (net of subsidies) and equals aggregate domestic income.

Table 1.1 provides data on the Indian economy for the last few years for which we have firm data available on GDP, depreciation, indirect taxes less subsidies, and aggregate domestic income. As Table 1.1 shows, depreciation over the last few years was 9.5 per cent of the aggregate output on average and indirect taxes was 8.3 per cent of the GDP. Aggregate domestic income, thus, constitutes 82.2 per cent of GDP in India (see Figure 1.2).

Before we continue, it is appropriate to consider how macroeconomic data are made available in India. In January of every year, quick estimates of the national accounts are released by the Central Statistical Organisation for the financial year ending March of the preceding year—a lag of 10 months. At that time the estimates for the earlier financial years are also revised and the advance estimates for the current financial year are also released about two months before the close of the year. Thus, in January 2007, quick estimates of 2005–2006 are released along with advanced estimates for 2006–2007. The advanced estimates are then updated and published as revised estimates in June 2007. The tables in this chapter, therefore, do not report the more current advanced estimates as they are subsequently revised.

Finally, GDP is decomposed in another way on the basis of citizenship. Some Indian citizens work for foreign-owned companies and some people who work for Indian-owned companies are foreign citizens. This means that the revenues from the sales of output of Indian firms do not automatically become incomes in the hands of Indian citizens. Some of it becomes incomes earned by foreign citizens. So if we take the output of Indian firms to measure Indian aggregate income we commit an error by including incomes paid to foreign workers and we also mistakenly exclude the incomes earned by Indian citizens working for foreign-owned firms.

› Net factor incomes from abroad are the income paid to domestic factors of production by the rest of the world less the income paid to foreign factors of production by the domestic economy.

› Gross national product or GNP equals GDP plus net factor incomes from abroad.

To take care of this, we consider the net factor incomes from abroad. NET FACTOR INCOMES FROM ABROAD refer to the income paid to the domestic factors of production by the rest of the world minus the income paid to the foreign factors of production by the domestic economy. When the net factor incomes from abroad are added to the total value of output in India, we get the GROSS NATIONAL PRODUCT or the GNP. Thus,

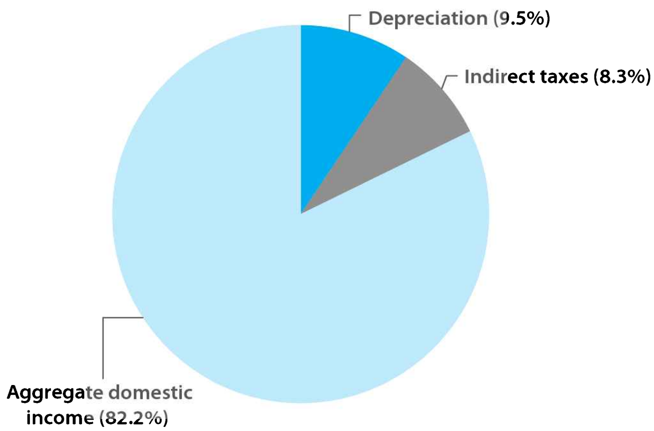
$$\text{GNP} = \text{GDP} + \text{Net Factor Incomes from Abroad}$$

The difference between the two measures is that the GDP is the output produced within a country's borders regardless of whether or not it is produced by the country's citizens. The GNP is the output produced by a country's citizens regardless of where in the world they work. The difference between GDP and GNP essentially boils down to how many of a country's citizens and

**› Table 1.1**
GDP and Its Components (INR Billion)

*Source:* Central Statistical Organisation, Government of India, National Accounts Statistics, 2006.

| | 1999–2000 | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|---|---|
| GDP (Market prices) | 19,520.35 | 21,023.75 | 22,810.58 | 24,580.84 | 27,654.91 | 31,265.96 | 35,671.77 |
| Depreciation | 1,855.93 | 2,062.95 | 2,324.52 | 2,547.67 | 2,847.02 | 3,324.90 | 3,792.00 |
| NDP (Market prices) | 17,664.42 | 18,960.80 | 20,486.06 | 22,033.16 | 24,807.88 | 27,941.05 | 31,879.76 |
| Indirect Taxes Less Subsidies | 1,655.10 | 1,769.60 | 1,808.71 | 1,927.80 | 2,160.73 | 2,706.63 | 3,162.45 |
| NDP (Factor cost) = Aggregate Domestic Income | 16,009.32 | 17,191.20 | 18,677.35 | 20,105.36 | 22,647.15 | 25,234.42 | 28,717.31 |
| Net Factor Incomes from Abroad | −154.31 | −227.33 | −200.68 | −166.90 | −182.50 | −223.75 | −249.69 |
| NNP (Factor cost) = National Income | 15,855.01 | 16,963.87 | 18,476.67 | 19,938.46 | 22,464.65 | 25,010.67 | 28,467.62 |




**› Figure 1.2**
Components of GDP at Market Prices. **The major component of measured GDP is the aggregate domestic income earned from the production of output within the nation.**

This figure is based on Table 1.1.

*Source:* Central Statistical Organisation, Government of India, National Accounts Statistics, 2006.

their assets work abroad. Payments to domestic factors of production by the rest of the world amounted to INR 250.76 billion in 2005–2006. A total of 3.1 per cent of this amount was in the form of compensation to Indian employees and 96.9 per cent comprised payments in the form of property and entrepreneurial incomes to Indians from the rest of the world. In the same year payments made out to factors of production from the rest of the world by India amounted to INR 500.45 billion of which 6.6 per cent was in the form of compensation to foreign employees. The net factor incomes from abroad in 2005–2006 was accordingly in deficit by INR 249.69 billion (see Table 1.1).

For countries such as Egypt and Turkey that have many citizens working abroad, the distinction between the two measures is significant because incomes earned by citizens abroad are part of a country's GNP and not its GDP. For the case of India, net factor payments abroad are a small percentage of the GNP in the last few years—0.83 per cent on average (see Table 1.1). The GDP and GNP are, therefore, roughly equal for India. Even though GNP is a better measure of the welfare of citizens, most countries are interested in how policies and programmes translate into effects within national borders and so GDP is the measure of output that is generally used.

Not surprisingly, the NATIONAL INCOME of a country comprises two components. One is the net output produced within a country's borders valued at production prices, which is the NDP at factor cost, $NDP_{FC}$, or the domestic income. The other component is the net factor incomes from abroad, NFI. The national income is also known as $NNP_{FC}$ and can be written as follows:

› The national income is the sum of the net domestic product valued at factor cost plus the net factor incomes from abroad.

$$\underset{\text{(National Income)}}{NNP_{FC}} = \underset{\text{(Aggregate Domestic Income)}}{NDP_{FC}} + NFI$$

As net factor incomes from abroad are a very small proportion of national income (see Table 1.1), there is not much harm in interpreting the NDP at factor cost, $NDP_{FC}$, as the measure of the national output and income in the country. GDP is thus treated as the aggregate measure of output in macroeconomics. Of course, given that output equals income, GDP is also the measure of income.

Output is also often referred to in terms of production from a sector—agriculture and allied activities—such as forestry and fishing, industry, and services.

› **Table 1.2**
GDP (Factor Cost) by Industry (INR Billion)
*Source:* Central Statistical Organisation, Government of India.

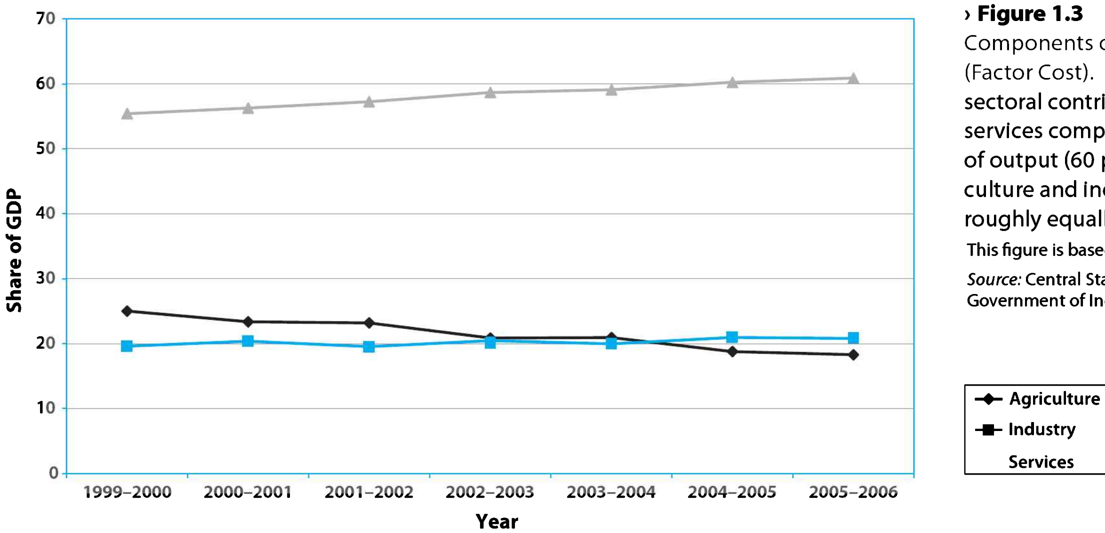
| | 1999–2000 | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|---|---|
| **Agriculture and Allied Activities** | 4,465.15 | 4,497.46 | 4,870.63 | 4,726.79 | 5,336.42 | 5,366.29 | 5,950.58 |
| **Industry** | 3,502.33 | 3,921.38 | 4,106.67 | 4,633.02 | 5,091.06 | 5,986.74 | 6,762.07 |
| Mining and Quarrying | 415.94 | 457.06 | 478.71 | 627.42 | 638.82 | 844.64 | 904.82 |
| Manufacturing | 2,641.13 | 3,003.92 | 3,153.14 | 3,460.29 | 3,885.49 | 4,536.03 | 5,197.46 |
| Electricity, Gas, and Water Supply | 445.26 | 460.40 | 474.82 | 545.31 | 566.75 | 606.07 | 659.79 |
| **Services** | | | | | | | |
| Construction | 1,020.07 | 1,119.99 | 1,208.65 | 1,351.72 | 1,568.06 | 1,856.69 | 2,221.10 |
| Trade, Hotels, Transport, and Communications | 3,875.14 | 4,288.55 | 4,778.36 | 5,275.69 | 6,102.39 | 7,145.47 | 8,249.36 |
| Financing, Insurance, Real Estate, and Business Services | 2,335.50 | 2,547.72 | 2,930.35 | 3,321.15 | 3,756.06 | 4,131.29 | 4,644.93 |
| Community, Social, and Personal Services | 2,667.07 | 2,879.05 | 3,107.21 | 3,344.66 | 3,640.19 | 4,072.85 | 4,681.28 |
| **GDP$_{FC}$** | 17,865.25 | 19,254.15 | 21,001.87 | 22,653.04 | 25,494.18 | 28,559.33 | 32,509.32 |

*Notes:* 1. For 2004–2005, quick estimates and for 2005–2006, revised advance estimate.
2. Agriculture and Allied Activities + Industry + Services = GDP$_{FC}$
3. GDP$_{FC}$ + Indirect Taxes less Subsidies = GDP$_{MP}$

The data of output by sector of production for the Indian economy are given in Table 1.2. As Figure 1.3 depicts, services is the largest production sector in the Indian economy—close to 60 per cent of GDP. Till recently, agriculture was the second-largest production sector and industry was a close third. Most economies develop from being primarily agricultural to industrial and eventually to service economies. The Indian economy has only just begun to have agricultural output as a smaller denomination than industrial output.

The incomes earned as a result of the output produced are usually attributed to the factor that is the source of earnings—wages to labour, rent to land, buildings and equipment, interest on loans, and profit to entrepreneurs.

In India, a predominant number of people are self-employed. In 1999–2000, 53 per cent of workers were self-employed, 14 per cent had regular salaried employment, and 33 per cent had casual employment (Government of India, 2001). Self-employed individuals do not attribute their incomes under pure categories like entrepreneurship or labour and their incomes are therefore classified as mixed income. The national income in India is thus classified under two account heads—compensation of employees and mixed income (see Table 1.3). In the last few years, the compensation of employees has, on an average, been 37 per cent of the aggregate domestic income. In countries like the United States, the compensation of employees is around 70 per cent of




› **Figure 1.3**
Components of GDP (Factor Cost). **In terms of sectoral contribution to output, services comprise the major share of output (60 per cent) with agriculture and industry contributing roughly equally to the remainder.**
This figure is based on Table 1.2.
*Source:* Central Statistical Organisation, Government of India.

the national income. A major reason for the lower share of labour in India is because a large part of the returns to labour gets counted in as the mixed income of the self-employed.

### *1.2.3* Components of Aggregate Expenditure

Now that we have looked at income, let us look at the way it is spent. Just as output is produced by the private sector and by the government, expenditure in an economy is subdivided into expenditure by the private sector and expenditure by the public sector or the government. The public sector constitutes approximately 25 per cent of the GDP (see Table 1.4).

Expenditures by the government are of three types—consumption, investment, and transfers. Consumption expenditure includes wages paid for services such as labour and expenditure on trash collection. Investment expenditure by the government includes actions such as the building of roads and the purchase of aircraft. Transfer payments are direct payments that agents in the economy receive without having to sell anything in return to the government. These are mainly welfare payments such as subsidies on food and fertilizers, and interest payments on previous borrowings that the government made from citizens. Transfers are simple redistributions of income—the government collects taxes from some people and distributes to others. They are not counted as part of a country's expenditures as they do not represent any additional economic production. By counting those, we would have counted the incomes earned that were taxed and then redistributed back among citizens twice.

› **Table 1.3**
$NDP_{FC}$ by Factor Incomes (INR Billion)
*Source:* Central Statistical Organisation, Government of India.

| | 1998–1999 | 1999–2000 | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 |
|---|---|---|---|---|---|---|---|
| **Compensation of Employees** | 5,122.67 | 6,037.53 | 6,637.52 | 7,034.72 | 7,471.34 | 8,194.18 | 9,242.07 |
| **Mixed Income** | 9,177.94 | 9,971.80 | 10,553.68 | 11,642.62 | 12,634.03 | 14,452.99 | 15,992.34 |

*Note:* Compensation of Employees + Mixed Income = $NDP_{FC}$

› **Table 1.4**
GDP of the Public Sector (INR Billion)
*Source:* Central Statistical Organisation, Government of India.

| | 1999–2000 | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|---|---|
| $GDP_{FC}$ | 17,865.25 | 19,254.15 | 21,001.87 | 22,653.04 | 25,494.18 | 28,559.33 | 32,509.32 |
| GDP (Public sector) (% share) | 4,570.06 (25.6) | 4,818.61 (25.0) | 5,263.26 (25.1) | 5,788.51 (25.6) | 6,262.19 (24.6) | 6,863.33 (24.0) | 7,485.28 (23.0) |

Government expenditures include spending on consumption and investment and transfers but only the spending on consumption and investment is included in expenditure by the public sector. Table 1.5 shows that in fact for central government expenditures, on average, transfers constituted 55.5 per cent of it in the last few years.

The expenditure by the private sector, in turn, is subdivided into expenditure by households and businesses. Most household expenditure is on consumption whether of non-durable goods such as food or durable goods such as refrigerators and television sets. Even though expenditure on college education or health is fundamentally an investment, the national accountants count it as consumption expenditure. Similarly, most investment expenditure, such as the purchase of machinery and factory space and additions to inventory, is usually deemed to be done in most countries by businesses or firms rather than individuals or households. In India, however, given the high extent of self-employment, investment expenditure or capital formation in the private sector is by firms and households, with households having the largest share in capital formation until recently (see Figure 1.4).

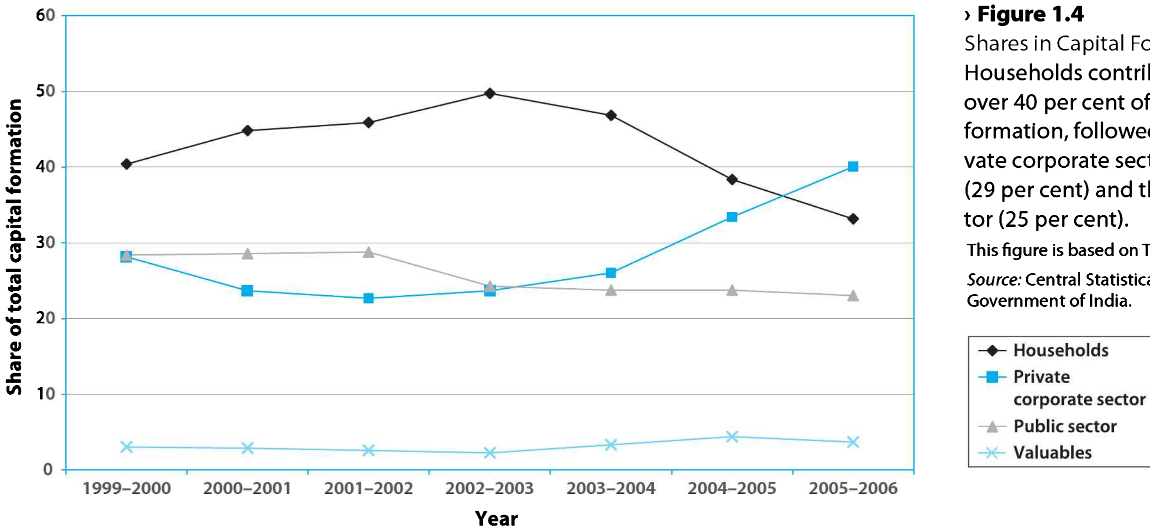
But before looking at the data on expenditure, there is a factor we have to account for which is that most countries have open economies—citizens sell and buy goods from foreigners. If an Indian citizen buys goods worth INR 100 from another Indian citizen, the value of Indian income is INR 100—the same as expenditure. But if the Indian citizen spends INR 75 on Indian goods and imports or spends INR 25 on foreign goods, then the Indian expenditure is INR 100 but the Indian income is INR 75. To maintain equality between expenditure and income, the total expenditure by Indian citizens should be reduced by the value of imports. Similarly, some goods produced by Indian business are sold to foreigners or exported. If

› **Table 1.5**
Economic Classification of Central Government Expenditures (INR Billion)
*Source*: Economic and Functional Classification of Central Government Budgets or Economic Survey, Table 2.6.

| | 1997–1998 | 1999–2000 | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 (RE) | 2005–2006 (BE) |
|---|---|---|---|---|---|---|---|---|
| Consumption Expenditures | 530.90 | 688.31 | 719.77 | 773.24 | 853.89 | 871.70 | 1,135.74 | 1,256.07 |
| Transfers* | 1,115.77 (49.6%) | 1,615.49 (52.5%) | 1,836.96 (56.0%) | 2,011.88 (55.8%) | 2,285.01 (57.3%) | 2,484.36 (58.3%) | 2,569.59 (54.3%) | 3,001.56 (58.5%) |
| Capital Formation | 601.99 (26.8%) | 771.29 (25.1%) | 725.92 (22.1%) | 821.04 (22.8%) | 834.86 (21.3%) | 911.26 (21.2%) | 1043.29 (21.7%) | 872.32 (17.0%) |

*Transfers = Subsidies + Interest Payments + Grants

*Note*: Figures in brackets are percentages to Total Central Government Expenditures
RE: Revised Estimates, BE: Budget Estimates.




› **Figure 1.4**
Shares in Capital Formation. Households contribute to over 40 per cent of capital formation, followed by the private corporate sector (29 per cent) and the public sector (25 per cent).
This figure is based on Table 1.6.
*Source:* Central Statistical Organisation, Government of India.

a foreigner buys INR 15 worth of Indian goods then the Indian income is INR 15 but the Indian domestic expenditure for those goods is INR 0. Again, to maintain the equality between expenditure and income, we must add the value of exports to expenditure.

For any country, then, we must add the value of exports and subtract the value of its imports (exports less imports is called net exports) from the domestic expenditure to obtain the equality

$$\text{Output} = \text{Income} = \text{Domestic Expenditure} + \underbrace{\text{Exports} - \text{Imports}}_{\text{Net Exports}}$$

However, domestic expenditure is the expenditure by households, business, and the government, not inclusive of transfers. Given that household and business expenditures are mainly on consumption and on investment goods, we may write

$$\text{Income} = \text{Consumption} + \text{Investment} + \text{Government Expenditure} + \text{Export} - \text{Imports}$$

Note that government expenditure also has two components: consumption expenditure and investment expenditure on capital formation. In symbols, this may be written as

$$Y = C + I + G + X - M$$

where, $Y$ is the country's output, income, or GDP;
$C$ is private consumption expenditure;
$I$ is private investment expenditure;
$G$ is government expenditure on consumption and capital formation;
$X$ is exports; and
$M$ is imports.

$G$ could further be written as

$$G = G_C + G_I,$$

where, $G_C$ is the government's consumption expenditure; and
$G_I$ is the government investment expenditure.

However, conventionally, these two components are clubbed together into a single value for government expenditure, $G$.
For the year 2005–2006, the numbers for this equation for the Indian economy were as follows:

$$\underset{(35{,}672)}{Y} = \underset{(20{,}646)}{C} + \underset{(8{,}828)}{I} + \underset{(6{,}689)}{G} + \left(\underset{(-1{,}056)}{X - M}\right)$$

where the numbers are (approximately) in billions of Indian rupees.

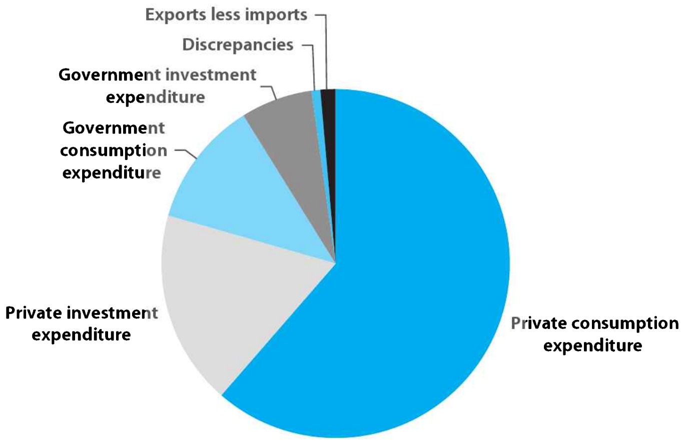
The data for the Indian economy showing the various components of final expenditures by domestic citizens and the government on consumption, capital formation and imports, and foreign nationals on exports are given in Table 1.6.

Figure 1.5 reveals the salience of these components from 1999–2000 to 2005–2006 in terms of a pie diagram. Private consumption expenditure on average is 61.6 per cent of the GDP, followed by private investment expenditure, which is 20.4 per cent of the GDP. The government consumption expenditure is 11.8 per cent of the GDP and its investment expenditure is 6.9 per cent of the GDP. Indians spent 17.3 per cent of GDP on foreign goods via imports and foreigners spent 15.6 per cent of GDP on Indian goods, which comprise Indian exports.

› **Table 1.6**
Components of Aggregate Expenditure (INR Billion)
*Source:* Central Statistical Organisation, Government of India.

| | 1999–2000 | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|---|---|
| Private Final Consumption Expenditure | 12,536.43 | 1,3401.09 | 14,632.47 | 15,438.73 | 17,093.89 | 18,656.45 | 20,646.38 |
| Government Final Consumption Expenditure | 2,527.44 | 2,654.01 | 2,821.23 | 2,913.20 | 3,106.35 | 3,425.42 | 4,045.11 |
| Gross Capital Formation | 5,095.18 | 5,061.81 | 5,435.62 | 6,146.79 | 7,345.85 | 9,276.29 | 11,472.53 |
| a. Households | 2,059.14 | 2,269.17 | 2,494.82 | 3,058.19 | 3,440.67 | 3,560.43 | 3,806.55 |
| b. Private Corporate Sector | 1,434.75 | 1,199.01 | 1,233.49 | 1,455.79 | 1,913.49 | 3,100.45 | 4,597.15 |
| c. Public Sector | 1,446.10 | 1,446.39 | 1,565.44 | 1,493.24 | 1,745.97 | 2,204.87 | 2,644.26 |
| d. Valuables | 155.19 | 147.24 | 141.87 | 139.57 | 245.72 | 410.54 | 424.57 |
| Exports | 2,276.97 | 2,781.26 | 2,907.57 | 3,555.56 | 4,078.03 | 5,690.51 | 7,251.24 |
| Imports | 2,657.02 | 2,975.23 | 3,110.50 | 3,799.81 | 4,433.98 | 6,259.45 | 8,306.78 |
| Discrepancies | −258.65 | 100.81 | 124.19 | 326.37 | 464.77 | 476.74 | 563.29 |
| GDP (Market prices) | 19,520.35 | 21,023.75 | 22,810.58 | 24,580.84 | 27,654.91 | 31,265.96 | 35,671.77 |

*Notes:* 1. Valuables excludes works of art and antiques.
2. Gross Capital Formation = Gross Fixed Capital Formation + Change in Stocks
3. Gross Domestic Capital Formation = Gross Capital Formation + Errors and Omissions




**› Figure 1.5**
Components of Aggregate Expenditure. **Private consumption is the dominant component of expenditure (61.6 per cent). This is followed by private investment (20.4 per cent), government consumption (11.8 per cent), and government investment (6.9 per cent) expenditure. Net exports are 1.6 per cent of GDP.**
This figure is based on Table 1.6.
*Source:* Central Statistical Organisation, Government of India.

Net exports or the trade balance is measured in various ways. The BALANCE OF (MERCHANDISE) TRADE, which is the value of exported goods minus the value of imported goods, is the commonly used measure. This balance, however, does not include services of any kind. A broader measure of the trade balance, which includes trade in banking services, travel, insurance, incomes earned from abroad, and transfers, is the balance on goods and services.

› The balance of (merchandise) trade, otherwise known as net exports, is the value of exports less the value of imports of goods.

The net receipts from services such as from travel, transportation, insurance, and software exports are also referred to as the non-factor incomes from the exports and imports of these services. A second component of exports and imports of services is the net factor incomes from abroad. This, as we saw earlier, includes the compensation of employees and investment income, or, property and entrepreneurial income from the rest of the world. A third component of services exports and imports

MACROFOCUS 1.2

## Chinese and Indian GDP: Re-emerging Economies?

Definitions of emerging economies vary. The United Nations counts Hong Kong, Singapore, South Korea, and Taiwan as emerging economies. The International Monetary Fund, however, counts all four as "developing economies" in its International Financial Statistics but as "advanced economies" in its *World Economic Outlook*. In policy discussions, these are labelled as emerging economies. By 2005, the emerging economies accounted for about 30 per cent of the world GDP and have begun to dominate the global economy in terms of exports, energy consumption, and foreign-exchange reserve holdings. Currently, two fast growing emerging market economies are China and India which together contributed to about 25 per cent of the increase in emerging market economies' GDP in 2005.

In a study of the world economy the economic historian, Angus Maddison* found that until the late 19th century, China and India were the world's two biggest economies. These two economies dominated the world output before the steam engine and the industrial revolution placed Britain up in front. Maddison estimates that in the eighteen centuries up to 1820, the current emerging market economies produced on average a staggering 80 per cent of the world's GDP. After a gap of about 180 years they are now back with amongst the fastest growth rates of GDP.

Given this long perspective it may be more appropriate to label these economies as re-emerging economies. Two important features of economies such as China, Hong Kong, Indonesia, Malaysia, Oman, Malta, Singapore, and Thailand that have seen sustained high increases in the GDP in the past couple of decades have been (1) the high levels of investment—attempting high growth without investment in infrastructure and skills is a sure recipe to ignite inflation rather than increase the output of goods and services; and (2) the large fraction of GDP associated with exports and imports—an indicator of a country's ability to leverage the demand and resources of the global economy. In 2004 for instance, China's total investment expenditure was 43.2 per cent of its GDP compared with India's 27.2 per cent. Similarly, international trade (exports plus imports) as a percentage of GDP was 58.2 per cent in China whereas it was 27.4 per cent in India.

*Angus Maddison, *The World Economies: A Millennial Perspective*, OECD, 2001.

comprises transfers—private and official transfers between governments. Transfer receipts in India comprise mainly remittances sent by Indian workers abroad.

The current account balance is the sum of the balance of trade and the net receipts from the exports and imports of the three components of services—non-factor services, factor incomes from abroad, and transfers. The balance of trade and the CURRENT ACCOUNT BALANCE for India are given in Table 1.7. In addition to depicting some of the major items of exports and imports, Table 1.7 also reveals how private transfers and software exports have in the recent past sometimes turned a negative balance of trade into a positive current account balance such that foreign citizens' spending on domestic goods and services exceeds Indian citizen spending on foreign goods and services.

› The current account balance is the statement of a country's payments to and from the rest of the world for goods and non-factor services transacted in, for factor incomes from abroad, and for transfers.

The current account balance is a broad measure, but it is not exhaustive and many major services are not included in it. For instance, expenditure by Indian students going to the United States for education is not included as an

**› Table 1.7**
India's Foreign Trade (INR Billion)
*Source:* Reserve Bank of India, *Handbook of Statistics on Indian Economy*.

| | 2000–2001 | 2001–2002 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|---|
| **Exports** | 2,078.52 | 2,133.45 | 2,600.79 | 3,039.15 | 3,817.85 | 4,657.05 |
| Agriculture | 272.88 | 281.44 | 324.73 | 346.16 | 380.78 | 452.20 |
| Gems and Jewellery | 337.33 | 348.45 | 437.01 | 485.86 | 618.34 | 687.53 |
| Chemicals and Related Products | 268.89 | 288.62 | 360.80 | 434.05 | 559.11 | 653.90 |
| Engineering Goods | 311.50 | 331.83 | 437.15 | 570.05 | 779.49 | 961.57 |
| Textiles | 515.55 | 486.77 | 562.21 | 587.79 | 609.06 | 726.18 |
| **Imports** | 2,645.89 | 2,683.00 | 3,117.76 | 3,673.01 | 5,335.50 | 6,951.31 |
| Petrol, Oil, and Lubricants | 714.97 | 667.70 | 853.67 | 945.20 | 1,340.94 | 1,946.40 |
| Capital Goods | 408.47 | 471.30 | 653.25 | 839.94 | 1,129.36 | 1,667.61 |
| Gold and Silver | 211.89 | 218.54 | 207.53 | 315.06 | 500.99 | 501.08 |
| Chemicals—Organic and Inorganic | 111.65 | 133.52 | 146.40 | 185.27 | 256.10 | 309.21 |
| Pearls, Precious and Semi-precious Stones | 219.64 | 220.46 | 293.41 | 327.57 | 423.38 | 404.41 |
| **Trade Balance** | −567.37 | −549.55 | −516.97 | −633.86 | −1,517.65 | −2,294.26 |
| Invisibles (Net) | 451.39 | 713.81 | 823.57 | 1,273.69 | 1,395.91 | 1,887.04 |
| Foreign Travel | 79.05 | 158.89 | 176.44 | 65.20 | 62.87 | 61.98 |
| Transport | −68.08 | −61.60 | −35.67 | 40.26 | 6.58 | −68.72 |
| Insurance | 2.30 | 0.35 | 0.96 | 2.50 | 6.64 | 0.69 |
| Software Exports (Net) | 263.08 | 328.36 | 428.59 | 566.06 | 758.25 | 986.78 |
| Net Factor Incomes from Abroad | −227.33 | −200.68 | −166.90 | −182.50 | −223.75 | −249.69 |
| Transfers (Net) | 599.67 | 755.60 | 814.03 | 1,016.96 | 931.35 | 1,076.73 |
| **Current Account Balance** | −115.98 | 164.26 | 306.60 | 639.83 | −121.74 | −407.22 |

*Notes:* 1. Only major components of exports, imports, and invisibles have been depicted. The list is not exhaustive.
2. Chemicals and pearls, precious and semi-precious stones are mainly for further processing and re-export.

| Year | Per Capita Income | Percentage Increase (Per Decade) |
|---|---|---|
| 1960–1961 | 395.6 | 42.9 |
| 1970–1971 | 844.3 | 113.4 |
| 1980–1981 | 2,117.3 | 150.8 |
| 1990–1991 | 6,778.0 | 220.1 |
| 2000–2001 | 20,650.6 | 204.7 |

› **Table 1.8**
Per Capita GDP (INR)
*Source:* **Central Statistical Organisation, Governement of India.**

import of a service and meals purchased by foreign travellers at a *dhaba* in Rajasthan are not reported as an export of cuisine.

# 1.3 Real and Nominal Incomes

The percentage change in a country's output or income from one year to the next is the GROWTH RATE of the country's economy. Table 1.8 shows the per capita income for the past few decades. This table reveals that in the decade of the 1990s, an average person's income increased by over three times from INR 6,778 to INR 20,651. In the previous decade of the 1980s income increased by 3.2 times. Most people comparing this with their own family income history would find these statistics unbelievable. Did anybody's income increase 9.8 times between 1980–1981 and 2000–2001? May be some business houses, like the Ambani family's, did but on average, incomes surely did not increase by that amount. The reason the data of Table 1.8 is misleading is that the rupees we earn as income are only as valuable as the goods we can purchase with them. If our rupee incomes rise and the prices of the things we buy also rise, then, our REAL INCOMES may not be increasing as fast as our rupee incomes or NOMINAL INCOMES.

› The growth rate of GDP is the percentage change in aggregate income or output over a unit of time, usually a year.

› Real incomes are incomes measured in physical terms, that is, in terms of the quantities of goods and services.

› Nominal incomes are incomes measured in terms of current rupee values.

To understand this, suppose that there is only one good in the economy, which happens to be pizza. Table 1.9 gives the nominal income and real income in year 0 (the base year) and year *j* (the current year). The table shows that despite your rupee income increasing by 12.5 per cent your real income declined by 10 per cent. Even though the nominal income increased we could buy less pizza with it.

## 1.3.1 Comparing Income Across Time

If there were more than one good (we cannot live only on pizzas) in the economy, to calculate the real income we proceed as follows:

1. Pretend that prices in the current year are the same as in the base year.
2. Calculate the value that the output would have in the current year at these constant prices of the base year.

| | Year 0 | Year *j* |
|---|---|---|
| Nominal Income | INR 5,000 | INR 5,625 |
| Price (of pizza) | INR 100 | INR 125 |
| Real Income | 50 pizzas per year | 45 pizzas per year |

› **Table 1.9**
Real Incomes with One Good

› **Table 1.10**
Real Incomes with Two Goods (1 and 2)

| | Pizzas | | Books | |
|---|---|---|---|---|
| | Quantity | Price | Quantity | Price |
| Year 0 | 30 | 100 | 4 | 500 |
| Year 1 | 30 | 85 | 5 | 615 |

To see how such a comparison is done across time we will calculate the price index for a student who spends all the pocket money for the year totalling INR 5,000 on two goods—pizzas and books. The student's expenditure is on 30 pizzas a year for INR 3,000 (each pizza costs INR 100), and on four books a year for INR 2,000.

In a subsequent year, the competition increases and the price of pizzas decline to INR 85 per pizza. However, there is no new entrant in the booksellers market and the price of books goes up to INR 615 per book because of the increase in the cost of paper. The student's income and expenditure goes up to INR 5,625 in the subsequent year and he purchases the same quantity of pizzas and an additional book over the previous year—the additional book being purchased with the intention to study more and do better in the examinations. This situation is depicted in Table 1.10.

› Nominal expenditure or GDP or aggregate income is the sum of the quantities of goods produced times their current price.

NOMINAL EXPENDITURE is the expenditure in a year in the prices of that year. The nominal income in Year 1, denoted as $C_{11}$, will then be

$$C_{11} = (30 \times 85) + (5 \times 615) = 5{,}625$$

› Real expenditure or GDP or aggregate income is the aggregate income measured in the prices of a base year. It measures aggregate income in constant rupees or adjusted for inflation.

REAL EXPENDITURE is the expenditure in a year in the prices of the base year. The real income in Year 1, denoted as $C_{10}$, will then be

$$C_{10} = (30 \times 100) + (5 \times 500) = 5{,}500$$

› The GDP deflator is the nominal GDP divided by the real GDP.

The inflation rate of the economy is found by deriving the implicit GDP DEFLATOR, which is the current price estimate divided by the constant price estimate. Thus,

$$\text{GDP Deflator} = \left(\frac{C_{11}}{C_{10}}\right)100 = \left(\frac{\text{Nominal GDP}}{\text{Real GDP}}\right)100$$

For our example with just two goods—pizzas and books—the deflator is given by

$$\left(\frac{C_{11}}{C_{10}}\right)100 = \left(\frac{5{,}625}{5{,}500}\right)100 = 102.27$$

The obtained deflator indicates that inflation was of the order of 2.27 per cent in the economy with two goods. This summary indicator makes sense if we note that the proportion of expenditure devoted to pizzas in real terms is the share of real expenditure on pizzas, $w_1$, where

$$w_1 = \frac{30 \times 100}{(30 \times 100 + 5 \times 500)} = \frac{3{,}000}{5{,}500} = 0.545455$$

The proportion of expenditure devoted to books in real terms is the share of real expenditure on books, $w_2$, where

$$w_2 = \frac{5 \times 500}{(30 \times 100 + 5 \times 500)} = \frac{2{,}500}{5{,}500} = 0.454545$$

The price of pizzas declined to 85 per cent of its original price from INR 100 to INR 85. The price of books increased by 23 per cent from INR 500 to INR 615. The weighted average of price increase in the overall economy is then given by

$$\begin{aligned}[0.545455(0.85) + 0.454545(1.23)]100 &= (0.463636 + 0.559091)100 \\ &= (1.022727)100 \\ &= 102.27\end{aligned}$$

This is exactly what the summary indicator calculated. (See Appendix 1.1 for a derivation of the formula for the GDP deflator for the case of $n$ goods.)

### *1.3.2* Measures of Inflation

The GDP deflator is the only measure of inflation that takes into account the inflation in all the goods and services produced in the economy. Other measures of inflation are the wholesale price index (WPI) and the consumer price index (CPI). These are calculated by selecting a fixed set of goods rather than all goods available in an economy and then tracking the prices of these specific goods and services from one time period to another. The bundle of goods used in calculating the WPI is different from that incorporated into the calculation of the CPI.

The CONSUMER PRICE INDEX (CPI) is a weighted sum of prices of the set of goods and services that the statistical agency of the government determines through surveys of the typical domestic consumer purchases each year. The government conducts a survey to determine the importance of different expenditures by the average household. Weights are assigned to the different expenditure categories to reflect the importance of each in the household spending. For example, expenditure on fuel and light by industrial workers receive a weight of 6.43 out of 100, and clothing, bedding, and footwear receives a weight of 6.58 (see Table 1.11). In addition to a consumer price index for industrial workers, the government also calculates alternative consumer price indices for agricultural workers and urban non-manual employees.

› The consumer price index (CPI) is the cost of a standard basket of goods and services consumed by a typical domestic consumer.

The WHOLESALE PRICE INDEX (WPI) measures the average level of prices of goods sold by producers. Unlike the CPI, the WPI measures price changes from the perspective of the seller. (Table 1.11 indicates the weighting pattern for this index.) The WPI is useful because it indicates what is likely to happen to consumer prices in the proximate future. If producers are receiving higher prices from their sales to wholesalers we could expect that sooner or later retailers would have to charge higher prices. This will then be reflected in a higher CPI. Both the WPI and the CPI are measured using base year quantities as weights. Price indices that do this are called LASPEYRE INDICES. The expenditure proportions in a LASPEYRE INDEX are those prevailing in the base year. Both the WPI and the CPI represent only a subset of all the goods and services produced by the entire economy and so may not reflect all changes in price levels.

› The wholesale price index (WPI) is a price index of goods and services domestically produced.

› A Laspeyre price index is calculated from a basket of fixed quantities of a given list of goods.

› We first calculate the total price of the basket of goods in the base period. Let the price index in the base period be 100. Then we calculate the total price of exactly the same basket of goods in a subsequent year $j$. Next we compute the ratio of the total price of the basket of goods in period $j$ divided by the total price of the basket in the base period. The price index in year $j$ then is this ratio multiplied by 100. The **Laspeyre index** tells us the expenditure that must be incurred in year $j$ as a fraction of the expenditure in the base year in order that we may obtain the same bundle of goods as that in the base year.

In price indices such as the CPI, we are interested in finding out how much it costs in the subsequent year to purchase the same basket of goods that were purchased in the base year. If the student whose expenditure pattern is represented in Table 1.10 were to purchase the same basket of goods the subsequent year he would have to pay INR 2,550 for pizzas (30 pizzas at INR 85 per pizza) and for books the payment would be INR 2,460 (4 books at INR 615 per book). The expenditure on the same basket of goods in the subsequent year is then equal to INR 5,010. In the subsequent year then the student's cost of living has increased. Setting the value of the index equal to 100 in the base year the price index in the subsequent year is calculated as $100 \times (5{,}010/5{,}000) = 100.2$. The inflation rate is none other than the percentage change in the price index. The price index rose from 100 in the base year to 100.2 in the subsequent year—an increase of 0.2 per cent. Inflation as measured by the price index with base year quantities as weights is 0.2 per cent in the subsequent year.

The advantage of an index is that once calculated for a base year we can compare it with any other year. For instance, if the index were to rise to 103.04 in a later year, then, between the subsequent year and the next year the index rose by $103.04 - 100.2 = 2.84$. Hence, the inflation for this time period is the percentage change in the index or $100 \times (2.84/100.2) = 2.83$ per cent.

Before looking at the weighting pattern of WPI and CPI, it is worth pointing out that all price indices could overstate the true inflation due to two factors:

1. Price indices cannot account for price changes in newly invented goods as the indices are calculated on a bundle of fixed goods and undergo revision with a lag. For example, the prices of new computers tend to start high and come down rapidly as output expands and so the price index in the initial years would overstate inflation.
2. Existing goods change in quality over time. For example, automobiles attributes over time have improved significantly. Apart from improved engine performance, there have been a number of innovations in luxury,

› **Table 1.11**
WPI and CPI: A Comparison

| WPI | | CPI | |
|---|---|---|---|
| ***Properties of WPI and CPI (Industrial workers)*** | | | |
| Movement in wholesale prices of 435 commodities in all trade and transactions | | CPI for industrial workers measures retail prices of goods and services for 260 items | |
| Index available on a weekly basis with a short time lag of 2 weeks | | Index constructed on a monthly basis with lag of 1 month | |
| Widely used in business and industry and by government as the indicator of inflation | | | |
| Base year is 1993–1994 | | Base year is 2001 | |
| ***Weighting Diagram of WPI and CPI (Industrial workers)*** | | | |
| Primary products | 22.02 | Food | 46.19 |
| Food articles | 15.40 | *Pan*, *supari*, tobacco, | |
| Minerals | 0.49 | and intoxicants | 2.27 |
| Fuel group | 14.23 | Fuel and light | 6.43 |
| Coal mining | 1.75 | | |
| Electricity | 5.48 | | |
| Manufactured products | 63.75 | Housing | 15.27 |
| Food products | 11.54 | Clothing, bedding, | |
| Textiles | 9.80 | and footwear | 6.58 |
| Chemicals | 11.93 | Miscellaneous | 23.26 |
| Basic metals | 8.34 | | |
| Machinery and machine tools | 8.36 | | |
| ***Differences Between WPI and CPI*** | | | |
| Fuel group gets a much higher weightage | | Food gets maximum weightage | |
| Services not included | | Miscellaneous group includes services such as transport, education, and health | |

comfort, and safety features. As time goes by, a given brand and vehicle of a given engine size (small, mid-size, or luxury segment) may be available with additional features such as a colour-coded bumper, disc brakes, driver airbags, and a digital odometer. The price index overstates the inflation in the price of the commodity because it records a higher price which is the sum of the old, unimproved good plus the premium for higher improved quality.

To account for differences in consumption patterns across households, the CPI is compiled for industrial workers, urban non-manual employees, and agricultural labourers. In Table 1.11, we list some properties of the CPI for industrial workers, and the WPI, which covers the prices of goods purchased and transacted in mainly by producers of goods. The WPI excludes the prices of all services. The CPI is published by the Labour Bureau of the Ministry of Labour, Government of India.

Table 1.12 provides statistics for the inflation rate calculated using the three indicators of inflation—GDP deflator, WPI, and CPI (industrial workers). As Figure 1.6 (which is drawn on the basis of the data in Table 1.12) depicts, inflation in India has been on the decline since 1994–1995. The CPI (industrial workers) shows the steepest decline in inflation compared with the other two indicators of inflation.

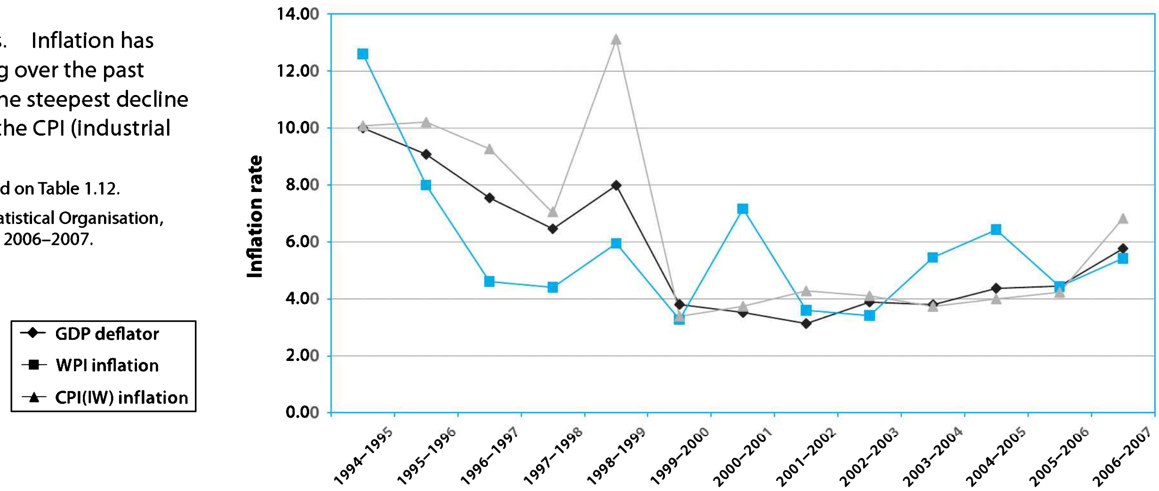
Index numbers do help us to understand whether or not we are better off in terms of real incomes today. However, they are computed by fixing a base year in which the cost of a fixed bundle of goods and services is determined and the cost of this bundle over time is then tracked. But we also have to account for the changes in the available goods and services.

› **Table 1.12**
Inflation Rates based on GDP Deflator, WPI, and CPI (Industrial Workers)

| Year | GDP Deflator | WPI | CPI (IW) |
|---|---|---|---|
| 1994–1995 | 10.00 | 12.60 | 10.08 |
| 1995–1996 | 9.08 | 7.99 | 10.21 |
| 1996–1997 | 7.55 | 4.61 | 9.27 |
| 1997–1998 | 6.46 | 4.40 | 7.02 |
| 1998–1999 | 7.98 | 5.95 | 13.11 |
| 1999–2000 | 3.80 | 3.27 | 3.38 |
| 2000–2001 | 3.52 | 7.16 | 3.74 |
| 2001–2002 | 3.13 | 3.60 | 4.28 |
| 2002–2003 | 3.89 | 3.41 | 4.10 |
| 2003–2004 | 3.80 | 5.46 | 3.73 |
| 2004–2005 | 4.37 | 6.42 | 4.00 |
| 2005–2006 | 4.45 | 4.43 | 4.23 |
| 2006–2007 | 5.76 | 5.42 | 6.83 |
| Average (from 1994–1995 to 2006–2007) | 5.68 | 5.75 | 6.46 |
| Standard Deviation | 2.31 | 2.51 | 3.23 |
| Trend Growth Rate | −6.75 | −3.87 | −7.74 |

*Source*: Central Statistical Organisation, Economic Survey, 2006–2007.

› **Figure 1.6**
Inflation Rates. **Inflation has been declining over the past decade with the steepest decline measured by the CPI (industrial workers).**
This figure is based on Table 1.12.
*Source:* Central Statistical Organisation, Economic Survey, 2006–2007.




One reason why life is better today than it was, say, 25 years ago is not just because average real incomes have risen, but that there are many more goods available. In 1980, people were still using typewriters instead of computers; they were lucky if they got to watch just one channel on a black-and-white television set; and they could not speak with their families and loved ones while driving back from work. There are a greater variety of ways to spend incomes today and measuring your income by correcting for inflation is not going to capture that very important transformation in our lives. Even if someone today has a real income that is lower than that of a person who was in a comparable position 50 years ago, the person has a larger range of products on which to spend their income. As was emphasized in the beginning of this chapter, the way in which we measure income is not a perfect manifestation of the welfare we experience.

## SUMMARY

- Choices made in markets result in macroeconomic outcomes such as aggregate income. Macroeconomics concerns itself with the market for goods, the market for bonds, the market for labour, and the market for money.
- Macroeconomists' interest in aggregate income stems from it being an approximate measure of welfare.
- Measures of aggregate income are incomplete due to under-reported income and participation in underground and non-market activities.
- Gross domestic product, GDP, is a measure of aggregate income and output in an economy. It is the total value of the current production of final goods and services within the national territory during a given period of time.
- In macroeconomics, aggregate output equals aggregate income, which, in turn, equals aggregate expenditure.
- As some citizens work for foreign-owned business organizations and some foreigners work for domestically owned organizations, gross national product (GNP) measures aggregate output by citizens regardless of where in the world they work.
- National income is the sum of aggregate domestic income and net factor incomes from abroad.
- Aggregate expenditure ($Y$) is the sum of domestic expenditure plus net exports (or exports less imports).
- Aggregate domestic expenditure equals aggregate expenditure on consumption ($C$) by households, on investment ($I$) by business units, and aggregate expenditure by government ($G$).

$$Y = C + I + G + X - M$$

- In order to understand whether income increases represent an increase in command over goods and services we differentiate between real income (income measured in constant prices) and nominal income.
- The rate of change in prices or inflation in the price of goods and services is measured by a producer price index such as the wholesale price index (WPI), the consumer price index (CPI), or the GDP deflator.

## NOTES

1. Money includes not just currency but also other assets used for purchases such as savings and current account bank balances. In Chapter 2, we adopt this more inclusive definition of money but here for simplicity we are referring to money as currency.
2. Robert Lane, *The Market Experience* (New York: Cambridge University Press, 1991).

## TEST YOURSELF

1. Economics is the study of choices. In macroeconomics, which markets are studied in relation to these choices?
2. Why does GDP, as reported in the national accounts, not correctly measure total output?
3. Explain how cutting down trees to make furniture may increase a country's GDP but cause economic welfare to decline at the same time.
4. GDP only counts the production of goods in their final form during a given period of time. What are the various components of the final demand for goods?
5. What is the difference between GDP and GNP?
6. Is the balance of merchandise trade a good indicator of a nation's trade position, if the nation has a comparatively large service sector?
7. Why is the output of the private sector measured in market prices and that of the government measured at cost?
8. Why may the GDP deflator overstate the average annual inflation rate for the economy?
9. How can we measure inflation when the prices of different goods change at different rates?
10. Why is the wholesale price index used more widely than the consumer price index as a measure of inflation?

## ONLINE APPLICATION

1. (a) Go to the Web site of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
   (c) Then, under the classification of annual data, click on "Handbook of Statistics on Indian Economy".
   (d) Click on "Table 1: Macroeconomic Aggregates (at Current Prices)" and create a spreadsheet of data on GDP for the 13 data points corresponding to the period from 1993–1994 to 2005–2006.
2. (a) Click on "Table 2: Macroeconomic Aggregates (at Constant Prices)" and access data for GDP at constant prices for the same time period.
   (b) Calculate the GDP deflator for this time period using the formula from the text.
3. Inflation is the percentage change in the price index. Calculate the inflation based on the GDP deflator by taking the percentage change in the GDP deflator price that you have calculated.
4. (a) Click on "Table 39: Wholesale Price Index—Annual Average". Access the data on this price index for the period from 1993–1994 to 2005–2006 with base 1993–1994 = 100.
   (b) Again click on "Table 40: Consumer Price Index—Annual Average". Access the data on this price index for the period from 1993–1994 to 2005–2006 for the consumer price index for industrial workers with base 1982 = 100.
5. Calculate the inflation rate based on the wholesale price index and the consumer price index for industrial workers.
6. Plot the three series of inflation rates over time. Which price index results in higher average inflation rates over time? Why do you think this is so?
7. Plot the data on GDP at constant prices and GDP at current prices over time. Observe whether these curves cross one another. Why do you think that occurs? What do you observe in the graph you have plotted about the relationship between real GDP and nominal GDP?

# *Appendix 1.1* The GDP Deflator

If there were two goods, 1 and 2, and $p$ stands for price, and $q$ for output, then, the prices of the two goods in the base year and current year (year 1) could be as depicted in Table A1.1.1.

**› Table A1.1.1**
Real Incomes with Two Goods (1 and 2)

| | $q_1$ | $p_1$ | $q_2$ | $p_2$ |
|---|---|---|---|---|
| Year 0 | $q_{10} = 2$ | $p_{10} = 1$ | $q_{20} = 4$ | $p_{20} = 1$ |
| Year 1 | $q_{11} = 4$ | $p_{11} = 3$ | $q_{21} = 12$ | $p_{21} = 2$ |

Let $i$ stand for a good ($i$ = 1, 2) and let $j$ denote the years where $j$ takes the value 0 if it is the base year and any positive number for a year that comes later.

Define $C_{jx}$ as the expenditure in year $j$ in the prices of year $x$. Then, the figure for nominal income in year 1 is given by

$$C_{11} = q_{11}p_{11} + q_{21}p_{21} = (4 \times 3) + (12 \times 2) = 36$$

The real income in year 1 is given by

$$C_{10} = q_{11}p_{10} + q_{21}p_{20} = (4 \times 1) + (12 \times 1) = 16$$

Note that the common weights of the prices $p_{i1}$ and $p_{i0}$ in the above expression are the current year's expenditures on the commodities, $q_{ij}$. When prices are weighted by current year quantities, the price index generated is said to be a *Paasche Index*.

The inflation rate of the economy is found by deriving the implicit GDP deflator, which is the current price estimate divided by the constant price estimate. Thus,

$$\text{GDP Deflator} = \left(\frac{C_{11}}{C_{10}}\right)100 = \left(\frac{\text{Nominal GDP}}{\text{Real GDP}}\right)100$$

Now,

$$C_{11} = q_{11}p_{11} + q_{21}p_{21}$$

Adding and subtracting $p_{10}$ in the first part of the right-hand side of the expression for $C_{11}$ and $p_{20}$ in the second part of the expression,

$$C_{11} = q_{11}(p_{10} + p_{11} - p_{10}) + q_{21}(p_{20} + p_{21} - p_{20})$$

This may be rewritten as

$$C_{11} = q_{11}p_{10}\left(1 + \frac{p_{11}}{p_{10}} - 1\right) + q_{21}p_{20}\left(1 + \frac{p_{21}}{p_{20}} - 1\right)$$

Inflation in the price of commodity $i$ measured in year $j$ ($\pi_{ij}$) is defined as the percentage change in prices. We can define this as

$$\pi_{ij} = \frac{\text{Change in price}}{\text{Original price}} = \frac{p_{ij} - p_{i0}}{p_{i0}}$$

**Then,** $$\pi_{11} = \frac{p_{11}}{p_{10}} - 1 \text{ and } \pi_{21} = \frac{p_{21}}{p_{20}} - 1$$

We can then write

$$C_{11} = q_{11}p_{10}(1 + \pi_{11}) + q_{21}p_{20}(1 + \pi_{21})$$

Dividing throughout by the expression for $C_{10}$,

$$\frac{C_{11}}{C_{10}} = \frac{q_{11}p_{10}}{q_{11}p_{10} + q_{21}p_{20}}(1 + \pi_{11}) + \frac{q_{11}p_{10}}{q_{11}p_{10} + q_{21}p_{20}}(1 + \pi_{21})$$
$$= w_1(1 + \pi_{11}) + w_2(1 + \pi_{21})$$

where $w_i = q_{i1}p_{i0}/(q_{11}p_{10} + q_{21}p_{20})$ is the proportion of commodity $i$'s contribution to GDP in the base year, or the weight that commodity $i$ has in the overall output produced. (Note that the sum of the weights equals unity.)

Generally, the GDP deflator involving the expression $C_{jj}/C_{j0}$ in year $j$ is defined as:

$$\text{GDP Deflator} = [w_1 (1 + \pi_{1j}) + w_2 (1 + \pi_{2j}) + \cdots + w_n(1 + \pi_{nj})]100$$

where,

$$w_i = \frac{q_{ij}p_{i0}}{(q_{1j}p_{20} + \cdots + q_{nj}p_{n0})}$$

For the example above:

$$w_1 = \frac{4 \times 1}{(4 \times 1) + (12 \times 1)} = \frac{4}{16} = \frac{1}{4}$$
$$w_2 = \frac{12 \times 1}{(4 \times 1) + (12 \times 1)} = \frac{12}{16} = \frac{3}{4}$$

Also,

$$\pi_{11} = \frac{p_{11}}{p_{10}} - 1 = \frac{3}{1} - 1 = 2 \text{ and } \pi_{21} = \frac{p_{21}}{p_{20}} - 1 = \frac{2}{1} - 1 = 1$$

Plugging these values into the expression for $\frac{C_{11}}{C_{10}}$,

$$\left(\frac{C_{11}}{C_{10}}\right)100 = \left[\frac{1}{4}(1 + 2) + \frac{3}{4}(1 + 1)\right] 100$$
$$= \left(\frac{3}{4} + \frac{6}{4}\right) 100$$
$$= \left(\frac{9}{4}\right)100$$
$$= 225$$

Note that the price of good 1 increased by [(3 − 1)/1] or 200 per cent. The price of good 2 increased by [(2 − 1)/1] or 100 per cent. The weighted average of price increase in the economy is

$$\frac{1}{4}(100 + 200) + \frac{3}{4}(100 + 100) = \frac{300 + 600}{4}$$
$$= \frac{900}{4}$$
$$= 225 \text{ per cent}$$

# 2

# Savings, the Balance of Payments, and the Money Supply

CRITICAL QUESTIONS

- *How is the aggregate savings of an economy used?*
- *What is the balance of international payments and what are its components?*
- *How is the exchange rate determined and how does it affect the financing of the balance of payments?*
- *What is the difference between the monetary base, or reserve money, and the money stock?*
- *How does the interaction between the public, the commercial banks, and the central bank determine the money supply?*
- *What is the money multiplier?*
- *What is an open-market operation and how does it affect the money supply?*

## 2.1 The Sources and Use of Savings

In Chapter 1 we derived the fundamental macroeconomic identity that output equals income equals expenditure (domestic expenditure plus net exports). This is a truism from the perspective of an aggregate economy. You might think that at an individual level income is not equal to expenditure because some of it gets taxed and it is only after some of it gets saved that we get to spend on goods like cars and computers. Yet what is true at an individual level is not true in the aggregate, which is an insight that macroeconomics provides.

From your income a certain amount leaks out as taxes but that is what the government spends. What happens when you save and put your money into a bank account? The bank lends your money to firms who plan to spend on setting up new factories, or to the government which plans to spend on, say, the salaries of professors at the Indian Institutes of Technology and the Indian Institutes of Management. So in the final analysis, even if we do not spend all our income it turns out that all of our income does get spent. From the perspective of an aggregate economy, therefore, all income is spent.

To probe this perspective further, we begin with the identity that income in the aggregate equals expenditure, which we wrote in Chapter 1.

$$Y = C + I + G + X - M$$

The term $X - M$ here denotes the balance of merchandise trade plus net receipt from non-factor service income. As our interest is currently on savings, which is due to the current income (from all sources) being set aside rather than being spent, we expand the identity above to include the net factor incomes earned by residents of the country from abroad, NFI, and transfers (mainly remittances) from abroad. Adding net factor incomes and transfers from abroad to both sides,

$$Y + \text{NFI} + \text{Transfers from Abroad} = C + I + G + X - M + \text{NFI} + \text{Transfers from Abroad}$$

Taking the terms in consumption and government expenditure on the left-hand side,

$$Y + \text{NFI} + \text{Transfers from Abroad} - C - G = I + (X - M + \text{NFI} + \text{Transfers from Abroad})$$

However,

$$(X - M + \text{NFI} + \text{Transfers from Abroad}) = \text{CA}$$

where the left-hand side is the net exports of goods and services (or non-factor incomes) plus the net factor incomes and transfers from abroad and this is the definition (see Chapter 1) of the term on the right-hand side, the current account surplus (CA). We may then write,

$$Y + \text{NFI} + \text{Transfers from Abroad} - C - G = I + \text{CA}$$

To connect the savings in the economy with this identity, we need to define private savings and public savings.

› Private savings is the value of the private sector's disposable income minus its consumption.

PRIVATE SAVINGS is the income left over after the private sector spends whatever income is available for spending. The private disposable income is the amount of income the private sector has available to spend. This equals income received from private sector activities plus payments received from government less taxes paid to government.

$$\begin{aligned}\text{Private Disposable Income} = \text{GDP} &+ \text{Net Factor Incomes Abroad}\\ &+ \text{Transfers from Abroad}\\ &+ \text{Transfers Received from Government}\\ &- \text{Taxes}\end{aligned}$$

Recollect that the net factor incomes from abroad is the income paid to domestic factors of production by the international economy less income paid to foreign factors of production by the domestic economy. Transfers from abroad are mainly remittances by Indians working abroad. The transfers from government include subsidies and interest payments on government borrowings or the public debt.

As consumption is spending by the private sector to meet current needs, savings of the private sector is defined as private disposable income minus consumption.[1]

$$\text{Private Savings} = S_{\text{pvt}} = \text{Private Disposable Income} - \text{Consumption}$$

› Transfers by government include subsidies and interest payments on government borrowing from the private sector.

The income of the government is the taxes paid by the private sector to the government. The net government income is the tax revenues of government less the TRANSFERS it makes.

$$\text{Net Public Sector (Government) Income} = \text{Taxes} - \text{Transfers (to private sector) by Government}$$

› Government savings is equal to net government revenues less government spending.

GOVERNMENT SAVINGS then is the net government income less government purchases of goods and services for the current needs (government consumption expenditure).

In most macroeconomic models, the government purchases are the sum of the government consumption and investment expenditure and are treated as though they are entirely expenditures on current needs. In that case we need not differentiate between the kinds of government expenditure that we take to be current government outlays.

$$\begin{aligned}\text{Government Saving} = S_{\text{govt}} &= \text{Net Government Income} - \text{Government Expenditure}\\ &= \text{Taxes} - \text{Transfers by Government}\\ &\quad - \text{Government Expenditure}\end{aligned}$$

A commonly used term for government savings is the government fiscal surplus. The fiscal surplus equals the government non-debt related receipts comprising taxes net of transfers minus government outlays.[2]

› The sum of private and public savings is national saving.

NATIONAL SAVING or the saving of the economy is equal to private saving plus government saving.

$$\begin{aligned}\text{Saving} = S_{\text{pvt}} + S_{\text{govt}} = (\text{GDP} &+ \text{Net Factor Incomes Abroad}\\ &+ \text{Transfers from Abroad} + \text{Transfers from Government}\\ &- \text{Taxes} - \text{Consumption})\end{aligned}$$

$$+ \text{ (Taxes} - \text{Transfers by Government}$$
$$- \text{ Government Expenditure)}$$
$$= Y + \text{NFI} + \text{Transfers from Abroad} - C - G$$

Since we had written the aggregate income equals expenditure identity as $Y$ + NFI + Transfers from Abroad $-C - G = I +$ CA, we can substitute this into the expression for national savings,

$$\text{Savings} = I + \text{CA}$$

or,
$$S_{pvt} + S_{govt} = I + \text{CA}$$

or,
$$S_{pvt} + S_{govt} - \text{CA} = I$$

What this means is that the investment expenditure in the economy is financed through three sources:

1. Private savings
2. Government savings
3. Foreign savings.

A current account deficit ($M > X$ + NFI + Transfers from Abroad) implies that what foreigners receive from us (as monetary payment when we import goods and services from them) is more than the amount that we receive as payments to us on the Indian goods and services they demand (Indian exports). The money coming into the Indian financial system from abroad as payment for exports and returns to factors of production (NFI) as well as remittances is insufficient to pay for the import demand of residents. The country then requires to raise the money required to finance the excess imports from abroad. It does this by borrowing abroad. A CURRENT ACCOUNT DEFICIT thus requires borrowing from abroad, which means that domestic residents are utilizing the savings of foreigners to finance expenditures. A current account deficit is thus financed by an increased indebtedness abroad and represents the savings of foreigners.

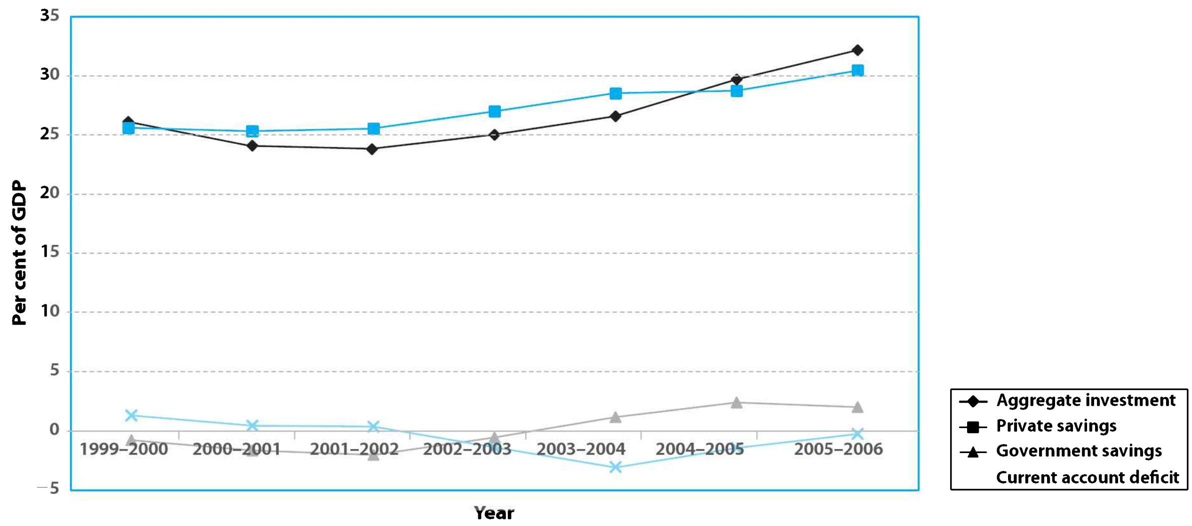
› A current account deficit implies that a country's payments to the rest of the world exceed the payments received from the rest of the world. This deficit in the payments to be made to the rest of the world increases the country's indebtedness abroad and is plugged via the use of savings of foreigners to finance it.

Aggregate investment in the economy is, therefore, equal to aggregate savings. Aggregate savings, in turn, is the sum of national savings—private savings ($S_{pvt}$) plus government savings ($S_{govt}$)—and foreign savings ($-$CA).

The aforementioned identity that investment is financed through three sources of savings is always true by definition. Being an identity, a reduction in any component of the identity must result in an increase in another component or a combination of components so that the identity continues to hold true. Hence, if the government decides to splurge and reduces its savings—the government deficit increases—at least one of three things or a combination of them must occur: private saving rises, investment decreases, and/or the current account balance deteriorates.

The accompanying Figure 2.1 graphs data on the sources of savings and aggregate investment and depicts two phases of the Indian economy. Prior to 2002–2003 government dis-saving was accompanied with a current account surplus and private savings were larger than aggregate investment. Indian savings were being used by the rest of the world to finance its current account deficit with India.

After 2002–2003, public-sector savings improved and were positive. Private savings also increased. However, this increase in the national savings was insufficient to finance the sharp rise in investment in the economy from




› **Figure 2.1**
Sources of Savings and Investment. **Aggregate investment is financed by private savings, government savings, and the current account deficit (foreign savings). Despite national savings increasing the increased investment has resulted in the economy increasingly accessing foreign savings to finance capital formation in the economy.**

25 per cent of the GDP in 2002–2003 to 32.2 per cent of the GDP in 2005–2006. The economy resorted to foreign savings to fill the gap and the current account went into deficit from 2002–2003. This means that in the recent past Indians were borrowing from foreigners to finance the capital formation in the economy.

A change in the current account of an economy is associated with a change in the capacity to hold assets of the residents of a country as compared with the residents of other nations. With each flow of goods, services, and income[3] that comprises the transactions reflected in the current account, there will be an associated outflow of domestic assets abroad and inflow of foreign assets into the domestic economy. For instance, the export of software services by India will result in an inflow of the foreign exchange as payment received for such exports.

These flows of financial assets are recorded separately in an account that depicts the income and asset transactions between domestic residents and residents of other nations and is called the balance of payments accounts of a country. The increase in the stock of foreign exchange by the economy will also get reflected in the asset and liabilities position of the banking system since that is where these assets will be held. That, in turn, will affect the money supply in the economy. We thus turn to a study of the balance of payments and the money supply of an economy.

## 2.2 The Balance of Payments

› The balance of payments summarize a country's transactions—monetary payments for goods and services, and acquisition of financial assets—with the rest of the world.

The BALANCE OF PAYMENTS accounts are a summary of the nation's financial relationship with the rest of the world. It records the market value of goods, services, and financial assets that domestic residents exchange with foreigners from other nations during an accounting period. There are three balance of payments concepts—the balance on current account, the balance on capital account, and the overall balance.

› The current account is the summary of the country's payments to and from the rest of the world.

The CURRENT ACCOUNT measures the flow of goods, services, and income across national borders. The flow of income recorded in this account is of the interest and dividend payments received by residents from the holding of financial assets abroad less similar payments made to foreign residents, and also international transfers.

The CAPITAL ACCOUNT measures the capital outflow and the capital inflow into the economy. A capital outflow is the purchase of foreign assets by domestic agents such as households and firms. A capital inflow is the purchase of domestic assets by foreign households and firms.

› The capital account is a summary of the country's financial asset transactions with the rest of the world.

The OVERALL BALANCE account records the transactions of reserve assets such as gold and foreign currencies by official government agencies.

› The overall balance records the changes in a country's stock of international reserves during the year. International reserves are the foreign currency and other assets held by the government for the purpose of purchasing the domestic currency in the foreign-exchange market.

The balance of payments is a bookkeeping system that records both sides of any transaction involving a domestic resident and a resident abroad. A debit entry is made if a transaction results in a domestic resident making a payment abroad. A credit entry is made if a transaction results in a domestic resident receiving a payment from abroad.

If the total of debit entries or imports recorded in the current account exceeds (is less than) the sum of the credit entries or exports and net factor incomes from abroad, then, there is a deficit (surplus) in the current account.

Similarly, the purchase of domestic financial assets (bonds, stocks) by foreign residents is a capital inflow, which is a credit entry. A purchase of foreign financial assets by domestic residents is a capital outflow and is a debit entry. When the debit entries exceed (are less than) the credit entries—capital outflows are greater (lesser) than capital inflows—there is a net capital outflow (inflow) or a deficit (surplus) in the capital account.

The sum of the current account balance and capital account balance is the OVERALL BALANCE that states a net debit or credit. If it is a net credit, as has been the case in some years in India in the recent past, then, domestic residents are receiving more payments from abroad. This surplus in the overall balance requires the central bank to accumulate the excess foreign assets and add to its holding of foreign currencies as foreign-exchange reserves.

› The **overall balance** is the sum of the current account balance and the capital account balance. A surplus overall balance is the net increase in the country's stock of international reserves over the year. A deficit overall balance is the net decline in the country's stock of international reserves over the year.

The accounting of the balance of payments thus is

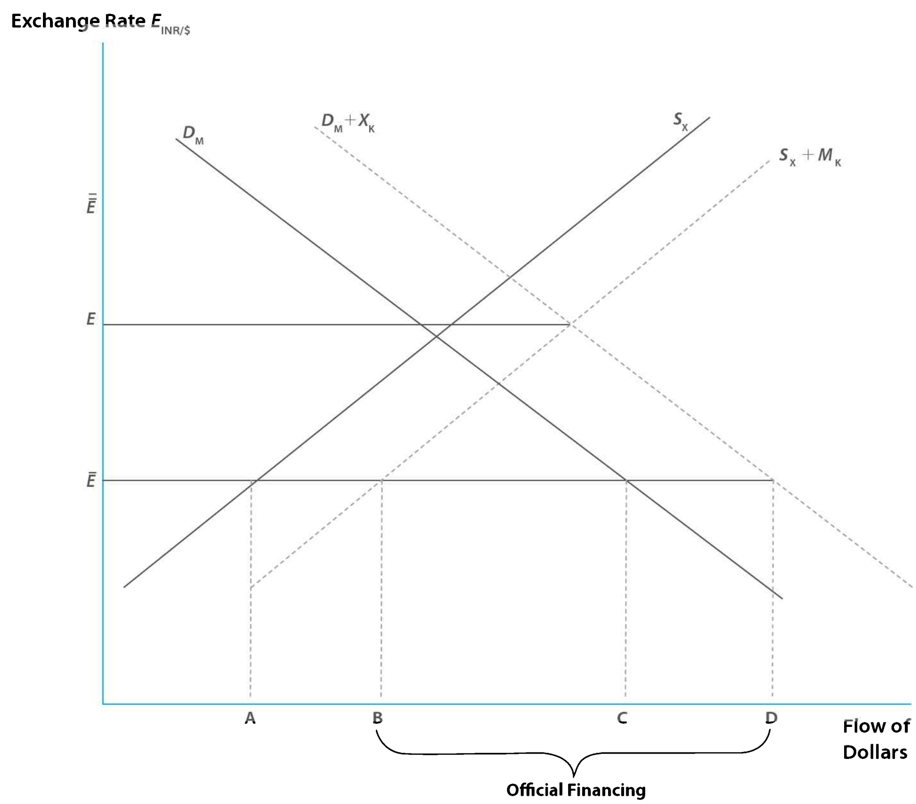
$$\begin{aligned}
\text{Current Account Balance} &= \text{CA}\\
\text{Capital Account Balance} &= \text{KA}\\
\hline
\text{Overall Balance} &= (\text{CA} + \text{KA})\\
\hline
\text{Official Reserves, R} &= -(\text{CA} + \text{KA})
\end{aligned}$$

The overall balance is achieved through official financing—the authorities absorb the finance required for overall balance by accumulating foreign-exchange reserves or making a deposit with monetary authorities overseas such as the International Monetary Fund when there is an overall surplus. This accumulation of reserves is shown as a negative value of the sum of the capital and current accounts. A negative value to the reserves thus represents an increase in the official reserves held and a positive value is a decrease in the holding of official reserves by the central bank.

To understand this role of the government whereby it absorbs or releases foreign exchange in the foreign-exchange market, we employ the following diagram—Figure 2.2. A foreign-exchange market is where commercial banks, foreign-exchange dealers, and the central bank buy and sell the currency of other nations. The exchange rate is the market price at which foreign exchange is bought and sold. The EXCHANGE RATE expresses the value of a given currency relative to another. In Figure 2.2, the vertical axis depicts the exchange rate measured as the number of units of domestic currency (INR) per US dollar, $E_{\text{INR/\$}}$. The exchange rate thus expresses how many domestic currency units

› The exchange rate is the value of a given currency relative to another.

› **Figure 2.2** Exchange Rate and Balance of Payments. **The demand for foreign exchange from importers and due to capital outflows and the supply of foreign exchange by exporters and due to capital inflows determines the exchange rate and the balance of payments.**




› A **spot exchange rate** is the market price for immediate purchase and delivery of foreign exchange. The forward-exchange rate and real exchange rate are studied in Chapter 5.

› Depreciation occurs when the price of the foreign currency in terms of the domestic currency rises—an increase in the exchange rate.

› Appreciation occurs when the price of the foreign currency in terms of the domestic currency decreases—a decrease in the exchange rate.

(INR) it takes to purchase one dollar, or, rupees per US dollar. The exchange rate we are concerned with here is the SPOT MARKET EXCHANGE RATE—the rate for immediate purchase and delivery of the foreign currency.

When the number of rupees required to purchase one US dollar increases, we say that the rupee price of the dollar rose, or, the rupee has DEPRECIATED relative to the US dollar. When the number of rupees required to purchase a dollar decreases, we say that the rupee price of the dollar has fallen, or the rupee has APPRECIATED relative to the US dollar.

In Figure 2.2, higher values of $E$ represent a depreciation of the currency and lower values an appreciation. On the horizontal axis, we measure the flow of dollars per period (a year).

Now, take the case of an economic agent who imports goods and services and is thus required to pay the supplier of those imports in a foreign currency, let us say, dollars. The importer goes to the foreign-exchange market with his rupees that he wants converted into dollars—he has a demand for dollars in the foreign-exchange market. A depreciation of the currency requires the importer to give up more rupees for the dollars that are required to pay for the imports of goods and services. We should then expect that depreciation by raising the foreign-exchange cost of imports would reduce the demand for imports. An appreciation would have the opposite effect. The demand for dollars as a result of the demand for imports of goods and services is, therefore, downward sloping and depicted as $D_M$. Similarly, exporters supply dollars that they receive from customers abroad in exchange for rupees in the foreign-exchange market. As the rupee depreciates, foreigners will find Indian goods cheaper, exports will increase in volume and value, and the supply of dollars by exporters will be an upward-sloping line as depicted by $S_X$. If the exchange rate is fixed by the government at $\bar{E}$, then the current account balance is the distance AC and is in deficit as drawn.

We now take into consideration the capital account. If an economic agent wishes to export capital, say, to buy shares on the New York Stock Exchange, he supplies rupees to be converted to dollars. From the viewpoint of the impact on currency demand, this is like an import of goods and services in that there is a demand for dollars. We thus add the capital export $X_K$ to the $D_M$ schedule and denote the total demand for dollars as $D_M + X_K$. Conversely, we may have agents overseas who want to buy stock on the National Stock Exchange in India—they supply dollars in exchange for the rupees that will allow them to buy the shares. The impact on foreign currency supply is like that of an export of goods and services. We add the capital import, $M_K$, to the $S_X$ schedule and denote the total supply of dollars as $S_X + M_K$. The amount of capital export at exchange rate $E$ is CD and the distance AB measures the capital import. The capital account deficit is therefore CD − AB.

Combining the deficit on current account (AC) and capital account (CD − AB) we obtain an overall deficit of AC + (CD − AB), or, BD. This is the amount that government acting through its central bank—the Reserve Bank of India (RBI) in the case of India—is required to finance from reserves or by borrowing from overseas monetary authorities. In Figure 2.2, the excess demand for foreign exchange, BD, arising from transactions on the current and capital account is what the RBI would be required to finance. It does so by selling rupees in exchange for dollars, by decreasing its holdings of foreign-exchange reserves, or by asking a foreign monetary authority such as the Federal Reserve Board in the United States to buy rupees on its behalf and by incurring a debt to the Federal Reserve. If the exchange rate were at

**MACROFOCUS 2.1**

### Forex Reserves: Index of Economic Strength?

Towards the end of 2005, Japan was the largest foreign-exchange reserve holder in the world, valued at US $843.5 billion, with China in second place. However, China's ratio of foreign-exchange reserves to GDP is 37 per cent, whereas that of Japan is 18 per cent. India holds the fifth largest stock of reserves among emerging market economies. It is often claimed that a greater stock of foreign-exchange reserves reflects a country's economic strength and an enhancement of its international status. Is such thinking justified?

In order to determine a rational amount of foreign-exchange reserves, a policymaker must take into account the extent of foreign trade, the utilization of foreign investment and the capacity to finance internationally, the level of economic opening up, and the associated efficiency of adjustment and regulation, among other things. A developed economy may exhibit a lower demand for foreign-exchange reserves because its currency is convertible and can be used for international payments. In contrast, an emerging economy's currency is not generally acceptable for international use because of doubts about the efficacy of their macroeconomic regulatory framework and such countries usually have a stronger demand for foreign-exchange reserves. Besides, an emerging economy needs reserves to handle contingencies such as financial risks due to insufficient liquidity with financial institutions and due to its high requirement of oil that is subject to price fluctuations in the international market.

Foreign-exchange reserves are accumulated by continual surpluses in the current and capital accounts of the balance of payments. Reserves emanating from a surplus in the current account are mainly due to the capacity of domestic firms to export products in international markets and compete with other countries. This reflects economic supremacy only to the extent that exports do not comprise resource-based primary products that give a quantity-based advantage, and include more products that are of high quality and have high value added. To some extent, reserves are accumulated to keep the exchange rate stable and protect the export competitiveness of domestic firms, as has happened in China. In such cases, expectations are generated towards revaluation of the currency and that results in an influx of "hot money" from speculators betting on the revaluation. Foreign-exchange reserves, then, are the result of a policy objective of weakening exchange rates so as to steer the economy's external demand. To the extent that it results in speculative flows, it increases the likelihood of financial crisis and increases the riskiness of the economy rather than strengthening it.

To sum up, reserves that come from the capital account surplus are largely debt or equity based (foreign direct investment or FDI and foreign institutional investment, or FII flows) and result in an increased indebtedness to the rest of the world. They are an indicator of economic power only to the extent that they finance investments that raise the productivity and efficiency of the economy sufficiently such that the returns after compensating these capital inflows increases the incomes received by domestic citizens.

› **Table 2.1**
India's Overall Balance of Payments (INR Billion)
*Source:* Reserve Bank of India, Annual Reports.

| | 1999–2000 | 2002–2003 | 2003–2004 | 2004–2005 | 2005–2006 |
|---|---|---|---|---|---|
| **Current Account Balance** | −203.31 | 306.60 | 639.83 | −247.86 | −468.56 |
| Capital Account | | | | | |
| Foreign Investment (Net) | 225.01 | 200.98 | 628.42 | 542.60 | 807.59 |
| (a) FDI | 93.96 | 155.94 | 109.44 | 146.39 | 254.24 |
| (b) Portfolio Investment | 131.05 | 45.04 | 518.98 | 396.21 | 553.35 |
| Commercial Borrowings and External Assistance | 69.51 | −185.70 | −196.13 | 479.38 | 204.23 |
| Banking Capital (Including NRI Deposits) | 92.17 | 503.33 | 277.82 | 170.83 | 57.95 |
| **Total Capital Account Balance**[a] | 481.01 | 523.66 | 772.27 | 1,383.68 | 1,085.21 |
| **Overall Balance**[b] | 277.70 | 820.37 | 1,439.93 | 1,159.07 | 658.96 |
| Monetary Movements | −277.70 | −820.37 | −1,439.93 | −1,159.07 | −658.96 |
| (1) IMF, Net | −11.22 | 0 | 0 | 0 | 0 |
| (2) Foreign-Exchange Reserves (Increase −, decrease +) | −266.48 | −820.37 | −1,439.93 | −1,159.07 | −658.96 |

[a]Inclusive of errors and omissions.
[b]Overall Balance = Current account balance + Capital account balance
Figures for 2004–2005 are partially revised and those for 2005–2006 are preliminary.

› A managed or dirty float is a system of flexible exchange rates but with periodic intervention in the foreign exchange market by official agencies.

$\bar{\bar{E}}$ instead of $E$, then the overall balance BD would be in surplus and the RBI could just accumulate the excess supply of foreign exchange as reserves. The alternative to decreasing the holding of foreign-exchange reserves or running up a debt to a foreign monetary authority when the exchange rate is $\bar{E}$ is to stop controlling the exchange rate and to let it float. The exchange rate would depreciate to $E$ where the overall balance would be zero. Most economies, however, do not "free float" but pursue a "MANAGED" or "DIRTY FLOAT" in which official financing of the overall balance goes on.

Table 2.1 depicts India's balance of payments in the last few years. Some items in the capital account such as rupee debt servicing have not been reported here. The current account swung from a deficit into a surplus from 2001 to 2002. Direct-investment flows and portfolio flows provided support to the balance of payment with foreign-exchange reserves rising sharply in the last few years. In 2004, for instance, India's share of global portfolio flows to emerging market economies was about 25 per cent and by December 2004, India's foreign-exchange reserves were the fifth largest in the world.

The purchase and sale of foreign exchange is often mediated through the banking system and affects the asset and liability position of commercial banks and the central bank. We turn to the balance sheet of the banking system and the determination of the money supply in Sections 2.3 and 2.3.1.

## 2.3 The Money Supply

The money supply and its formation rest on a set of identities. The money-supply process deals with the interface between the central bank and the commercial banks.

The creation of the MONETARY BASE is dominated by the central bank. The central bank exercises control over the economy because of its unique position as a monopolistic supplier of the monetary base. In all countries, this is guaranteed by the central bank's monopoly over the supply of banknotes and coins.

› The money issued by the central bank is called the monetary base or high-powered money.

The creation of the MONEY STOCK is determined by the interplay between the central bank, the commercial banks, and the non-bank private sector. We can think of the monetary base as an input factor, controlled by the central bank, for the process by which the money stock is created as an output by the commercial banks. The three major participants involved in the determination of the money supply are

› The money stock is the market value of a sum of liquid assets that are used to make transactions in the economy.

1. The central bank, which determines the monetary base, the reserve requirements, and sets the discount rate at which it lends to commercial banks.
2. The public, which determines its currency holdings relative to its demand deposits.
3. The commercial banks, which for a given required reserve ratio, determine their actual holding of reserves as against their demand deposit liabilities.

### 2.3.1 Bank Balance Sheets and the Money Supply

The money-supply process is best understood by looking at the balance sheets of the central bank and the commercial banks. If we assume a closed economy, a typical central bank balance sheet would be as given in Table 2.2.

Some points that we need to consider include the following:

1. Banks often borrow from each other bilaterally and often do so in the context of an overnight loan market when there is a requirement for liquidity to meet RESERVE REQUIREMENTS. Borrowing by a bank from other banks within the system does not change the monetary base and is therefore ignored in the determination of the money supply. However, when commercial banks increase their borrowing from the central bank as a whole—credits to domestic banks by the central bank—the monetary base increases and the money supply expands. In lending to the commercial banks as a whole, the

› Bank reserves are the holdings of central bank money by banks. This is the difference between the amount of deposits in a bank and what it lends or holds as bonds. Central banks require banks to hold a minimum amount of reserves as a proportion of their deposits.

**› Table 2.2**
A Central Bank Balance Sheet

| Assets | Liabilities |
|---|---|
| Credits to Domestic Banks (Credit Facility + Open Market Operations) | Currency in Circulation = Currency with Public ($C^P$) + Currency with Banks ($C^B$)<br>Currency in Circulation $C = C^P + C^B$ |
| Central Bank Credit to Government = Holdings of Government Debt (Including Loans to Government) *less* Government Deposits $G_{b/CB}$ | Reserves and Central Bank Balances of Domestic Banks |
| Advances (Credits) to Commercial Private Sector $A_{CB}$ | Net Non-monetary Liabilities of Central Bank[a] $NML_{CB}$ |

[a]Non-monetary liabilities include contingent liabilities, guarantees, and other claims, *less* fixed and other assets.

› The central bank as a lender of last resort stands ready to lend to any temporarily illiquid bank that is in requirement of funds.

› The direct lending of reserves by the central bank to commercial banks is the credit facility window of the central bank or the discount window as it is sometimes known.

› An open-market operation is the purchase or sale of government bonds by the central bank for the purpose of increasing or decreasing the money supply. For more details see Section 2.3.3.

central bank is said to act as a LENDER OF LAST RESORT when the banking system is not in a position to obtain additional funds from its own internal borrowing and lending. Since borrowing from the central bank under a CREDIT FACILITY—as occurs under the refinancing facility—often triggers greater supervision from the central bank into the borrowing bank's asset management and other practices, it is rarely practised. However, the central bank lends to the commercial banks via an auction procedure or OPEN-MARKET OPERATION (an issue we discuss a little later in Section 2.3.3).

2. Currency in circulation ($C$) is the sum of currency with the public ($C^P$) and currency with banks ($C^B$), that is, $C = C^P + C^B$. Note that money that the central bank has in its own account is neither part of reserves nor of the money supply.
3. When we make a deposit into a commercial bank, it can use this deposit to purchase financial assets such as government bonds or lend it to fund projects. Reserves are the deposits the bank has not used to make loans or to purchase financial assets. The cash kept by a bank in its vault to meet daily business needs is one part of its reserves. Another part is held as a deposit with the central bank. Reserves then comprise the cash in the vault and the reserves held as deposits at the central bank. Commercial banks are required to hold reserves as a fraction of deposits, at levels imposed by the central bank. By requiring banks to hold larger reserve requirements, the central bank can affect the level of economic activity as this makes lesser deposit money available for lending or purchasing financial assets. REQUIRED RESERVES are essentially all deposits that are required to be held by banks in the country with the central bank as a statutory requirement. The demand for central bank balances by the commercial banks originates primarily because the central bank imposes statutory minimum-reserve requirements. Any balances held above these requirements are regarded as excess reserves.
4. Net non-monetary liabilities ($NML_{CB}$) include contingent liabilities, guarantees and other claims, less fixed and other assets. These, as we shall see, are netted out from the balance sheet as the focus is on net monetary liabilities that affect the money supply.

› Required reserves are the result of the central bank regulation requiring banks to hold specified fractions of their deposits as funds on deposit at the central bank.

› The monetary base or reserve money is the central bank's liabilities of currency and commercial bank reserves held with it, or, from the assets perspective, the credit given to commercial banks, government, or the private sector less non-monetary liabilities.

The MONETARY BASE of the central bank is also referred to as RESERVE MONEY and can be written in terms of its application (liabilities side) as

**Reserve Money = Currency in Circulation + Reserves and Central Bank Balances of Domestic Banks**

The monetary base could also be written in terms of its source (assets side) as

**Reserve Money = Credits to Domestic Banks**
**+ Central Bank Credit to Government and Private Sector**
**− Net Non-monetary Liabilities**

**› Table 2.3**
Commercial Banks' Balance Sheet

| Assets | Liabilities |
|---|---|
| Reserves and Deposits with Central Bank | Deposits $D_R$ |
| Commercial Banks Credit to Government = Holdings of Government Debt (Including Loans to Government) *less* Government Deposits $G_{b/B}$ | Loans from Central Bank |
| Advances (Credit) to Private Commercial Sector $A_B$ | Net Non-monetary Liabilities $NML_B$ |
| Currency with Banks (Vault Cash) $C^B$ | |

› **Table 2.4**
Private Sector Balance Sheet

| Assets | Liabilities |
|---|---|
| Bank-related Financial Assets | Wealth (= equity) |
| (a) Deposits ($D_R$) | |
| (b) Currency with Public ($C^P$) | |
| Other Financial Assets | Advances (Loans) from Banks $A_B + A_{CB}$ |
| Real Assets | Other Financial Liabilities |

Similarly, a typical commercial bank's balance sheet would be as given in Table 2.3.

A few properties of the commercial bank's balance sheet that need to be noted are:

1. The central bank requires commercial banks to meet a certain minimum ratio—called the required reserve ratio—of their reserves to their deposits. Central banks put in place required reserves with an eye to controlling the money supply and through that the level of economic activity. Central banks usually assess reserve requirements of banks on an average basis with banks being required to meet their reserve requirements only on average over a period of a fortnight or so.
2. A part of the reserves is held as deposits with the central bank. Another part of the reserves is held in cash at the tills, in the automatic teller machines, or in the banks' vault ($C^B$). These reserves are held because demand deposits may be withdrawn on demand at any time. In anticipation of its deposit liabilities, the commercial bank makes a decision on its reserve holdings in the form of cash holdings. As these reserves are at the banks' disposal, they are sometimes referred to as free reserves. Free reserves are decided on by the commercial bank depending on the withdrawal patterns of its depositors while required reserves are imposed by the central bank.

For the sake of completeness, the balance sheet of the private sector, which we do not discuss in detail here, is as depicted in Table 2.4.

If we CONSOLIDATE the balance sheets of the central bank and all the commercial banks, we will get the consolidated balance sheet of the banking system showing only claims and liabilities with respect to non-banks. The aggregated balance sheet of the banking system is as given in Table 2.5.

› Consolidating requires that all claims and liabilities between individual commercial banks and between commercial banks and the central bank are netted out. For instance, loans from central bank in Table 2.3 nets out with credits to domestic banks in Table 2.2.

In terms of an identity, this may be written as

$$C^P + D_R = G_b + A_R - \text{NML}$$

Let us now relax the assumption that we are dealing with a closed economy. Then, the *additional* terms on the balance sheets of the central bank, commercial banks, and the consolidated balance sheet of the banking system will be as depicted in Table 2.6.

› **Table 2.5**
Consolidated Balance Sheet of a Banking System

| Assets | Liabilities |
|---|---|
| Bank Credit to Government $G_b = G_{b/CB} + G_{b/B}$ | Currency in Circulation with Public $C^P$ |
| Advances (Credit) to Private Commercial Sector $A_R = A_{CB} + A_B$ | Deposits $D_R$ |
| | Net Non-monetary Liabilities $\text{NML} = \text{NML}_{CB} + \text{NML}_B$ |

**› Table 2.6A**
Additional Items on Central Bank Balance Sheet

| Assets | Assets |
|---|---|
| Foreign-exchange Reserves (Including Gold Holdings) $R_{FE}$ | |

**› Table 2.6B**
Additional Items on Commercial Banks' Balance Sheet

| Assets | Assets |
|---|---|
| Advances (Credit) to Non-residents $A_{NR}$ | Deposits of Non-residents $D_{NR}$ |
| Net Foreign Currency Assets of Banks $B_{FE}$ | |

**› Table 2.6C**
Additional Items on Consolidated Balance Sheet of Banking System

| Assets | Liabilities |
|---|---|
| Foreign-exchange Reserves (Including Gold Holdings) $R_{FE}$ | |
| Advances (Credit) to Non-residents $A_{NR}$ | Deposits of Non-residents $D_{NR}$ |
| Net Foreign Currency Assets of Banks $B_{FE}$ | |
| Foreign-exchange Assets (Net) of Banking Sector $\mathrm{FE} = R_{FE} + B_{FE} + (A_{NR} - D_{NR})$ | |

In Table 2.6C, the net position is stated in the last row. The holding of foreign-currency deposits have to be netted off ($A_{NR} - D_{NR}$) because the conventional definition of money supply includes only domestic (rupee) bank deposits.

Adding the additional items for an open economy on the balance sheet to the items in the balance sheet for a closed economy (consolidating Tables 2.5 and 2.6C) gives us the consolidated balance sheet of the banking system in an open economy—Table 2.7.

In terms of an identity, the consolidated balance sheet may be written as

$$C^P + D_R = G_b + A_R + \mathrm{FE} - \mathrm{NML}$$

› The money supply is the currency with the public and the deposits at commercial banks. Alternatively, it is the bank credit given to the government and the private sector, and the foreign-exchange assets held by banks minus their non-monetary liabilities.

The MONEY SUPPLY is thus equal to

$$M^S = C^P + D_R = G_b + A_R + \mathrm{FE} - \mathrm{NML}$$

In the event that there is an expansion of money supply due to foreign-exchange market intervention or huge capital inflows due to foreign

**› Table 2.7**
The Consolidated Balance Sheet of a Banking System in an Open Economy

| Assets | Liabilities |
|---|---|
| Bank Credit to Government $G_b = G_{b/CB} + G_{b/B}$ | Currency in Circulation $C^P$ |
| Advances (Credit) to Private Commercial Sector $A_R$ | Deposits $D_R$ |
| Foreign-exchange Assets (Net) of Banking Sector $\mathrm{FE} = R_{FE} + B_{FE} + (A_{NR} - D_{NR})$ | Net Non-monetary Liabilities $\mathrm{NML} = \mathrm{NML}_{CB} + \mathrm{NML}_B$ |

› **Table 2.8**
Sources of Money Stock (INR Billion)
*Source:* Reserve Bank of India, Annual Reports.

| | 31 March 2002 | 31 March 2003 | 31 March 2004 | 31 March 2005 | 31 March 2006 |
|---|---|---|---|---|---|
| Net Bank Credit to Government ($G_b$) | 5,895.65 (39.3%) | 6,765.23 (39.4%) | 7,429.04 (38.6%) | 7,579.06 (33.6%) | 7,690.93 (28.2%) |
| Net Bank Credit to Commercial Sector ($A_R$) | 7,596.47 (50.7%) | 8,989.81 (52.3%) | 10,161.51 (50.7%) | 12,805.40 (56.8%) | 16,909.61 (62.0%) |
| Net Foreign-exchange Assets (FE) | 3,110.35 (20.8%) | 3,937.15 (22.9%) | 5,265.86 (26.3%) | 6,492.55 (28.8%) | 7,261.94 (26.6%) |
| Net Non-monetary Liabilities (NML)[a] | 1,618.92 (10.8%) | 2,512.59 (14.6%) | 2,799.65 (14.0%) | 4,337.62 (19.2%) | 4,567.12 (16.7%) |
| Money Supply | 14,983.55 | 17,179.60 | 20,056.76 | 22,539.38 | 27,295.36 |

[a]Government's currency liabilities to the public have been netted out of these figures.
Figures in parentheses are percentages to total money supply. Data from 2003 are provisional.

investment—portfolio or FDI—the central bank can offset this expansion in the money supply (brought about by a rise in FE) by reducing its lending to the government ($G_{b/CB}$). This procedure is known as STERILIZATION. Over the last few years, net Reserve Bank of India (RBI) credit to the central government has been declining sharply as the central bank has sterilized the accumulated foreign-exchange reserves from capital inflows. In the late 1980s, for instance, central bank lending to the government was 90 per cent of the reserve money and by the end of March 2004, this had declined to 8.5 per cent. The sources of money stock for the last few years are given in Table 2.8.

› Sterilization is the central bank policy of altering the domestic credit extended by it in an equal and opposite direction to the variation in foreign exchange reserves so that the monetary base remains unchanged.

MACROFOCUS 2.2

## Does Money Have a Future?

The central bank has a monopoly in the supply of bank reserves. In order to transact, individuals require money and banks can only cater to this if they hold reserves in parallel to their outstanding deposits. For many goods and services today neither cash in one's wallet nor an adequate balance in the bank account is necessary at the time of purchase. Credit cards and other electronic cash have contributed to this trend. In Hong Kong you can buy fast food with an Octopus store value card and in Finland you can pay for a car wash by using your mobile. Even if you do not have the cash up front when undertaking a transaction, you need it for the final settlement of the transaction. If you have used a credit card for instance, you must make a payment by transferring money out of your bank account. You still have a need for money balances. The central bank, which influences the money supply through the control of reserves, is still relevant.

What if a telephone service provider (TSP) sells you an advance payment smart card? You would still pay for it using cash or a bank cheque. But suppose now the telephone service provider convinces merchants to accept its smart cards in payment for purchases, which is a growing possibility according to Friedman.* You could then walk into a store and pay for merchandise using the telephone service's smart card. If the TSP settled with merchants by transferring balances with banks, then they would be providing a service akin to that of a credit-card provider. However, if it could convince merchants that its smart card is widely used, then the merchants would be willing to accept balances on the books of the TSP as payments for transactions undertaken with the smart card. The merchant runs up a credit with the TSP and uses this credit to square off debits it incurs elsewhere. As the acceptance of the smart card issued by the TSP becomes widespread, the merchant's balances on the TSP's books are accepted in payment for transactions. The TSP in effect swaps balances on its books as settlement of transactions and does not have to maintain bank balances to meet its payment liabilities. Transactions can then be settled independently of any bank money. The monopoly of the central bank loses relevance and it is unable to affect interest rates and economic activity through the control of bank reserves.

This is an emerging reality. For those bidding on eBay, the biggest Internet auction site, PayPal is the most popular method of payment. It can be used for payment at over a million other business Web sites as well as for person-to-person payments. PayPal has even encouraged buyers to turn away from using credit cards by paying interest on deposits in a PayPal account that can be used for payments to other PayPal accounts. Banks are losing their hold over the start and finish of each payment transaction.

*B. M. Friedman, "The Future of Monetary Policy: The Central Bank as an Army with Only a Signal Corps?" *International Finance* 2 no. 3(1999): 321–338.

› **Table 2.9**
Components of Reserve Money (INR Billion)
*Source:* **Reserve Bank of India, Annual Reports.**

| | 31 March 2002 | 31 March 2003 | 31 March 2004 | 31 March 2005 | 31 March 2006 |
|---|---|---|---|---|---|
| Currency in Circulation[a] | 2,509.74 | 2,824.73 | 3,270.28 | 3,686.61 | 4,306.76 |
| Bankers' and Other Deposits with RBI | 869.97 | 865.88 | 1,094.84 | 1,204.74 | 1,423.90 |
| Reserve Money ($M_0$) | 3,379.71 | 3,690.61 | 4,365.12 | 4,891.35 | 5,730.66 |

[a]Currency in circulation = Currency with public + Cash with banks.

Often, instead of looking at the sources (assets) side of the balance sheet money is written in terms of its application (liabilities) side. The applications side shows the forms in which money is held.

First, there is the central bank holding, which comprises currency in circulation plus deposits of the domestic banking sector (Table 2.9). This is referred to as the base money or reserve money ($M_0$)

**Reserve Money ($M_0$) = Currency in Circulation + Reserves and Central Bank Balances of Domestic Banks**

The reserve money is the input controlled by the central bank for the process by which the money stock is created as output by the commercial banks. Economists measure money beginning with the currency people carry in their wallets and then expand this measure to include other components. $M_1$ is narrow money and comprises basically of currency with the public and demand deposits. Demand deposits are checking accounts or deposits we make with banks which allow us to get our money back on demand. From the perspective of the non-bank private sector, a good substitute to currency in circulation is a demand deposit such as a savings or current account that can be transferred into currency at a very low transaction cost. Broader definitions of money stock include components that are less close or perfect substitutes to currency in circulation. $M_2$, for instance, adds to $M_1$ time deposits with a maturity of less than a year. $M_3$ includes all time deposits regardless of their maturity. To sum up:

› $M_1$ is the sum of currency with the public and demand deposits—the most liquid assets that can be used in transactions. It is also sometimes called narrow money.

› $M_2$ is $M_1$ plus time deposits, with a maturity of less than a year. It is sometimes called broad money.

› $M_3$ is a monetary aggregate—$M_1$ plus all time deposits regardless of their maturity.

**$M_1$ = Currency with Public + Demand Deposits + Other Deposits with RBI**
**$M_2$ = $M_1$ + Time Deposits with Maturity Less than a Year**
**$M_3$ = $M_2$ + Time Deposits with Maturity of 1 Year and More**

Most central banks use more broadly defined monetary aggregates such as $M_3$ as the reference value for the money stock. The components of the money stock are given in Table 2.10.

There are, therefore, a variety of measures of the money stock. In a fractional reserve system, central banks require that commercial banks maintain a fraction of deposits as reserves. This fraction determines the increase in the money stock that a change in the reserves of the banking system can create. This is the subject of Section 2.3.2.

› **Table 2.10**
Applications of Money Stock (INR Billion)
*Source:* Reserve Bank of India, Annual Reports.

| | 31 March 2002 | 31 March 2003 | 31 March 2004 | 31 March 2005 | 31 March 2006 | Average Share |
|---|---|---|---|---|---|---|
| Currency with the Public (A) | 2,407.94 | 2,715.81 | 3,149.71 | 3,557.68 | 4,131.43 | 15.64% |
| Demand Deposits (B) | 1,791.99 | 1,987.57 | 2,586.26 | 2,840.17 | 4,052.24 | 12.99% |
| Time Deposits (C) | 10,755.12 | 12,443.79 | 14,269.60 | 16,076.75 | 19,042.90 | 71.13% |
| Other Deposits with RBI (D) | 28.50 | 32.42 | 51.19 | 64.78 | 68.79 | 0.24% |
| Narrow Money $(M_1)$[a] | 4,228.43 | 4,735.81 | 5,787.16 | 6,462.63 | 8,252.45 | |
| Broad Money $(M_3)$[b] | 14,983.55 | 17,179.60 | 20,056.76 | 22,539.38 | 27,295.35 | |

[a]$M_1$ = (A) + (B) + (D).
[b]$M_3$ = (A) + (B) + (C) + (D).

Time deposits = Deposits with maturity less than a year + Deposits with maturity of 1 year and more
$M_2 = M_1$ + Deposits with maturity less than a year

## 2.3.2 The Money Multiplier

In its simplest form, the MONEY MULTIPLIER is a coefficient that describes the relationship between the volume of money stock and the monetary base or reserve money.

› The money multiplier is the increase in the money supply as a result of an increase by one rupee in central bank base money.

$$\text{Multiplier} = m = \frac{\text{Money Supply}}{\text{Reserve Money}} = \frac{M^S}{M_0} \text{ where } M^s = M_3.$$

For the year 2006, the multiplier is $\frac{2,729,535}{573,066} = 4.76$ .

We can calculate the money multiplier in two different ways:

*Alternative* 1: Another way to obtain the multiplier is to note that $M^S = C^P + D_R$. Now, assume that the public holds $\theta$ per cent of its money as currency and the remainder $(1 - \theta)$ as demand and time deposits with banks.

$$C^P = \theta M^S$$

$$D_R = (1 - \theta)M^S$$

The public's demand for currency relative to its holding of demand deposits depends on the extent of financial development. In economies with few bank branches in the rural areas and where lower-income groups are excluded from the formal banking system, people tend to hold currency for the purposes of transactions, which has a clear advantage over cheques that authorize payments from a demand deposit for transactions purposes.

As economies advance, however, cash is still accepted as a means of payment for transactions of small amounts whilst the use of cheques is restricted to payments where the issuer's creditworthiness can be established, or where the delivery of goods can be delayed until after the clearance of the cheques through the banks. With advances in payment technologies—electronic purses (smart cards) or credit cards—transactions rely even less on currency. The currency ratio thus depends on the availability of alternative modes of payment such as debit and credit cards and the ease of transacting in currency.

The central bank sets a required reserve ratio which determines the fraction of deposits that the commercial banks must hold with the central bank. Let $x_1$ be the fraction of deposits that are held as reserves and deposits with the central bank. Thus, we have

$$R^r = x_1 D_R$$

where $R^r$ is the reserves and deposits required to be held by commercial banks with the central bank. The commercial banks are assumed to keep some constant fraction $x_2$ of their demand deposit liabilities in the form of coin and currency in their vaults. Thus,

$$C^B = x_2 D_R$$

› The monetary base or reserve money is the central bank's liabilities of currency—currency with the public and with banks, $C^P$ and $C^B$, respectively, plus commercial bank required reserves held with it, $R^r$.

Now, recall that the MONETARY BASE is the sum of reserves plus currency with the public, or,

$$M_0 = R + C^P = R^r + C^B + C^P$$

or, $$M_0 = (x_1 + x_2)D_R + C^P$$

However, the public holds a fraction $\theta$ of its money as currency and a fraction $(1 - \theta)$ as demand deposits. Hence,

$$M_0 = (x_1 + x_2)(1-\theta)\, M^S + \theta M^S$$

or, dividing throughout by $M_0$ and solving for the money multiplier $M^S/M_0$,

$$\frac{M^s}{M_0} = \frac{1}{(x_1 + x_2)(1 - \theta) + \theta} = m$$

The money multiplier is therefore

$$m = \frac{1}{(x_1 + x_2)(1 - \theta) + \theta} = \frac{1}{1 - (1 - \theta)\,[1 - (x_1 + x_2)]}$$

For the year 2006,[4] from Tables 2.9 and 2.10, $\theta = 413{,}143/\,2{,}729{,}535 = 0.15136$, $x_1 = 142{,}390/\,2{,}316{,}393 = 0.061471$ and $x_2 = 17{,}533/\,2{,}316{,}393 = 0.007569$. Thus the aggregate reserve ratio is $x = x_1 + x_2 = 0.06904$. This on substitution in $1/\,[(x_1 + x_2)(1 - \theta) + \theta]$ gives the result derived in the beginning of the section above, that is, $m = 4.76$.

*Alternative 2*: The multiplier formula is also derived for an alternative cash-holding ratio which indicates the proportion of cash holdings to deposits (instead of the money supply). In this case, the cash to deposit ratio is given by the fraction $\alpha$, where

$$C^P = \alpha D_R$$

and, $0 \leq \alpha \leq 1$. The reserves held by a bank are as before a fraction $x = x_1 + x_2$ of deposits or, $R = xD_R$. In turn, recall that reserves is the sum of required reserves and cash with the bank, $R = R^r + C^B$. Hence, $R = R^r + C^B = xD_R$. The multiplier may then be written as

$$m = \frac{M^S}{M_0} = \frac{C^P + D_R}{C^P + R}$$

This may be written as (see Appendix 2.1 for the derivation)

$$m = \frac{1 + \alpha}{\alpha + x}$$

For the year ended in 2006, the currency to deposit ratio is $\alpha = 0.178356$ and with the reserve to deposits ratio at $x = 0.06904$, we get $(1 + 0.1784)/(0.1784 + 0.0690) = 4.76$, which is the money multiplier as before.

The multiplier as an explanation of the process of money creation is actually a relation derived from pure accounting equations. In Figure 2.3, we plot the sum of required reserves and cash balances with banks, or, $R = R^r + C^B = (x_1 + x_2)D_R$ on the horizontal axis to the right of the origin and the money supply, $M^S$, to the left of the origin. On the vertical axis, we plot $D_R$ above the origin and $C^P$ below the origin. Let us look at the various relationships now:

First quadrant of Figure 2.3: The following relation between reserve holding and deposits at banks is presented in the first quadrant:

$$R = R^r + C^B = (x_1 + x_2)D_R$$

Dividing throughout by the sum of the reserve ratios, we get a relationship between deposits and reserves, which is the straight line in the first quadrant that represents the following equation:

$$D_R = \frac{1}{(x_1 + x_2)}(R^r + C^B)$$

In the second and third quadrants, we depict the public propensity to hold money in the form of currency versus demand deposits.

Second quadrant of Figure 2.3: The deposit-money supply behavioural ratio, $D_R = (1 - \theta)M^S$, is presented.

Third quadrant of Figure 2.3: The currency-money supply behavioural ratio $C^P = \theta M^S$ of the public is presented.

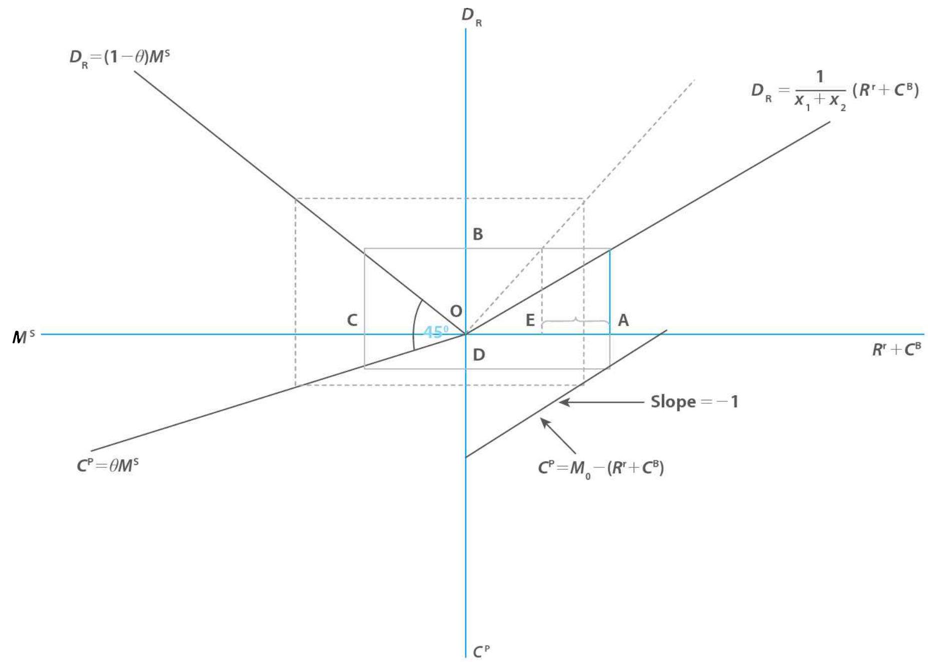
Fourth quadrant of Figure 2.3: Finally, reserve money is given by

$$M_0 = R^r + C^B + C^P$$

This is rewritten with the currency in circulation, $C^P$, on the left-hand side and all other terms on the right-hand side: $C^P = M_0 - (R^r + C^B)$. Thus, for a constant amount of reserve money $M_0$ the relationship between $C^P$ and $(R^r + C^B)$ is a straight line with a slope of $-1$, which is presented in the fourth quadrant.

Given the shapes and positions of the four functions in the diagram, the sum of vault cash and required reserves is equal to the distance OA, demand deposits are equal to the distance OB, the currency in circulation with the public is the distance OD, and the money supply is equal to OC. Any change in the shape or position of any of these four functions in the diagram will lead to a change in the values of the variables.

**› Figure 2.3**
Money Supply Determination. Reserves are held by banks as a ratio of deposits. The public holds a fraction of money as deposits and the remainder as currency. The money supply is the outcome of the reserve holding ratio and the behaviour of the public with respect to the holding of deposits and currency.




Suppose, for instance, the central bank reduces the required reserve ratio $x_1$. The slope of the line $D_R = (R^r + C^B)/(x_1 + x_2)$ increases and the new $D_R$ curve is the dashed line. We trace out the implications for the money supply as follows:

1. The central bank reduces the reserve ratio. The banks now hold excess reserves equal to distance EA.
2. The banks attempt to get rid of their excess reserves. It is assumed that the banks decide to increase their lending to the private commercial sector by the same amount as the increase in excess reserves. Then, $\Delta A_R$ = distance EA. (The multiplier model assumes that banks do not reduce their lending from the central bank. In such a case the expansionary effect would be lost immediately.)
3. The private sector is willing to increase its holdings of loans from commercial banks (say, to buy consumer or investment goods). The private sector needs only a fraction of this increase to conduct transactions. Thus, cash holdings in the economy increase in accordance with $\theta\Delta A_R$, and at the same time there is an increase in deposits $\Delta D_R = (1-\theta)\Delta A_R$. The banks, in turn, must hold higher minimum reserves at the central bank on their additional deposits in accordance with the reserve ratio, $x_1$.
4. After the first round the banks still have excess reserves. They use these again for more lending to the private sector. This process comes to an end only when the excess reserves have been fully used up in cash holdings and in the additional minimum reserves that banks hold as a consequence of the growth of deposits. (It is assumed in the mechanistic interpretation that the change in money supply is identical with the change in total lending, $\Delta M^S = \Delta A_R$.)

Let us take a numerical example. Suppose the central bank reduces the minimum-reserve ratio and the commercial banks hold excess reserves of INR 150 million at the time this policy is implemented (distance EA is INR 150 million).

› **Table 2.11**
The Money Multiplier Model

| Period | Excess Reserves at Beginning of Period | Credits Provided to Private Sector ($\Delta M^S = \Delta A_R$) | Increase in Currency in Circulation with $\frac{C^P}{M^S} = \frac{1}{3} = \theta$ | Increase in Deposits ($\Delta D_R$) | Increase in Minimum Reserves ($\Delta R$) when $R = \frac{1}{4} D_R$ | Excess Reserves at End of Period |
|---|---|---|---|---|---|---|
| 1 | 150 | 150 | 50 | 100 | 25 | 75 |
| 2 | 75 | 75 | 25 | 50 | 12.5 | 37.5 |
| 3 | 37.5 | 37.5 | 12.5 | 25 | 6.25 | 18.75 |
| — | — | — | — | — | — | — |
| Sum Over All Periods | — | 300 | 100 | 200 | 50 | — |

The commercial banks immediately lend this amount of INR 150 million to the private sector (column 3 of Table 2.11). The private sector holds one-third of its money as currency, causing an increase in the circulation of INR 50 million (column 4 of the Table 2.11) and the remainder of INR 100 million is held as deposits (column 5).

Banks after the reduction in the reserve ratio hold required and free reserves equal to one-fourth of their deposits as reserves. Minimum reserves to be held on the INR 100 million are $(1/4)100$ = INR 25 million and excess reserves are INR 75 million.

In Period 2, the commercial bank lends out to the private sector these excess reserves of INR 75 million from the previous period, which the private sector then holds as INR 25 million in currency and INR 50 million in deposits. The commercial banks end up in Period 2 with excess reserves of $[1 - (1/4)](50)$ = INR 37.5 million.

These are re-lent in Period 3 and the process repeated. If we sum across all periods, this series gives a value of credit provided and money creation of INR 300 million.

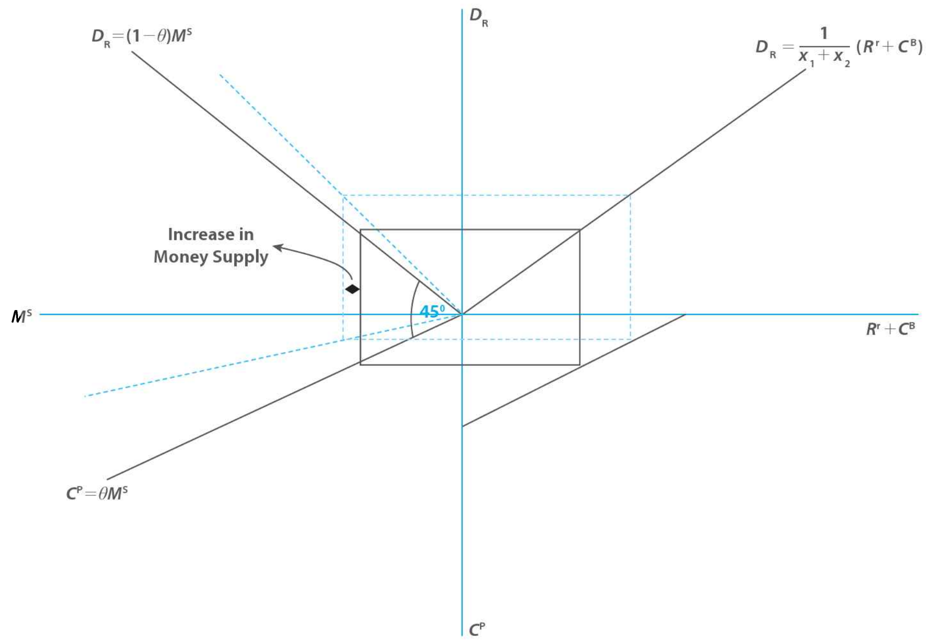
Alternatively, we could calculate this by noting that the value of the MULTIPLIER is

› The money **multiplier** is the amount by which a new deposit into the commercial banking system is multiplied as it is loaned out by banks, redeposited, loaned out again, and so on.

$$\frac{1}{x(1-\theta)+\theta} = \frac{1}{\frac{1}{4}\left(1-\frac{1}{3}\right)+\frac{1}{3}} = 2$$

This outcome is because in a fractional reserve system, the government permits banks to loan out a fraction of their deposits, as only a small percentage are withdrawn at any given time. In the real world, when we pay for our shopping with a cheque at the grocery store, the store cares only whether we have that amount in our account, not whether the bank has enough vault cash to cover all deposits. Hence, the bank can lend out a fraction of our deposits, which is practical, profitable, and also officially authorized by the central bank.

Of course, excess reserves may not be deployed solely as loans advanced to the private sector as has been presumed in the example. The banks may lend to the government instead or hold government securities. Additional resources of the banks are not automatically converted into additional loans to the private sector. In fact, as we shall see when studying financial liberalization in India, this was the outcome of reducing controls and reserve ratios on banks.




› **Figure 2.4**
Increase in Propensity to Hold Demand Deposits. **An increase in the proportion of money held in deposits by the public increases reserves with the banks. This is relent and the process of multiple deposit creation increases the money supply.**

If the public decides to change the proportions in which it holds currency and deposits, the lines in the second and third quadrant shift. For example, suppose the propensity to hold deposits rises, as has been occurring in the recent past. Then, the slope of the line in the second quadrant increases, while the slope of the line in the third quadrant would decrease by the same amount—the dashed lines in Figure 2.4.

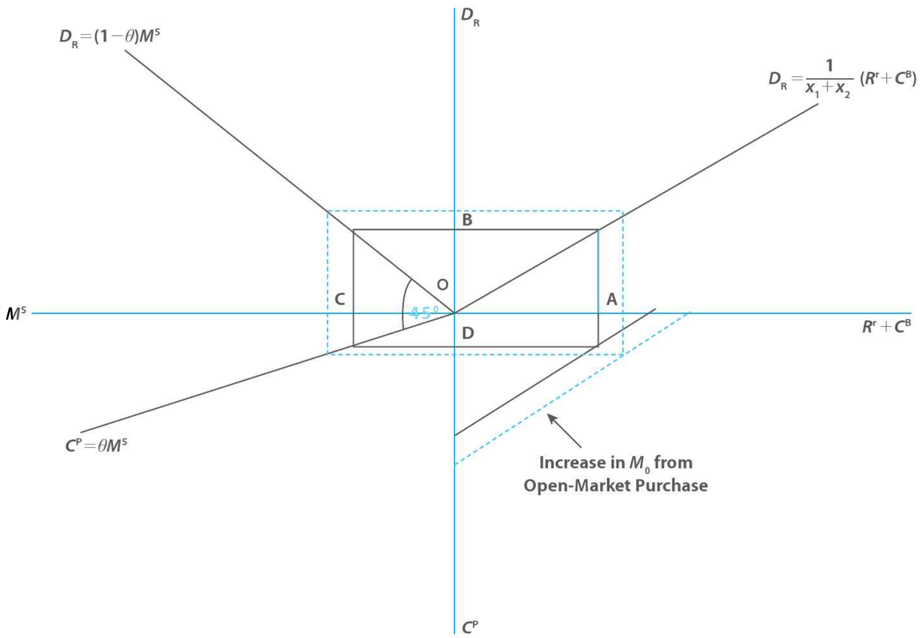
As can be seen in Figure 2.4 the money supply has increased. This is due to the additional liability that commercial banks match with the creation of additional assets via lending. This, in turn, results in a process of multiple-deposit creation that increases the money supply.

As we noted earlier, the central bank affects economic activity by controlling the level of reserves. Apart from establishing reserve requirements, the central bank alters the stock of bank reserves through the use of *open-market operations* when it buys or sells financial assets, mainly government securities (for further details, see Section 2.3.3).

Open-market operations are usually conducted in the market with securities dealers that are designated as primary dealers. Let us see this through a numerical example.

Suppose the central bank purchases INR 1 million of government bonds from a primary dealer. The dealer deposits the cheque of INR 1 million in his bank, which credits the dealer's account for INR 1 million. The dealer's bank then presents the cheque to the central bank, which credits the bank's reserve account with INR 1 million. The bank now has a new liability of a deposit of INR 1 million and a matching new asset of INR 1 million in new reserves. Given a reserve ratio of 5 per cent ($x_1 = 0.05$), the bank can lend out an additional INR 950,000 and hold the remaining INR 50,000 in order to meet its mandatory 5 per cent reserve requirement. The money multiplier now operates with this additional loan to the private sector and the increase in money supply will be a multiple of the initial INR 1 million in deposits.

An open-market purchase by the central bank thus increases the reserves at the central bank and its corresponding liabilities to commercial banks increases. The commercial banks then increase the asset side of their balance sheets by




› **Figure 2.5**
Money Supply Increase from Open-Market Purchase.
An open-market purchase of government bonds by the central bank increases the deposits of the banks. After keeping aside a ratio as reserves the remainder is relent resulting in a multiplier process that expands the money supply.

increasing lending to the private sector and the process of multiple-deposit creation again results in an increase in the money supply. An open-market purchase causes the line in the fourth quadrant to shift away from the origin and increases the money supply—the dashed line intercepts the $M^S$ axis in Figure 2.5—further away from the solid line which represents the situation prior to the open-market purchase.

### ****2.3.3 Open-Market Operations and Reserves**

Central banks operate through procedures that are based on the market for bank reserves. Central banks engage in open-market operations. An OPEN-MARKET OPERATION refers to central bank purchases or sales of government securities. These affect the supply of reserves in the banking system and the interest rate on inter-bank funds—the interest rate that the banks in need of reserves pay to borrow them from banks with surplus. There are two types of open-market operations:

› An open-market operation is a purchase (or sale) of government bonds from (or to) the public by the central bank for the purpose of increasing (or decreasing) the supply of bank reserves and the money supply.

a. An outright operation—a definitive sale or purchase of securities by the central bank with no obligation to resell or repurchase securities at a later date.
b. A repurchase agreement is an operation committing the seller of the security to repurchase it at a later date. In this case, the central bank buys securities from dealers under an agreement for the dealers to repurchase them at a later date. The central bank also uses reverse repurchase agreements when it conducts open-market sales, which are obligations to repurchase the securities from dealers at a later time. Central bank interventions are chiefly intended to smooth temporary fluctuations and conducted mainly through repurchase and sale–repurchase agreements. When the central bank intends to permanently change the level of bank reserves in the aggregate, it engages in outright purchases or sales.

Table 2.12 shows the central bank's balance sheet and the commercial banks' balance sheet.

**› Table 2.12**
Simplified Balance Sheets of the Central Bank and Commercial Banks

| Central Bank | | Commercial Banks | |
|---|---|---|---|
| **Assets** | **Liabilities** | **Assets** | **Liabilities** |
| Foreign Exchange | Currency | Loans | Deposits |
| Government Securities<br>— Loans to Government<br>— Open Market Operations (+INR1) | Bank Reserves (+INR1) | Bank Reserves (+INR1) | |
| Monetary Base | Monetary Base | | |

Bank reserves are rules specifying portions of savings (transactions) and term deposits that commercial banks must hold either as vault cash (cash reserve ratio) or as funds on deposit at the central bank.

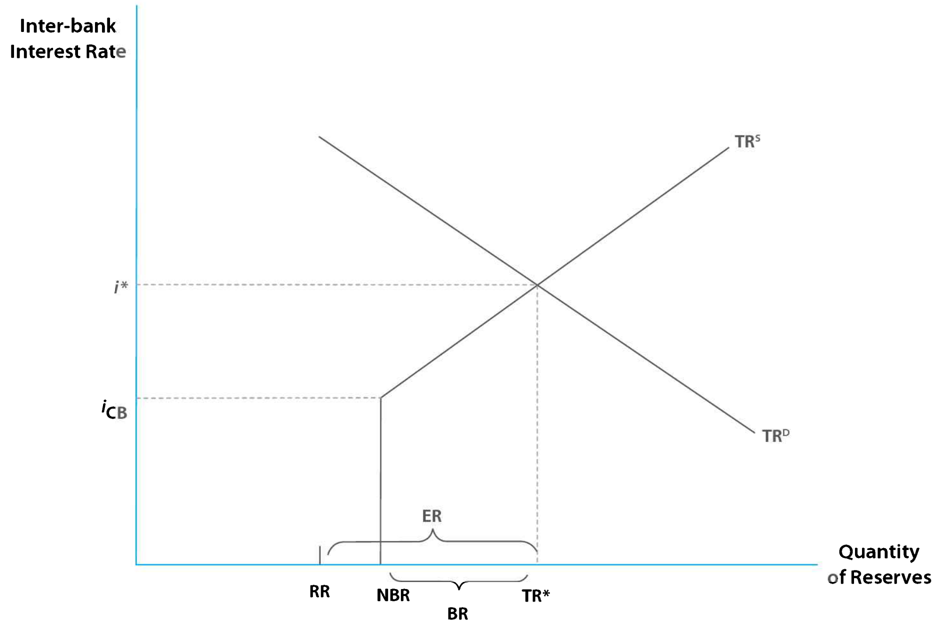
Suppose the central bank conducts an open-market operation involving a repurchase agreement wherein it purchases securities for INR1 (Table 2.12). As a result, the commercial bank's balance sheet at the central bank increases by INR1 as recorded on the liability side of the central bank's accounts. Of course, that rupee also appears on the asset side of the balance sheet of the commercial banks. The commercial bank lends this amount in the inter-bank loan market for the length of the residual maturity of the security transferred to the central bank. The inter-bank loan is repaid on the date the security is transferred back to the commercial bank and the commercial bank uses the proceeds from this inter-bank loan to repurchase the security. The purchase of an open-market security by the central bank is thus equivalent to the creation of a loan on the inter-bank market for the same residual maturity as the security. From the commercial bank's perspective, the cost of funding via open-market operations is the interest gain that occurs by exchanging non-interest bearing central bank money for an interest-bearing security. The central bank in a modern world influences the monetary base not by producing and distributing additional currency, but rather by inducing changes in bank-reserve balances through open-market operations.

Reserve demand and supply interact to determine the inter-bank market interest rate. The demand for reserves arises largely because of a requirement that banks hold reserves equal to a specified fraction of their deposit liabilities. Hence, as the public's demand for deposits varies there will be revisions in the banks sector's demand for reserves. *Required reserves* are the bulk of reserve holdings and are equal to a fraction of the deposit liabilities of a bank (with the fractions set by the central bank). The banking system does hold, however, EXCESS RESERVES[5] to meet the daily unpredictable net inflow or outflow of funds that each bank faces due to deposit withdrawals or due to sudden prospects for gainful loans or purchases of securities. These excess reserves are held over and above required reserves to ensure that the bank can always settle its accounts given the unpredictable nature of daily payments and receipts.

› Excess reserves are the cash assets held in the vault of banks as a contingency against a need for liquidity due to unpredictable withdrawals and payments.

What happens if banks do not have excess reserves? They routinely lend or borrow reserves to fund their extensions of credit for loans or security purchases from the inter-bank funds market.

The inter-bank interest rate then represents the opportunity cost of holding excess reserves. If this interest rate is low, then banks will be prone to hold larger quantities of excess reserves as the opportunity cost of doing so is negligible.




› **Figure 2.6**
The Demand and Supply of Reserves. The inter-bank interest rate is the opportunity cost of holding excess reserves. As it declines, banks hold more excess reserves and the total reserve demand increases. If the inter-bank interest rate rises above the discount rate that the central bank sets for borrowing reserves, then banks tend to go to the central bank to borrow reserves and the supply of reserves increases with the inter-bank interest rate. The equilibrium interest rate in the inter-bank money market is where the demand for reserves equals the supply of reserves. Here, $TR^S = BR + NBR$.

As the inter-bank interest rate rises, however, the opportunity cost of holding excess reserves increases and banks would lend these reserves to others in the inter-bank funds market. The demand for excess reserves by commercial banks will accordingly be inversely related to the inter-bank interest rate as depicted in Figure 2.6. Note that in Figure 2.6, the minimal reserves that banks demand is the amount of required reserves, RR. The demand for excess reserves schedule slopes downwards to the right from a point vertically above the inelastic demand RR.

The total supply of reserves ($TR^s$) held by the banking system is the quantity of reserves that the banks have borrowed from the central bank (BR) plus non-borrowed reserves (NBR)

$$TR^s = BR + NBR$$

NON-BORROWED RESERVES are reserves supplied by the central bank through instruments other than by the direct advance of credit, such as open-market operations. By purchasing and selling government securities, the initiative for which comes from the central bank, the stock of non-borrowed reserves is affected. Open-market transactions are usually based on competitive bids or offerings by market participants and the central bank determines the amount of securities purchased or sold. The quantity of non-borrowed reserves thus does not vary with the inter-bank interest rate. If non-borrowed reserves were the only reserves supplied to the banking system, the reserve supply schedule would be vertical.

› Non-borrowed reserves are the reserves that the central bank supplies to commercial banks through open-market operations.

The individual banks in an economy still require a lender of last resort when they find themselves short at the end of the day after all inter-bank transactions have taken place. Thus, central banks have a DISCOUNT WINDOW to handle the fluctuations of bank reserves in relation to requirements. However, the discount window is not a vehicle for major adjustments in reserves and the central bank typically sets terms on which individual banks can deal directly with it for borrowing reserves. The elasticity of the supply of BORROWED RESERVES depends on the extent of central bank rationing. The central bank uses non-price methods to ration bank borrowing and affects the degree to which banks turn to the discount window to borrow. Banks then weigh the benefits

› The direct lending of reserves by the central bank to commercial banks is called discount window lending.

› Borrowed reserves are the reserves that central banks supply directly to commercial banks in the form of advances of credit.

of borrowing reserves at a particular time against the possible cost in terms of reduced future access to the discount window. Banks will reduce their current borrowing if they expect the inter-bank interest rate to be higher in the future as they may prefer to preserve their future access to the discount window when the inter-bank interest rate is high.

› The discount rate is the interest rate that the central bank charges commercial banks to borrow reserves.

If the inter-bank interest rate is below the DISCOUNT RATE set by the central bank, $i_{CB}$ , banks will tend to borrow from other banks-rather than directly from the central bank. As there is an increase in the inter-bank interest rate relative to the central bank discount rate the amount of borrowed reserves banks desire are supplied less expensively through the central bank. Commercial banks then increase their borrowings from the central bank and the supply of reserves through direct credit advances (borrowed reserves) increases with an increase in the inter-bank interest rate. The total reserve supply schedule $TR^s$ is, therefore, upward sloping above the discount rate with the slope of the schedule depending on the extent of non-price rationing the central bank employs (Figure 2.6)

› The monetary base is the **amount of reserves TR*** in Figure 2.6 plus the currency with the public.

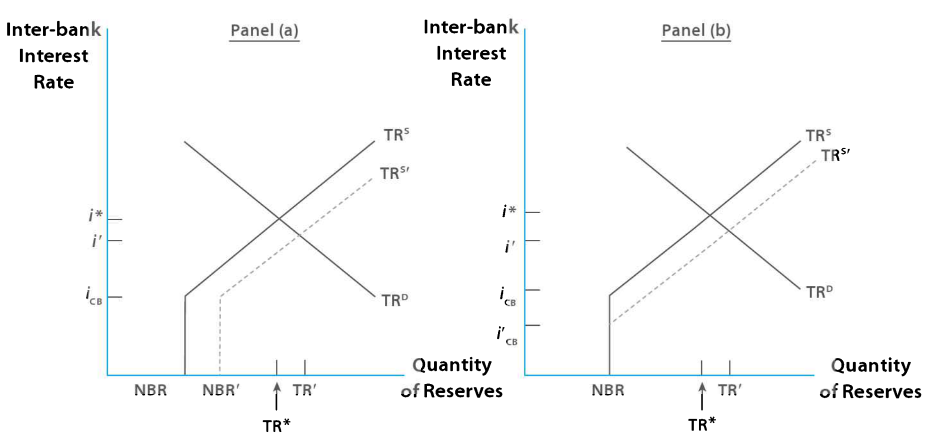
In Figure 2.6, the total quantity of reserves demanded by banks ($TR^D$) equals the supply of reserves through open-market operations (NBR) plus direct central bank advances (BR), at TR*. The equilibrium inter-bank interest rate is $i^*$ and ER is the excess reserves demanded at that rate. Thus, the AMOUNT OF RESERVES TR* would be reserves appearing on the central bank's balance sheet, which along with the amount of currency with the public, establishes the magnitude of the monetary base. It must be the case that

$$TR^D = TR^S = TR^*$$

or,

$$\underbrace{TR^D}_{RR + ER} = \underbrace{TR^S}_{NBR + BR}$$

The central bank is accordingly in a position to impact on the magnitude of the monetary base by engaging in actions that alter the equilibrium level of bank reserves and inter-bank interest rate. For instance, it could initiate an open-market purchase of securities, which introduces additional non-borrowed reserves immediately into the financial system, causing non-borrowed reserves to rise to NBR′ and the supply of total reserves to $TR^{S'}$ Figure 2.7(a). The increased supply of reserves by the central bank bids down the inter-bank interest rate to $i'$ and the increase in bank reserves is accompanied by a rise in base money.




› **Figure 2.7**
Central Bank Policy to Increase the Monetary Base. **There are two ways the central bank can increase the supply of reserves and reduce the interest rate in the inter-bank money market. In panel (a) it conducts an open market purchase and in panel (b) it reduces the discount rate it charges banks for borrowing reserves from it.**

An alternative way of increasing the monetary base is by reducing the discount rate on borrowed reserves or advances to $i'_{CB}$ [Figure 2.7(b)]. This induces banks to increase their holdings of excess reserves and the increased bank reserves produce a rise in the monetary base.

The inter-bank market provides continuing information on the relationship between the demand for reserves and their availability in the banking system. Commercial banks bid for funds in the inter-bank market if they are short and lend funds there if a long position is building up. The inter-bank interest rate indicates whether economic activity is building up. As the object of liquidity management is to maintain reasonable stability in the overnight inter-bank interest rate, the central bank intervenes when the rate moves above or below the range it desires. It uses repurchase agreements to inject reserves on a temporary basis, to counter unusual fluctuations, or the reverse of such operations if the money market becomes undesirably slack. These short-term operations are an important supplement to the outright purchase or sale of securities addressed to longer lasting developments in the demand for total reserves.

## SUMMARY

- Private saving is private disposable income less consumption expenditure.
- Government saving is tax revenues less transfers and government consumption expenditure.
- Total national saving is the sum of private and government saving, or, GNP (GDP plus net factor incomes from abroad) less private consumption expenditure and government expenditure.

$$\text{Saving} = Y + \text{NFI} - C - G$$

- The private-sector savings of an economy is used to finance investment expenditures by firms, to finance a current account surplus (equivalent to a purchase of foreign financial assets), or to finance the government's budget deficit (equivalent to dissaving by the government).

$$S_{pvt} = I + \text{CA} - S_{govt}$$

- The sum of the current account balance and capital account balance gives the overall balance of payments, which is financed mainly by the stock of foreign-exchange reserves held by the central bank.
- The demand for foreign exchange is by importers and by domestic resident decisions on capital outflows. The supply of foreign exchange is by exporters and capital inflows from abroad. The exchange-rate regime in a free float is determined where the demand equals the supply of foreign exchange.
- In a managed exchange-rate regime, foreign-exchange reserves are used to meet the excess demand for foreign exchange and reserves are accumulated when there is an excess supply of foreign exchange.
- The money supply in an economy is the outcome of the interplay between the central bank that influences the monetary base, the public's behaviour in holding currency relative to deposits, and the commercial banks' demand for bank reserves.
- Reserve money or the monetary base ($M_0$) is the currency in circulation plus bank reserves.
- Money supply is the sum of bank credit to government ($G_b$) plus advances to the commercial sector ($A_R$) plus foreign-exchange assets of the banking sector (FE) less net non-monetary liabilities (NML).

$$M^S = G_b + A_R + \text{FE} - \text{NML}$$

- The money multiplier is the ratio between the money supply stock and reserve money.
- The money supply increases with a decline in the mandated reserve ratio that commercial banks are required to hold against deposits and with an increase in the propensity to hold deposits relative to currency by the public.
- An open-market operation is the purchase or sale of government securities by the central bank. This could be an outright sale/purchase of a security or a repurchase agreement. Open-market operations are used to influence the monetary base.
- The demand for reserves is the sum of required reserves (RR) mandated by the central bank plus excess reserves (ER) held as a result of the daily net inflow or outflow of funds. $\text{TR}^D = \text{RR} + \text{ER}$. The demand for excess reserves is negatively related to the inter-bank interest rate.
- The supply of bank reserves is the sum of non-borrowed reserves (NBR) plus borrowed reserves (BR).

$$\text{TR}^S = \text{NBR} + \text{BR}$$

- The central bank influences the monetary base and the inter-bank interest rate by varying the discount rate on borrowed reserves and/or the supply of non-borrowed reserves (open-market operations). The stock of money supply is influenced by such open-market operations and adjustments to the discount rate.

## NOTES

1. Investment is a part of private-sector expenditure and is not deducted from private disposable income. This is because investment is expenditure on capital goods that are long-lived or are purchased to enhance future productive capacity rather than to satisfy current needs.
2. Since government outlays are devoted partly to meet current needs (government consumption) and partly on long-lived capital goods (government investment), the fiscal surplus equals government savings only when government purchases are made up entirely of government consumption—government investment equals zero.
3. These are income flows due to the holding of financial assets. They include interest and dividend payments to foreigners who hold domestic financial assets and similar payments received by domestic residents who hold financial assets abroad. See the components of the current account in Chapter 1.
4. For those of you who check the numbers as given in the Annual Reports of the Reserve Bank of India, it is important to clarify that required reserves as taken here includes apart from bankers deposits with RBI, Other deposits with RBI, for simplicity. Also, for parity, aggregate deposits with Banks include other deposits with the RBI. Other deposits with RBI were a miniscule INR 68.79 billion on 31 March 2006.
5. Excess reserves are essentially the vault cash held by commercial banks. While deriving the money multiplier, we treated this as a fraction of demand deposits. Here we are analysing how the fraction is determined by behaviour.

## TEST YOURSELF

1. Why is it that despite an individual saving (spending less than his/her income), the aggregate income equals expenditure?
2. Explain how aggregate investment in the economy is supported from savings by (a) the private sector, (b) by the government, and (c) by foreigners.
3. Why should a nation's overall balance of payments always equal its foreign-exchange reserves?
4. Which actions by agents result in the demand and supply of foreign exchange in the economy? Explain how this influences the determination of exchange rates when they are freely floating.
5. What is the role of foreign-exchange reserves in the determination of the balance of payments?
6. What are the components of the monetary base of the central bank in terms of its sources and applications?
7. Write out the consolidated balance sheet of the banking system in an open economy and then define the money supply on the basis of this.
8. What is the money multiplier and what are the alternative currency holding behavioural ratios that are considered in explaining the magnitude of the multiplier?
9. What are the components of the demand and supply of reserves and how is the inter-bank interest rate determined?
10. Explain how a reduction in the required reserve ratio or an open-market operation of purchasing government securities affects the money supply in the economy.

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India Internet (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon
   (c) Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy".
   (d) Click on "Table 43: Components of the Money Stock" and create a spreadsheet of data on reserve money for the period 1993–1994 to 2005–2006.
2. Then from the same Table 43, download data for the same time period for broad money.
3. Calculate the money multiplier for this time period and create a graph of the money multiplier for this time period. What does the slope of this graph indicate?
4. Now click on "Table 44: Sources of Money Stock" and download data on the net foreign-exchange assets of the RBI. Calculate the fraction of the central bank's monetary base that is sourced from foreign-exchange assets over time. Create a graph of this ratio.
5. Create a graph of the net foreign-exchange assets of the RBI and of broad money over time. Which graph is rising at a faster rate? What does this indicate about how the central bank is offsetting expansions in the money supply?

# *Appendix 2.1* The Formula for the Money Multiplier

Derivation of the formula for the multiplier when the cash holding ratio is the cash to deposit ratio:

The multiplier is written as

$$m = \frac{M^s}{M_0} = \frac{C^P + D_R}{C^P + R}$$

Treat the term $C^P$ as common to the numerator and denominator

$$m = \frac{C^P\left[1 + \frac{D_R}{C^P}\right]}{C^P\left[1 + \frac{R}{C^P}\right]}$$

Substituting from $C^P = \alpha D_R$ gives us $D_R / C^P = 1/\alpha$ in the numerator. For the denominator we have $R = R^r + C^B = x D_R$ and again $C^P = \alpha D_R$, $R / C^P = x D_R / \alpha D_R = x/\alpha$. Hence, we may write,

$$m = \frac{1 + \frac{1}{\alpha}}{1 + \frac{x}{\alpha}}$$

$$= \frac{1 + \alpha}{\alpha + x}$$

# 3

# Consumption

CRITICAL QUESTIONS

» *What are the marginal and average propensity to consume and how are they related?*

» *What is the savings rate and how does it evolve over time?*

» *How are rational decisions about consumption over time (intertemporal consumption) made?*

» *What is permanent income?*

» *What is the effect of a shock to the income on the consumption decision?*

» *What is the effect of a change in the interest rate on the current consumption?*

» *How is consumption affected by the age cohort that an individual belongs to?*

# 3.1 Keynes on Consumption

We began by stating that macroeconomics deals with the relation between choices in four markets—the markets for goods, money, bonds, and labour. In the goods market in the aggregate we demonstrated the truism that output is always equal to income, which, in turn, is always equal to expenditure. Aggregate expenditure comprises domestic expenditure and net exports. In symbols we had written,

$$Y = C + I + G + X - M$$

In this chapter, we will focus on the determinants of aggregate private consumption expenditure, *C*.

Recall that consumption expenditure by households is a very significant part, 61.6 per cent, of the GDP and so an examination of the determinants of this component of GDP is crucial to understanding aggregate expenditure.

The first formal attempt to understand consumption was that of J. M. Keynes who related it to the current income by adopting a behavioural approach. It is best to quote Keynes,[1] "The fundamental psychological law, upon which we are entitled to depend with great confidence both a priori from our knowledge of human nature and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption as their income increases, but not by as much as the increase in their income."

This was the basis of the model of consumption that linked current income and current consumption. This can be written algebraically as

$$C = a + cY^d \tag{3.1}$$

where $Y^d$ is personal disposable income. Recall that the private disposable income[2] is the GDP plus net factor incomes from abroad plus transfers from abroad as well as payments from the government less direct taxes paid. The constants $a$ and $c$ in this equation are taken to represent Keynes' psychological law which requires that $0 < c < 1$. Here, $c$ is referred to as the MARGINAL PROPENSITY TO CONSUME (MPC), and MPC $< 1$ is another way of saying that for any increment to income, the incremental consumption resulting from this will be less than the incremental income. More formally,

› The marginal propensity to consume (MPC) is the amount by which consumption rises when disposable income rises by INR 1. It is assumed that the increase in consumption is less than the increase in disposable income or, MPC $< 1$.

$$\text{MPC} = \frac{\Delta C}{\Delta Y^d} = c < 1 \tag{3.2}$$

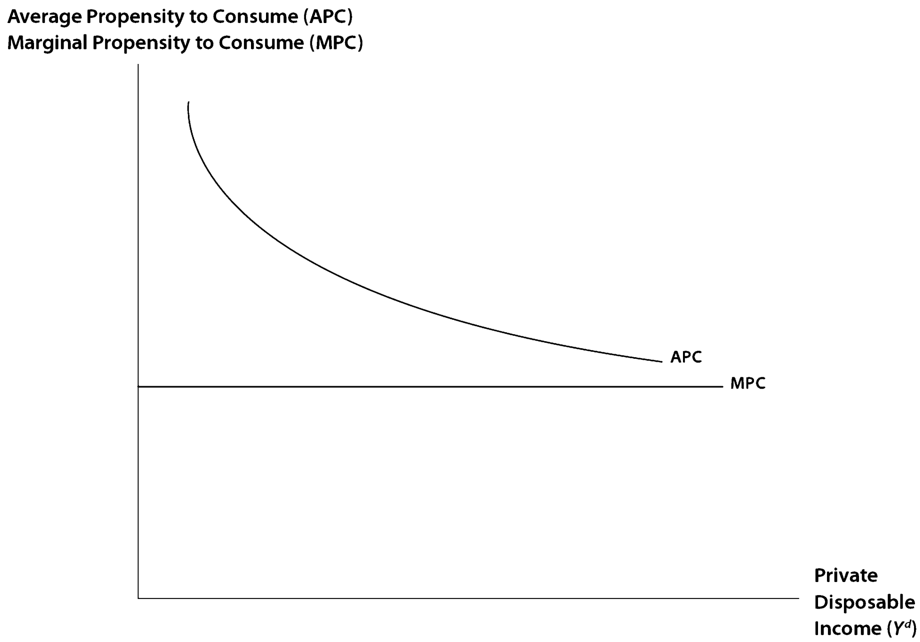
In Eq. (3.2), the change in consumption has been represented algebraically as $\Delta C$ and the marginal change in disposable income as $\Delta Y^d$. The ratio of the change in consumption to a marginal change in disposable income is the MPC. The AVERAGE PROPENSITY TO CONSUME (APC) is defined as the average consumption per unit of disposable income. More formally,

› The average propensity to consume (APC) is the ratio of consumption to disposable income.

$$\text{APC} = \frac{C}{Y^d} = \frac{a}{Y^d} + \text{c} \tag{3.3}$$

The graphs of MPC and APC are shown in Figure 3.1. As can be seen, an implication of Keynes' theory is that MPC $<$ APC. Also, since $Y^d = C + S$, private savings can be written as

**› Figure 3.1**
Average and Marginal Propensities to Consume. **The average propensity to consume declines with an increase in disposable income. The marginal propensity to consume is constant.**




$$\begin{aligned} S &= Y^d - C \\ &= Y^d - (a + cY^d) \\ &= -a + (1 - c)Y^d \end{aligned} \tag{3.4}$$

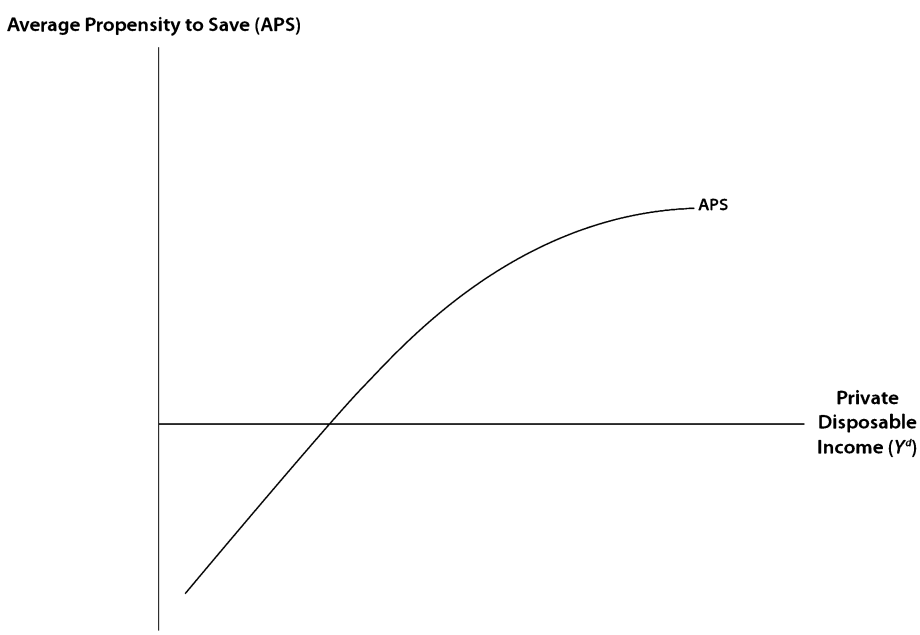
› The average propensity to save (APS) or the savings rate is the ratio of savings to disposable income.

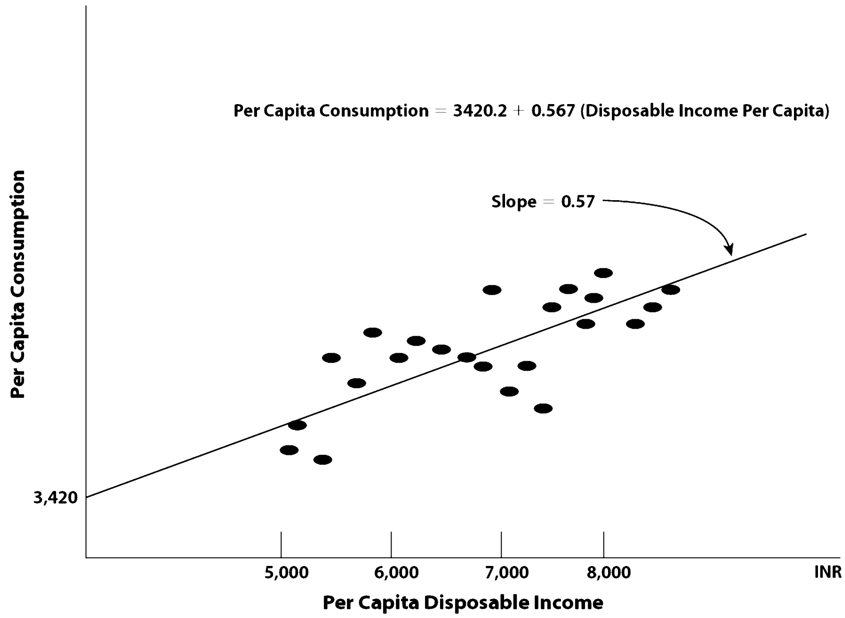
The AVERAGE PROPENSITY TO SAVE (APS) is given by Eq. (3.5) and is depicted in Figure 3.2. Mathematically,

$$\text{APS} = s = \frac{S}{Y^d} = -\frac{a}{Y^d}d + (1 - c) \tag{3.5}$$

The APS is also known as the *savings rate*. Figure 3.2 implies that as disposable incomes rise the savings rate rises, or individuals in the aggregate save increasing proportions of their incomes.

**› Figure 3.2**
Average Propensity to Save. **The average propensity to save or the savings rate increases with disposable income.**






› **Figure 3.3**
Relation between Consumption and Disposable Income. **The relationship between consumption per capita and disposable per capita income is linear.**

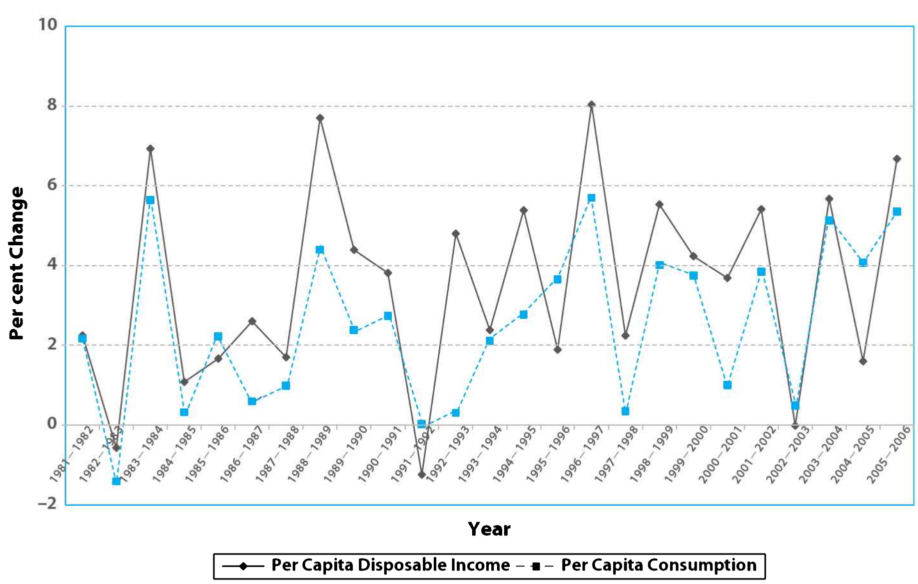
*Source:* Data taken from CSO, National Accounts Statistics.

Figure 3.3 depicts the measured consumption function for India from 1980–1981 to 2005–2006. The estimates[3] for the parameters of the consumption function are as follows:

$$C = 3420.2 + 0.567\, Y^d \qquad \textbf{(3.6)}$$

where the variables in Eq. (3.6) have been stated in per capita terms. From this, we can see that the measured value of the MPC is $c = 0.57$. The APC in this time period was 0.83. The APC declined from an average of 0.90 during 1980–1981 to 1992–1993 to an average of 0.77 during 1992–1993 to 2005–2006.

The dotted points in Figure 3.3 are the actual values of consumption and disposable income, and Eq. (3.6) gives the best fit for a straight line through these points. Figure 3.3 would lead us to contend that changes in disposable incomes in a given period are closely associated with changes in consumption. To check out this contention, we can examine Figure 3.4 that graphs the changes in consumption and disposable income.[4]




› **Figure 3.4**
Percentage Change in Real Per Capita Consumption and Disposable Income. **The variability of disposable income is greater than the variability of consumption. Also, an increase in income does not automatically translate into an increase in consumption.**

*Source:* Data taken from CSO.

The graph of Figure 3.4 indicates a connection between the two variables—it seems that when disposable incomes increase (decrease), this is accompanied by an increase (decrease) in consumption. However, this is not entirely accurate an account of what the graph depicts. There are periods, for instance, 1986–1987, when despite an increase in the percentage change in disposable incomes, the percentage change in consumption declined. In some periods such as 1993–1994 and 1995–1996, despite a drop in disposable income, consumption actually increased. More importantly, the percentage change in disposable income shows much more variation than that in consumption. This can be seen from the larger variation in movement in the graph of the disposable income compared with that of consumption. One statistical measure of variation that indicates this is the coefficient of variation. The coefficient of variation in consumption is 18.8 and in disposable income it is 27.4 for the period under consideration.[5]

› The data on consumption reveals that it is not subject to as much variability as is income—consumption over time has a flat or **smooth** profile.

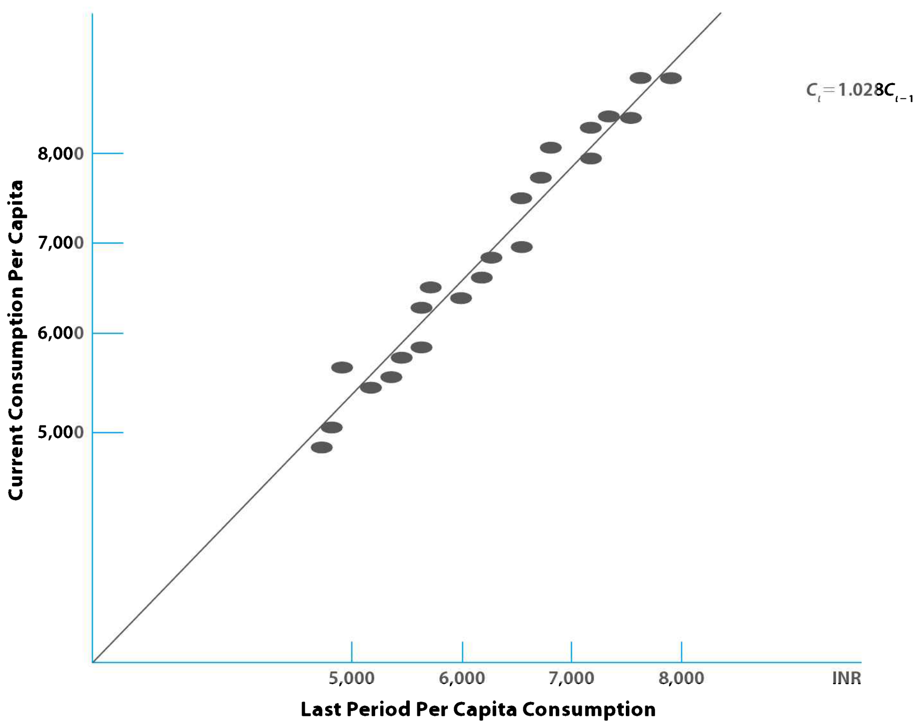
This implies that consumption is not subject to as much variability as is disposable income. In other words, the time path of consumption is SMOOTH. Figure 3.5 indicates just this. If we graph current consumption against the previous period's consumption, then we get the graph indicated in Figure 3.5.

In fact, consumption in a given year is 1.028 times the consumption in the previous year as estimated by the straight line through the scatter of points in Figure 3.5.

$$C_t = 1.028\, C_{t-1} \tag{3.7}$$

So, we seem to continue our consumption patterns inherited from the previous period with, of course, some growth in consumption over time which is implied by the growth factor that increases $C_{t-1}$ to $C_t$ next period. Also, the scatter of points in Figure 3.5 is clustered closer to the straight line than the scatter in Figure 3.3, indicating a much better fit.

› **Figure 3.5**
Current and Lagged Consumption (from 1981–1982 to 2005–2006).
**The time path of consumption is smooth with consumption today proportional to consumption yesterday.**



That consumption demonstrates smoothness in its pattern calls for an explanation. Also, an implication of a Keynes-type consumption function is that savings rates tend to rise over time as shown in Figure 3.2. Simon Kuznets[6] tested the data for proof of this and found instead that in the United States, for over the period of a century, the savings rate had not risen. This contradicted what the Keynesian consumption function was proclaiming. An important research area in economics was to find an explanation for the smoothness in consumption and static nature of savings rate. Two economists who led the research in this area are Franco Modigliani of MIT and Milton Friedman of the University of Chicago. Their theories are, respectively, the life-cycle theory and the permanent income theory of consumption. With sincere apologies to the two Nobel laureates, we will treat the two theories as providing analogous explanations for the smoothness of consumption and identify their contribution as the consumption smoothing theory of consumption.

## 3.2 Consumption Smoothing

There are two hallmarks of the consumption smoothing approach. One is that unlike Keynes, who appealed to behavioural inference, the consumption smoothing approach uses the standard tool of economics, which is that agents optimize subject to constraints. Thus, in this theory consumption is the consequence of rational agents seeking to maximize their preferences subject to the opportunities that they face. The second differentiating factor in the CONSUMPTION SMOOTHING THEORY is that agents are not just reactive and tied in solely to the present—where current consumption moves in step with current income—but rather are treated as forward-looking proactive agents. They have foresight about what is in store for them in the future and keep that in mind when they take decisions today. That the future casts its shadow on the present requires that a successful portrayal of economic behaviour involves taking into consideration at least two periods—a present and a future—in our analysis. This type of analysis is known as INTERTEMPORAL ANALYSIS. A lot of economics is concerned with trade-offs between alternatives in the present—the choice between eating Thai food or an *idli-dosa* meal, or between buying the latest Harry Potter book and treating your girlfriend or boyfriend to a dinner and a movie. In intertemporal analysis, by contrast, the fundamental consideration is that alternatives chosen today affect the menu of alternatives available in the future.

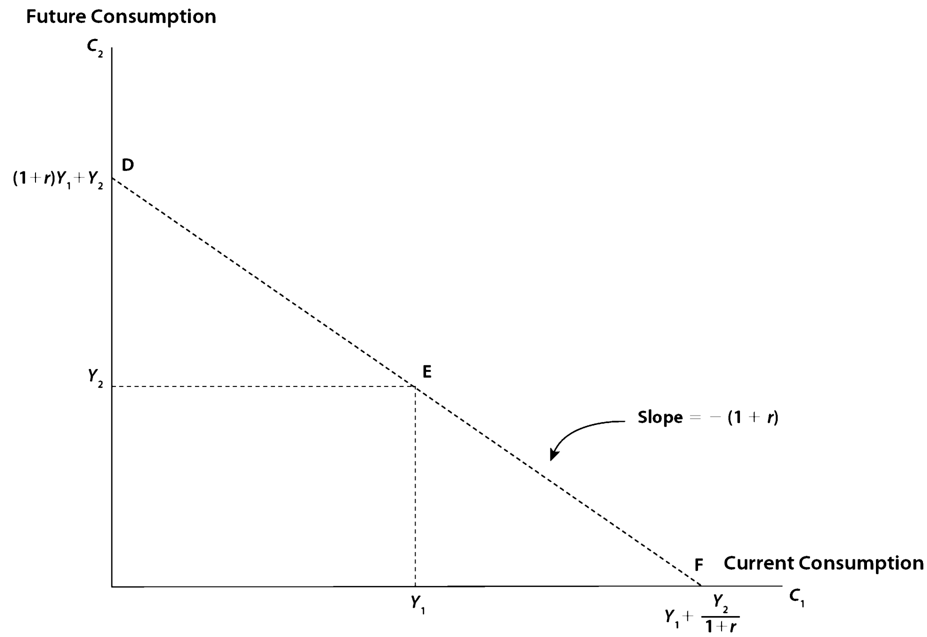
› Consumption smoothing theory is not a behavioural theory but a preference maximization theory where individuals are forward looking.

› Intertemporal analysis is about the trade-offs when the present choices affect the alternatives available in the future.

### *3.2.1* Intertemporal Constraints and Preferences

When deciding what to do with their incomes, people may consume them all now, for instance, or save a part for the future. So the question that the intertemporal approach will focus on is: How do rational consumers distribute their consumption over time? To keep the analysis manageable, and also to allow us to draw diagrams, we suppose that there are only two time periods, namely, current and future. So the alternatives facing the consumer are current consumption ($C_1$) versus future consumption ($C_2$). To derive the intertemporal budget constraint facing a consumer, we begin by noting that a consumer always has the choice of consuming their entire income in each period. So with $Y_1$ and $Y_2$ denoting current and future income, the consumer sets $C_1 = Y_1$ and $C_2 = Y_2$. This is the point E in Figure 3.6.

**› Figure 3.6**
Intertemporal Budget Constraint. **The intertemporal budget constraint depicts the current and future consumption possibilities from the stream of income available across time.**




Another option open to the consumer is to deposit all the income earned in period 1 in a bank and with the interest rate of $r$ per cent they would be able to get $(1 + r)Y_1$ as the principal plus interest payment in the future. They would not consume any income in the present ($C_1 = 0$) and by saving it all derive an income from their savings of $(1 + r)Y_1$. Since there are only two periods and no opportunities for leaving bequests to heirs, they must consume their first period savings along with their second period income $Y_2$ so that $C_2 = Y_2 + (1 + r)Y_1$. This alternative of consuming only in the future is represented by point D in the diagram. Of course, we might presume that if there is no consumption today, the consumer might not live to see another day, and so point D does not seem like a practical alternative. Though extreme behaviour for a consumer, it is still a theoretical possibility and as we want to be exhaustive about the possibilities open for consumption, we include this point in our analysis.

The other extreme behaviour is to forget altogether that there is a future and consume all the future and present income today. Consuming tomorrow's income requires borrowing that future amount today and in present value terms $Y_2$ of future income is equivalent to $Y_2/(1 + r)$ income today. This follows from the observation that if we were to deposit $Y_2/(1 + r)$ in a bank today, we would get back $[Y_2/(1 + r)](1 + r) = Y_2$ in the next period as repayments on our deposit. So consuming both tomorrow's and today's income right away today amounts to $C_2 = 0$ and $C_1 = Y_1 + [Y_2/(1 + r)]$, which is represented by point F. Obviously, it is not necessary to save or borrow the maximum amounts possible. If from the current income $Y_1$ the individual consumes only a part of it such that $C_1 < Y_1$ and saves $(Y_1 - C_1)$, then the earnings on that savings in the next period amounts to $(1 + r)(Y_1 - C_1)$, which must be spent along with next period's income $Y_2$ on next period's consumption $C_2$ (given that there are only two periods). Thus,

$$(1 + r)(Y_1 - C_1) + Y_2 = C_2$$

Dividing throughout by $(1 + r)$ gives

$$Y_1 - C_1 + \frac{Y_2}{1+r} = \frac{C_2}{1+r}$$

Rearranging this expression gives

$$C_1 + \frac{C_2}{1+r} = Y_1 + \frac{Y_2}{1+r} \tag{3.8}$$

This is the equation for the INTERTEMPORAL BUDGET CONSTRAINT drawn as the line DEF in Figure 3.6.

› The intertemporal budget constraint states the present value of consumption possibilities over time from the present value of the flow of income over that period.

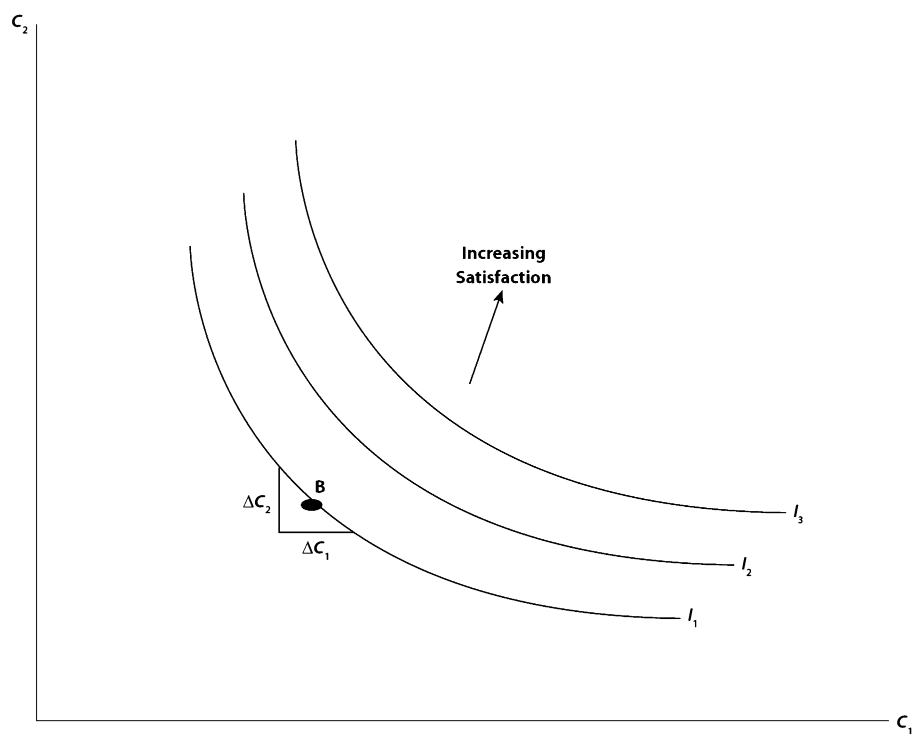
Another way of interpreting Eq. (3.8) is that with only two periods and no heirs to worry about, the present value of life-cycle consumption $C_1+[C_2/(1 + r)]$ must equal the present value of life-cycle income $Y_1 +[Y_2/(1 + r)]$. As depicted, the slope of the intertemporal budget constraint is $-(1 + r)$, as this is the opportunity cost of the interest foregone when money is spent rather than saved.

Which point will the consumer select on this intertemporal budget constraint? We need to know their preferences over current and future consumption to answer this question, which is given by their preference function. The preferences are defined by a utility function as is done in the atemporal case in microeconomics texts and can be written as

$$U = U(C_1, C_2) \tag{3.9}$$

This can be represented by an indifference curve as in Figure 3.7 where the consumer is indifferent between two consumption bundles that lie on the locus of any given indifference curve. Also, the indifference curves that are more to the north-east represent more desirable consumption for the consumer and provide the consumer with greater satisfaction or utility.

The slope of an intertemporal indifference curve in Figure 3.7 measures the marginal rate of substitution (MRS) between future and current consumption. It is the rate at which the present consumption can be substituted for the future consumption with no change in the utility obtained. A marginal change




› **Figure 3.7**
Intertemporal Indifference Curves. **The preferences of an individual for consumption today versus tomorrow are depicted by indifference curves. Along any indifference curve the individual is indifferent between the consumption possibilities depicted by any two points on the curve.**

in the consumption results in a change in utility as expressed by the marginal utility of consumption or $MU_C$. If at point B in Figure 3.7 we reduce the present consumption by an amount $\Delta C_1$, then the reduction in utility due to this reduction in consumption will be $MU_{C_1}(\Delta C_1)$. To compensate the individual for the loss in the utility from the decline in the present consumption would require the future consumption to be augmented. If we raise the future consumption by an amount $\Delta C_2$, then the increase in utility arising from the increase in the consumption will be $MU_{C_2}(\Delta C_2)$.

An individual's time preference for consumption is usually such that the utility from consumption received in the future is valued less than the utility from consumption today. Let the subjective discount rate at which an individual values future consumption relative to current consumption be given by $\delta$. We assume that $\delta > 0$. Accordingly, the value of the gain in the utility from an increase in the future consumption is given by

$$MU_{C_2}\frac{(\Delta C_2)}{1+\delta}$$

For the individual to be on the same indifference curve, the value of the gain in the utility from the increased future consumption must exactly offset the loss in the utility from a reduction in present consumption:

$$MU_{C_2}\frac{(\Delta C_2)}{1+\delta} = -\,MU_{C_1}(\Delta C_1)$$

or,

$$-\frac{\Delta C_2}{\Delta C_1} = (1+\delta)\,\frac{MU_{C_1}}{MU_{C_2}}$$

The slope of the intertemporal indifference curve is called the marginal rate of time preference (MRTP), to differentiate it from the slope of an atemporal indifference curve, which is the MRS. The MRTP is

$$\text{MRTP} = -\frac{\Delta C_2}{\Delta C_1} = (1+\delta)\frac{MU_{C_1}}{MU_{C_2}} \qquad \textbf{(3.10)}$$

› The slope of an intertemporal indifference curve measures the rate of time preference for present consumption versus future consumption.

(From your reading of microeconomics you would recall that MRS $= MU_C / MU_{C_2}$.) The slope of an INTERTEMPORAL INDIFFERENCE CURVE is a measure for the amount of $C_2$ that must be obtained when a unit of $C_1$ is given up and there is no change in the level of satisfaction or utility (that is the consumer is on the same indifference curve). There are three possible cases:

(i) If $|\Delta C_2/\Delta C_1| > 1$, at point B the consumer exhibits positive time preference at that point and requires more than one unit of future consumption to compensate them for the loss of a unit of current consumption.

(ii) If $|\Delta C_2/\Delta C_1| < 1$, the consumer exhibits negative time preference and is willing to forgo a unit of current consumption in return for less than a unit of future consumption.

(iii) If $|\Delta C_2/\Delta C_1| = 1$, the consumer has neutral time preferences with present and future consumption traded off against one another with a ratio of 1:1. The more the current consumption of a person, the more that person will be willing to give up in order to obtain an additional unit of future consumption—the MRTP declines as one moves downward along an indifference curve. Then, the time preference depends on where one is currently on the indifference curve.

A farmer who has had a bumper harvest this year, which by the law of averages should turn into a poor harvest next year, is likely to have negative time preferences currently. Similarly, the son of a tycoon who is due to inherit the empire in the future will have a very small current income compared with the future income and is likely to have positive time preference.

### *3.2.2* Time Preference and the Permanent Income

Exactly as is the case in the standard choice theory, the consumer's problem is to maximize their utility subject to the budget constraint

$$\text{Max } U = U(C_1, C_2)$$

such that

$$C_1 + \frac{C_2}{1+r} = Y_1 + \frac{Y_2}{1+r} = W \qquad \textbf{(3.11)}$$

This is akin to the consumer selecting the point along their budget constraint (as depicted in Figure 3.6) that corresponds to the highest attainable indifference curve (from those depicted in Figure 3.7). This implies (in the normal case without corner solutions) a tangency solution to the maximization problem stated in Eq. (3.11) where the slope of the indifference curve is equal to the slope of the budget line as depicted in Figure 3.8.

The consumer consumes $C_1^* < Y_1$ of the current income and saves $(Y - C_1^*)$ so as to consume $C_2^* = Y_2 + (Y_1 - C_1^*)(1 + r)$[7] in future. Different consumers will of course have different preferences and so will allocate current and future consumption differently. In Figure 3.9(A), for instance, we have a consumer whose preferences are heavily in favour of future consumption—we could consider this person to be patient as the person is willing to postpone the bulk of consumption until the future period. Figure 3.9(B) depicts a consumer who

**MACROFOCUS 3.1**

#### Emotions and Intertemporal Decisions

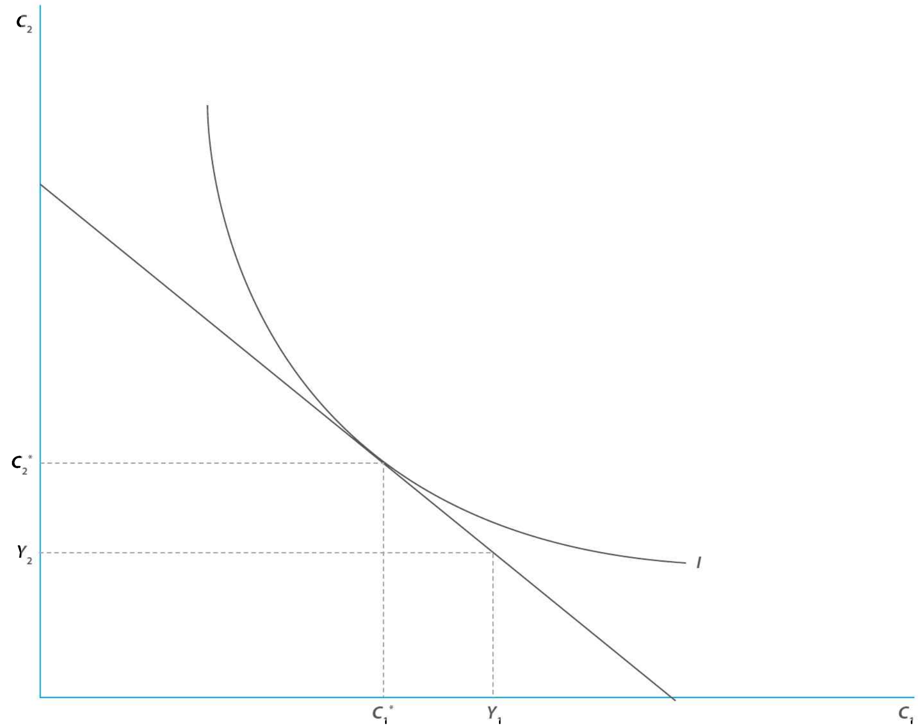
Is decision-making based solely on rationality? An emerging field in economics that goes by the name of neuroeconomics* argues that decision-making draws on the emotions even when reason is clearly involved. As humans evolved, new capabilities such as the ability to deliberate about the broader consequences of our actions were gradually added to the basic prehistoric brain system. The neural anatomy essential for full rationality—the prefrontal cortex—emerged only in the last 150,000 years of humanity's six-million-year existence.

Neuroeconomists conceive human behaviour as a struggle between affective and cognitive systems. Affective systems embody not only emotions such as anger and jealousy, but also drive states such as hunger and sexual desire, and motivational states such as anger and physical pain that motivate us to aggress and take steps to ease the pain. Cognitive systems assess the options available to an individual with a broad, goal-based perspective, such as posited in the standard rational choice approach. The affective system has primary control over behaviour, but the deliberative system chooses which behaviour to implement. It does this by trading off its objectives against the cost of exerting cognitive effort or willpower so as to influence the affective system's choice.

With regard to intertemporal choice, people are often of two minds—they often find themselves encouraged to take myopic actions such as eating high-calorie foods or eschewing contraception while concurrently realizing these activities are not in their self-interest. This is because even though the deliberative system recognizes the desirability of certain actions motivated from the affective system—that focuses on near-term environmental stimuli and payoffs—it also incorporates the willpower effort it must exert to control these affective motivations. The individual can then make myopic choices such as extravagant spending by individuals who buy with credit cards, which defers the pain of paying to the future unlike payment by cash. However, the interaction between affect and cognition can help us understand not just impulsivity but also the opposite problem of high self-control—people with dominant deliberative systems can exhibit miserliness and cannot get themselves to spend enough. Workaholics who cannot take a break are in the same category.

* C. Camerer, G. Loewenstein, and D. Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature* 43 (March 2005): 9–64.

**› Figure 3.8**
Equilibrium of the Individual. **Optimal consumption is where an indifference curve is tangential to the intertemporal budget line. In this case, the individual saves some of their current income with a view to increasing the consumption possibilities in the future.**




cares much more about present consumption—and who is much more impatient and prefers consuming more today rather than waiting for tomorrow.

If the utility function is written in additively separable form,

$$U = U(C_1, C_2) = U(C_1) + \frac{U(C_2)}{1+\delta} \tag{3.12}$$

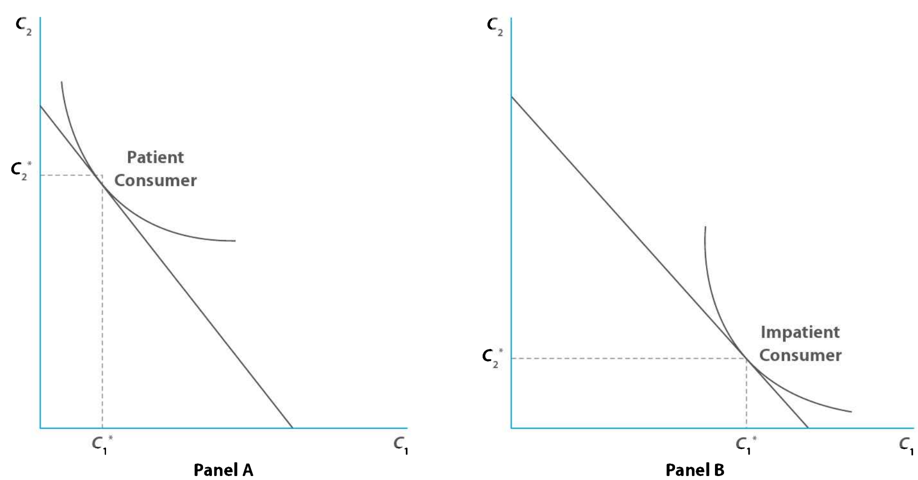
› $\delta$ is the subjective discount rate for valuing future consumption relative to current consumption. The **higher the $\delta$** the more impatient is the individual for current consumption relative to future consumption.

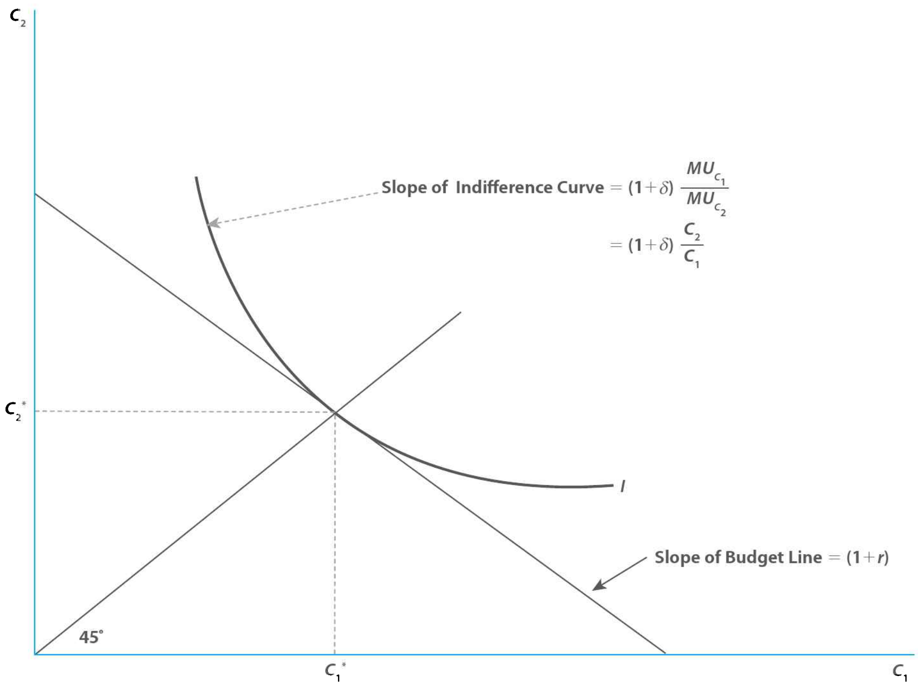
where $\delta > 0$ is the measure of impatience of a consumer. HIGHER VALUES OF $\delta$ imply higher impatience and a consumer preference for consumption today rather than tomorrow [see Figure 3.9(B)].

In order to fix the idea of impatience further, let us give a form to the utility function. Suppose the utility of consumption is logarithmic in consumption,

$$U = U(C_1, C_2) = U(C_1) + \frac{U(C_2)}{1+\delta} = \log C_1 + \frac{\log C_2}{1+\delta} \tag{3.13}$$

**› Figure 3.9**
Time Preference and the Individual. **A patient individual has a high preference for consumption in the future. An impatient individual prefers consumption today at the expense of less consumption in the future.**






› **Figure 3.10**
Consumption Smoothing.
Consumption smoothing, $C_1^* = C_2^*$, results when the marginal rate of time preference given by the slope of the indifference curve equals the slope of the intertemporal budget line and the subjective rate of time preference equals the market rate of interest, or $\delta = r$.

then, the slope of an indifference curve is given by

$$-\frac{dC_2}{dC_1} = (1+\delta)\frac{C_2}{C_1} \qquad \textbf{(3.14)}$$

As maximizing preferences requires tangency—the slope of the indifference curve equals the slope of the budget line—we can write the condition for optimal consumption as

$$\underbrace{(1+\delta)\frac{C_2}{C_1}}_{\text{Slope of indifference curve}} = \underbrace{1+r}_{\text{Slope of budget line}} \qquad \textbf{(3.15)}$$

or,

$$\frac{C_2^*}{C_1^*} = \frac{1+r}{1+\delta} \qquad \textbf{(3.16)}$$

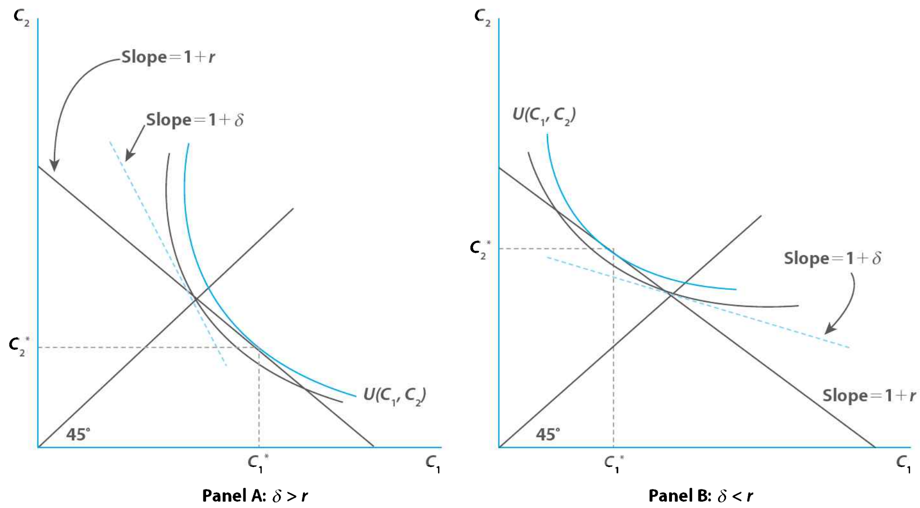
If $\delta = r$, then $C_2^* = C_1^*$, or the consumer equates consumption in the two time periods. This is referred to as CONSUMPTION SMOOTHING. In Figure 3.10, consumption smoothing with $C_1^* = C_2^*$ occurs at the point where a 45-degree line from the origin intersects the point where the slope of the indifference curve equals the slope of the intertemporal budget constraint.

› If the utility function is separable in consumption across time periods and if the MRTP equals one plus the interest rate, **consumption smoothing** occurs.

CONSUMPTION TILTING by contrast is the case where the consumer is impatient and tilts their consumption towards the present—$C_1^* > C_2^*$ and $\delta > r$ as depicted in Figure 3.11(A), or the consumer is patient and tilts their consumption towards the future—$C_1^* < C_2^*$ and $\delta < r$ as depicted in Figure 3.11(B).

› Consumption tilting is the case where the consumer's impatience for the present versus future consumption is not equal to the interest rate at which income is discounted.

The essence of this approach is that current consumption depends not only on the current income but also on the future income. It is also dependent on the interest rate, which determines the slope of the budget constraint, and on the preferences of individuals, which determine the shape of the indifference curves. Suppose the individuals' preferences are such that they smoothen their consumption over time, $C_1^* = C_2^*$. Substituting this allocation of lifetime consumption into the intertemporal budget constraint $C_1 + [C_2/(1+r)] = Y_1 + [Y_2/(1+r)] = W$ and solving for current consumption, we obtain




**› Figure 3.11**
Consumption Tilting. Individuals who are impatient and prefer more consumption today (panel A) and those who are patient and prefer more consumption in the future (panel B) depict consumption tilting behaviour.

$$C_1^* + \frac{C_1^*}{1+r} = W$$

or,

$$\frac{2+r}{1+r} C_1^* = W$$

or,

$$C_1^* = \frac{1+r}{2+r} W = \frac{1+r}{2+r}\left(Y_1 + \frac{Y_2}{1+r}\right) \tag{3.17}$$

› Permanent income is that constant stream of income which equals the present discounted value of the fluctuating stream of current and future incomes.

Equation (3.17) makes it clear that current consumption depends on current and future incomes and the interest rate. It indicates that current consumption is some kind of weighted average of current and future incomes. Milton Friedman[8] in his study on consumption pursued this by defining PERMANENT INCOME as a kind of average of present and future incomes. More precisely, we may define permanent income ($Y_p$) as that constant stream of income ($Y_p, Y_p$) in periods 1 and 2, which gives the same lifetime income stream as does the fluctuating income stream ($Y_1, Y_2$). In short, $Y_p$, the annuity value of ($Y_1, Y_2$), is given by

$$Y_p + \frac{Y_p}{1+r} = Y_1 + \frac{Y_2}{1+r}$$

or,

$$\frac{2+r}{1+r} Y_p = Y_1 + \frac{Y_2}{1+r}$$

or,

$$Y_p = \frac{1+r}{2+r}\left(Y_1 + \frac{Y_2}{1+r}\right) \tag{3.18}$$

It turns out that permanent income $Y_p$ in Eq. (3.18) is the same expression as the right-hand side of Eq. (3.17) and noting this we can then write

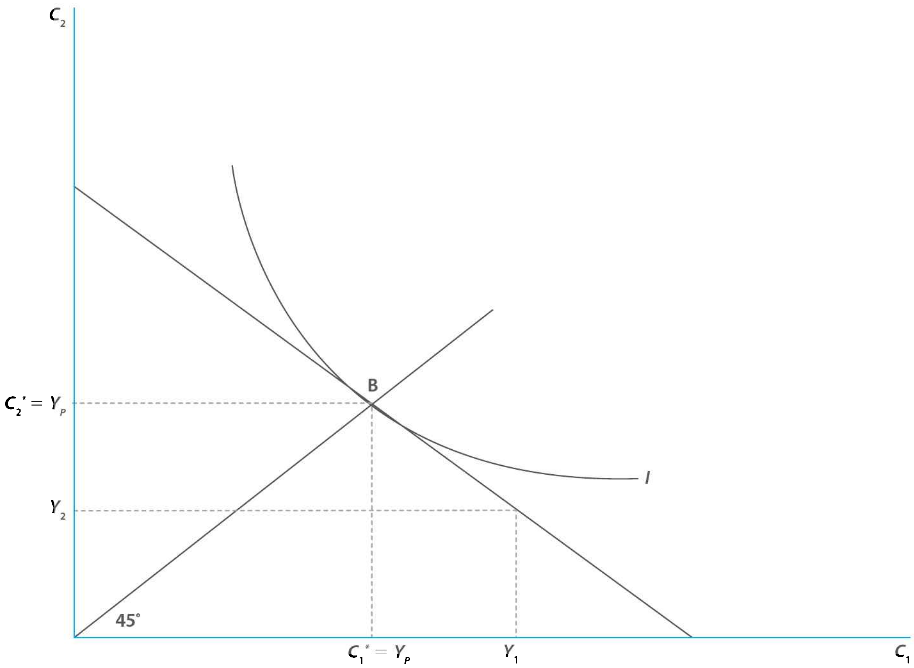
$$C_1^* = C_2^* = Y_p \tag{3.19}$$

› When the **consumption** in each period is set equal to the permanent income, it is called consumption smoothing.

› Consumption smoothing is identical to consumption each period being equal to **permanent income**.

Thus, when the individual possesses preferences, which lead them to maintain a stable consumption path, consuming the same every period, we can say that CONSUMPTION is set equal to permanent income.

In other words the individual decides on their consumption levels on the basis of their PERMANENT INCOME, and not current income as in Keynes. In graphical terms, permanent income is a constant stream across periods and so its value is represented by the point where a 45-degree line from the origin intersects the intertemporal budget constraint—point B in Figure 3.12.




› **Figure 3.12**
Permanent Income and Consumption. **Consumption smoothing is identical to the individual consuming permanent income and saving the difference between actual and permanent income.**

Consumption smoothing is akin to the individual setting current consumption equal to permanent income. SAVING is, thus, the gap between current and permanent income or

› An individual who smoothens consumption across time periods has a saving given by the difference between permanent and actual income.

$$S_1 = Y_1 - C_1 = Y_1 - Y_P \quad \textbf{(3.20)}$$

Recall that the data in Figure 3.4 indicate that individuals pursue consumption smoothing. This implies that the savings of individuals in India is the amount of the difference between current and permanent income.[9]

# 3.3 Temporary and Permanent Shocks

Up to now we have treated future income as a known variable. In fact, Ando and Modigliani[10] demonstrated that, on the average, future income is proportional to current income, or $Y_2 = \beta Y_1$, where $\beta > 0$. However, there are always shocks to income and to expect it to grow smoothly by a proportional amount within each period is not realistic. Accordingly we need to discuss the effects on consumption and saving of SHOCKS to income.

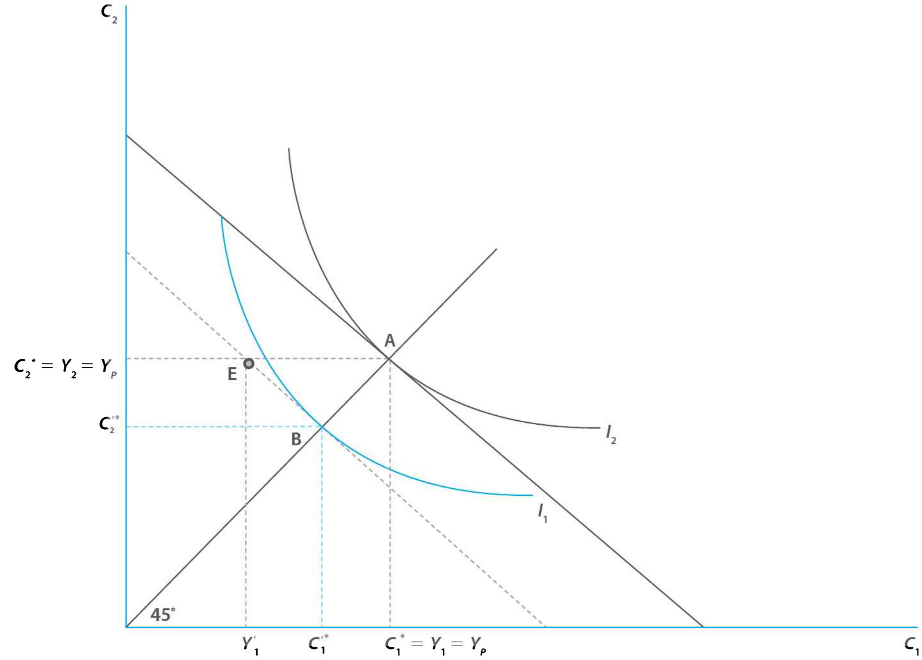
› A shock is a disturbance that changes the value of a variable by a random amount.

To do this in a manner that is amenable to non-messy graphical analysis, we assume first that the individual's preferences are such that they allocate income across time so as to smoothen consumption, or $C_1^* = C_2^*$. Second, we presume that the profile of income is identical to that of permanent income so that $Y_1 = Y_2 = Y_p$. Both these assumptions imply [see Eq. (3.19)]

$$C_1^* = C_2^* = Y_1 = Y_2 = Y_P \quad \textbf{(3.21)}$$

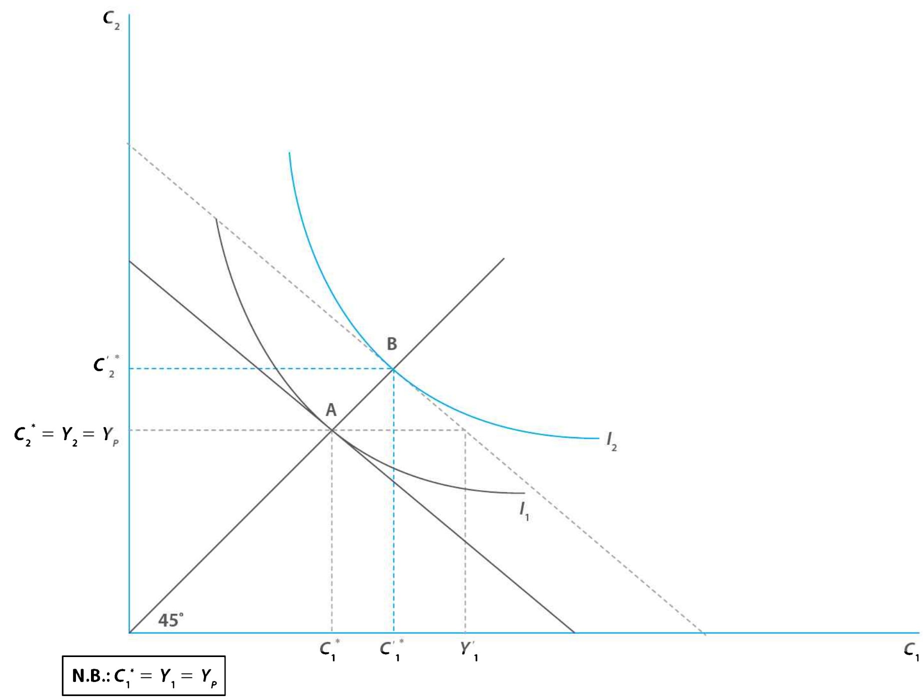
In Figure 3.13, this is equivalent to the individual being on their intertemporal budget constraint at point A. There are five possible situations arising from this scenario:

**› Figure 3.13**
Temporary Current Negative Shock. **Individuals adjust to a temporary current negative shock that reduces income by dissaving and maintaining the smoothness of consumption.**




*Case 1*: Suppose there is a temporary *current shock*, say a drought, and current income falls from $Y_1$ to $Y'_1$. $Y_2$, however, remains the same. The individual's lifetime income profile is now represented by point E rather than point A. Through E draw a line with slope $(1 + r)$ to obtain the alternative consumption possibilities from the new income stream. The budget line shifts inwards parallel to itself as $Y_2$ is unchanged, and the dashed line represents the new budget line resulting from the temporary current negative shock. As the individual prefers to smooth consumption, consumption will now be at the point where the 45-degree line intersects the new dashed budget line—point B in the figure. As the figure reveals, the individual consumes more than the current income,

**› Figure 3.14**
Temporary Current Positive Shock. **A temporary current positive shock that increases income results in an individual increasing saving while smoothening consumption.**



that is $C'^*_1 > Y'_1$, and this is possible because the shock is temporary and the individual borrows part of future income to finance current consumption.

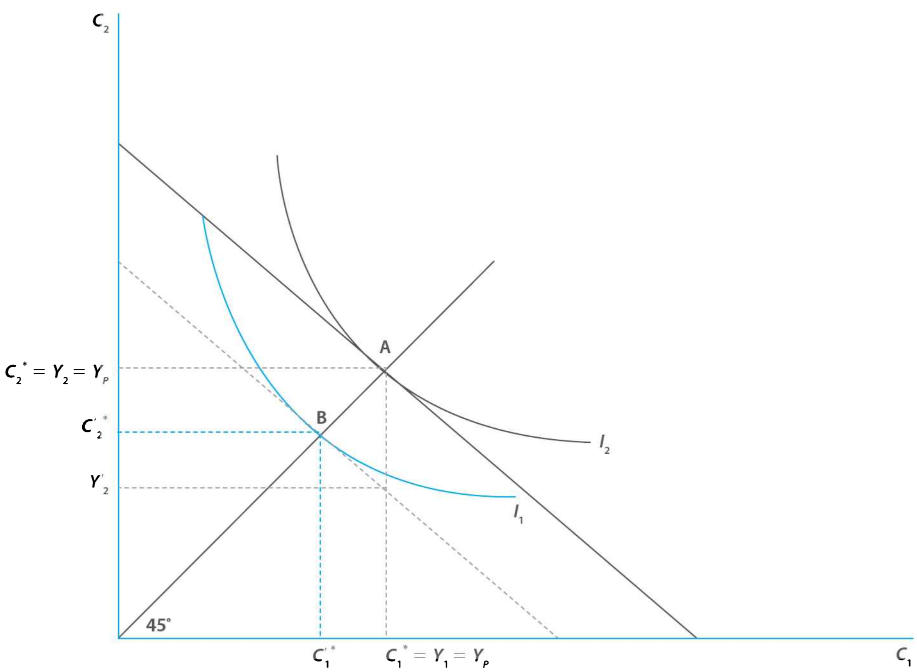
Savings were zero before the shock $C^*_1 = Y_1 = Y_p$, and after the shock $C'^*_1 > Y'_1$ and there is dissaving. A temporary current shock that results in a decline in current income is thus accompanied by a decline in savings. This is because individuals prefer to smoothen out their consumption and absorb the decline in income by reducing consumption a bit but allowing the shock to hit their savings considerably more.

*Case 2*: In contrast, if the temporary shock to current income is a favourable positive shock (the individual wins a lottery, for instance), then, as Figure 3.14 depicts, this results in an increase in current savings and an increase in consumption. With temporary current shocks, therefore, increases in consumption and savings absorb the impact of the shock.

*Case 3*: A second type of shock is an *anticipated future shock*. In this case, current income is unchanged at $Y_1$, but there is an expectation, say, of a future decline in incomes to $Y'_2$, say, due to a war that could break out. Figure 3.15 depicts this situation, with the dashed budget line indicating the post-shock situation.

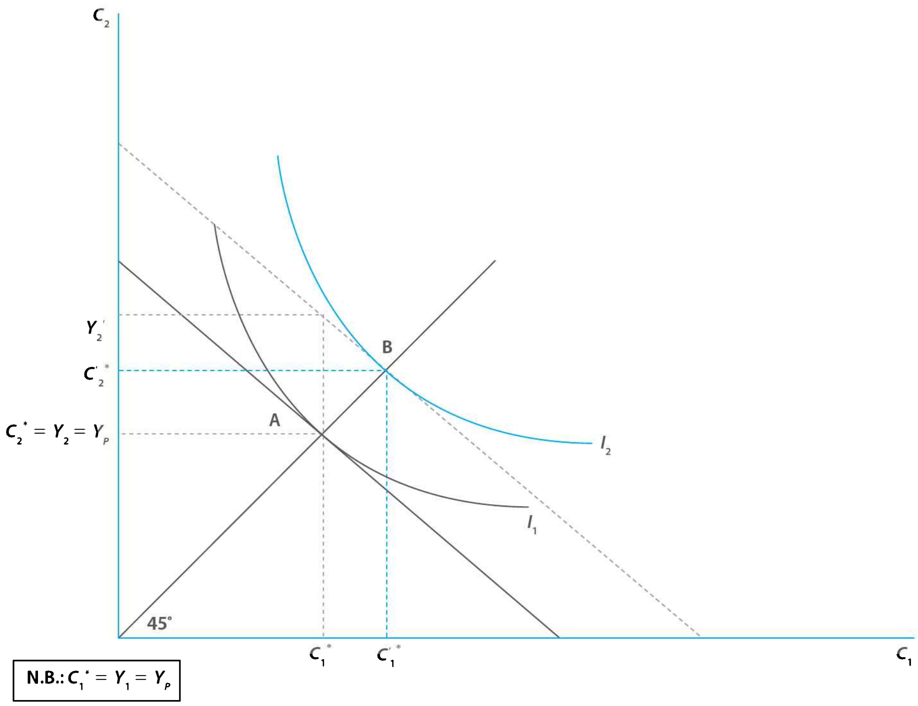
The consumption path moves from point A to point B with the individual adopting consumption smoothing. Current consumption declines to $C'^*_1$ despite no change in current income $Y_1$. With current consumption declining, savings must rise—anticipating a decline in future income, the individual is transferring the current income to the future by saving more and financing a smoothened consumption flow. Savings, thus, increase in response to an anticipated future decline in income.

*Case 4*: If the anticipated future shock to income is positive (say a discovery of an oil field), then (see Figure 3.16) the individual borrows some of that income for consumption in the present; and current consumption rises to $C'^*_1 > Y_1$ which amounts to a decline in savings with no change in current income. The individual cuts down savings because the future is




› **Figure 3.15**
Anticipated Future Negative Shock. **Anticipating a decline in future income, an individual increases savings today to smoothen consumption.**

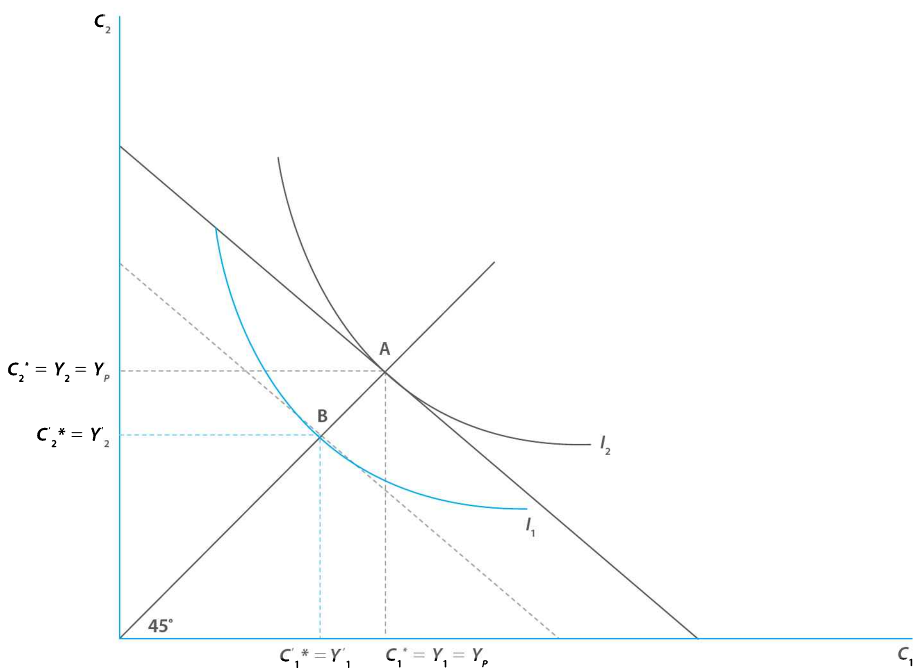
› **Figure 3.16**
Anticipated Future Positive Shock. **Anticipating an increase in future income, the individual decreases savings today and increases present consumption, as with an improved future income there is no need to save as much as before.**




now anticipated to be better and so they do not need to save as much as before.

*Case 5*: A third type of shock is a *permanent shock* to incomes—current and future incomes change by the same amount. If the permanent shock is a negative shock (say a dictator assumes power and property rights are no longer secure), then individuals adjust to this fully (see Figure 3.17) by reducing consumption by about the same amount as the decline in income in both periods. There is no change in savings. If the permanent shock is positive—the commercial introduction of newly invented computers or steam engines—then analogously,

› **Figure 3.17**
Permanent Negative Shock. **If the decline in income will persist in the future then the individual adjusts by reducing consumption by the extent of the decline in income.**



› **Table 3.1**
Effect of Shocks on Savings

| Outcome | Type of Shock to Income |
|---|---|
| Decline in Savings | Temporary Negative Current Shock |
| | Anticipated Positive Future Shock |
| Increase in Savings | Anticipated Negative Future Shock |
| | Temporary Positive Current Shock |

the change in income is consumed entirely and again there is no change in savings as seems to have been the case in the US economy, as discovered by Kuznets.

To summarize the discussion (see Table 3.1), if we see a decline in current savings, we are informed by the theory, we have just developed, that there are two possible reasons for this to occur—there is a temporary negative current shock to income (see Figure 3.13), or there is an anticipated positive shock to future income (see Figure 3.15).

On the other hand, if there is an increase in current savings, then the two possible reasons for this to have happened are—there is a temporary positive current shock to income (see Figure 3.14), or there is an anticipation of a negative shock to future income (see Figure 3.16). Permanent shocks to income are absorbed by changes in consumption with no effect on savings (see Figure 3.17).

## **3.4 Stochastic Income Expectations

So far we have proceeded on the assumption that the future income is known completely. More plausible than this assumption of point expectations about future income is the association of different probabilities with different income levels. In such a stochastic environment, an individual maximizes the expected utility and not the utility as we have been presuming up to now. In this case, consumption smoothing requires an individual to not just equate consumption across periods $C_1^* = C_2^*$ but rather to set current consumption equal to the expectation of next period's consumption:[11]

MACROFOCUS 3.2

### Joint Families and Borrowing Constraints

An implication of life-cycle theory is that saving is a vehicle for transferring income from high- to low-income earning phases of the life cycle. In developing countries, however, households tend to be larger and there is also a tendency for quite a few generations to live together. Such households resort to internal transfers between workers and dependents and in the process take care of expenditures that are health related or for maintenance during old age. To the extent that individuals belonging to such households are ensured of such transfers, they would not require savings to compensate for shortfalls in income in any time period. Aggregate savings in societies, which are less individualistic, may be, therefore, lower.

The life-cycle model also assumes that individuals are free to borrow and lend at a given rate of interest. In developing economies, however, financial institutions are often unwilling to lend because individuals have little or no collateral, and even if they do, legal difficulties in securing the transfer rights of assets to creditors in case of defaulting makes them impose constraints on borrowing. Knowing that the possibility of borrowing is limited, individuals would forsake high consumption even when it is possible out of their current income, as a way of getting ready for periods of misfortune when incomes would be low. This is a *precautionary motive* for saving where individuals take precautions against bad times in which they know they will be unable to borrow.

This motive also provides us an insight expounded on by Deaton* as to why many households save even when the return on their assets is negative. Poor households often save in the form of grain stocks or cash.

The stocks of grain often have a real annual rate of return that is negative as stocks depreciate due to pests and desiccation. Similarly, cash has negative return due to inflation. Yet, grain and cash are major vehicles for savings especially by poor households as a precaution against dramatic drops in income due to bad harvests.

* A. Deaton, "Household Saving in LDCs: Credit Markets, Insurance and Welfare," *Scandinavian Journal of Economics* 94 no. 2 (1992): 253–273.

$$C_1^* = E(C_2^*) \tag{3.22}$$

However, consumption in any time period deviates from its expected value by a random error factor, or

$$C_2 = E(C_2) + \epsilon_2$$

where $\epsilon_2$ is a surprise or innovation error that is non-forecastable. We could thus write the expected value of next period's consumption as

$$E(C_2) = C_2 - \epsilon_2$$

Substituting this expression into Eq. (3.22) gives us the result for consumption smoothing in a stochastic environment

$$C_1^* = C_2^* - \epsilon_2$$

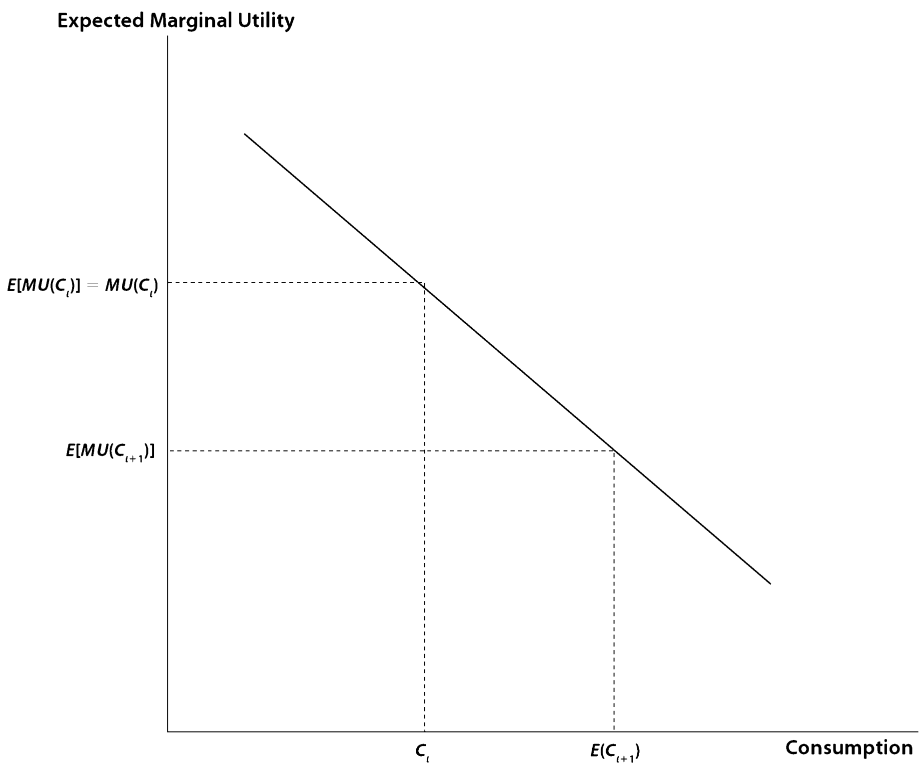
or,

$$C_2^* = C_1^* + \epsilon_2$$

Generally,

$$C_t^* = C_{t-1}^* + \epsilon_t \tag{3.23}$$

This is Hall's[12] famous result for consumption, which has the revolutionary interpretation that the best forecast of the next period's consumption is this period's consumption. No other currently available information is of any use. The intuition for the result is that if consumption is expected to change, the individual can do a better job of smoothing consumption. Suppose consumption is expected to rise—$E(C_{t+1})$ increases. This means (see Figure 3.18) the marginal utility of consumption today $MU(C_t) = E[MU(C_t)]$ is greater than the




› **Figure 3.18**
Consumption Smoothing Under Stochastic Income. **If there is uncertainty about future income, the individual sets consumption today to equate the expected marginal utility of consumption now and in the future.**

expected marginal utility of consumption in the future, $E[MU(C_{t+1})]$, whereas optimization requires that they be equal.

An individual seeking to maximize utility would then be better off raising current consumption and as the result shown in Eq. (3.23) informs us, the individual should adjust current consumption to the point where consumption is not expected to change.

HALL'S RESULT ran counter to existing views on consumption. The traditional view of consumption over the business cycle is that when output and income declines, consumption also declines but is expected to recover—there are predictable movements in consumption. Hall's result by contrast is that when output declines unexpectedly, consumption declines only by the amount of the fall in permanent income and so it is not expected to recover.

› When **Hall** presented his paper, it is said that a prominent macroeconomist commented that he must have been on drugs when he wrote the paper.

# 3.5 Effect of Interest Rates

So far we have examined how shocks to income affect consumption and savings. We now examine the effects of interest rate changes. It is often thought that as the interest rates rise, the rate of return to savings rises, and so savings would also increase. However, the relation between interest rates and savings is a bit more complex than that. Consider an individual with income stream ($Y_1$, $Y_2$) represented by point E in Figure 3.19.

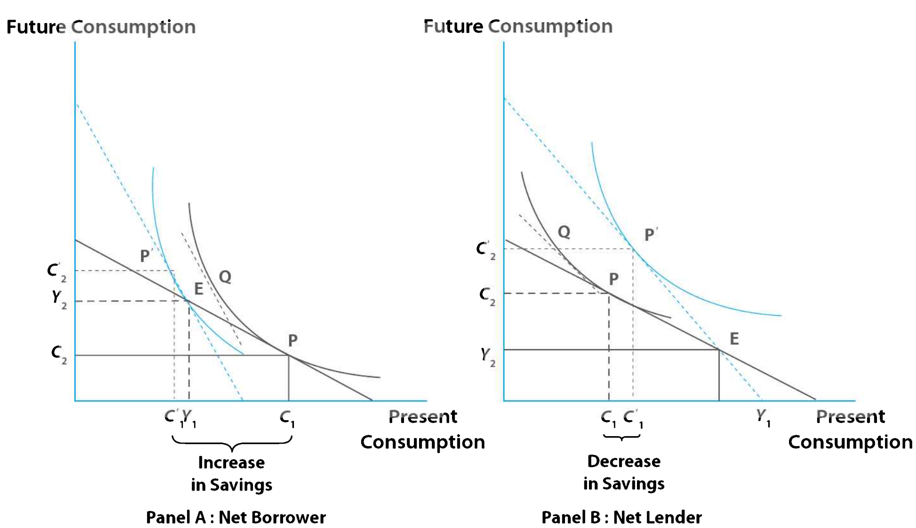
A rise in the interest rate will make the slope of the budget line given by $(1 + r)$ steeper—the intertemporal budget constraint rotates through point E in a clockwise direction and is represented in the figure by the dashed budget line. The individual who was originally consuming at point P in Figure 3.19(A) as a result of the rise in interest rates changes their allocation of consumption to point P′ where present consumption has declined and future consumption has risen relative to their values at point P. This is because higher interest rates have reduced current consumption and raised current savings.

Figure 3.19(B), however, portrays a contrasting picture. The very same rise in interest rates results in an increase in consumption and a fall in savings. The effect of the rise in the interest rates on saving is therefore ambiguous.

Economists explain the ambiguity by dividing the effect of the interest rate increase into two parts—a SUBSTITUTION EFFECT and an INCOME EFFECT.

› A substitution effect indicates the change in consumption solely due to a change in the interest rate and is measured along the initial indifference curve.

› An income effect is the change in consumption due to the change in purchasing power or income implicit in the change in the interest rate.




**› Figure 3.19**
Effect of Rise in Interest Rate. **An increase in the interest rate increases savings only if the individual is a net borrower or if the substitution effect dominates the income effect.**

Let us consider the substitution effect. As consumption plus savings exhaust income, a decline in consumption must be matched by an equal rise in savings for a fixed income, $\Delta S_1 = -\Delta C_1$. The increased savings earns a rate of return $r$ which adds to consumption in the next period, $\Delta C_2 = (1 + r)\Delta S_1$. A rise in the interest rate thus enables more future consumption and in effect has made future consumption cheaper relative to current consumption. Individuals, thus, substitute away from current consumption and increase future consumption. More technically, the substitution effect measures the change in the individual's desired allocation of present and future consumptions when interest rates rise, assuming that individuals remain on the initial indifference curve. Along the original indifference curve, this implies a shift from point P to point Q where the slope of the indifference curve equals the new higher interest rate. As can be seen at point Q, the higher interest rate unambiguously leads to a reduction of $C_1$ and a rise in $C_2$, which means that savings rises.

› The substitution effect of a rise in the interest rate causes current consumption to decline and savings to rise. The income effect is ambiguous. A net lender reduces and a net borrower increases savings.

The SUBSTITUTION EFFECT always tends to raise saving but the same is not true of the income effect. The impact of the income effect depends on whether the individual was a net lender or a net borrower. In panel A of Figure 3.19, the individual is a net borrower as consumption in the first period exceeds individual's income, whereas in panel B the individual is a net lender as consumption is less than the first period income and lends that savings which contributes to individual's future consumption. If the individual is initially a net lender [see Figure 3.19(B)], the rise in interest rates makes the individual richer. Note that in the situation when current consumption is unchanged, the individual can afford a higher level of future consumption due to the higher interest rate—the dashed new budget line around point P is above the solid budget line that indicates the lifetime income prior to the change in the interest rate. The income effect of this increased allocation of future consumption possible due to an expanded budget set makes the individual richer and the individual tends to raise consumption levels of both $C_1$ and $C_2$. This is an income effect which is shown in Figure 3.19(B) as a shift from point Q to P′, which raises current consumption $C_1$ and reduces savings.

If the individual is initially a net borrower [see Figure 3.19(A)], the rise in interest rates makes that individual poorer as evidenced by the fact that with current consumption unchanged the individual can no longer afford the original level of $C_1$ as the new dashed budget line representing the rise in the interest rate is below the solid budget line prior to the change in interest rate. The income effect of this diminished allocation of current consumption, possible due to a contracted budget set, makes the individual poorer, and the individual tends to reduce consumption levels of $C_1$. This is an income effect which is shown in Figure 3.19(A) as a shift from point Q to P′ which reduces current consumption $C_1$ and raises savings.

The substitution effect, then, always causes individuals to transfer incomes to the future by increasing savings as the rise in interest rate makes future consumption cheaper. The income effect reduces savings for net lenders and raises savings for net borrowers. The total effect of a rise in interest rates is to raise

› **Table 3.2**
Effect of Interest Rate Increase on Savings

| | Net Lender | Net Borrower |
|---|---|---|
| Substitution Effect | + | + |
| Income Effect | – | + |
| Total Effect | ? | + |

or lower savings for an individual who is a lender depending on whether the substitution or income effect dominates. For an individual who is a borrower, the rise in interest rates unambiguously raises saving. The aggregate economy comprises a mix of borrowers and lenders. If the income effects of net borrowers and net lenders tend to cancel each other at the aggregate level, the substitution effect will dominate. It is this presumption about the offsetting income effects that generates the expectation that a rise in interest rates will reduce current consumption and raise aggregate savings (see Table 3.2). The empirical evidence backs up the validity of this presumption for developing countries.[13]

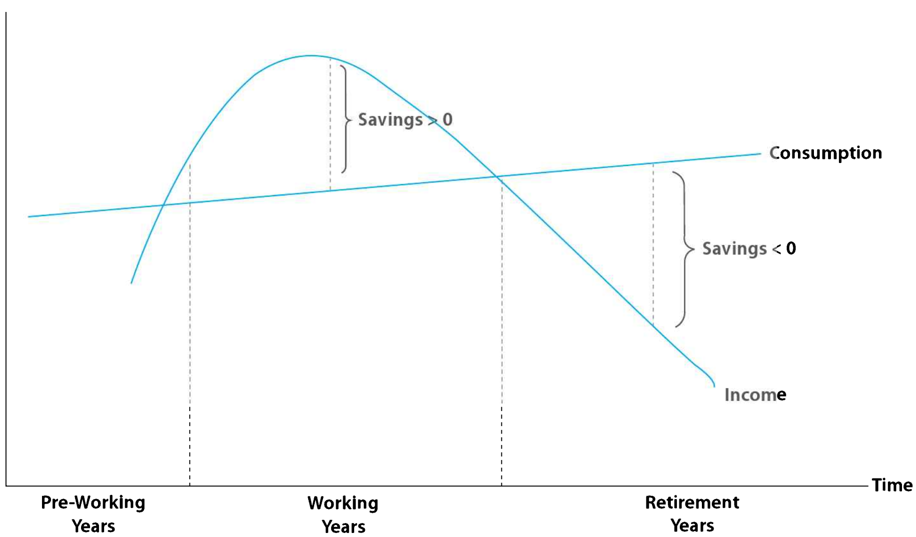
## 3.6 Aggregating Across Individuals

We have looked at the decision to split the income between consumption and savings from the perspective of an individual's decision-making so far. In the aggregate, individuals not only have different incomes but also belong to different age groups. When they are young, individuals have low incomes and tend to dissave (or incur debt) as they know they will be earning more later in their lives. During working years, incomes rise to reach a peak at maybe just after middle age and individuals repay their debts incurred earlier and save for their retirement years. When individuals finally reach the retirement phase of their lives (see Figure 3.20), their incomes due to them from past work (pensions) are much lower and they top this up with incomes accruing from past savings to maintain a steady level of consumption. It was MODIGLIANI[14] who pointed out to this observation that in an individual's life there are two periods of dissaving—early years prior to working and retirement years—and that savings is, therefore, determined by one's stage in the life cycle.

› **Franco Modigliani** emphasized that the planning horizon of individuals is their lifetime, which in a two-period set-up includes a working period and a retirement period.

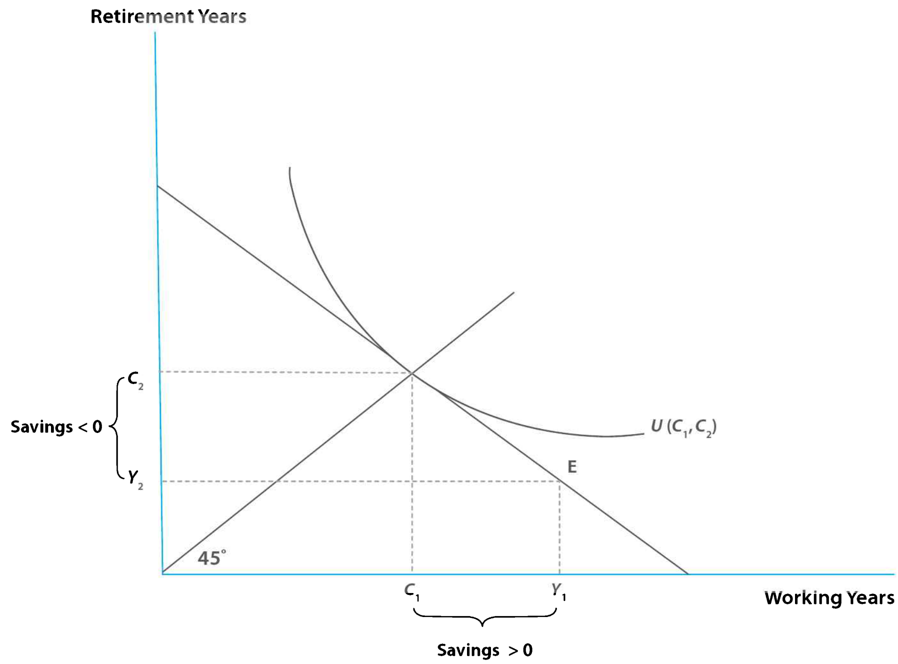
Figure 3.20 depicts the typical patterns of consumption, savings, and income across a person's life cycle. If we temporarily ignore the fact that an individual has a younger period of life (because we cannot use two-dimensional graphs to depict three periods of life as shown in Figure 3.20), then our two-period model with the current period as working years and the future period as retirement years can be used to understand the right-hand two-thirds of Figure 3.20.

In Figure 3.21, higher incomes during working life relative to retirement life is depicted by the point E. Consumption smoothing makes individuals transfer income from working years, when incomes are larger, to provide for




› **Figure 3.20**
Life-Cycle Income and Consumption. **Over the life cycle, individuals start becoming net savers around their middle age. Dissaving is typical prior to joining the workforce and during retirement.**

› **Figure 3.21**
Life-Cycle Income and Savings. Individuals save during working years so as to smooth consumption as income is low during retirement years.




retirement, when incomes are smaller. The hump-shaped pattern of income over the life cycle as depicted in Figure 3.20 and represented by point E in Figure 3.21 generates a time path of consumption and savings that arises due to individuals pursuing consumption smoothing across time.

Aggregation across individuals is, therefore, complicated. Individuals differ in their preferences for consumption today versus the future and they also differ in their stage in the life cycle. Young savers coexist with even younger and older dissavers. The aggregate savings in the economy is the outcome of the balance of saving and dissaving averaged over the entire population. In a static economy with no population growth and no growth in incomes per person, and with the young and old population evenly distributed, we could expect aggregate savings to be very low as the older generation would be dissaving at about the rate at which the young are saving for retirement.

› The dependency rate is the ratio of dependents (the young and the old population) to the working population.

In most economies, however, overall savings is positive—savings by the young exceeds dissaving by the old—due to population growth and a growth in output per person stemming from technological change that fosters improvements in productivity. So, demographic profile (as proxied by the DEPENDENCY RATE in a country)—the proportion of the very young and the retired to the working population—and technical change are important determinants of aggregate savings. However, why some countries save at a very high rate while others save so little is still a topic that is being extensively researched. In 1987, for instance, aggregate savings to GDP was 34 per cent in Japan while it was just 14 per cent in the United States which prompted research into the causes for this.[15] Currently, the same question is being asked about why is the savings rate in China, at over 40 per cent of GDP, so high. China is ranked below 100 in the world according to its per capita income. Its savings rate, however, has been one of the highest worldwide in recent decades. Just before Modigliani passed away, he along with his co-author, Shi,[16] put together data to show (predictably) that income growth as it turned into a market-oriented economy has been the dominant factor behind the dramatic increase in China's savings rate with demographic structure and inflation also having significant impacts.

## **3.7 Savings and Portfolio Choice

Let us now concern ourselves with a more complicated situation—individuals having diversified portfolios. Up to now, we allowed the individual to put their savings into just one asset which had a fixed rate of return. In reality, savers do invest in more than one asset and this would be related to the uncertainty of the yield associated with assets. That savings portfolios are generally diversified requires us to drop the assumption of one asset and to enquire into how savings is allocated among different assets. This requires an explanation of the determinants of the optimal allocation of a portfolio of given size. Of course, with uncertainty in asset yields, future income is now subject to capital risk. The decision on the total amount of saving to be made out of the current income is, therefore, interdependent with the portfolio decision about how to allocate savings between various kinds of assets. After all, savings in assets that yield high returns increase future income which in turn affects future (and current) consumption. A satisfactory theory should, therefore, be able to explain

**MACROFOCUS 3.3**

### House Prices and Consumption

Housing is a dominant component of wealth for many households with more households owning homes than stocks. In many economies, house price rises have been accompanied by a consumption boom. For example, Barata and Pacheco* state, "the housing market appears to be more important than the stock market as a factor influencing consumption" and find that a 1 per cent increase in housing wealth leads to a 0.1–0.2 per cent increase in consumption for various European countries. In Taiwan, the savings rate reached a peak of 38.5 per cent of GDP in 1987 and, subsequently, dropped to a low of 24 per cent by 2001. The real average housing prices in Taiwan increased almost three times from 1987 to 1990 and then remained stable. Hsueh** found that the rise in house prices had a windfall effect on consumption in Taiwan.

An increase in house prices may not impact positively on the consumption expenditure of all households through a wealth effect. (A wealth effect is the tendency of changes in asset prices to affect household wealth and their spending on consumption goods. As wealth is the value of assets minus liabilities, such as outstanding debt, an increase in house prices being an increase in the value of real estate assets is an increase in wealth.) A young household with limited savings is typically "short" in housing services and must purchase these services in a spot market—it must rent in housing services—and/or it must save to make the initial down payment on the house before it can take out a mortgage. Banks usually insist that prospective homeowners seeking a loan pay upfront a percentage of the cost of the house. An older household that has already purchased a house is "long" in housing services as a house is a long-lived asset that delivers a stream of housing services over time. With the value of housing services per year being the rent for leasing these services, the price of a house can be thought of as the present value of future rents. When house prices rise relative to rents, potential homebuyers will choose to rent thereby reducing the demand for houses and increasing the demand for leasing on rent, and in the process, bring house prices back in line with rents. A rise in the price of housing is thereby accompanied by a commensurate rise in rent or an increase in the consumption of housing services.

(1) For young homeowners who remain in their houses over the long term, the increase in house prices is roughly offset by the increase in the cost of housing services and there should be no impact on consumption. For older homeowners, the increase in their wealth is more than the cost of housing services over their remaining lives. If real estate can be used as a collateral for borrowing, then an increase in the value of real estate will increase the value of collateral and with it the availability of funds for consumption. Consumption by older homeowners will increase. Campbell and Cocco† find for the UK that there is a large positive effect of house prices on consumption for old households who are homeowners and a close to zero effect for young households. In the aggregate then consumption may be responsive to house prices as older homeowners become an increasing fraction of the population.

(2) If homeowners cannot borrow by pledging their ownership of real estate as collateral because banks are reluctant to accept such assets due to the difficulty of securing such assets in a weak legal enforcement regime, consumption will be unaffected by housing prices. Even those homeowners who want to trade down and consume less housing may find barriers in their way in terms of high moving and transactions costs. When transactions costs pertaining to services charges for brokers and document preparation by lawyers, stamp duty, and registration charges are significant (estimated as varying between 8 and 27 per cent of housing costs in India—NHB††), homeowners are unable to realize windfall capital gains and will not treat them as an increase in wealth or savings.

(3) By contrast, young households who intend purchasing a housing unit will have to save more to make a higher down payment for the house when house prices increase and their consumption declines.

(4) Those who view housing price increases to be so drastic that it makes it impossible for them to save enough money for the down payment to buy a housing unit—those discouraged from homeownership—may respond by increasing their consumption.

* J. M. Barata and L. M. Pacheco, "Asset Prices and Monetary Policy: Wealth Effects on Consumption," *The Twentieth Symposium on Banking and Monetary Economics*, University of Birmingham, June 2003.
** L. M. Hsueh, "The Relationship Between Housing Price, Tenure Choice, and Saving Behaviour in Taiwan," *International Real Estate Review* 3 (2001): 11–33.
† J. Y. Campbell and J. F. Cocco, "How Do House Prices Affect Consumption? Evidence from Micro Data," NBER Working Paper Series, Report No. 11534, August 2005.
†† NHB, "Transaction Cost of Housing in India: An Analysis," Occasional Paper No. II (New Delhi: National Housing Bank, May 2007).

› The amount of savings out of income constitutes the size of the portfolio.

› The decision on the forms (assets) in which to hold an individual's savings is the composition of the portfolio.

how the SIZE of the portfolio—the amount of savings out of income—which is residually explained by the consumption theory, and the COMPOSITION of the portfolio—the allocation of the savings among different assets—are simultaneously determined.[17] In this section, we will consider the portfolio choice when there is more than one asset into which savings can be put.

Earlier income was deployed either to consumption or to savings which we identified as a bank deposit with a known interest rate, $r$. Now, we allow income in the first period to be used for consumption in the same period, or to be saved in an asset. This could be a secure asset $M$, such as a demand deposit with a known rate of return $r$, or a risky asset $X$, which has a random rate of return $x$. If $C_1$ is the first period's current consumption, the savings or the initial asset holdings of the individual in time period 1 is

$$S_1 = A_1 = Y_1 - C_1 = M + X \tag{3.24}$$

The risky asset has a random rate of return. To enable graphical representation, we focus on just two states of the world or outcomes from investing in the risky asset:
State 1: The risky asset yields more than the safe asset, or $x = x_1 > r$; and State 2: The risky asset yields less than the safe asset, or $x = x_2 < r$. Note that in order that the risky asset is considered worth holding, we would, of course, require that the average return on the risky asset dominates the return on the safe asset, or $\bar{x} > r$.

Now, the income received in the second period from the first period savings, $A_2$, is the initial savings, $A_1$, plus the return from the investments in the safe asset and the risky asset:

$$\begin{aligned} A_2 &= A_1 + Mr + Xx \\ &= (M + X) + Mr + Xx \\ &= M(1 + r) + X(1 + x + r - r) \\ &= (M + X)(1 + r) + X(x - r) \\ &= A_1(1 + r) + X(x - r) \end{aligned} \tag{3.25}$$

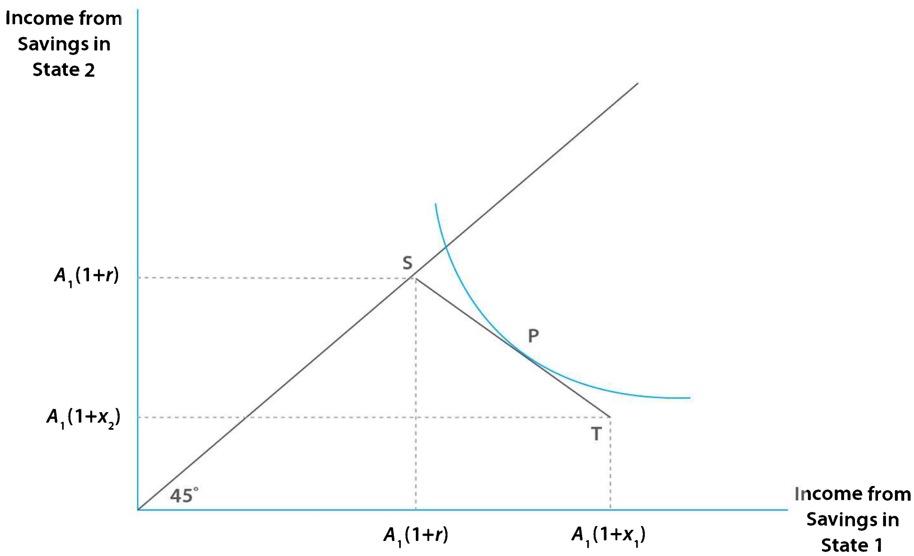
where $(x - r)$ is the excess return on savings in the risky asset. If all the savings is used to buy the safe asset, we have the situation just as before when there was only one sure return asset

$$A_2 = A_1(1 + r) = M(1 + r)$$

As the asset is safe, it provides the same income to the individual regardless of which state of the world occurs. In Figure 3.22, which measures income from savings in State 1 on the horizontal axis and income from savings in State 2 on the vertical axis, putting all savings into the safe asset implies a second period income from assets on the 45-degree line such as at point S. If all the savings is put into the risky asset, we have

$$A_2 = A_1(1 + x) = X(1 + x)$$

With a random return on the risky asset, $A_2$ can be:
(Case i) As low as $A_1(1 + x_2) < A_1(1 + r)$, if State 2 occurs; and
(Case ii) As high as $A_1(1 + x_1) > A_1(1 + r)$, if State 1 occurs.




› **Figure 3.22**
Portfolio Possibilities. Savings can be allocated to a risky or a safe asset. Any combination of these assets is represented by the budget line ST. The individual chooses point P to maximize the expected utility of the return from the amounts invested in the safe and the risky asset.

In Figure 3.22, point T represents the outcome when the individual puts all their savings into the risky asset. It is evident that points S and T are end points where either all savings is in the safe asset (point S) or all savings is in the risky asset (POINT T). The individual can hold a mix of the safe and risky asset and can attain any point on the line ST—the budget constraint the individual faces. Given that there is a risk, the individual is taken to maximize an expected utility function (rather than just a utility function) with income from savings as its argument, that is, the individual chooses the amount of savings to be invested in the risky asset to maximize the expected return from savings, $E[U(A_2)]$, subject to the budget constraint in Eq. (3.26) that is depicted as line ST in the figure.

› The individual can, in principle, extend their budget line beyond **point T** by borrowing. However, we ignore shorting here.

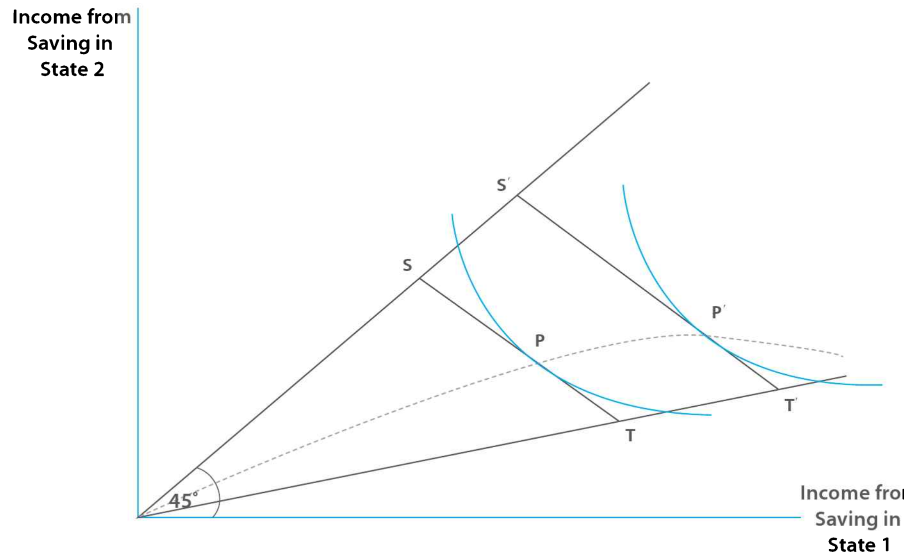
If $p$ is the probability of State 1 occurring, the objective of the individual is

$$\underset{\{X\}}{\text{Max}}E[U(A_2)] = pU[A_1(1+r) + X(x_1 - r)] + (1-p)U[A_1(1+r) + X(x_2 - r)] \tag{3.26}$$

The above expected utility function generates indifference curves, each indifference curve representing combinations of income from savings in the two states of the world which result in the same constant expected utility. The individual maximizing their expected utility selects the highest possible indifference curve, which is given by the point of tangency between the budget line and an indifference curve at point P in the figure. At this point P, the proportion of savings in the risky asset is the ratio of the distance SP to ST. Thus, ratio SP/ST of savings is put into the risky asset with return $x$ and [1− (SP/ST)] is put into the safe asset with return $r$.

If the individual decides to save more in period 1, $S_1 = A_1$ increases and this moves the budget line outwards in a parallel manner from line ST to S′T′ in Figure 3.23.

When this occurs, the selected portfolio mix of assets saved is now P′. As savings increases and the budget line shifts outwards, we can trace out the locus of points of tangency between indifference curves and successive budget lines which represent the individual's allocation of savings between the safe and the risky asset such that they maximize the expected utility of income from savings. In Figure 3.23, this locus of points is given by the dashed line which bends down indicating that as savings increases, the proportion allocated to the risky asset increases.[18]




› **Figure 3.23**
Increased Savings Allocated to Risky Asset. As savings increases, the budget line shifts out and the individual depicted here increases the proportion of savings allocated to the risky asset.

It is obvious that an individual's preference for risk will determine the effect of savings variations on portfolio allocation. Figure 3.24 depicts three other alternative loci. Panel A shows that as savings increases, the proportion allocated to the risky assets decreases.[19] Panel B depicts the situation where as savings increases, the proportion of savings allocated to the risky asset remains unchanged.[20] Finally, panel C depicts the situation where all the increase in savings goes into the safe asset such that there is an equal increment in income from savings regardless of which state of the world occurs.[21] We can conclude that in a situation of risk the purchase of a risky asset in a savings portfolio depends on the individuals' preferences which indicate their aversion to risk.

So far we have sought to understand how the individual will select their portfolio of assets given that they have already determined their level of savings. As savings changed exogenously, Figures 3.23 and 3.24 depicted how the portfolio allocation between the safe and risky assets might change for an individual. In general, the individual chooses the level of consumption—and thus, savings which is the residual income after consumption is determined—and the composition of their portfolio of savings simultaneously.

Future consumption, however, is a stochastic variable and is the income from savings in the earlier period ($A_2$), plus other earnings ($Y_2$) in the later period.

$$C_2 = Y_2 + A_2$$

$$= Y_2 + A_1(1 + r) + X(x - r) \qquad \textbf{(3.27)}$$

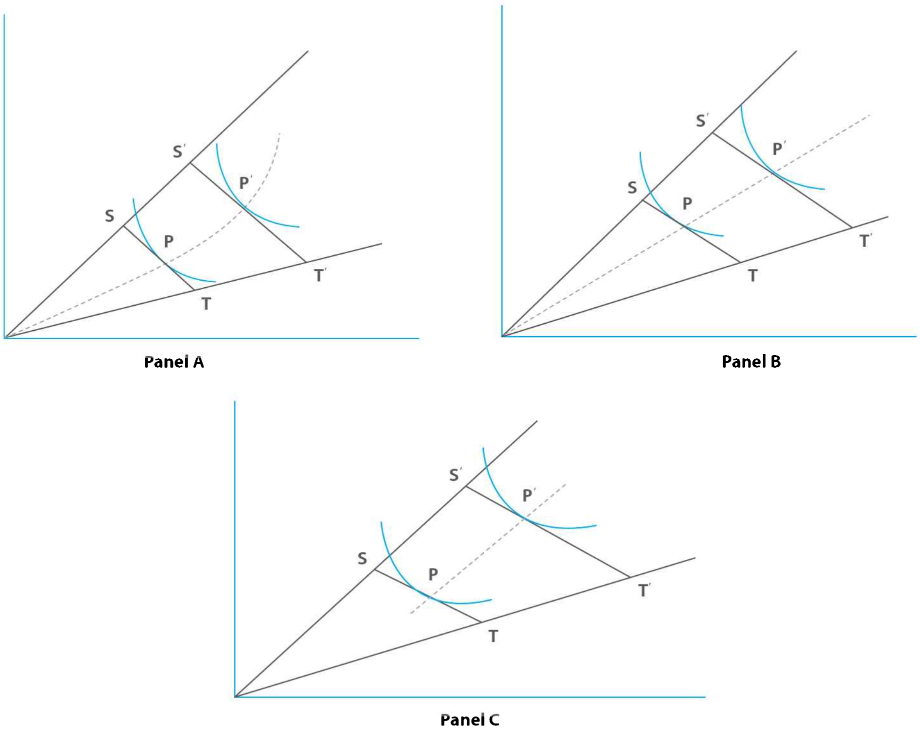
$A_1$, however, is the savings of the present period—Eq. (3.24)—and so we may write $C_2$ as

$$C_2 = Y_2 + (Y_1 - C_1)(1 + r) + X(x - r) \qquad \textbf{(3.28)}$$

The preferences of the consumer is over consumption profiles represented by their utility function $U(C_1, C_2)$. However, in a situation of risk consumers maximize expected utility which is given by

$$E[U(C_1, C_2)] = pU(C_1, C_2) + (1 - p)U(C_1, C_2) \qquad \textbf{(3.29)}$$

Keeping in mind that in State 1 the random variable $X$ takes the value $x_1$ with probability $p$ and in State 2 it takes the value $x_2$ with probability $(1 - p)$, we substitute the expression for $C_2$ into the expected utility function to obtain




› **Figure 3.24**
Alternative Portfolio Allocations as Savings Increases. **As savings increases, the individual may allocate less of their savings in the risky asset–panel A; keep the proportion of their savings in the risky asset unchanged–panel B; or put all of their increased savings in the safe asset–panel C.**

$$E[U(C_1,C_2)] = pU[C_1,Y_2 + (Y_1 - C_1)(1+r) + X(x_1 - r)] \\ + (1-p)U[C_1,Y_2 + (Y_1 - C_1)(1+r) + X(x_2 - r) \qquad \textbf{(3.30)}$$

Maximizing the expected utility of the stream of consumption thus requires the individual to select $C_1$ and the amount saved in the risky asset, $X$, for a maximum in Eq. (3.30). This leads to the following two first-order conditions:

$$p[U_1 - U_2(1+r)] + (1-p)[U_1 - U_2(1+r)] = 0$$
$$pU_2(X_1 - r) + (1-p)U_2(X_2 - r) = 0$$

The above expressions are the sum of outcomes across the probability distribution and can thus be written in terms of the expectations operator as

$$E[U_1 - U_2(1+r)] = 0 \qquad \textbf{(3.31)}$$

and
$$E[U_2(x - r)] = 0 \qquad \textbf{(3.32)}$$

As can be seen, the former first-order condition is a generalization of the certainty case to the case where there is a random return which requires the use of the expectations operator:

$$E(U_1) - E(U_2)(1+r) = 0$$

or
$$\frac{E(U_1)}{E(U_2)} = (1+r)$$

The left-hand side of the above expression is the slope of the indifference curve along which the expected utility is constant and the right-hand side is the slope of the budget line.

Similarly, the second first-order condition is a direct counterpart of the first-order condition in the portfolio model as can be seen by selecting $X$ to maximize Eq. (3.26). Thus, consumption and the allocation of a savings portfolio across different assets can be understood as interdependent decisions.

## SUMMARY

- Keynes' fundamental psychological law of consumption states that an increase in disposable income is accompanied by an increase in consumption but of a smaller magnitude.
- The marginal propensity to consume is the incremental consumption arising from the increment to disposable income and is less than unity.
- The average propensity to consume is consumption per unit of disposable income. The marginal propensity to consume is less than the average propensity to consume.
- The average propensity to save or the savings rate rises with income.
- The variation in consumption is significantly less than the variation in disposable income.
- The consumption smoothing approach to consumption employs the decision framework that people choose rationally from among the opportunities available that one which best furthers their preferences for consumption today versus consumption in the future and that people are forward looking.
- The intertemporal budget constraint represents the present discounted value of current and future incomes (life-cycle income).
- An individual exhibits positive (negative) time preference if they require more (less) than one unit of future consumption to compensate themselves for the loss of a unit of current consumption. The rate of time preference declines as one moves downwards along an indifference curve.
- If the utility function is separable in consumption across time periods and if the marginal rate of time preference equals the interest rate, then consumption smoothing occurs.
- Consumption tilting is because the consumer's impatience for present versus future consumption is not equal to the interest rate at which income is discounted.
- Permanent income is that constant stream of income that gives the same lifetime income as the fluctuating income stream that occurs.
- Savings is the difference between the current income and the permanent income when consumption smoothing occurs.
- Savings declines when there is a temporary negative current shock or an anticipated positive future shock.
- Savings increases when there is an anticipated negative future shock or a temporary positive current shock.
- A rise in the interest rate increases savings for an individual who is a net borrower and has an ambiguous effect on an individual who is a net lender.
- A hump-shaped pattern of income exists over the life cycle causing individuals to transfer incomes via savings from working years to provide for retirement.

## NOTES

1. John M. Keynes, "The General Theory of Employment, Interest and Money," *The Collected Writing of John Maynard Keynes* (London: Macmillan, 1992).
2. In the national accounts, there is a distinction between personal disposable income and private disposable income, which we ignore here.
3. The data on consumption and disposable income contain a unit root but co-integration is rejected. Even when using instruments to correct for endogeneity the GMM estimates are no different from the OLS estimates.
4. The national accounts in India do not report real personal disposable income. We have estimated this series using a comparison with real net national product at factor cost.
5. We should expect this because with a linear relation between $C$ and $Y^d$, we should have Var $(C) = c^2$Var $(Y^d)$.
6. Simon Kuznets, *National Income: A Summary of Findings* (New York: National Bureau of Economic Research, 1946).
7. Note that as the slope of the budget line is $(1 + r)$, an indifference curve tangent at this slope implies a positive marginal rate of time preference.
8. Milton Friedman, *A Theory of the Consumption Function* (Princeton, NJ: Princeton University Press, 1957).
9. $Ct = 1.028C_{t-1}$ does not strictly indicate a marginal tilting of consumption and the patience of waiting for slightly more consumption in the future because of high dependency ratios (ratio of old to young and economically active population) and a higher share of a younger population in the demographic profile of the country. This is an issue that requires a more advanced treatment and we leave it at this.
10. Albert Ando and Franco Modigliani, "The 'Life Cycle' Hypothesis of Saving: Aggregate Implications and Tests," *American Economic Review* 53, no.1 (March 1963): 55–84.
11. Two problems arise when deriving an exact analytical solution for consumption in a stochastic environment. First, the real interest rate is stochastic and the expectation of the product of two stochastic variables is not the product of expectations- $E[(1 + r) / (1 + \delta)]C_2 \neq E[(1 + r) / (1 + \delta)]E(C_2)$. Second, in general the marginal utility $(1 / C_2)$ is non-linear in consumption and so the expectation of next period's marginal utility of consumption is different from that of expected consumption—$E[U'(C_2)] \neq U'[E(C_2)]$. One case where an exact solution is possible is when preferences are quadratic, so that marginal utility is linear in consumption. If to that we allow the market interest rate and subjective discount rate to coincide $(\delta = r)$, as Hall did, we get the result as shown in Eq. (3.23).
12. Robert E. Hall, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy* 96 (1978): 971–987.

13. Alberto Giovannini, "Savings and the Interest Rate in LDCs," *World Development* 11, no. 7 (July 1983): 601–607.
14. Franco Modigliani, "Life Cycle, Individual Thrift and the Wealth of Nations," Nobel Lecture, *American Economic Review* 76, no. 3 (June 1986): 297–313.
15. Fumio Hayashi, "Why is Japan's Saving Rate So Apparently High?" in Stanley Fischer (ed.) *Macroeconomics Annual* (New York: National Bureau of Economic Research, 1986).
16. Franco Modigliani and Larry Cao Shi, "The Chinese Saving Puzzle and the Life-Cycle Hypothesis," *Journal of Economic Literature* 42, no. 1 (March 2004): 145–170.
17. Larry Selden, "A New Approach to the Joint Consumption-Portfolio Problem," *Journal of Money, Credit, and Banking* 12, no. 3 (August 1980): 429–447.
18. Technically, this is referred to as the case when the wealth elasticity of the demand for the risky asset is greater than unity.
19. This is the case of a wealth elasticity of demand for the risky asset greater than zero, but less than unity.
20. Here the wealth elasticity of demand for the risky asset equals unity.
21. The wealth elasticity of demand for the risky asset is zero in this case.

## TEST YOURSELF

1. Suppose the consumption function (INR billion) for a closed economy is $C = 20 + 0.8Y^d$. Also suppose that disposable income increases from INR 50 billion to INR 150 billion. By how much would the consumption rise? By how much would the savings rise? What has happened to the savings rate?
2. If current income is INR 100 billion and future income is INR 150 billion and the rate of interest is 20 per cent, draw the intertemporal budget constraint associated with this stream of income. What is the slope of this constraint? Indicate the point on this constraint that depicts the income stream.
3. When does an individual use consumption smoothing and when does their consumption tilt?
4. What is permanent income and how can we determine savings behaviour on its basis?
5. Demonstrate how a temporary current positive shock to income and an anticipated negative future shock affect consumption behaviour.
6. Show how a decline in the interest rate affects the savings behaviour of a net borrower and a net lender.
7. If the labour ministry decides to raise the age of retirement by a few years because of an increase in longevity, what would be the impact on the current savings in the economy?
8. What sort of marginal rate of time preference would you expect a cricketer like Sachin Tendulkar—who is towards the end of his career—to have and why?
9. Suppose there are two time periods. In period 1, income is given by INR 200,000 and in period 2, it is INR 110,000. The rate of interest is 10 per cent. If consumption in period 1 is INR 150,000 is the individual consumption smoothing or consumption tilting?
10. Two youngsters A and B have similar qualifications and earn the same income. However, A comes from a family with a longer history of coronary problems and expects to die early. Who will have a higher present consumption and why?

## ONLINE APPLICATION

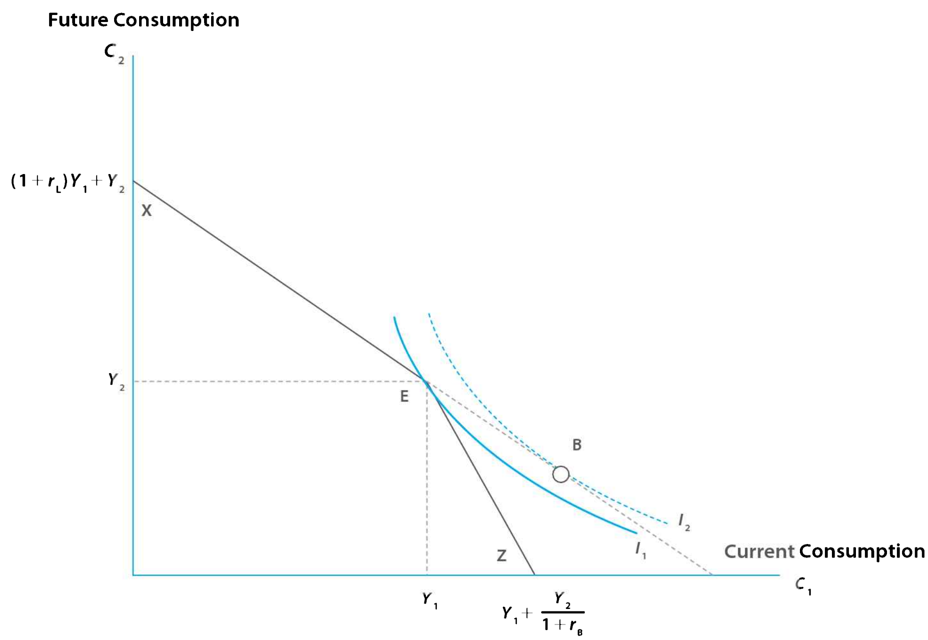
1. Go to the home page of the Central Statistical Organisation, National Accounts Division, Press Release and Statements. This page can only be viewed with the browser Internet explorer (URL: www.mospi.nic.in/mospi_cso_rept_pubn.htm).
2. Click on the "Macroeconomic aggregates at constant (1993–1994) prices, 1950–1951 to 2003–2004" icon. This will open up Statement S-1.2. You will be required to login as a new user to access this statement.
3. Create a spreadsheet of data on NDP at factor cost, Private Final Consumption Expenditure in the domestic market and population (which are in millions unlike the monetary data which are in crores) from 1990–1991 to 2003–2004.
4. In what way is NDP at factor cost different from Personal Disposable Income? See Statement 4 of the National Accounts Statistics which shows this accounting relationship. Unfortunately, personal disposable income data are only available in current prices and so cannot be used unless we deflate them appropriately. For our purposes, we will use NDP at factor cost as an indicator of private disposable income. This understates private disposable income by about 3.8 per cent.
5. Calculate per capita consumption expenditure and per capita income (net domestic product at factor cost divided by population).
6. Find out the average propensity to consume and draw a graph of this. Interpret this graph in terms of what it states about the behaviour of savings in India.
7. You have calculated the average propensity to consume on the assumption that direct taxes do not vary with income—as it is consumption out of income and not disposable income that you have calculated. To what extent would your calculation be incorrect if this assumption were not true?

# *Appendix 3.1* Borrowing Constraints

In the life-cycle model of consumption, individuals synchronize their income and preferred consumption profiles by borrowing and lending at a single rate of interest $r$. However, capital markets are not perfect and usually impose quantitative restrictions on borrowing or charge higher interest rates (on borrowing) than the rate at which individuals lend to the capital markets when they place their savings in various financial instruments. For instance, if you place your savings in a bank deposit, the interest rate you will receive is much less than that you would pay if you were to borrow an identical amount from the bank. Borrowing constraints prohibit individuals from consuming today the income they would receive in the future.

To understand the role played by capital market imperfections, assume that loans and savings instruments are available at constant but unequal rates of interest. The interest rate at which the individual lends to the capital market when they put their savings into financial instruments, $r_L$, or the interest rate received on lending, will be taken to be smaller than the interest rate at which they can take a loan or borrow from the capital market, $r_B$. When the individual is a net lender, their consumption in period 1 must be less than income, $C_1 < Y_1$, and $(Y_1 - C_1)$ is the amount loaned to the capital market at interest rate $r_L$.

**› Figure A3.1.1**
Borrowing Constraints and Consumption. **The presence of borrowing constraints with the interest rate on borrowing larger than the interest rate on lending results in the budget constraint XEZ. An individual who would have preferred to consume at point B, where they would have borrowed against future income, is now forced to limit their consumption by the amount of current income and consume at point E.**




If the individual is a net borrower, consumption is more than their current income, $C_1 > Y_1$, and $(C_1 - Y_1)$ is the amount of borrowing from the capital market at interest rate $r_B$. The interest rate faced by an individual can then be written as

$$r = \begin{cases} r_L & \text{if } C_1 \leq Y_1 \\ r_B > r_L & \text{if } C_1 > Y_1 \end{cases}$$

This implies that the intertemporal budget constraint is truncated about the point $(Y_1, Y_2)$—the coordinates representing current and future incomes. All points to the left of point E that depicts incomes $(Y_1, Y_2)$ in Figure A3.1.1 are points where current consumption is given by $C_1 \leq Y_1$ and $r_L$ is the operative interest rate.

All points to the right of point E depict current consumption where $C_1 > Y_1$ and the operative interest rate is the borrowing rate $r_B > r_L$. The locus XEZ then depicts the budget constraint of the individual with the segment XE having a slope given by $1 + r_L$ and the slope of the segment EZ is $1 + r_B$.

Let the indifference curves of the individual be given by $I_1$ and $I_2$ in Figure A3.1.1. The individual's maximizing choice of consumption would be at point B where the slope of the indifference curve equals the slope of the dashed budget line that represents borrowing and lending at a uniform interest rate in perfect capital markets of $r = r_L$.

When capital markets are imperfect, individuals are forced to operate along the locus XEZ and their preferences are maximized at a lower utility level at point E. The individual is then able to consume only up to their current income and cannot finance consumption beyond this by borrowing against the prospect of future income. Of course, if the indifference curves of the individual were tangential along the segment XE, then the presence of borrowing constraints does not affect the optimal consumption decision of the individual. The presence of borrowing constraints implies that consumption is driven by current income and that the life-cycle income may have a marginal impact on consumption behaviour.

# 4

# Investment

CRITICAL QUESTIONS

» *What is the cost of employing factor inputs such as physical capital and labour for a firm?*

» *How does a firm select the optimal stock of capital?*

» *What is Tobin's marginal Q?*

» *How does the cost of external finance affect the investment decision?*

» *What are the determinants of investment in housing and in inventory?*

» *How does the irreversibility of investment affect the decision to invest now rather than to wait and invest later?*

» *What are the other circumstances applicable to developing countries that affect investment behaviour?*

# 4.1 Profit Maximization and the Optimal Capital Stock

If consumption is more or less stable, then in a closed economy with $Y = C + I$, where the components of expenditures are an aggregate of private and public sector expenditures, it follows that investment must be the principal source of cyclical fluctuations. Given its inherent dynamics, however, theories have not been as successful in providing explanations for this variable. That is why, we start out with a static exposition and gradually let the dynamics creep in to the analysis. At the outset, it is advisable to make a distinction between financial and physical capital. *Financial capital* is essentially money or some other form of paper asset or security that function as contracts for the exchange of commodities or money. *Physical capital*, on the other hand, is a piece of productive equipment, such as a lathe or a Xerox machine that generates a flow of productive services over time. In this chapter, we will directly be concerned with real or physical capital—capital as a factor of production.

A firm or business requires capital inputs along with other inputs such as labour ($N$) so that it can produce an output that is sold on the market. Suppose the firm was to hire capital equipment and use it in the production process. For simplicity let us assume that the firm operates in competitive markets where it can sell all its output at the going market price, $p$, and it can hire any amount of capital it wishes at a rental rate of $k$ per year. To see how this rental rate could be determined, put yourself in the position of a business organization that rents out machines.

Let us say that the purchase price of a particular machine is INR 100 and the interest rate at which money can be borrowed or loaned is 10 per cent per annum. In order to cover the opportunity cost of the INR 100 tied up in the machine, you would have to charge INR 10 per year for it. There will be additional costs as well. A machine loses some of its value each year. This is partly because newer and more efficient machines are designed each year and existing machines lose economic value when this happens. This is a situation referred to as technological obsolescence. Also, as the machine is used, it undergoes wear and tear and needs maintenance, and sometimes, parts could even require replacement. The effect of this physical wear and tear and technological obsolescence could result in the price of the machine to fall by, say, INR 20 per year. At the very minimum, then, you would be willing to supply the machine in the market for services of capital inputs at a cost of INR 30 per year—INR 10 in forgone interest, and INR 20 in lost market value.[1] Generalizing, let $\delta$ stand for physical and technological depreciation, expressed as a fraction of the price of the capital good, and $r$ denote the market rate of the interest expressed in decimal form. Then, the annual rental rate on capital, $K$, will be given by

$$k = r + \delta$$

The unit cost to the firm for hiring capital equipment is then given by

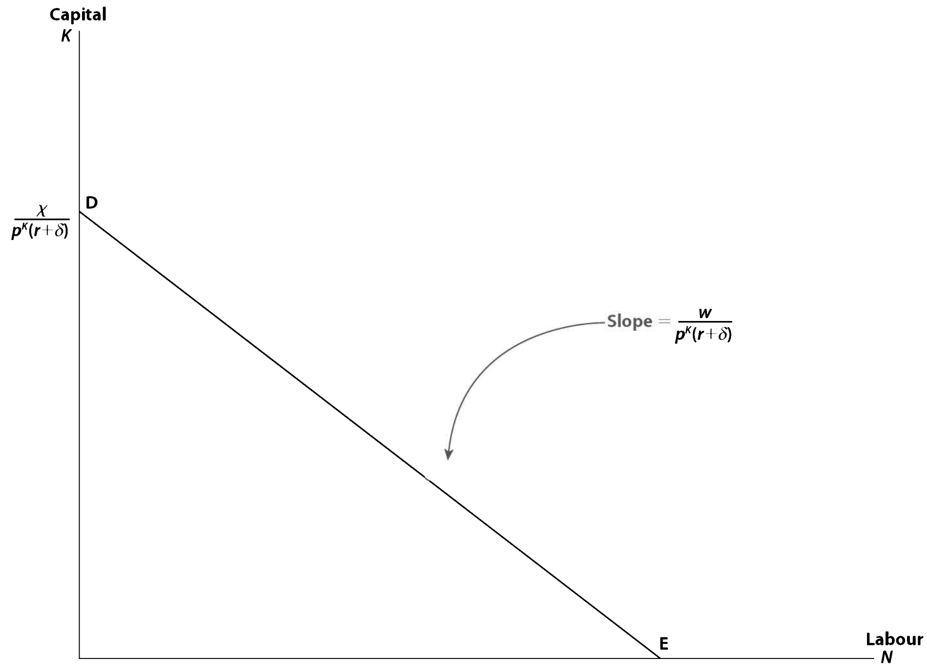
$$p^K k = p^K(r + \delta)$$

› The rental cost of capital, also referred to as the user cost of capital, is the cost of using capital over a given period of time (a year). This is the sum of the real interest rate and depreciation rate times the unit price of the capital good.

where $p^K$ is the supply price of the capital good and $p^K(r + \delta)$ is the UNIT CAPITAL RENTAL.

The other factor of production that the firm hires is labour ($N$ man-days), which is paid a compensation in the form of a wage rate of INR $w$

› **Figure 4.1**
Resource Constraint for Factor Inputs. The locus of inputs that can be purchased for a given level of resources is depicted as a resource constraint that the firm faces.




per man-day. The total input costs, $\chi$, incurred by the firm can then be written as

$$\chi = wN + p^K(r+\delta)K \quad \textbf{(4.1)}$$

In what follows, we simplify the discussion by assuming that the firm has static expectations regarding all the prices and interest rate, and technology is constant.[2] For a fixed $\chi$, the firm can select how much of the capital and labour to hire so as to exhaust this resource. Solving Eq. (4.1) for $K$ we obtain,

$$p^K(r+\delta)K = \chi - wN$$

or,

$$K = \frac{\chi}{p^K(r+\delta)} - \frac{w}{p^K(r+\delta)}N \quad \textbf{(4.2)}$$

› The resource constraint depicts for given input prices—the wage and the rental on capital—the locus of all possible input combinations that can be purchased for a given level of resources available to spend on purchasing these inputs.

› The production function describes how inputs like capital and labour are transformed into output.

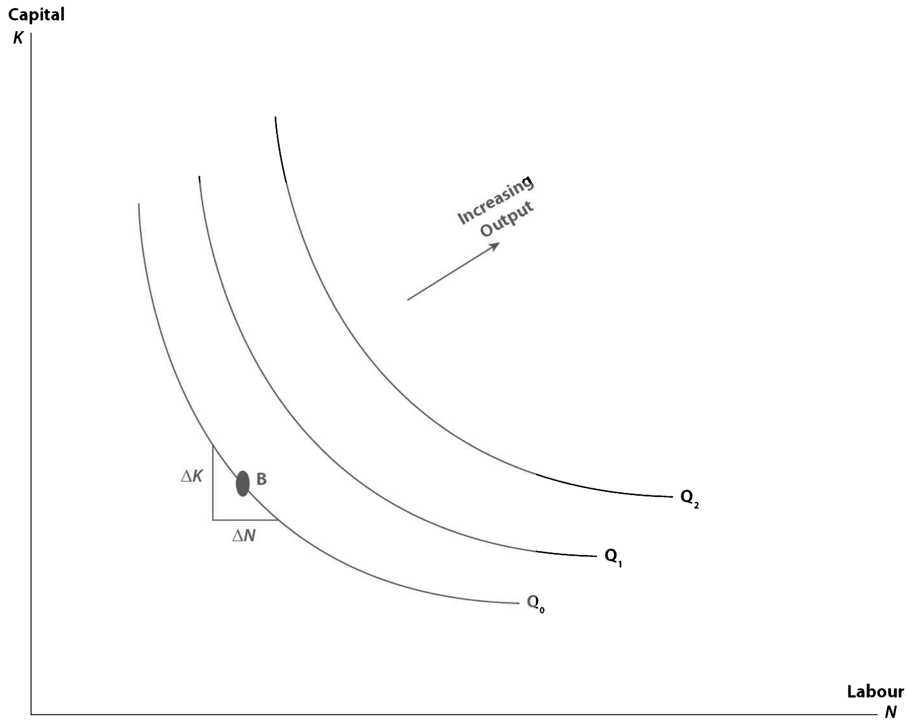
The graph of this RESOURCE CONSTRAINT is depicted in Figure 4.1, where, the slope of the resource constraint is $w/p^K(r+\delta)$, or the ratio of the relative price of a unit of labour to the rent on a unit of capital. As Figure 4.1 depicts, many alternative combinations of capital and labour can be hired as given by the points on the line DE. Which point on DE should the firm select? To answer this question we need to specify the output that can be produced using alternative combinations of capital and labour. This is given by the PRODUCTION FUNCTION that specifies how much output $Y$ we get when we employ specific quantities of $K$ and $N$.

This may be expressed as:

$$Y = F(K, N) \quad \textbf{(4.3)}$$

› The marginal product is the change in output that occurs in response to a unit change in the application of an input.

It is common to express that the production function is subject to diminishing marginal productivity—as we add more of a factor of production, with the other inputs held fixed, the resulting increments to output will eventually diminish. This property of a production function can be stated in terms of the MARGINAL PRODUCT ($MP$) of a factor of production, defined as the resulting change in output when an additional unit of the input is hired (with no change in the




**› Figure 4.2**
Isoquants. Each isoquant depicts the possible combinations of inputs that produce the same level of output.

amount of the other input used). For instance, the marginal product of labour, $MP_N$, is the increment to output when an additional unit of labour is hired, which can be written as:[3]

$$MP_N = \left.\frac{\Delta Y}{\Delta N}\right|_{\Delta K=0} = \frac{\partial Y}{\partial N} = F_N$$

Combinations of capital and labour that yield a given level of output are represented by an ISOQUANT. An isoquant for the production of output $Q_0$ for instance is given in Figure 4.2.

› An isoquant is the set of input combinations that yield a given level of output.

The slope of an isoquant is called the MARGINAL RATE OF TECHNICAL SUBSTITUTION or MRTS. It is the rate at which one input can be substituted for another without altering output. At point B in the diagram, the $MRTS$ is the absolute value of the slope, or,

› The marginal rate of technical substitution is the rate at which an input can be substituted or exchanged for another without altering the level of output.

$$MRTS = -\left.\frac{\Delta K}{\Delta N}\right|_{\Delta Q_0=0} = -\left.\frac{dK}{dN}\right|_{dQ=0} \quad \textbf{(4.4)}$$

If at point $B$ we reduce the amount of labour input marginally by an amount $\Delta N$, then the reduction in output caused by the loss of $\Delta N$ is equal to $MP_N(\Delta N)$. To compensate for the loss in output, if we augment $K$ by an amount $\Delta K$ that is just sufficient to maintain the original level of output, then the gain in output would equal $MP_K(\Delta K)$. As the reduction in output from having less $N$ is exactly offset by the gain in output from having more $K$, it follows that

$$MP_K \Delta K = -MP_N \Delta N$$

or,

$$-\frac{\Delta K}{\Delta N} = \frac{MP_N}{MP_K}$$

or,

$$MRTS = -\frac{\Delta K}{\Delta N} = -\frac{dK}{dN} = \frac{MP_N}{MP_K} \quad \textbf{(4.5)}$$

› The slope of an isoquant is the marginal rate of technical substitution which in turn is the ratio of the marginal products of labour to the marginal product of capital.

Hence the SLOPE OF THE ISOQUANT is the ratio of marginal products of the factor inputs.[4]

The shape of the isoquant tells us how one input can be substituted for another without changing the output produced. The bowed outward shape of the isoquant reflects that the more intensive the use of an input in production, the lower its marginal productivity. To maintain constant output, the firm needs to give up that input in exchange for the other input. Just like indifference curves, an isoquant map is a set of isoquants with isoquants further away from the origin representing higher levels of output.

The firm selects that level of output which maximizes profits. Given the firms' objective is profit ($\Pi$) maximization, this requires

$$\text{Max } \Pi = \text{Total Revenue} - \text{Total Cost} = pY - \chi$$

Substituting the production function given by Eq. (4.3) and the input cost function given by Eq. (4.1) into the profit function, we obtain:

$$\underset{\{K,N\}}{\text{Max}}\ \Pi = pF(K, N) - wN - p^K(r + \delta)K$$

In this equation for a given output price ($p$) and given input prices ($w$, $p^K$) and rental rate on capital ($r$, $\delta$), maximizing profits is akin to selecting the amount of employment and capital stock that produces the output that when sold gives the firm maximum profits.

An additional unit of capital will cost the firm $p^K$ ($r + \delta$) and this additional capital will produce an additional output given by the marginal product of capital, $MP_K$. This additional output generates a revenue to the firm given by the price times the additional output sold, $pMP_K$. So long as the additional revenue $pMP_K$ exceeds the additional cost $p^K$ ($r + \delta$), the input is adding more to the revenue than the cost of the input, and the firm will continue to hire additional units of the input. Thus, capital will be hired till the point where

$$pMP_K = p^K(r + \delta)$$

Similarly, labour will be hired till the additional revenue it generates will equal the additional cost of the labour, or,

$$pMP_N = w$$

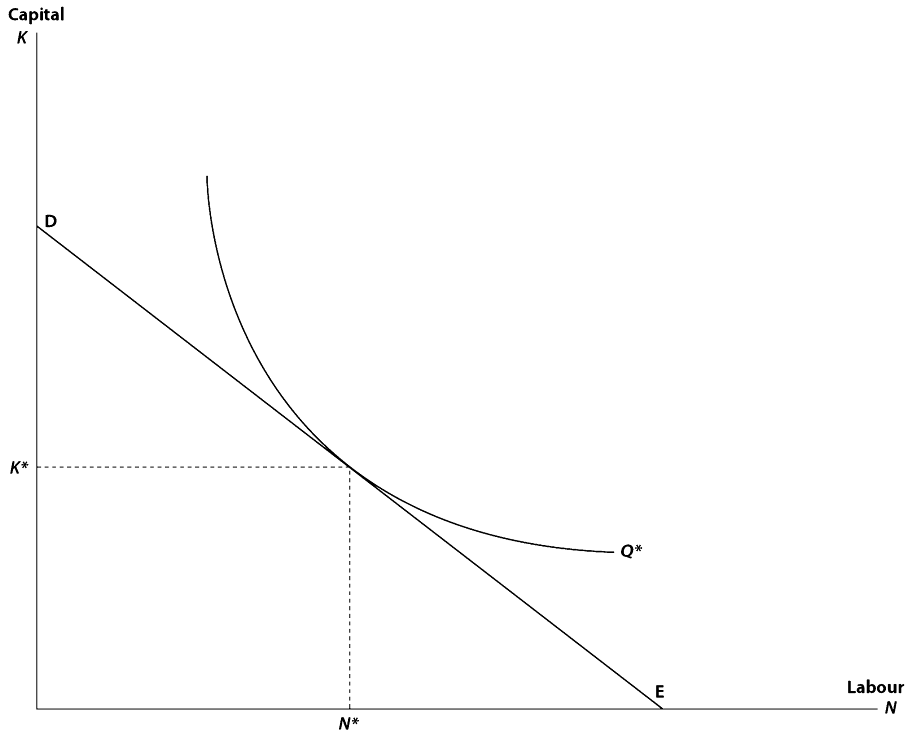
Dividing this equation by the previous equation

$$\overbrace{\frac{MP_N}{MP_K}}^{\text{slope of isoquant}} = \underbrace{\frac{w}{p^K(r + \delta)}}_{\text{slope of resource constraint}} \tag{4.6}$$

The left-hand side is the slope of an isoquant and the right-hand side the slope of the resource constraint. Hence, profit maximization by the firm requires it to select that capital stock and employment levels where the slope of the isoquant equals the slope of the resource constrain.[5]

› For the interested reader, Appendix 4.1 derives the **optimal** capital stock when the production function takes a particular form known as the Cobb–Douglas production function.

The firm that is a profit maximizer and asks itself how much of an input to hire would get the answer $K^*$ of capital and $N^*$ of labour (as depicted in Figure 4.3). Alternatively, the solution to Eq. (4.6) gives the OPTIMAL amount of capital and labour the firm should hire.




› **Figure 4.3**
Profit Maximization. **Profit maximization occurs where the slope of an isoquant equals the slope of the resource constraint. The set of inputs that maximizes profits is $K^*$ and $N^*$.**

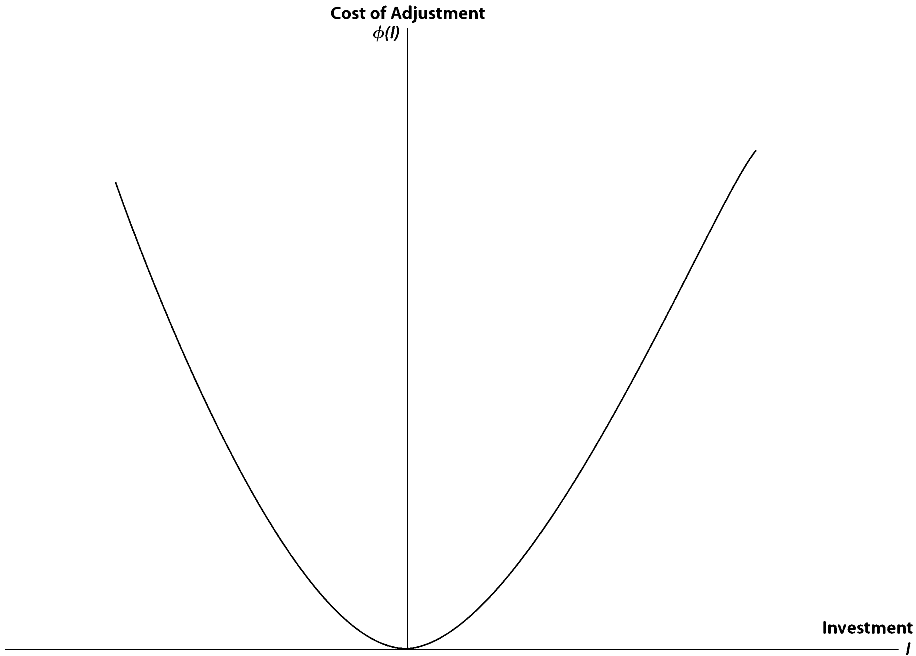
# 4.2 Adjustment Costs and Investment Decisions

Up to now we have proceeded by considering how much capital the firm should hire when it seeks to maximize profits. However, a feature that sets capital apart from other inputs is that whereas other inputs are normally hired on a period-by-period basis, capital equipment is usually owned outright by the firm. Rather than paying a rental of equal instalments per period over the capital equipment's life, the firm purchases the capital equipment outright at a cost of $p^K$.[6]

The second thing to note about our discussion is that so far, it was concerned solely with finding out the optimal capital stock that the firm should acquire in order to maximize profits. If the actual capital stock at any moment in time differs from the optimum capital stock, $K_t \neq K^*$, at what speed should the firm then go about acquiring new capital stock? Should it acquire capital slowly or should it adjust its capital stock to the optimal level immediately? Recall that investment is the change in capital stock and so another way of stating this is that with a divergence between actual and optimum capital stock, we can have investment at any rate between zero and infinity. The theory we have followed so far advises us about the adjustment to the stock of capital and the direction of that adjustment. It does not, however, prescribe the speed at which the process should evolve. It is a static theory that helps us locate equilibrium positions and it is silent about movements between equilibrium. How to adjust the CAPITAL STOCK is a topic that falls within the jurisdiction of investment theory which, accordingly, must be dynamic. It informs us about the speed at which we should acquire capital when a new stock of capital other than the current one is required in the pursuit of profit maximization.

› **Capital theory** informs us about the optimal stock of capital to be used as an input. It does not indicate the speed at which the stock of capital should be acquired, which is a dynamic problem that belongs to the realm of the theory of investment.

**› Figure 4.4**
Adjustment Costs of Investment.
**Adjustment costs rise more than proportionately with the level of investment.**




› Adjustment costs are the costs of integrating new capital goods into the production process. The larger the investment, the larger will be the cost of adjustment.

To get around the problem that firms should instantaneously adjust their capital stock to the desired level, a partial adjustment mechanism is added to the above theory about the choice of capital stock to yield a model of investment. Proponents of this approach such as Eisner and Strotz[7] argue that it takes time to put new machinery into place, to integrate it into the production process, and to train workers to use the new machinery. This is introduced into the theory as a convex COST OF ADJUSTMENT. For low levels of investment these adjustment costs are low, but the costs rise more than proportionately with the level of investment. This implies a cost of adjustment, $\phi(I)$, given by the following equation that is graphed in Figure 4.4:

$$\phi(I) = \frac{b}{2} p^K I^2$$

The cost of current investment activities thus has two elements— current expenditure on purchasing capital goods, $p^K I_t$, and the adjustment costs given by $\phi(I_t)$ incurred in absorbing increases in its productive capacity. Thus the total cost to the firm of investment activities at time $t$ is given by $p^K I_t + (b/2)\, p^K I_t^2$.

Recollect the motivation of the firm for increasing its capital stock is that the new capital will bolster the firm's rate of product not only in the current period, but also in the future. That extra output will enhance the firm's total revenue in the coming years. We ignore that there is a maintenance cost of the new capital and that at the end of its economic life it will have a scrap value.

To keep things simple, we proceed by assuming that the output is produced solely from the capital stock and that employment does not enter the picture. The output produced in this case is given by $Y = F(K)$. Our conclusion is not affected if we allow the production function to be linear—output is a linear function of the capital stock. Therefore, for simplicity, we allow the production function to be linear:

$$Y = F(K) = AK$$

where $A$ is a constant. In this case, the marginal product of capital is given by:

$$MP_K = F_K = A$$

We may alternatively write the production function then as:

$$Y = AK = MP_K K = F_K K$$

The firm is the usual profit maximizer—it maximizes the present value of the profits resulting from the sale of the output produced by the firm. The revenue that the firm earns when the price level is given by $p$ is $pY = pF_K K$. However, the capital stock depreciates at a constant rate of $\delta$ in each period, and the capital stock in period $t+1$, $K_{t+1}$, is the undepreciated capital stock plus gross investment, or,

$$K_{t+1} = (1-\delta)K_t + I_t$$

We begin the analysis in time period 0 where the new investment $I_0$ is undertaken at a cost of $p^K I_0 + (b/2)p^K I_0^2$ that contributes to the output of subsequent periods (see Table 4.1).

At time $t = 1$, the capital stock produces an output that results in a revenue of $pY_1 = pF_K K_1$ to the firm. As the investment decision is taken at $t = 0$, the present value of the discounted revenue to the firm when discounted at the rate $(1 + r)$ yields $pF_K K_1/(1 + r)$ of revenue in present value terms at time $t = 0$.

In the next time period, $t = 2$, the capital stock depreciates at the rate $\delta$, so that output in that period is produced from an input of the period $t = 1$ capital stock that now equals $K_1(1 - \delta)$. At time $t = 2$, the total revenue earned from the output produced by the capital stock is $pF_K K_1(1 - \delta)$ and we discount this at rate $(1 + r)^2$ to get the present discounted value of revenue earned in $t = 2$ of $pF_K K_1(1 - \delta)/(1 + r)^2$.

In a similar way, at $t = 3$, the capital stock $K_1(1 - \delta)$ depreciates again by $\delta$ so the output at that time is produced from a capital input given by $K_1(1 - \delta)^2$. The present discounted value of revenue earned at $t = 3$ is $pF_K K_1(1 - \delta)^2/(1 + r)^3$. Hence, the discounted present value of total revenue as a result of the investment in time period 0 is

$$\frac{pF_K K_1}{(1+r)} + \frac{pF_K K_1(1-\delta)}{(1+r)^2} + \frac{pF_K K_1(1-\delta)^2}{(1+r)^3} + \cdots$$

› **Table 4.1**
Investment, Capital Stock, and Output Over Time

| Time Period | Capital Stock | Investment | Cost of Investment | Value of Output |
|---|---|---|---|---|
| $t = 0$ | 0 | $I_0$ | $p^K I_0 + \frac{b}{2} p^K I_0^2$ | 0 |
| $t = 1$ | $K_1 = I_0$ | 0 | 0 | $pF_K K_1$ |
| $t = 2$ | $K_2 = (1-\delta) K_1$ | 0 | 0 | $pF_K K_2 = pF_K K_1(1-\delta)$ |
| $t = 3$ | $K_3 = (1-\delta) K_2 = (1-\delta)^2 K_1$ | 0 | 0 | $pF_K K_3 = pF_K K_1(1-\delta)^2$ |

*Note:* $K_{t+1} = (1-\delta)K_t + I_t$

We can write this[8] as:

$$\frac{pF_K K_1}{(1+r)} + \frac{pF_K K_1(1-\delta)}{(1+r)^2} + \frac{pF_K K_1(1-\delta)^2}{(1+r)^3} + \cdots = \frac{pF_K K_1}{r+\delta}$$

The total cost to the firm with only capital as an input in production is the cost of purchasing new capital goods, $p^K I_0$, and the adjustment costs, $(b/2)p^K I_0$, or,

$$\text{Total Cost} = p^K I_0 + \frac{b}{2} p^K I_0^2$$

The firm chooses the level of investment to maximize the present value of its profit $V_0$ which is the present value of total revenue less the cost of investment, or,

$$V_0 = \frac{pF_K K_1}{(1+r)} + \frac{pF_K K_1(1-\delta)}{(1+r)^2} + \frac{pF_K K_1(1-\delta)^2}{(1+r)^3} + \cdots - p^K I_0 - \frac{b}{2} p^K I_0^2$$

$$= \frac{pF_K K_1}{r+\delta} - p_K I_0 - \frac{b}{2} p^K I_0^2$$

Substituting the law of motion that links the capital stock with gross investment,

$$V_0 = \frac{pF_K[I_0 + (1-\delta)K_{-1}]}{r+\delta} - p^K I_0 - \frac{b}{2} p^K I_0^2$$

$$= \frac{pF_K I_0}{r+\delta} - p^K I_0 - \frac{b}{2} p^K I_0^2$$

as the analysis is from a period $t = 0$ in which $K_{-1} = 0$.

Maximizing this expression for the value of the firm with respect to the level of investment, $I = I_0$, we have,

$$\frac{pF_K}{r+\delta} - p^K - bp^K I = 0$$

which upon solving for $I$ gives us:

$$I = \frac{1}{b}\left\{\frac{pF_K}{p^K(r+\delta)} - 1\right\} \qquad \textbf{(4.7)}$$

› Tobin's marginal Q is the ratio of the marginal value of the capital stock (computed by adding the market value of the additional equity and the debt raised by the firm to finance the addition to capital stock) to the rental cost of capital.

$pF_K/p^K(r + \delta)$ is known as TOBIN'S MARGINAL Q or $Q^M$. $Q^M$ is the ratio of addition to the value of the firm or the marginal value of the firm, $(pF_K)$, to the cost of capital $[p^K(r + \delta)]$. Tobin had defined $Q$ as

$$Q = \frac{\text{Market Value}}{\text{Replacement Cost}} = \frac{pF_K K}{[p^K(r+\delta)]K}$$

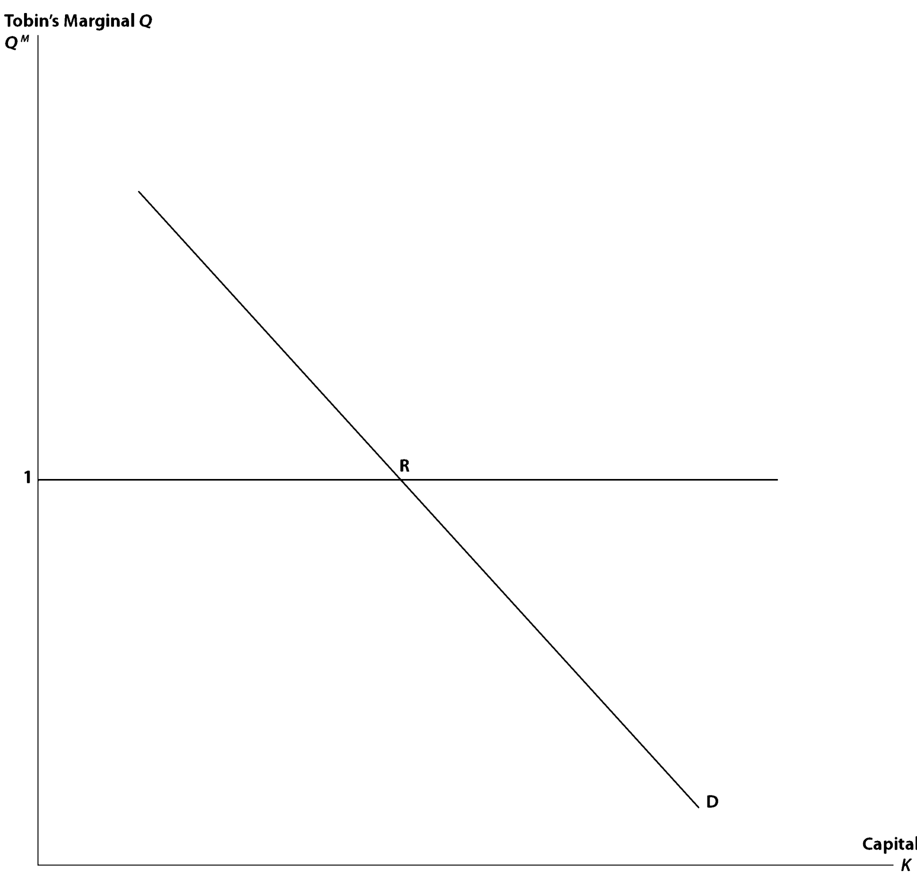
Our analysis implies that what is relevant for investment is marginal $Q$ or $Q^M$—the ratio of the market value of a marginal unit of capital to its replacement cost.[9] Equation (4.7) also reveals that investment is positive as long as $\{pF_K/[p^K(r+\delta)]\} - 1 > 0$ or $Q^M > 1$. The factors that have a bearing on investment according to Tobin's $Q$ theory are:

1. Value of marginal product of capital $(pF_K)$—A higher price for the produced output increases the revenue and raises investment. Also, technological progress that alters the marginal product of capital affects favourably the investment by the firm.
2. Price of capital goods $(p^K)$: A change in the price of the capital good changes the relative price of capital $(p^K/p)$ and induces a change in the stock of capital. For instance, a reduction in the relative price of capital causes the firm to increase the stock of capital.

3. Interest rate ($r$): A reduction in the interest rate decreases the opportunity cost of the stock of capital that has been invested and causes investment to rise.
4. Depreciation rate ($\delta$): Even though we took this to be a constant, if there is technological progress, there would be a reduction in the profits from current investments when that occurs.
5. Adjustment cost ($b$): The adjustment cost function depicts that adjustment costs increase more than proportionately with the amount of investment. A rise in the adjustment cost parameter will thus slow down the pace of investment towards the desired level of the capital stock.

Recall that $F_K$ is the marginal product of capital which due to diminishing marginal productivity gives us a downward-sloping curve as the stock of capital increases.[10] This implies that for a given price, $pF_K$ is downward sloping. The cost of capital, $p^K(r + \delta)$, at any moment of time is a known constant. Hence, $pF_K/[p^K(r + \delta)]$ is a downward-sloping curve that we can interpret as the investment-opportunities curve, which we represent by point D in Figure 4.5. To the left of the point R, investment is positive as the investment opportunities are greater than the marginal market value to replacement cost ratio of unity. An improvement in investment opportunities shifts the D curve to the right, increasing the desired capital stock. A decline in investment opportunities shifts the D curve to the left decreasing the desired capital stock.

Empirical applications require that we have a measure of $Q^M$. If financial markets are efficient, the value of an extra unit of capital would be given by the financial market value of that extra unit. $Q^M$ is related to the change in market value which would be the financial market's valuation of the profit stream from a small change in the firm's stock of capital. Though $Q$ can be




**› Figure 4.5**
Investment Opportunities.
Tobin's marginal $Q$ is positive when the revenue from a unit of output produced by hiring additional capital—$pMP_K$—is greater than the rental on capital. As the marginal product of capital decreases with an increase in the capital stock—diminishing marginal productivity—$Q^M$ is a downward-sloping line.

observed, $Q^M$ cannot be directly observed[11] and is much more difficult to measure. The market value of a firm is the stock market value of its shares (price of a share times the number of shares) plus its net debt or total value of bonds outstanding. Dividing this total market value by the value of the capital stock of the firm at replacement cost—the price firms have to pay to replace their current productive capacity of plant, machines—gives us a measure of Tobin's Q.[12] We also need an estimate of the adjustment cost parameter $b$ in Eq. (4.7) to estimate the investment function. Summers, Bosworth, Tobin, and White[13] found an estimated value of $b = 32$, but more sophisticated econometric analysis has shown that the cost of adjustment parameter is varying through time, whereas in the theory developed here it is time invariant.[14] Given the problems of measurement in the real world it is never an easy task to verify a theory.

## **4.3 Financial Structure and Investment

So far we have looked at firms responding to prices set in centralized securities markets such as the market interest rates of Tobin's marginal $Q$. The decision on investment and production in this approach was derived on the assumption that this is independent of the firms' financing decision. Firms are presumed to choose the mix of finance among internal funds, debt, and new equity independently of the choice of how much physical capital to acquire—the availability of finance does not limit their investment.[15] That financial structure can matter for investment decisions is an important

MACROFOCUS 4.1

### Overconfidence and Investment

Are investment decisions solely due to the characteristics of the firm such as the market value of assets? Or do the personal characteristics of decision makers in firms affect such decisions? The literature on behavioural finance has increasingly documented evidence of biases in beliefs such as overconfidence in managerial decision making. A large body of experimental literature has recorded the tendency of individuals to consider themselves above average on positive characteristics. A substantial fraction of entrepreneurs have been found to rate their chances of success as above average and consider themselves unlikely to experience failure. Corporate executives and management students have been found to be especially prone to this self-serving bias.*

The tendency of certain individuals to expect their behaviour to produce success leads them to attribute a successful outcome to their actions and if they fail, it is attributed to bad luck and is deemed to be unconnected to their endeavours. Such a self-serving attribution of outcomes reinforces overconfidence. The tendency of individuals to be optimistic about their own future prospects also makes them prone to overestimate outcomes to which they are highly committed.

Malmendier and Tate** construct a measure of CEO overconfidence. Many CEOs receive extensive stock-based compensation and since most of their income is in this form one would expect them to diversify their portfolio by selling company stock beyond the vesting period especially when it is profitable to exercise the stock option—the option is in-the-money. Many CEOs, however, do exactly the opposite and hold options till their date of expiration even when they would have been better off by exercising earlier and investing the proceeds in a diversified portfolio. Such CEOs display overconfidence and bet their personal wealth on the company's performance in the future.

Overconfident CEOs, as identified by their option holding behaviour, are sensitive to cash flows. But it is for a different reason than that provided by the asymmetric-information literature where external finance is restricted to avoid diluting the undervalued shares of the company (because good types are taken for bad types). Overconfident CEOs overestimate the returns on investment, and if they have sufficient internal funds, then they overinvest. If internal funds are not sufficient, overconfident CEOs do not issue new equity because they perceive their company's stock to be undervalued. This is perceived asymmetric information by overconfident CEOs who restrict their investment and use additional cash flows to invest closer to their desired level. Overconfident CEOs display investment behaviour that is sensitive to cash flows as is the case in the asymmetric-information approach.

* L. Larwood and W. Whittaker, "Managerial Myopia: Self-serving Biases in Organizational Planning," *Journal of Applied Psychology* 62 (1977): 194–198.
** U. Malmendier and G. Tate, "CEO Overconfidence and Corporate Investment," *Journal of Finance* 60 no. 6 (December 2005): 2661–2700.

insight of the literature on ASYMMETRIC INFORMATION. This can cause the cost of internal finance to differ substantially from external finance, and makes the availability of finance a constraint on investment. Up to now we have proceeded as though decision makers in firms and external suppliers of funds have the same information about the firm's investment opportunities. However, asymmetric information can generate cost disadvantages of external finance for some firms.

› Asymmetric information is prevalent when investors in the firm are not as equally well-informed about the investment opportunities of the firm as the managers of the firm.

The arguments for this view draw on an idea called the "LEMONS PROBLEM" that was first expounded on by George Akerlof.[16] The essence of the argument is that some sellers with inside information about the quality of an asset or a security will be unwilling to accept the terms offered by a less-informed buyer. This can cause the sale of an asset at a lower price than it would command if all buyers and sellers had full information. It is this idea that has been applied to the problem of equity finance by Myers and Majluf[17] and Greenwald, Stiglitz, and Weiss.[18] They explain how asymmetric information causes suppliers of new equity to demand large premia, or can eliminate the possibility of new equity financing altogether.

› The lemons problem arises due to asymmetric information that tends to reduce the average quality of goods offered for sale.

In the PECKING ORDER or *financing hierarchy* theories, the managers of a firm are assumed to have full information about the value of the firms' existing assets and the returns from new investment projects. Hence, to the extent that managers control sufficient internal funds to finance all profitable investment projects, investment-demand models, based on firms operating in perfect capital markets, are valid.

› The pecking order in corporate finance is that internal funds are cheaper than debt, which in turn is cheaper than equity.

What happens when a firm exhausts all its internal funds and requires external finance to undertake a desirable project? In the Myers–Majluf model, external investors cannot distinguish the quality of firms and so value all of them at the population average. New shareholders then implicitly demand a premium to purchase the shares of relatively good firms to offset the losses that will arise from funding lemons. This premium can raise the cost of new equity finance—faced by managers of relatively high-quality firms—above the opportunity cost of internal finance faced by existing shareholders. Since investors cannot distinguish between good and bad investment opportunities faced by the firm, every issue of a new financial security or asset is priced assuming an average project outcome, which implies that the securities backing good projects are undervalued. Given this undervaluation, the cost of financing such projects with external funds exceeds that of financing with internal funds. This difference in costs represents the lemons premium associated with external finance. With a lemons premium $\Omega$ on new share issues the marginal cost of funds for a firm in an asymmetric information capital market[19] is $1 + \Omega$.

### *4.3.1* The Equity and the Value of the Firm

To account for external finance in investment decisions, we can presume that investors maximize the real return on their portfolios. This means that they will hold shares only if the expected return on equity composed of the dividend yield and capital appreciation equals the return on comparably risky assets. This required real return, denoted $\rho$, is assumed to be unaffected by changes in the quantity of equity. Also, share prices are proportional to the outstanding value of a firm's equity. If new shares are issued, this dilutes the holdings of the current equity holders, who in order to continue earning their required rate of return would have to adjust for this. This is equivalent to the required rate of return being adjusted by the value of the new share issues, $V^N$, which would have to be bought by the current or original equity holders in order to maintain their proportional claim on the firm. Thus, the asset market arbitrage condition governing a shareholder portfolio decision[20] is that the required rate of return is given by

$$\rho = \frac{\text{Dividends} + \text{Capital Gains} - \text{Value of New Share Issues}}{\text{Value of Firms' Equity}}$$

$$= \frac{D_t + (V_{t+1} - V_t) - V_t^N}{V_t} \quad \textbf{(4.8)}$$

where, $D_t$ is the dividend payment by the firm;

$V_t$ is the value of firms' equity;

$V_{t+1} - V_t$ is the capital gain.

$V_{t+1}$ is given by the following expression:

$$V_{t+1} = V_{t+1}^0 + V_t^N$$

where $V_{t+1}^O$ is value in period $t + 1$ of shares outstanding in period $t$ of original shareholders;

$V_t^N$ is New share issues.

Thus,

$$\rho V_t = D_t + (V_{t+1} - V_t) - V_t^N$$

This is a difference equation that we solve[21] to obtain the value of the firm's equity as:

$$V_t = \sum_{i=0}^{\infty} \frac{1}{(1+\rho)^{i+1}} (D_{t+i} - V_{t+i}^N) \quad \textbf{(4.9)}$$

› The intrinsic value of the firm is the present value of all the future dividends. In the formula here, this has been adjusted for the reduction in dividends to the original shareholders due to the issue of new shares.

Hence, the total VALUE OF THE FIRM is the present value of the dividend stream, adjusted for the current value of the issued new shares, which would have to be bought by the current equity holders to maintain their proportional claim on the firm. This is what we should expect intuitively since the value of the stocks of a firm is the discounted value of the dividends, the future cash flows, which are expected to be paid throughout the future to those who have bought the stock of the firm. Here, that future cash flow is adjusted for the reduction in the claims on the firm when new shares are issued. Further, incorporating the lemons PREMIUM ASSOCIATED WITH NEW EQUITY issues, we would need to reduce $V_t$ in the above equation by an amount $\Omega$ per rupee of the new equity issued so that in effect the value of the firm's equity is:

› Due to new shareholders facing asymmetric information where they may be funding lemons, a premium is charged for new equity issued.

$$V_t = \sum_{i=0}^{\infty} \frac{1}{(1+\rho)^{i+1}} [D_{t+i} - (1 + \Omega)\, V_{t+i}^N] \quad \textbf{(4.10)}$$

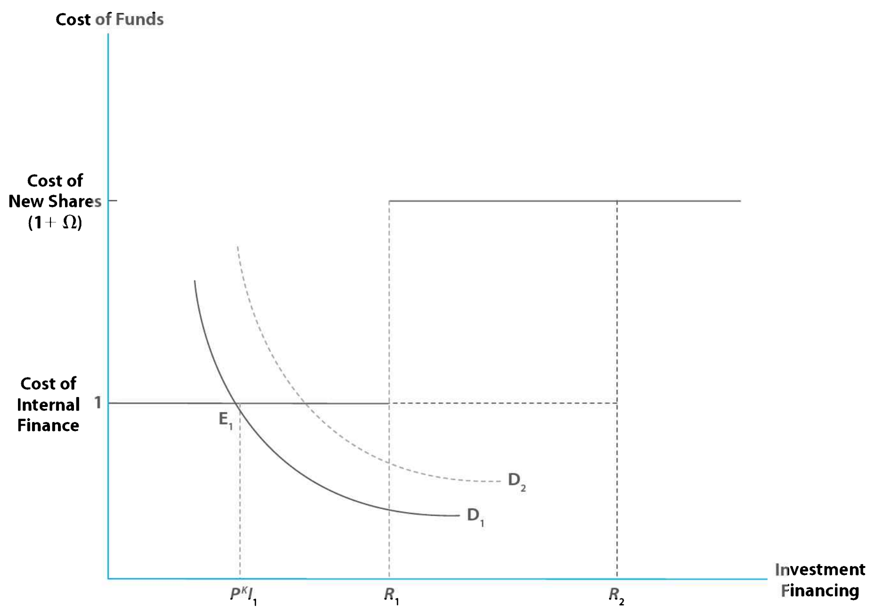
Equity holders, while maximizing the value of the firm, are choosing both the financial policy and the physical investment. They will be constrained by the requirement that the sources of funds equal the uses of those funds.

Sources of Funds = Uses of Funds

or, $\Pi_t + V_t^N = D_t + p^K I_t$

where, $\Pi_t$ are short-run profits. In turn, short-run profits are revenues less factor payments and adjustment costs, or,

$$\Pi_t = pF(K_t, N_t) - wN_t - \frac{b}{2} p^K I_t^2$$




**› Figure 4.6**
Cost of Internal and External Equity Finance. **External finance is costlier due to a lemons premium Ω. A firm with available internal finance $R_1$ and facing investment opportunities given by $D_1$ will maximize the value of the firm by internally financing $p^K I_1$ of investment and paying out a dividend of $R_1 - p^K I_1$.**

The firm then faces the following decision problem:

$$\left.\begin{aligned}
&\max_{\{I,K,V^n,D\}} V_t = \sum_{t=0}^{\infty} \frac{1}{(1+\rho)^{t+1}}\left[D_t - (1+\Omega)V_t^N\right] && \textbf{(4.11a)}\\
&\text{Subject to} \quad K_{t+1} = (1-\delta)K_t + I_t && \textbf{(4.11b)}\\
&\Pi_t + V_t^N = D_t + p^K I_t && \textbf{(4.11c)}\\
&V_t^N \geq 0 && \textbf{(4.11d)}\\
&D_t \geq 0 && \textbf{(4.11e)}
\end{aligned}\right\} \textbf{(4.11)}$$

In other words, the firm chooses $I$, $K$, $V^N$, and $D$ so as to maximize the value of the firm's equity (4.11a) subject to the capital stock accounting identity (4.11b); the cash flow constraint that the sources of funds equals the uses of those funds (4.11c); the constraint that it cannot repurchase shares (4.11d); and the constraint that it cannot pay negative dividends (4.11e).

The financing hierarchy of the firm is depicted in Figure 4.6. There is a discontinuous differential in the costs of internal and external equity finance. Suppose a firm with internal finance of $\Pi_t = R_1$ is faced with investment opportunities, as represented by an investment demand schedule of projects ranked by Tobin's $Q$, of amount $D_1$. The firm will invest up to point $E_1$ where marginal $Q$ is equal to 1—its earnings are sufficient to fully finance this level of investment with $R_1 - p^K I_1$ left to pay out as dividends. The firm has no reason to issue new shares $V^N = 0$ as none of the marginal projects has a $Q$ of at least $1 + \Omega$.

A small improvement in investment opportunities given by a rightward shift to $D_2$, which corresponds to an increase in marginal $Q$, would lead to an increase in retentions and a reduction in dividends.

On the other hand, an increase in cash flow—depicted by a rightward shift of the firm's supply of finance schedule from $R_1$ to $R_2$ has no effect on the firm's investment spending.

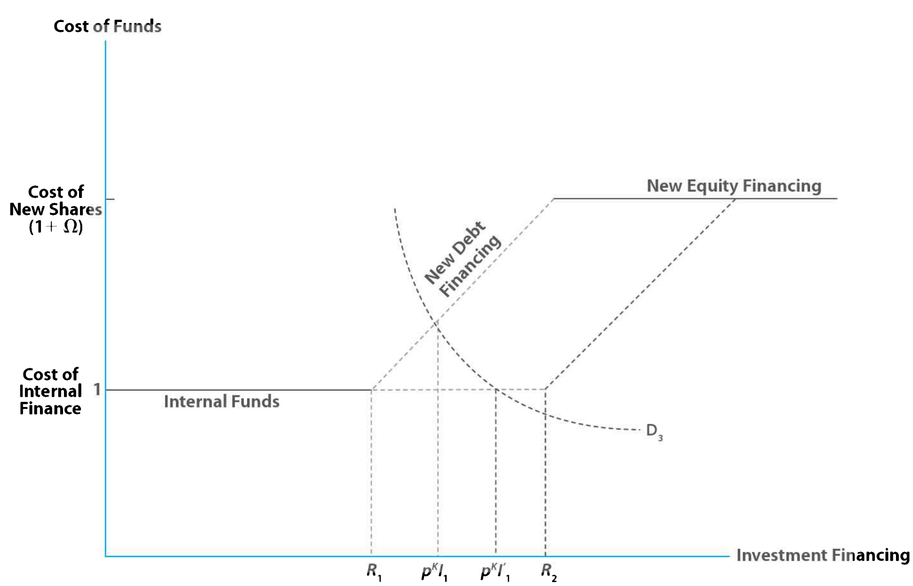
### *4.3.2* Debt and Investment Financing

So far, we have ignored the possibility that in addition to internal and external equity the firm may resort to debt finance. To the extent that debt can be secured from lenders such as commercial banks that specialize in monitoring the borrower, information problems in debt markets will be less severe than those in external equity markets.[22] However, the marginal external finance premium that creditors typically demand is increasing with the amount borrowed or, with leverage as an increase in debt, raises the chance of default by the borrower. Thus, instead of the vertical section of the financing supply schedule there will be an upward-sloping section between the flat segments of the schedules that depict the cost of internal and external equity finance. This middle section of the financing supply schedule is upward sloping due to the marginal cost of borrowing (debt) increasing with the increase in the amount borrowed. Let the external finance premium be $\eta_t = \eta_t(B_t)$ so that the gross return on borrowing ($B$) is $1 + r[1 + \eta_t(B_t)]$. The premium[23] increasing on the amount borrowed implies $\partial\eta/\partial B > 0$. The cash flow constraint of the firm as given in Eq. (4.11c) needs to be replaced now by Eq. (4.11c′) where $B_{t+1} - \{1 + r[1 + \eta_t(B_t)]\}B_t$ represents the net of repayments borrowing[24]

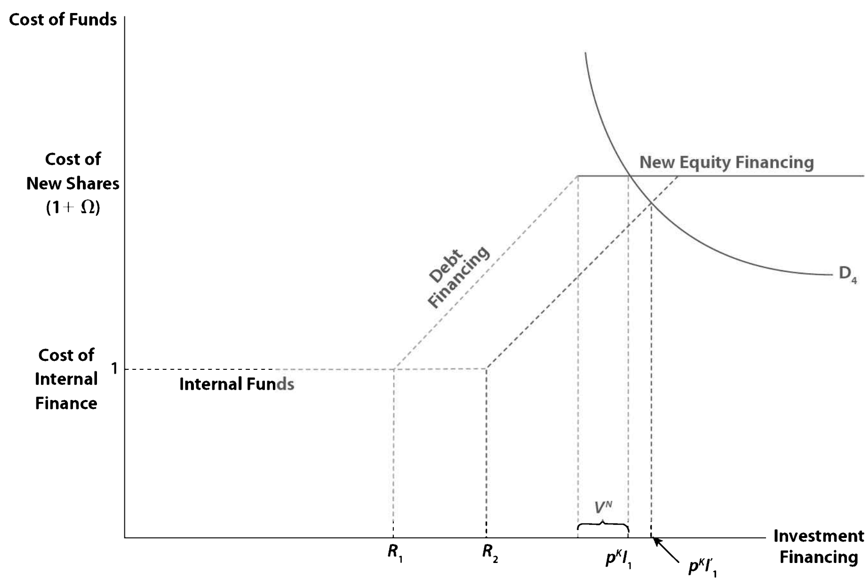
$$\Pi_t + V_t^N + B_{t+1} - \{1 + r[1 + \eta(B_t)]\}B_t = D_t + p^K I_t \qquad \textbf{(4.11c′)}$$

Information asymmetry problems in debt markets are less severe than in equity markets, which implies that a financing hierarchy where the cost of new borrowing is increasing, and is represented by the dashed line that connects the two horizontal segments in the middle in Figure 4.7.

Now consider a firm with substantially better investment opportunities than that depicted in the previous diagram (see Figure 4.7, schedule $D_3$). The firm after exhausting its internal funds resorts to debt finance, which is more costly due to it being an external source of finance. The amount of borrowing is $(p^K I_1 - R_1)$ and the firm pays no dividends. An increase in cash flow from $R_1$ to $R_2$ shifts the upward-sloping dotted line to the right increasing investment in all cases to $p^K I_1'$ and leading to positive dividends equal to $(R_2 - p^K I_1')$. The investment spending of the firm is clearly sensitive to its cash flows. Movements in internal finance, therefore, lead to fluctuations in investment spending.




**› Figure 4.7**
Hierarchy of Financing. A firm will exhaust internal funds before resorting to debt finance to fund an investment opportunity. If internal funding improves then the firm increases its investment expenditure as lower cost internal finance is available.




**› Figure 4.8**
Marginal New Equity Financing. As the investment opportunities improve to $D_4$ the firm issues new shares of the amount $V^N$ and finances the investment. An increase in the lower cost internal cash flow again increases the investment expenditure by the firm.

Finally, consider a firm with very strong investment opportunities as shown by schedule $D_4$ in Figure 4.8. Despite the lemons premium, it is worthwhile for such a firm to issue new shares of the amount, $V_t^N \geq 0$ . Under asymmetric information, its equilibrium marginal $Q$ will be $1 + \Omega$. A substantial increase in cash flow to $R_2$ will again increase this firm's investment spending to $p^K I_1'$.

In this case too the availability of internal financing has the potential to influence investment spending. $Q$, of course, varies both across firms and over time, depending on the severity of information problems that the firm faces. Typically, $\Omega$ is lower for a mature firm as compared to a young firm. Since under asymmetric information the availability of internal finance influences the investment spending of some firms, the more general from of the investment equation is

$$I_t = \beta_0 + \beta_1 Q_t + \beta_2 C_t + \mu_t \tag{4.12}$$

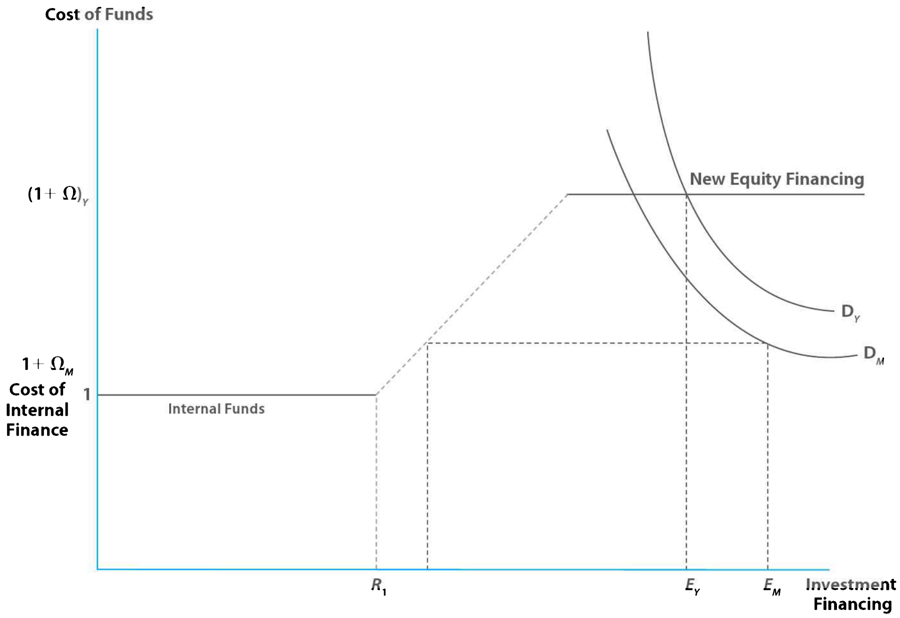
where $C_t$ is the cash flow of the firm and $\mu_t$ are shocks to adjustment costs.

### 4.3.3 Mature Firms and External Financing

It is typically observed that younger firms have better investment opportunities and higher rates of sales growth than mature firms. Better investment opportunities ought to be funded by external financing but due to liquidity constraints arising from asymmetric information the ratio of new share issues to cash flow is larger in mature firms. Figure 4.9 depicts the case where the lemon premium $\Omega_Y$ is larger for the younger firm than for the mature firm ($\Omega_M$). The younger firm has better investment opportunities given by $D_Y$ than the mature firm. Yet, the mature firm with inferior investment opportunities is able to depend more heavily on external finance with $E_M > E_Y$.

### 4.3.4 Collateral and Liquidity Constraints

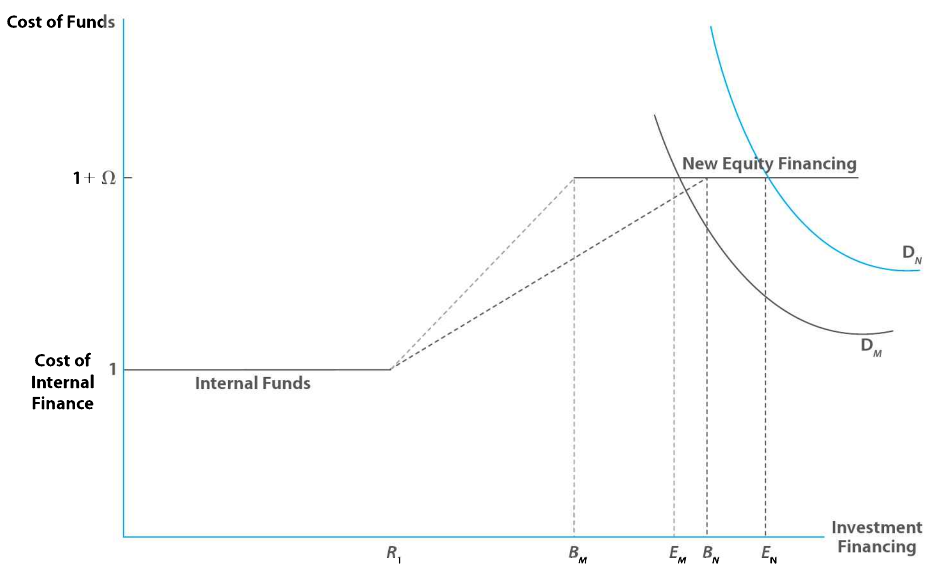
Again, some firms can face poorer investment opportunities than other firms, but may have good collateral (firm N in Figure 4.10), which enables them to secure debt financing at lower cost. Thus, unlike Figure 4.9, where the firm issued less equity because of a high lemons premium, here they issue less equity because of cheaper access to debt finance. In Figure 4.10, the firm N,



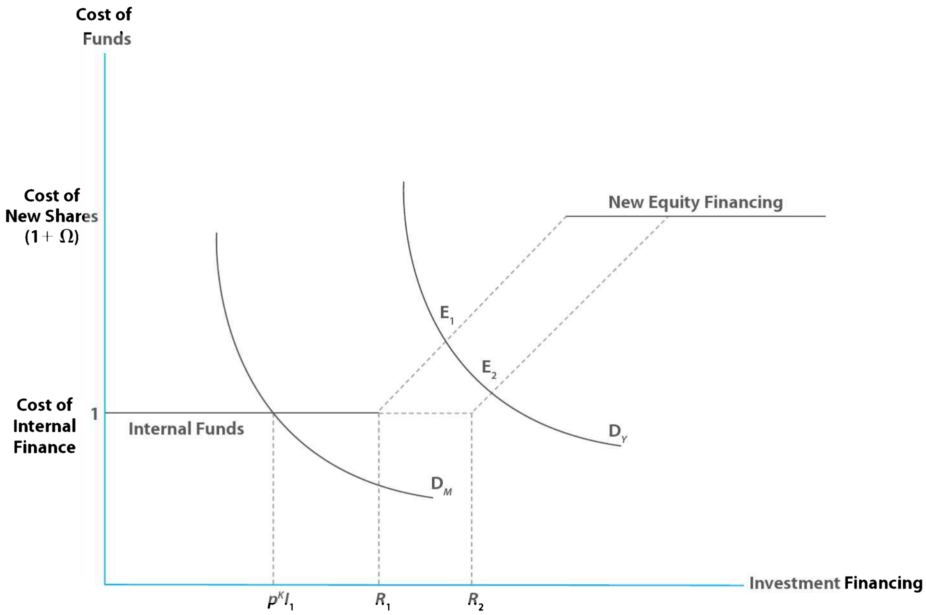

**› Figure 4.9**
Maturity and External Finance. Even though a young firm may have better investment opportunities $D_Y$, it may not be able to fund larger investments as compared to mature firms because it is subject to a larger lemons premium $\Omega_Y$.

which faces a lower cost of debt issues new equity amounting to $(E_N - B_N)$ despite having better investment opportunities. In contrast, firm M issues an equity of $(E_M - B_M)$, which is larger than firm N's equity.

There are also implications of this asymmetric information model for firm mergers. Take the case of a mature firm with substantial cash flow but relatively poor investment prospects—$D_M$ in Figure 4.11—that combines with a young, liquidity-constrained firm with good investment prospects such as $D_Y$ in the same figure. The new combined firm is able to relax the liquidity constraint on the young firm and carry out projects with a present positive net value that would otherwise have been set aside. The mature firm can use its internal resources equal to $R_1 - p^K I = R_2 - R_1$ to increase the cash flow of the younger firm and increase its investments from $E_1$ to $E_2$. The merger, therefore, creates value and this is an explanation for the premium received by shareholders of target firms that has often been documented.[25]




**› Figure 4.10**
Cost of Debt and External Finance. Mature firms may have better collateral that allows them to access debt finance at a lower cost than new firms. This gets leveraged into the mature firm's accessibility to higher levels of equity financing.




**› Figure 4.11**
Liquidity Constraint and Mergers. A young firm with better investment opportunities $D_Y$ can merge with a mature firm with substantial internal funds $R_1 > R_2$, and fund the profitable investment that otherwise would not have been undertaken.

This also explains why studies have often found that post-merger the profitability of the merged target often declines when the target was liquidity constrained; the marginal profitability of its investment was higher at $E_1$ than the unconstrained firms' cash flow. The removal of the constraint through a merger with a cash-rich firm brings marginal profitability down to $E_2$. The popular perception is that the decline in observed profitability is welfare reducing. However, this welfare implication is not warranted as investment at $E_2$ is at its full-information optimal level. The profitability of a firm acquired may go down post-merger, but its rate of asset growth will be larger.[26]

Empirical evidence supports this view of financing behaviour.[27] Small firms are more likely to face asymmetric-information problems which cause them to have high retention ratios and to pay no dividend at all for long periods of time. Many profitable small firms also exhaust internal finance but do not borrow or issue new equity due to a large lemons premium associated with external finance. Empirical studies relying on firm-level panel data find that investment is significantly correlated to changes in net worth and that the correlation is stronger for firms facing financial-market imperfections.[28] Empirical studies have been confronted by three types of challenges:

i) Identifying a proxy for investment opportunities ($Q$)
ii) Finding sorting criteria that are related to the firm's information problems such as firm size or dividend policy—to discriminate between constrained and unconstrained firms
iii) Identifying a proxy for net worth or changes in net worth, which is uncorrelated to investment opportunities.

# 4.4 Residential and Inventory Investment

So far, we have discussed business investment. There are two other COMPONENTS OF INVESTMENT EXPENDITURE in the macroeconomy that we now discuss. These are residential or housing investment and inventory investment.

› The total investment expenditure in the economy comprises (1) business investment, (2) residential investment, and (3) inventory investment.

Investment in housing and residential structures refers to expenditure on new housing units and on the maintenance and improvement of existing units. Residential investment spending constitutes a large fraction of total private investment spending in an economy and a dominant component of savings is in the form of housing. Our understanding of the determinants of investment in housing is developed with reference to the market for rental apartments. The market for owner-occupied housing is similar in nature because the owner of a house implicitly pays a rental price which has been paid up front at the time of the purchase.[29]

At any point of time, there is a fixed supply of housing units that result from the accumulated past investment. The supply of housing at any moment is the stock of housing units in existence at that time. The flow of investment in housing by builders and developers, the suppliers of housing, adds to the stock of housing units and increases the supply of housing in a later time period.

› The supply of housing units at any point in time is the result of accumulated past investment in housing and is a given constant stock of housing.

Let the initial total stock of housing units be denoted as $K_h$. Lags in housing construction result in an INELASTIC SUPPLY OF HOUSING units with respect to the rental rate—the supply of housing is fixed and can only increase after the interval of time it takes to construct new housing stock.

› The **demand for housing** decreases as the rental on housing increases.

The DEMAND FOR HOUSING is like the demand for any good. As the rental on housing—the price paid for purchasing housing services—increases, the demand for the stock of housing declines, a downward-sloping demand curve. What is the rental price of housing? Consider a housing unit that can be purchased at time $t$ at a price $p_t^h$. The rate of depreciation of the housing unit is $\delta$, and there is a resale value to the house which fetches a price $p_{t+1}^h$—one period hence. If you purchase the house at time $t$, rent it for a unit time period at a rental rate $R$, and then sell the house in the next time period, your return would be the rental plus the value of the house in next period. As the housing unit is subject to depreciation, a unit of housing in the next period is reduced to $(1 - \delta)$ of a unit of today's housing. Then, the value of the house next period is $p_{t+1}^h(1 - \delta)$. Your return next period is then given by $R + p_{t+1}^h(1 - \delta)$. The rate of return on purchasing a housing unit, earning a rent on that unit, and reselling it next period, is then given by

$$\text{Rate of return} = \frac{R + p_{t+1}^h(1 - \delta)}{p_t^h}$$

To purchase the housing unit in period $t$, your cost of funds would be the interest rate on the loan market. If the rate of return in the market for housing was higher than the interest rate on taking out a loan, you would see a profit opportunity and keep taking out loans and purchasing housing units. Of course, your ability to continue earning such a profit opportunity or free lunch cannot continue indefinitely due to ARBITRAGE, which will ensure that the rate of return in the market for housing must equal $(1 + r)$. Thus,

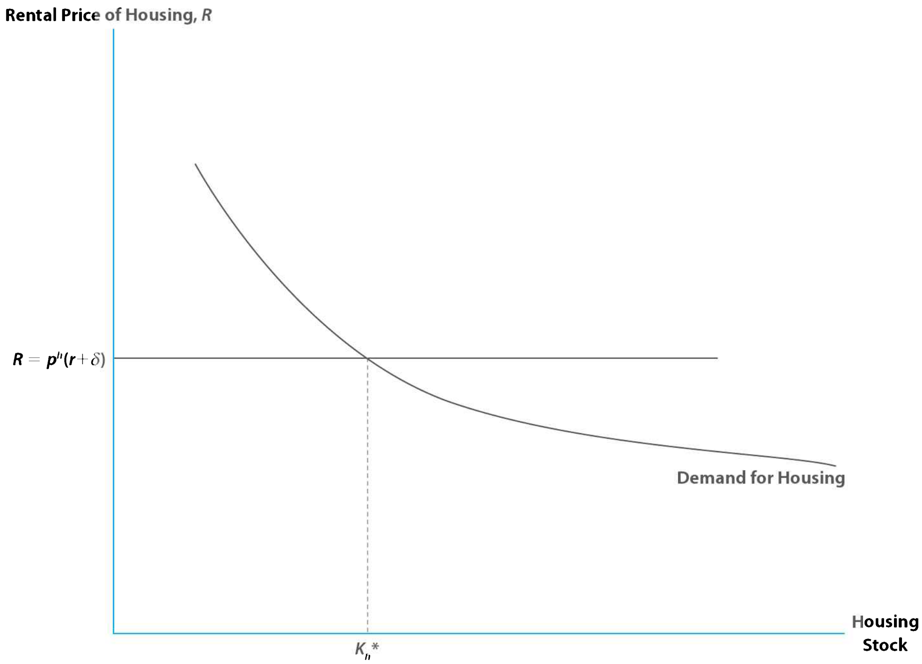
› Arbitrage is the proposition that the expected rates of returns on two assets must be equal.

$$1 + r = \frac{R + p_{t+1}^h(1 - \delta)}{p_t^h}$$

For the stationary simple case where price appreciation does not occur, $p_{t+1}^h = p_t^h$, and the RENTAL RATE can be solved out as

› The rental on housing is the price of a unit of housing times the rate of interest plus depreciation.

$$R = p^h(r + \delta)$$




› **Figure 4.12**
Demand for Housing Services.
The demand for housing is inversely related to the rental on housing.

This is the formula for the rental, on the capital stock as before but now with reference to the housing capital stock. Compare this equation with the equation in Section 4.1 for the rental on capital equipment. As before the rental price is the interest rate ($r$) plus a rate of depreciation ($\delta$) times the price of housing ($p^h$). With the demand for housing inversely related to the rental price, the desired demand for the stock of housing units, or, the amount of housing agents in the economy would like to consume, is then given by $K_h^*$ in Figure 4.12.

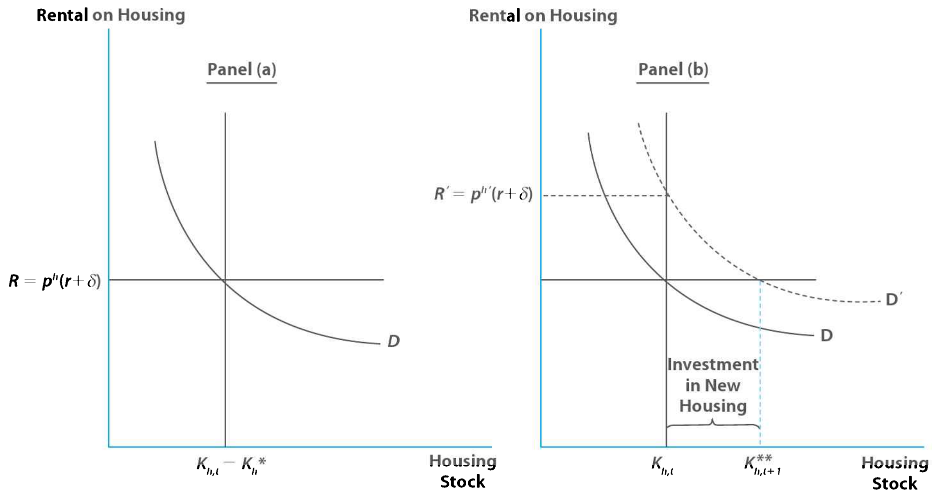
Now, suppose the stock of housing units in the market is $K_h = K_h^*$. As the supply of housing at any point in time is the result of accumulated past investment in housing, it is the vertical line through $K_h = K_h^*$ in Figure 4.13(a). The DEMAND FOR HOUSING STOCK equals the supply of housing stock and the housing market is in equilibrium. A disturbance to this equilibrium position can arise due to, say, a shift in the demand for housing. An increase in incomes associated with an economic boom acts through an income effect to increase the demand for housing. Alternatively, a rise in the population also causes a rise in the demand for housing. Both these factors cause the demand for housing to shift to the right to $D'$ . See Figure 4.13(b).

› The demand for housing stock increases with incomes and with an increase in the population.

The current stock of housing is now less than the demand for housing: $K_{h,t} < K^{**}_{h,t+1}$. With excess demand the rental rate in the housing market will rise to $R'$. With no change in $r$ and $\delta$, the higher rental rate is capitalized into a higher selling price for housing units and the price of housing increases to $p^{h\prime} > p^h$. As the price of existing homes has been driven up, the market valuation of the existing housing stock has increased. Builders and developers respond to this demand for housing by building new homes when the existing home prices are high relative to the price of new homes. The amount of investment in building new homes, $I_h$, is

$$I_h = K^{**}_{h,t+1} - K_{h,t}$$

and this amount of investment expenditure is depicted in Figure 4.13(b).




**› Figure 4.13**
Investment Expenditure on Housing. **The demand for housing increases with an increase in income and population. This makes existing homes expensive relative to the cost of building new homes and induces an increase in investment in housing.**

› Investment expenditure in housing results when the price of existing homes is expensive relative to the price of new homes. Alternatively, investment in housing occurs when the market valuation of housing stock exceeds the replacement cost, or, Tobin's $Q > 1$.

An alternative way of thinking about what drives the INVESTMENT EXPENDITURE is to ask what the replacement cost of the existing stock of housing assets is. How much would it cost to buy the existing housing stock again as new, off the production line? That would be the cost of new housing, which is $p^h$. As the current market valuation at $p^{h'}$ is larger than the replacement cost, Tobin's $Q$ is greater than unity and investment takes place. So, housing investment is the outcome of decision making akin to the decision behind a firm's investment in plant and machinery. The difference between housing investment and business investment is that housing depreciates at a low rate of around 2 per cent per year compared to business depreciation rates in plant and machinery that is typically well over 6 per cent. Hence, the interest rate is the dominant component of the rental on housing. Housing investment is, therefore, typically more sensitive to changes in interest rates as a result of shifts in monetary policy than business fixed investment.

The framework we have developed for analysing investment applies to the third component of investment expenditures in an economy, that is inventory investment. Inventories are stocks of goods in the process of production as well as finished goods waiting to be sold. Inventories are usually held in order to maintain a buffer stock, which can be used to cater to unexpected changes in demand. Such inventories are held to equate the marginal benefit to the rental cost. The benefit of having a certain number of goods in process or finished goods on hand is that potential buyers will not have to wait for delivery and the firm is able to sell a greater output. The expected increased sales from this is the expected marginal product of the increased inventory. The cost of holding the inventory is the rental price, which as usual, comprises two components—the depreciation of the stock of goods that are being held in excess of current sales as inventory and the interest that must be paid out on the loan that finances the inventory. The firm will make an INVENTORY INVESTMENT when the expected benefit of holding the inventory, which is the expected increase in sales, is greater than the depreciation and interest cost of holding that stock of goods. The principles that guide investment decision making are thus common across the three major categories of investment expenditures—business-fixed capital formation, investment in housing, and inventory investment.

› Inventory investment occurs when the marginal productivity of the inventory exceeds the rental on the inventory, or, Tobin's $Q > 1$.

The general form of the investment function that the $Q$ approach suggests [see Eq. (4.7)] is that it depends positively on the marginal product of capital and NEGATIVELY ON INTEREST RATES (when factors affecting the price of capital and depreciation rates can be ignored). The marginal product of capital, in turn, depends positively on output produced (see the definition for the Cobb–Douglas case in Appendix 4.1) and negatively on $K$ (a large capital stock base diminishes the marginal productivity of capital). In short,

› Appendix 4.5 formally shows how Tobin's $Q$ is inversely related to the **rate of interest**.

$$I = I\left(\underset{(-)}{r}, \underset{(+)}{Y}, \underset{(-)}{K}\right)$$

where $Y$ refers to output.

Most macroeconomic models hold the capital constant while allowing for the flow effects of investment. Also, with an increase in output resulting in an increase in cash flows, we could write an alternative equation to Eq. (4.12) for investment as:

$$I_t = \beta_0 - \beta_1 r_t + \beta_2 Y_t \tag{4.12'}$$

# 4.5 Irreversibility and Investment

An alternative theory of investment is the theory of irreversible investment. IRREVERSIBILITY refers to the fact that many firms may not be able to sell their capital stock. A firm faces total irreversibility when resale markets are absent, which is equivalent to the situation where the resale price of capital is zero. In such a case, once the firm invests, it cannot get rid of additional capital stock even if economic conditions warrant a lower desired capital stock. As it cannot

› Irreversibility of investment occurs when the cost associated with the investment is not completely recoverable if subsequently the firm wants to sell the capital stock.

MACROFOCUS 4.2

## India: An Inverted Pecking Order?

The costs of external capital are greater than for internal capital due to asymmetric-information problems between managers and the capital market, or because of the transactions costs of issuing bonds and equity. If the costs of availing external capital markets are such that firms in developed countries are constrained by the availability of their cash flows from investing, one would presume that cash flow would be an even more important source of finance in developing countries such as India where the transactions costs are more prohibitive and asymmetric information is more severe. The finding of Singh and Hamid* that firms in developing countries finance their growth from external capital sources more than firms in developed countries is, therefore, surprising. Moreover, Indian firms rely much more heavily on external debt as a source of finance than do developed country firms such as US firms. Gugler, Mueller, and Yurtoglu** find that Indian firms earn returns on their investments that are less than their cost of capital implying that for Indian firms marginal $Q$ is less than unity. They attribute this to the existence of weak corporate governance structures in India. The low efficacy of legal institutions in enforcing contracts, and weak accounting standards prevent shareholders from protecting themselves against self-serving managers who feel secure that low returns investments will not cost them their jobs. Moreover, most debt financing of investment in the Indian context has been by public sector financial institutions which have implicit government guarantees that they cannot fail. This imparted a confidence to depositors and borrowers that banks are safe and served to weaken financial discipline by bankers who did not pursue efficient screening of projects and resorted to imprudent lending. With government shouldering the loss arising in the event of bank failure, and banks absorbing the full returns otherwise, banks in India had an incentive to over lend.† Also, as interest rates in the financial system prior to liberalization were administered, banks competed through non-price competition by overlending so as to gain market share. Thus, the financial system in India for those medium and large firms that had access to bank finance was characterized more by a credit-liberal regime rather than a credit-constrained one. Managers making poor return investments were not constrained by their internal cash flows. With financial liberalization, however, as banks seek prudent balance sheet positions, managers will find that credit constraints are akin to those in developed countries.

* A. Singh and J. Hamid, "Corporate Financial Structures in Developing Countries," International Finance Corporation, Technical Paper No. 1 (Washington, DC: IFC, 1992).
** K. Gugler, D. C. Mueller, and B. B. Yurtoglu, "The Impact of Corporate Governance on Investment Returns in Developed and Developing Countries," *The Economic Journal* 113 (November 2003): F511–F539.
† E. D'Souza, "The Structure of Corporate Finance and Corporate Governance in India," *Economic and Political Weekly* 35, no. 48 (November 25, 2000): 4196–4205.

access resale markets, it can only allow the excess capital stock to depreciate through time. The firm can always adjust its capital stock upward, but it cannot adjust it downward by disinvesting ($I_t \geq 0$). Irreversibility usually arises because capital is industry or FIRM SPECIFIC. Industry-level uncertainty tends to affect all firms simultaneously. As a result, if a firm wishes to sell excess capital in response to an adverse-demand shock, it may not be able to find buyers willing to purchase it. For example, a steel plant is industry specific in the sense that if the demand for steel falls, the market value of the plant will fall. Though the plant could be sold to another steel company there is little gain from doing so and the investment in the plant must be viewed as a SUNK COST.

› Firm-specific capital enhances productivity only in the firm where it is acquired and the value of the capital is lost if the capital is transferred out of the firm.

› A sunk cost is an investment in an asset with no alternative use, also referred to as a specific asset. A sunk cost is, therefore, an investment in an asset with no opportunity cost (the value of the next best alternative use of the asset).

Most investments in marketing and advertising are also firm specific and so are sunk costs. Even less specific capital, such as computers and office equipment and general-use machinery, are partially irreversible due to the "lemons" problem of adverse selection—secondary markets for such capital are inefficient. Finally, irreversibility also arises due to government regulation or institutional impediments. For example, capital controls may make it impossible for foreign investors to sell assets and reallocate their funds.

Not only does traditional investment theory assume that invested capital is reversible, it also assumes that each investment opportunity facing the firm is a unique opportunity. If the firm declines the opportunity, it will never have the choice to reconsider. In such a case the firm faces a now or never situation.

To get a flavour of what irreversibility and delay can result in, let us take a simple two-period example as expounded in Pindyck.[30] Assume investment is totally irreversible—the firm will not be able to recover any of the investment. The firm in question is a German firm contemplating investment in a new steel factory. For simplicity there is no gestation lag—efficient Germans can build the factory instantly at a cost of INR 800 (million) and that factory can produce a unit of steel forever, with zero-operating costs. The current price of a unit of steel, 100 tones, is INR 100 (million), but the price will change next year to either INR 150 (million) with one-half probability or to INR 50 (million) with one-half probability. The price will then remain at this new level forever.

The firm can discount future cash flows using the rate of return, which is the risk-free rate if the risk is fully diversifiable of, say, 10 per cent. Suppose the firm is deciding whether to invest now. Is this a good investment? One decision rule is that the net present value should be positive. Let us use this rule.

The expected future price of steel in any time period is:

$$\text{INR } 100 = (0.5)(150) + (0.5)(50)$$

Hence, the present value of the factory's revenue is $100 + \frac{100}{1.1} + \frac{100}{(1.1)^2} + \cdots$.

The factory costs INR 800 and so the net present value is:

$$NPV_{\text{invest now}} = -800 + \sum_{t=0}^{\infty} \frac{100}{(1.1)^t} = -800 + 1{,}100 = 300$$

› **Table 4.2**
Evolution of Steel Prices

| Time: $t = 0$ | $t = 1$ | $t = 2$ | $t = 3$ | — | — |
|---|---|---|---|---|---|
| $q = \frac{1}{2}$ | $P_1 = 150$ | $P_2 = 150$ | $P_3 = 150$ | — | — |
| Price: $P_0 = 100$ | | | | | |
| $(1 - q) = \frac{1}{2}$ | $P_1 = 50$ | $P_2 = 50$ | $P_3 = 50$ | — | — |

$q$ = Probability of price change

where $$\sum_{t=0}^{\infty} \frac{100}{(1.1)^t} = 100 + \frac{100}{1.1} + \frac{100}{(1.1)^2} + \cdots.$$

The factory should be built on this decision criterion. We can also use the value of Tobin's $Q$ to guide our decision. This can be calculated as:

$$Q = \frac{1{,}100}{800} = 1.375$$

which is greater than 1. Hence this criterion also supports the decision to invest.

These calculations however ignore a cost—the opportunity cost of investing now, rather than waiting and thereby keeping open the possibility of not investing should the price fall. As an alternative, suppose the firm's managers wait a year and build the factory only if the price increases from INR 100 to INR 150.

In year zero: there is no expenditure and no revenue.

In year 1: the cost of investment is incurred only if the price rises and the present value of that cost is 800/1.1. The present value of returns from investing when the price increases is $\sum_{t=1}^{\infty} \frac{150}{(1.1)^t} = \frac{150}{1.1} + \frac{150}{(1.1)^2} + \cdots.$

Hence, if we wait a year and invest only if the price goes up, an event that occurs with a probability of half, the net present value is[31]

$$NPV_{\text{delay}} = 0.5\left[\frac{-800}{1.1} + \sum_{t=1}^{\infty} \frac{150}{(1.1)^t}\right] = \frac{425}{1.1} = 386$$

The NPV is higher if we plan to wait a year so clearly waiting is better than investing now.

Likewise, Tobin's $Q$ in the delay case is given by:

$$Q_{\text{delay}} = \frac{\sum_{t=0}^{\infty} [(150)/(1.1)^t]}{800/1.1} = \frac{1{,}650}{727} = 2.27$$

and $$Q_{\text{delay}} > Q_{\text{invest now}}$$

It is important to note that if our only choices were to invest today or never invest, we would invest today. In that case there is no option to wait a year, and hence no opportunity cost of killing such an option. The real option: the value of the call option of delay is INR 386. The less completely irreversible the investment, the lower the VALUE OF WAITING. This description of changes in the value of the option to delay follows the connected effects on the value of financial call options.[32]

› The value of waiting is the expected present value of the irreversible mistake that would be revealed in the future should the returns fall short of the costs.

An irreversible-investment opportunity is akin to a financial-call option. A call option gives the holder the right, for some specified amount of time, to pay an exercise price and, in return, receive an asset, for example, a share of stock that has some value. As with a financial-call option, the firm's option to invest is valuable in part because the future value of the asset that the firm gets by investing is uncertain. If the asset rises in value, the pay-off from investing rises. If it falls in value, the firm need not invest and will only lose what it spent to obtain the investment opportunity.

There are situations where a firm cannot wait, or wait very long, to invest. For instance, if the firm anticipates the entry of a competitor into a market that can sustain one large firm, waiting is costly. Also, if there is no patent or a mineral resource lease that is about to expire, the less will irreversibility affect the investment decision. The example above focused on uncertainty over future product prices. Investment, however, is sensitive to uncertainty in various forms: uncertainty over operating costs that determine cash flows; uncertainty over future interest rates; and uncertainty over the cost and timing of the investment itself.

› Credibility refers to the degree to which individuals and markets believe that a policy announcement will be implemented even when there are short-term economic costs that may be imposed.

This has strong policy implications. If a goal of macroeconomic policy is to stimulate investment, stability and CREDIBILITY may be much more important than tax incentives and a low-interest rate regime. Indeed, if uncertainty over the economic environment is high, tax and related incentives may have to be very large to have any significant impact on investment.

The implication of irreversibility is that the cost of investing includes the discounted marginal call option in addition to the market price for capital goods, $p^K$. By investing one additional unit at any point in time, the firm loses the opportunity to invest that unit in the future and incurs the loss of a call option. This additional cost tends to discourage investment. Greater uncertainty increases the value of the call option and reduces the incentive to invest. In an equilibrium situation, the sum of the costs must equal the marginal benefit of investing an additional unit of capital. This marginal benefit of investing an additional unit can be shown to be the sum of two terms. First is the discounted expected marginal value of capital if the firm never invests or disinvests in the future and simply allows its capital stock to depreciate. The second component of marginal benefit is the expected value of the unused portion of the capital stock, assuming the firm does not invest in the next period, or the increase in the expected marginal value of capital due to future marginal call options available to the firm.

The intuition that is prevalent here is that the value maximizing investment decision must consider the value of the option of delay when comparing the marginal benefit and cost of investing.

## 4.6 Investment in Developing Countries

There has been very little empirical work on the determinants of private investment in developing countries.[33] Investment theories formulated for industrial countries need to be changed to fit the circumstances in developing countries. For instance, given that financial variables influence investment behaviour, investment is sensitive to the institutional environment in the financial system. Equity markets are typically small and underdeveloped and the prevalence of financial repression implies that Tobin's $Q$ must be applied cautiously. The cost of funds in informal markets has been found to influence private investment when the firms are credit rationed in the formal financial sector. Then again, in many developing countries with fledgling capital goods sectors that resort to imports of capital goods, foreign exchange rationing and the cost of foreign exchange in unofficial curb markets can be important determinants of private investment. Also, the existence of a debt overhang, which indicates the possibility of future taxation that will be used to finance future debt service, has been found to inhibit private investment.[34]

Developing countries also tend to have large public sectors and there is accordingly a need to incorporate complementarity–substitutability relationships between public and private investment in empirical work on investment in these countries. The relation between public and private investment is *a priori* positive or negative. Public sector investment can crowd out private-investment expenditure if it uses scarce physical and financial resources that would otherwise be available to the private sector. The financing of public investment through taxes, issuance of debt, or inflationary finance can reduce the resources available to the private sector and depress private investment. Alternatively, public investment that expands infrastructure and the provision of public goods can raise the productivity of private capital and is complementary to private investment. To sum up, public and private investment can be complementary or substitutes to one another.[35]

An important source of uncertainty for the private-investment decision has been macroeconomic instability associated with economic policymaking. Some studies have found that political uncertainty—government instability, rapid government turnover, and social unrest—vary negatively with investment.[36] Volatility of output growth is an indicator of the unpredictability of demand and the volatility of inflation is an indicator of macroeconomic uncertainty. Both these and other indicators of uncertainty have been found to adversely impact private investment.[37]

Finally, corruption has been found to have a significant adverse impact on investment. Corruption tends to increase the number of projects undertaken and expand their size. Thus, whilst increasing the share of public investment to GDP, corruption also lowers the quality of public investment put in place.[38]

Thus, when investigating the substitution–complementary effects of public investment, we need to capture the indirect effect of corruption via its impact on the quality of public investment. Empirical work usually includes an interaction term between the corruption index and the level of public investment to capture this effect. It has been found that higher corruption for a given level of public investment is associated with lower quality of public investment, which tends to lower private investment.[39] Clearly, the behaviour of aggregate investment is subject to many influences that empirical and theoretical work is slowly uncovering.

## 4.7 Investment in India

A recent attempt to understand the determinants of private corporate investments in India[40] gives us an indication of the importance of some of these variables. Drawing on the literature focusing on finance constraints on investment decisions, which argue that external finance may be more costly than internal finance because of transaction costs, contract enforcement problems, and asymmetric information, Athukorala and Sen include the availability of finance—measured in terms of real bank credit to the private sector, $BC$—as a determinant of investment. Similarly, public sector fixed capital formation $I_{\text{public}}$, as argued above, can affect investment from the supply side through infrastructure investments, or the demand side through competing for investible funds. Other variables that are included in investment functions are the cost of capital[41] ($CC$), GDP ($Y$), and the capital stock in the private sector ($K$).

The equation was estimated for 1954–1996, and all variables were measured in natural logarithms, except, of course, the cost of capital which is measured in proportional form. In addition, because the credibility and sustainability of reforms could affect the perception of firms about investment opportunities, a post-reform intercept dummy ($PRD$), which takes the value of one for the post-reform years and zero otherwise, is included in the estimation.

The estimated equation[42] for private corporate investment ($I$) is as follows:

$$\Delta I_t = -1.28 + 0.13\Delta BC_{t-1} + 2.98\Delta Y_t - 3.89\Delta CC_t + 0.68\Delta I_{\text{public}}$$
$$+1.06 I_{\text{public},t-1} - 0.50K_{t-1} - 0.45I_{t-1} + 0.75PRD$$

The equation clearly indicates that business investment in any given year accelerates with changes in GDP ($\Delta Y$), and there is a positive effect of bank credit ($BC$). Moreover, public investment has a strong complementary effect on business investment in India. As predicted by investment theories, the cost of capital ($CC$) and the initial capital stock $K_{t-1}$ have negative effects on investment. The positive coefficient on the $PRD$ indicates that the reforms have had an investment enhancing impact over and above the impact from other explanatory variables. When incorporating the effects of future uncertainty on investment through measures, such as a moving average of the standard deviation of GDP as well as the cost of capital, it was found that they had insignificant effects on investment. Uncertainty in the Indian context, however, may be more associated with the variability of inflation rather than the variability of output or the rental on capital. Despite this, that investment in India is responsive to bank credit and public investment indicates the importance of proper economic policymaking so that there may be salutary effects on private investment in the economy.

## SUMMARY

- The rental on the stock of capital is the sum of the market rate of interest, $r$, and the constant rate of depreciation, $\delta$, multiplied by the unit supply price of the capital good, $p^K$.
- Combinations of capital and labour that produce the same level of output are represented by an isoquant. The slope of an isoquant is the ratio of the marginal products of the factor inputs.
- The optimal capital stock is determined where the slope of an isoquant equals the relative price of labour to the rental on a unit of capital.
- Convex adjustment costs that rise more than proportionately with the level of investment determine the speed at which the stock of capital adjusts to its optimal level.
- The investment undertaken by a firm is positive as long as Tobin's $Q^M > 1$.
- As external investors do not have the same information about a firm's investment opportunities compared to the managers of the firm, they demand a premium to invest in new issues of financial instruments by the firm.
- The value of the firm is the present value of dividends adjusted for the present value of new shares issued and is inclusive of the premium associated with these new issues.
- A firm that resorts to external financing of investment projects will find that its investment is sensitive to its internal cash flows.
- The supply of housing stock with respect to the rental rate on housing is inelastic and increases with the flow of investment in housing by builders and developers. The demand for housing is inversely related to the rental on housing.
- When the rental price of housing is higher than the rental cost of building new homes, the market value of housing exceeds the replacement cost and investment in housing increases.
- Inventory investment increases when the expected value of the increased sales from holding inventory exceeds the depreciation and interest cost of holding the increased inventory.
- Irreversibility arises due to the investment being a sunk cost when it is industry or firm specific. This makes it very costly to redeploy the capital stock. With irreversibility it is sometimes worth waiting for future uncertainty to be resolved before investing.
- Investment in developing countries is influenced by institutional factors such as the existence of financial repression, foreign exchange rationing, macroeconomic uncertainty, the complementarity between public and private investment, and the prevalence of corruption.

## NOTES

1. Any additional costs incurred in running the rental business such as the wages of the staff would have to be added to this figure of INR 30.
2. $\delta$ then is just the physical wear and tear of the machine expressed as a fraction of the price of the capital good.
3. Diminishing marginal productivity, which is the amount by which marginal product changes $\{\Delta(MP_N)\}$ when there is an additional unit of labour hired ($\Delta N$) implies that

$$\frac{\Delta(MP_N)}{\Delta N} = \frac{\partial^2 Y}{\partial N^2} = F_{NN} < 0$$

4. Obviously, $MP_N/MP_K = F_N/F_K$.
5. There are indeed many points of tangency between isoquants and resource constraints (or cost lines as they are sometimes called) and the locus of such points is called an expansion path. The firm selects that point on the expansion path corresponding to that level of output $Q^*$ that maximizes profits in the product or output market. Solving $\underset{\{Q\}}{Max} = pY - \chi$ we get $d\pi/dY = p - (d\chi/dY) = 0$ or, $p = MC$, the firm selects $Q^*$ where price equals marginal cost ($MC$) in the output market. It then considers the input/factor *market as in Figure 4.3 and selects* ($K^*$, $N^*$) associated with $Q^*$.
6. The two acquisition schemes—renting and outright purchasing—will be equivalent for entrepreneurs when the cost of the two schemes are equal. Thus, when the present discounted value of rental payments is identical to the purchase price of the equipment, renting and purchasing are equivalent actions. Under perfect competition this would always be true.
7. Robert Eisner and Robert Strotz, "Determinants of Business Investment", Commission on Money and Credit, *Impacts of Monetary Policy* (Englewood Cliffs, NJ: Prentice-Hall, 1963), pp. 60–138.
8. In what follows we use the result that the sum of the infinite series

$$S_n = 1 + x + x^2 + \cdots = \frac{1}{1-x}$$

9. See Fumio Hayashi, "Tobin's Marginal q and Average q: A Neoclassical Interpretation", *Econometrica* 50 (January 1982): 213–224.
10. In the case of the linear production function that we employed for simplicity that was given by $Y = AK$, the marginal product of capital is constant and does not diminish. Generally, however, any input is subject to diminishing marginal productivity.
11. Fumio Hayashi, op. cit.
12. For the difficulties in measuring Tobin's $Q$ in developing countries, see Michael J. Athey and Wende D. Reeser, "Asymmetric Information, Industrial Policy, and Corporate Investment in India," *Oxford Bulletin of Economics and Statistics* 62, no. 2 (2000): 267–292.
13. Lawrence H. Summers, Barry P. Bosworth, James Tobin, and Philip M. White, "Taxation and Corporate Investment: A Q Theory Approach," Brooking Papers on Economic Activity, no. 1 (1981): 67–140.
14. See Fanny S. Demers, Michel Demers, and Huntley Schaller, "Irreversible Investment and Cost of Adjustment," CEPREMAP Working Paper No. 9416, 1994.
15. Modigliani and Miller (1958) demonstrated the irrelevance of financial structure for real investment. Their key insight was that a firm's financial structure will not affect its market value in perfect capital markets. However, if there is asymmetric information about a firm's prospects between firms and potential investors then financial policy is not irrelevant for investment. See Franco Modigliani and Merton H. Miller, "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review* 48 (June 1958): 261–297.
16. George A. Akerlof, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (August 1970): 488–500.
17. Stewart C. Myers and Nicholal S. Majluf, "Corporate Financing Decisions when Firms have Investment Information that Investors Do Not," *Journal of Financial Economics* (June 13, 1984): 187–220.
18. Bruce Greenwald, Joseph E. Stiglitz, and Andrew Weiss, "Information Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review Proceedings* 74 (May 1984): 194–199.
19. The implications for a $Q$ model of investment are as follows. An investment that requires new share issue will be undertaken only if it increases the wealth of existing shareholders. Let the true gross returns (known to managers but not investors) from a new investment project be $Y'$ and the true gross returns from assets in place be $Y$. if the firm has run out of internal funds it will issue new shares if

$$\frac{Y'}{I} \geq \frac{Y}{V}$$

where $V$ is the market value assigned to both good firms and lemons and $I$ is the cost of the new investment project. This is equivalent to requiring that the marginal $Q$ on the new project be greater than the ratio of the firm's true average $Q$—labelled $Q^*$—to the average $Q$ assigned to all firms by the market-$\overline{Q}$. If there was full information, $Q^* = \overline{Q}$ or, $Q^*/\overline{Q} = 1$, and the threshold $Q$ value for issuing new shares is unity. When a good firm initially cannot be distinguished from lemons, $Q^*/\overline{Q} > 1$ for such firms. The ratio indicates how much dilution occurs when such good firms issue new shares. Hence, $(Q^*/\overline{Q}) - 1$ is the lemons premium ($\Omega$). We can write $\Omega = (Q^*/\overline{Q}) - 1$.
20. In practice, the capital gain tax rate and the personal tax rate on dividend incomes are not identical so that a

firm can affect its shareholders tax liabilities by choosing whether to retain or to distribute corporate profits. We ignore the effect of taxes. See James M. Poterba and Lawrence H. Summers, "Dividend Taxes, Corporate Investment, and 'Q'," *Journal of Public Economics* 22 (1983): 135–167.

21. See Appendix 4.3. Of course, the solution is subject to a transversality condition which prevents the firm's value from becoming infinite in finite time.
22. See Toni M. Whited, "Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data," *Journal of Finance* 47, no. 4 (September 1992): 1425–1460.
23. In general, the external premium causes the shadow price of internally generated funds to be state dependent and time varying.
24. In effect short-run profits now includes deductions for interest payments on borrowings, or, $\Pi_t = pF(K_t, N_t) - wN_t - (b/2)\,p^K I_t^2 - r[1 + \eta_t(B_t)]B_t$ and the cash flow constraint is $\Pi_t + V_t N + B_{t+1} - B_t = D_t + p^K I_t$.
25. See Gregg A. Farrell, James A. Brickley, and Jeffry M. Netter, "The market for corporate control: the empirical evidence since 1980," *Journal of Economic Perspectives* 2 (1988): 49–82.
26. See David J. Ravenscraft and Frederic M. Scherer, *Mergers, Sell-offs and Economic Efficiency* (Washington, DC: The Brookings Institution, 1987).
27. See Stewart C. Myers, "The Capital Structure Puzzle," *Journal of Finance* 39, no. 3 (July 1984): 575–592.
28. See Robert Hubbard, "Capital Market Imperfections and Investment," *Journal of Economic Literature* 36 (1998): 193–225.
29. A key difference between renting and owning that we are overlooking here is that homeowners get an offsetting benefit to owning in the form of tax deductibility of mortgage interest and property taxes. This means that from the interest and maintenance (depreciation) costs we must subtract out mortgage and property tax payments that are tax exempt. See Charles Himmelberg, Christopher Mayer, and Todd Sinai, "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions," *Journal of Economic Perspectives* 19, no. 4 (Fall 2005): 67–92.
30. Avinash K. Dixit and Robert S. Pindyck, *Investment under Uncertainty* (Princeton: Princeton, NJ: Princeton University Press, 1994); and Robert S. Pindyck, " Irreversibility, Uncertainty and Investment," *Journal of Economic Literature* 3 (29 September 1991): 110–148.
31. $S_n = X + \frac{X}{(1+r)} + \frac{X}{(1+r)^2} + \cdots = \sum_{t=0}^{\infty} \frac{X}{(1+r)^t} = \frac{(1+r)X}{r}$

    and $S_n = \sum_{t=1}^{\infty} \frac{X}{(1+r)^t} = \frac{X}{r}$ for $n \to \infty$
32. See Richard A. Brealey and Stewart C. Myers, *Principles of Corporate Finance* (New York: McGraw Hill, 1991).
33. See Luis Serven and Andres Solimano, "Private Investment and Macroeconomic Adjustment: A Survey," *World Bank Research Observer* 7, no. 1 (1992): 95–114.
34. See Jeffrey A. Sachs, "The Debt Overhang of Developing Countries," in G. A. Calvo, R. Findlay, P. Kouri, and J. Branga de Macedo (eds.) *Debt, Stabilization, and Development* (Oxford: Basil Blackwell, 1989).
35. See Martin Rama, "Empirical Investment Equations in Developing Countries," in L. Serven and A. Solimano (eds.) *Striving for Growth After Adjustment* (Washington, DC: The World Bank, 1993).
36. See Robert Barro, "Democracy and Growth," *Journal of Economic Growth* 1 (January 1996).
37. See Luis Serven, "Macroeconomic Uncertainty and Private Investment in Developing Countries—an Empirical Investigation," Policy Research Working Paper, The World Bank, Report No. 2035 (December 1998).
38. Tanzi and Davoodi measure corruption from the Political Risk Service's International Country Risk Guide which provides a measure of corruption based on whether government officials are likely to demand special illegal payments. Infrastructure quality is measured through proxies such as paved roads in good condition as a percentage of total roads, electric power system losses as a percentage of total power output, and telephone faults per 100 mainlines per year. For more on this, see Vito Tanzi and Hamid Davoodi, "Corruption, Public Investment and Growth," Working Paper, International Monetary Fund, WP/37/139 (Washington, DC: IMF, 1997).
39. See Stephen S. Everhart and Mariusz A. Sumlinski, "Trends in Private Investment in Developing Countries—Statistics for 1970–2000," International Finance Corporation, Discussion Paper No. 44 (Washington, DC: IFC, 2001).
40. Prema C. Athukorala and Kunal Sen, *Saving, Investment and Growth in India* (New Delhi: Oxford University Press, 2002).
41. The cost of capital that is estimated is $CC = [p^K(r + \delta)]/p$, that is, relative to output price.
42. To guard against simultaneity bias the equation was estimated using two-stage least squares. We do not present the diagnostic tests here for simplicity. See Prema C. Athukorala and Kunal Sen (2002).

## TEST YOURSELF

1. Differentiate between savings, financial investment, and investment in physical capital. For a macroeconomist, is the purchase of shares in the stock market a saving or an investment?
2. Why is the investment-opportunities curve downward sloping? Explain the role of diminishing marginal productivity in determining the slope of the investment-opportunity schedule.
3. What factors govern a firm's decision to buy a given piece of capital equipment?
4. Explain the role of adjustment costs in the theory of investment.
5. Suppose the upfront cost of a two-year MBA programme is INR 200,000. You estimate that your salary in the current year if you work would be INR 165,000 and a year from now it would be INR 275,000. The interest rate is 10 per cent per annum. How much income are you forgoing by joining the MBA programme? Explain why pursuing the MBA might be a good investment.
6. What are the various factors that affect Tobin's marginal *Q* and thereby the investment in the economy.
7. Why is it that external finance is more costly than the internal finance for investment projects?
8. A firm is contemplating an investment project that requires external finance. As a result of a sudden boom in the economy the internal cash flow of the firm increases. What impact will there be on the desired level of investment?
9. What is the advantage of investing later rather than now? What does uncertainty require of macroeconomic policy?
10. In a developing economy what factors other than Tobin's marginal *Q* would you consider as important for the decision making process?

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
   (c) Under the classification of annual data, click on "Handbook of Statistics on Indian Economy".
2. Click on "Table 2: Macro-economic aggregates (at constant prices)". Download data for the last 15 years on Gross Domestic Capital Formation (GDCF) and GDP (at market prices).
3. Calculate the GDCF/GDP ratio and draw a graph of this ratio over time. If GDCF is increasing (or decreasing) as a proportion of GDP, what does this indicate about other components of GDP?
4. Calculate the growth rate of GDP and the growth rate of GDCF. Graph the growth rates. Which variable, GDP or GDCF, depicts greater variability? What could be the reason for this?

# Appendix 4.1 The Optimal Capital Stock When the Production Function Is Cobb–Douglas

To find an explicit solution for the optimal capital stock, allow the production function to be of the Cobb–Douglas type, or,

$$Y = AK^{\alpha}N^{1-\alpha} \qquad (0 < \alpha < 1) \qquad \textbf{(A4.1.1)}$$

where $A$ which represents technology is a constant.

Then
$$MP_N = \frac{\partial Y}{\partial N} = (1-\alpha)AK^{\alpha}N^{-\alpha} = (1-\alpha)\frac{Y}{N}$$

and
$$MP_K = \frac{\partial Y}{\partial K} = \alpha AK^{\alpha-1}N^{1-\alpha} = \alpha\frac{Y}{K}$$

Substituting these values for $MP_N$ and $MP_K$ into Eq. (4.6) in the text gives us

$$\frac{(1-\alpha)\frac{Y}{N}}{\alpha\frac{Y}{K}} = \frac{w}{p^K(r+\delta)}$$

or,
$$\frac{1-\alpha}{\alpha}\frac{K}{N} = \frac{w}{p^K(r+\delta)}$$

or,
$$K = \frac{\alpha}{1-\alpha}\frac{w}{p^K(r+\delta)}N \qquad \textbf{(A4.1.2)}$$

From the production function given in Eq. (A4.1.1),

$$N^{1-\alpha} = \frac{Y}{AK^{\alpha}}$$

or,
$$N = \left[\frac{Y}{AK^{\alpha}}\right]^{\frac{1}{1-\alpha}}$$

Substituting this into the expression (A4.1.2),

$$K = \frac{\alpha}{1-\alpha}\frac{w}{p^K(r+\delta)}\left[\frac{Y}{AK^{\alpha}}\right]^{\frac{1}{1-\alpha}}$$

$$= \frac{\alpha}{1-\alpha}\frac{w}{p^K(r+\delta)}\left[\frac{Y}{A}\right]^{\frac{1}{1-\alpha}}K^{-\frac{\alpha}{1-\alpha}}$$

or,
$$K^{1+\frac{\alpha}{1-\alpha}} = K^{\frac{1}{1-\alpha}} = \frac{\alpha}{1-\alpha}\frac{w}{p^K(r+\delta)}\left[\frac{Y}{A}\right]^{\frac{1}{1-\alpha}}$$

or,
$$K^* = \left[\frac{\alpha}{1-\alpha}\frac{w}{p^K(r+\delta)}\right]^{1-\alpha}\frac{Y}{A} \qquad \textbf{(A4.1.3)}$$

Given that A and $\alpha$ are constants as they represent the current technology of production, if we take the cost of factor inputs to also be given, then the optimal capital stock depends only on output $Y$. This has led some economists to argue that as the demand for output increases, there is an increase in the stock of capital by the firm. As the change in capital stock is investment ($I = \Delta K$)—investment is a function of the change in output—a proposition referred to as the ACCELERATOR THEORY of investment. Whenever there is an acceleration in the growth of output, firms respond by increasing investment. When factor prices are not constant, however, the optimal capital stock to hire is increasing in output, $Y$, and decreasing in the rental to be paid per unit of capital, $k = p^K(r+\delta)$. We may then write generally, that

› The effect of a change in output (or an acceleration in output) on investment is called the accelerator.

$$K^* = K\left[Y, \frac{w}{p^K(r+\delta)}\right] \qquad \textbf{(A4.1.4)}$$

# Appendix 4.2 The Firm's Optimal Decision for Hiring Labour and Accumulating Capital (Investment)

With no maintenance costs and zero-scrap value, the present value of the firm's stream of profits when contemplating an activity of investment is:

$$V = \sum_{t=0}^{\infty} 1/(1+r)^t \pi_t = \sum_{t=0}^{\infty} 1/(1+r)^t \left[ pF(K_t, N_t) - wN_t - p^K I_t - \frac{b}{2} p^K I_t^2 \right] \qquad \textbf{(A4.2.1)}$$

As the firm purchases new capital equipment—undertakes a gross investment activity that costs $p^K I_t + (b/2)\, p^K I_t^2$—the capital stock of the firm changes through time. The law of motion that links the capital stock with gross investment is:

$$K_{t+1} = (1-\delta)K_t + I_t \qquad \textbf{(A4.2.2)}$$

where $\delta$ as before is the constant rate of depreciation. This capital-accumulation identity implies a path for the capital stock once a path is chosen for investment. The firm then seeks to maximize the discounted present value of present and future cash flows subject to the capital accumulation identity

$$\underset{\{N_t, K_{t+1}, I_t\}}{\text{Max}}\; V = \sum_{t=0}^{\infty} 1/(1+r)^t \left[ pF(K_t, N_t) - wN_t - p^K I_t - \frac{b}{2} p^K I_t^2 \right]$$

such that
$$K_{t+1} = (1-\delta)K_t + I_t \qquad \textbf{(A4.2.3)}$$

This is a dynamic-optimization problem where the firm must formulate plans regarding production now and in the indefinite future and choose $Q_t$ for $t = 0, 1, \ldots, \infty$. It does so by choosing time paths for employment, $N_t$, investment, $I_t$, and the capital stock, $K_{t+1}$, such that it maximizes the present value of present and future cash flows subject to the law of motion for capital accumulation. The solution to this problem (see Appendix 4.3) is given by the following two first-order necessary conditions:

$$pF_N = w \qquad \textbf{(A4.2.4)}$$

$$pF_K = p^K (r+\delta)\{1 + bI_t - b(I_t - I^*)\} \qquad \textbf{(A4.2.5)}$$

where $I^*$ is the steady-state value of investment, and $I^* = I_{t+1} = I_t$. Solving Eq. (A4.2.5) for $I^*$ gives the steady-state solution for investment as:

$$I^* = \frac{1}{b}\left[ \frac{pF_K}{p^k(r+\delta)} - 1 \right] \qquad \textbf{(A4.2.6)}$$

The above expression makes clear the role of adjustment costs. If adjustment costs do not exist ($b = 0$), the firm has no well-defined optimal investment policy.

In fact, Eq. (A4.2.6) reduces to $pF_K = p^K(r+\delta)$, which is the static condition derived earlier for the optimal capital stock. So, in the absence of adjustment costs, the firm adjusts its capital stock instantaneously to the optimal level. For investment to take place adjustment costs must be prevalent.

Equation (A4.2.4) is the standard marginal productivity condition for the labour input that was derived in the static case. There is no change here as with labour a flexible factor of production the choice of how much labour to use is not subject to dynamic consideration. At each moment of time, the firm should hire labour up to the point where the value of its marginal product equals the wage. Equation (A4.2.5) by contrast illuminates the intertemporal nature of the firm's investment decision. It states that along the optimal path the value of the marginal product of capital (the left-hand side) equals the marginal user cost of capital (the right-hand side), or the value of the marginal cost attached to adding an additional unit of capital. This comprises three components.

The first is the purchase price of the additional unit of capital—the term $p^K(1)$.

The second component is the current marginal cost of adjustment when an additional unit of capital is acquired—the term $p^K(bI_t)$.

› This **component**, $p^K b(I^* - I_t)$ is the current cost of investing later rather than now.

The THIRD COMPONENT of the right-hand side is the future adjustment cost that is saved as a result of investing now instead of later—the term $p^K b(I_t - I^*)$.

Each expression on the right-hand side of Eq. (A4.2.5) is multiplied by the term $(r + \delta)$. Recollect from our earlier discussion that $(r + \delta)$ converts a stock price to a flow price so that $p^K(r + \delta)$ is the flow price of an additional unit of capital. $(r + \delta)$ is essentially the opportunity cost of a unit of capital. The unit of capital must earn the market rate of return on financial assets ($r$) plus the rate of physical depreciation of the capital stock ($\delta$) in order to make it worthwhile to be deployed in production. Hence, the benefit from an additional unit of capital, $pF_K$, must in equilibrium equal the cost of that additional unit which includes the purchase price, the current adjustment cost, and the cost of the investment postponed to a later date.

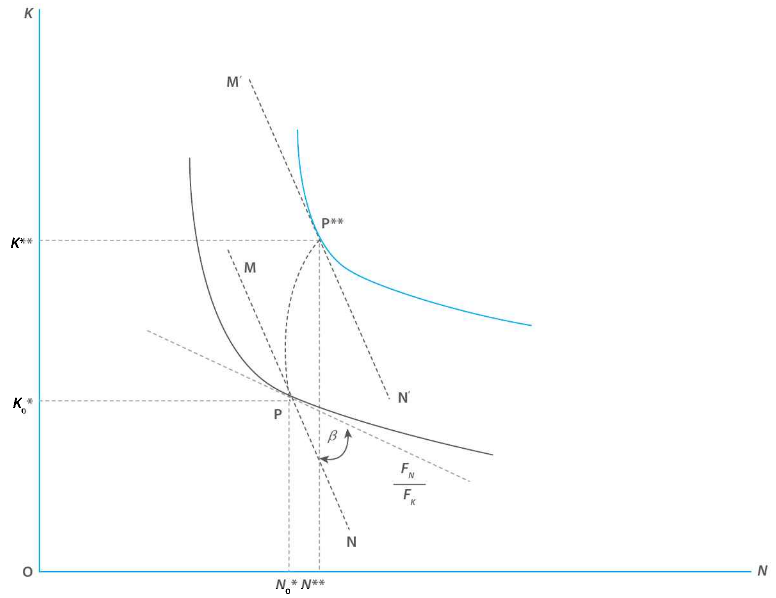
The Eqs. (A4.2.4) and (A4.2.5) solve recursively. The firm starts period $t$ with an inherited value of $K_t$. Given this predetermined value of $K_t$, it chooses $N_t$ so as to satisfy Eq. (A4.2.4). Once it hires $N_t$ of labour any expansion of production in the period from adding to capital will take place along a path characterized by a constant marginal product of labour $F_N$. Once we know $N_t$, and given the inherited $K_t$, the marginal product of capital $F_K$ on the left-hand side of Eq. (A4.2.5) is determined. That leaves the firm to choose the amount of investment $I_t$ that satisfies Eq. (A4.2.5). The marginal product of capital having already being determined from the value of $N_t$ in Eq. (A4.2.4), and the rental on capital—$p^K(r + \delta)$—being given exogenously, investment fills the wedge between $pF_K$ and $p^K(r + \delta)$. $pF_K$ and $p^K(r + \delta)$ are equal to one another only when the optimal capital stock is in existence (which obviates the need for investment) as described by the static optimization condition when the firm is a profit maximizer.

We now divide Eq. (A4.2.4) by (A4.2.5) to obtain

$$\frac{F_N}{F_K} = \frac{w}{p^K(r+\delta)[1 + bI_t - b(I_t - I^*)]} \qquad \textbf{(A4.2.7)}$$

The left-hand side of Eq. (A4.2.7) is the ratio of marginal products or the slope of the isoquant. The right-hand side is like a ratio of input prices—we refer to it as the *total input price ratio.* This ratio differs from the factor price ratio $w/p^K(r + \delta)$, which is the slope of the resource constraint for factor inputs, because it includes the marginal cost of adjustment as the firm acquires capital in order to achieve the optimal capital stock. When the stock of capital is optimal, the firm has no need for any further investment in which case $I_t = I^* = 0$, and the right-hand side of Eq. (A4.2.7) equals the factor price ratio.

Let us consider a firm which at time $t_0$ is in such a position. It possesses the stock of capital $K_0$* and hires the amount of labour $N_0$* that enables it to maximize its profits—the slope of the resource constraint $w/p^K(r + \delta)$ equals the slope of the isoquant $F_N/F_K$. Let us now allow this position of equilibrium to be disturbed by a fall in $r$, the interest rate in the financial market. As interest rates have declined, there is a fall in the cost of capital, and the firm would




**› Figure A4.2.1**
Path of Investment. A profit maximizing firm sets the slope of an isoquant equal to the input price ratio at point P. A decline in the interest rate reduces the rental cost of capital and induces an increase in investment along the path PP**.

seek to acquire more capital as it is cheaper by initiating an investment policy. The new cost line when it does this has a slope given by the right-hand side of Eq. (A4.2.7) and this total input price ratio is illustrated by the line MN in Figure 4.14, which is steeper than the factor-price ratio.

From Eqs. (A4.2.5) and (A4.2.7), the firm plans a positive rate of investment because its cost of capital at $(K_0^*, L_0^*)$ as given by $p^K(r + \delta)$ is less than the benefit from that capital stock, $pF_K$. In terms of the Figure A4.2.1, a wedge has been driven between the line with slope $F_N/F_K$ and the line MN. Investment fills this wedge between the factor-price ratio and the total input-price ratio and the investment undertaken depends on the size of this wedge as measured by the angle $\beta$ in the diagram. In the next time period due to investment taking place, $K$ will be larger than $K_0$, and from Eq. (A4.2.4), as the marginal product of labour increases, more labour is hired by the firm. As investment takes place, the wedge between the factor-price ratio and total input-price ratio will decline because investment takes place at a diminishing rate as the new optimal capital stock $K^{**}$ is achieved.* The firm therefore expands along the path PP** when the decline in interest rate has caused the new factor price ratio to be steeper and have the slope of the line MN. At points between P and P**, the cost line intersects isoquants and only at P** when the optimal capital stock has been acquired is it tangential to an isoquant.

At P**, the factor-price ratio and the total input price ratio are equal. This implies that $pF_N/pF_K = w/p^K(r + \delta)$ or $pF_K = p^K(r + \delta)$. The right-hand side is the flow value of a unit of capital and the left-hand side is the value of the output from a unit of capital, or the value of the marginal product of capital. At point P, the equality was broken by a drop in the interest rate which resulted in $PF_K > P^K(\delta + r)$. Investment or the acquisition of capital stepped in to fill this wedge between the value of the output and the cost of capital input. Investment thus takes place as long as

$$pF_K > p^K(r+\delta)$$

or,

$$\frac{pF_K}{p^K(r+\delta)} > 1$$

or,

$$Q^M > 1 \tag{A4.2.8}$$

*We are assuming here that the production function exhibits diminishing returns to scale (see Appendix 4.2, Section B). If there is constant returns to scale the firm will expand along a ray from the origin through point P and maintain a constant rate of investment.

# Appendix 4.3 The Dynamic Optimization Approach to the Firm's Factor Accumulation Decision

Section A

To find the optimal solution to

$$\underset{\{N_t, K_{t+1}, I_t\}}{\text{Max}} V = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t}\left[pF(K_t, N_t) - wN_t - p^K I_t - \frac{b}{2} p^K I_t^2\right]$$

such that $K_{t+1} = (1 - \delta)K_t + I_t$ set up the Lagrangean

$$L = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t}\left[pF(K_t, N_t) - wN_t - p^K I_t - \frac{b}{2} p^K I_t^2\right] + \sum_{t=0}^{\infty} \frac{\lambda_t}{(1+r)^t}\left[(1-\delta)K_t + I_t - K_{t+1}\right] \quad \textbf{(A4.3.1)}$$

Hence,

$$\frac{\partial L}{\partial N_t} = \frac{1}{(1+r)^t}\left[pF_N - w\right] = 0 \quad \textbf{(A4.3.2)}$$

$$\frac{\partial L}{\partial K_{t+1}} = \frac{1}{(1+r)^t}\left[\frac{pF_K(K_{t+1}, N_{t+1}) + \lambda_{t+1}(1-\delta)}{(1+r)} - \lambda_t\right] = 0 \quad \textbf{(A4.3.3)}$$

$$\frac{\partial L}{\partial I_t} = \frac{1}{(1+r)^t}\left[-p^K - bp^K I_t + \lambda_t\right] = 0 \quad \textbf{(A4.3.4)}$$

Equation (A4.3.2) is the marginal productivity condition for the labour input. Equations (A4.3.3) and (A4.3.4) can be combined to get the optimal path for investment. From Eq. (A4.3.4) we can derive expressions for $\lambda_t$ and $\lambda_{t+1}$:

$$\lambda_t = p^K + bp^K I_t$$

$$\lambda_{t+1} = p^K + bp^K I_{t+1}$$

Substituting these expressions into Eq. (A4.3.3), we obtain the condition for investment. First rewrite Eq. (A4.3.3) as:

$$pF_K(K_{t+1}, N_{t+1}) + \lambda_{t+1}(1-\delta) - \lambda_t(1+r) = 0$$

Hence, $$pF_K(K_{t+1}, N_{t+1}) + p^K(1 + bI_{t+1})(1-\delta) - (1+r)p^K(1 + bI_t) = 0$$

or, $$bp^K(1-\delta)I_{t+1} + p^K(1-\delta) - (1+r)p^K - (1+r)p^K bI_t + pF_K(K_{t+1}, N_{t+1}) = 0$$

or, $$bp^K(1-\delta)I_{t+1} - (1+r)p^K bI_t + pF_K(K_{t+1}, N_{t+1}) - p^K(r+\delta) = 0$$

Dividing throughout by the coefficient of $I_{t+1}$,

$$I_{t+1} - \frac{1+r}{1-\delta} I_t + \frac{pF_K(K_{t+1}, N_{t+1}) - p^K(r+\delta)}{bp^K(1-\delta)} = 0 \quad \textbf{(A4.3.5)}$$

As $(1 + r)/(1 - \delta) > 1$, this is an unstable difference equation for investment. The steady-state solution is given by $I^* = I_{t+1} = I_t$. Hence,

$$I^* - \frac{1+r}{1-\delta} I^* = -\frac{pF_K(K, N) - p^K(r+\delta)}{bp^K(1-\delta)}$$

$$-\frac{(r+\delta)}{1-\delta}I^* = -\frac{pF_K(K,N) - p^K(r+\delta)}{bp^K(1-\delta)}$$

or
$$I^* = \frac{pF_K(K,N)}{bp^K(r+\delta)} - \frac{1}{b}$$

And,
$$I^* = \frac{1}{b}\left[\frac{pF_K(K,N)}{p^K(r+\delta)} - 1\right] \qquad \textbf{(A4.3.6)}$$

The Eq. (A4.3.6) states that if the value of the marginal product of capital, $pF_K$, is greater than the opportunity cost of capital plus the depreciation charge, $p^K(r+\delta)$, the firm should invest. Also note that $b = 0$, or no adjustment cost implies no investment policy as the firm adjusts instantaneously to the optimal level of its capital stock.

To get the dynamics of investment, it is easier to proceed in continuous time rather than the discrete version adopted so far. In continuous time the firm maximizes the present value of the stream of profits now and in the future subject to the capital accumulation constraint, or

$$\text{Max } V = \int_0^\infty \left\{pF(K_t, N_t) - wN_t - p^K I_t - \frac{b}{2}p^K I_t^2\right\}e^{-rt}dt$$

such that
$$\frac{dK}{dt} = I_t - \delta K_t \qquad \textbf{(A4.3.7)}$$

Setting up the current valued Hamiltonian for the above,

$$H = pF(K_t, N_t) - wN_t - p^K I_t - \frac{b}{2}p^K I_t^2 + \mu(I_t - \delta K_t) \qquad \textbf{(A4.3.8)}$$

The firm chooses employment and the rate of investment at each instant, so that the first-order conditions are:

$$\frac{\partial H}{\partial N} = pF_N - w = 0 \qquad \textbf{(A4.3.9)}$$

$$\frac{\partial H}{\partial I} = -p^K - bp^K I + \mu = 0 \qquad \textbf{(A4.3.10)}$$

Note the similarity with Eqs. (A4.2.4) and (A4.2.5) in the discrete case. Also, the differential equation for the multiplier $\mu$ is

$$\dot{\mu} - r\mu = -\frac{\partial H}{\partial K} = -(pF_K - \mu\delta)$$

or,
$$\dot{\mu} = \mu(r+\delta) - pF_K \qquad \textbf{(A4.3.11)}$$

However, from the first-order condition for $I$, $\mu = p^K(1 + bI_t)$.

Hence,
$$\dot{\mu} = p^K b\frac{dI}{dt} \qquad \textbf{(A4.3.12)}$$

Substituting in Eq. (A4.3.11), the dynamic equation for $I$ is

$$p^K b\frac{dI}{dt} = p^K(r+\delta)(1+bI) - pF_K$$

Thus,
$$pF_K = p^K(r+\delta) + p^K b(r+\delta)I - p^K b\frac{dI}{dt} \qquad \textbf{(A4.3.13)}$$

In the steady state, $dI/dt = 0$, or, $I = I^*$ is a constant. Hence,

$$pF_K = p^K(r+\delta) + p^K b(r+\delta)I^*$$

or,
$$\frac{pF_K}{P^K(r+\delta)} = 1 + bI^*$$

or,
$$I^* = \frac{1}{b}\left[\frac{pF_K(K,N)}{p^K(r+\delta)} - 1\right] \quad \textbf{(A4.3.14)}$$

as in the discrete case. Also, rewrite Eq. (A 4.3.11) as follows:

$$r+\delta = \frac{\dot{\mu} + pF_k}{\mu} \quad \textbf{(A4.3.15)}$$

The interpretation of the multiplier $\mu$ is that it is the shadow price of capital. It measures by how much the objective function—the value of the firm—increases if the constraint—the initial capital stock—is relaxed marginally.

In short, $\mu$ can be thought of as the value of capital. $\dot{\mu}$ is the change in the value of capital, or, the capital gain or loss. Thus, the rate of return from incremental investment is the value of the output from that capital, $pF_K$ plus the capital gain or loss as given by the change in the value of capital, $\dot{\mu}$, expressed in terms of the shadow price of capital, $\mu$, or $(pF_K + \dot{\mu})/\mu$.

Thus, the shadow rate of return on investment must equal the market rate of return on financial assets, $r$, plus the rate of depreciation of the capital stock. Hence, Eq. (A4.3.15) is like an implicit arbitrage equation in that the intertemporal solution to the path of investment is akin to choosing an optimal path for the shadow price of capital such that Eq. (A4.3.15) holds. This is why $(r + \delta)$ is the required return on new capital.

Now, the left-hand side of Eq. (A4.3.13) is the value of the additional product from expanding the capital stock. The right-hand side is the value of the additional cost to expanding the capital stock. It comprises three terms as stated in Appendix 4.2—the actual price of acquisition of the capital stock, the adjustment cost of acquiring the capital, and the cost of adjustment saved by investing now rather than later.

Dividing throughout Eq. (A4.3.13) by $p^K(r + \delta)$, we obtain:

$$\frac{pF_K}{p^K(r+\delta)} = 1 + bI - \frac{b}{r+\delta}\frac{dI}{dt}$$

or
$$\frac{pF_K}{p^K(r+\delta)} - 1 = bI - \frac{b}{r+\delta}\frac{dI}{dt}$$

Dividing throughout by $b$:

$$\frac{1}{b}\left[\frac{pF_K}{p^K(r+\delta)} - 1\right] = I - \frac{1}{r+\delta}\frac{dI}{dt}$$

The left-hand side of the above equation is steady-state investment as given by Eq. (A4.3.14). Hence,

$$I^* = I - \frac{1}{r+\delta}\frac{dI}{dt}$$

and
$$\frac{dI}{dt} = (r+\delta)(I - I^*)$$

Substituting this expression into Eq. (A4.3.13) gives us:

$$pF_K = p^K(r+\delta) + p^K b(r+\delta)I - p^K b(r+\delta)(I - I^*) \quad \textbf{(A4.3.16)}$$

which is Eq. (A4.2.5) in Appendix 4.2.

Section B

A production function is considered to be homogeneous of degree $n$ if

$$F(\lambda K, \lambda N) = \lambda^n F = \lambda^n Q \quad \textbf{(A4.3.17)}$$

If $n > 1$, the production function is said to exhibit increasing returns to scale. Similarly, $n = 1$, and $n < 1$ describes a constant returns to scale and decreasing returns to scale production function. Euler's theorem for homogeneous production functions implies that

$$F_N N + F_K K = nQ \tag{A4.3.18}$$

Multiplying throughout by $p$ gives:

$$pF_N N + pF_K K = npQ \tag{A4.3.19}$$

The right-hand side is the degree of homogeneity multiplied by the total revenue of the firm, $pQ$. So, if factors of production $N$ and $K$ are paid the value of their marginal product, $pF_N$ and $pF_K$, then, the above equation allows us to relate factor payments to the total revenue of the firm. For a decreasing returns to scale production function ($n < 1$), we have,

$$pF_N N + pF_K K = npQ < pQ \tag{A4.3.20}$$

or the expenses on factor inputs is less than total revenue—referred to as under exhaustion of the product. If the production function exhibits constant returns to scale, $n = 1$, then total factor payments totally exhaust the firm's revenue. Labour is paid the value of its marginal product by the firm as given by Eq. (A4.3.9) in Section A, or,

$$pF_N = w \tag{A4.3.21}$$

For constant product and factor prices, this implies that the marginal product of labour $F_N$ is constant for the firm. Totally differentiating the above equation,

$$\frac{\partial}{\partial N}\left\{\frac{\partial F}{\partial N}dN + \frac{\partial F}{\partial K}dK\right\} = 0$$

or

$$\frac{dK}{dN} = -\frac{F_{NN}}{F_{NK}} \tag{A4.3.22}$$

If a production function is homogeneous of degree $n$, its marginal product function $F_N(K, L)$ is homogeneous to degree $(n - 1)$. Then, employing Euler's theorem for the marginal productivity function we can write

$$F_{NK}K + F_{NN}N = (n-1)F_N \tag{A4.3.23}$$

Dividing throughout by $F_{NK}K$

$$1 + \frac{F_{NN}N}{F_{NK}K} = \frac{(n-1)F_N}{KF_{NK}}$$

or

$$-\frac{F_{NN}N}{F_{NK}K} = 1 + \frac{(1-n)F_N}{KF_{NK}}$$

or

$$-\frac{F_{NN}}{F_{NK}} = \frac{K}{N} + \frac{(1-n)F_N}{NF_{NK}} \tag{A4.3.24}$$

Substituting Eq. (A4.3.24) into Eq. (A4.3.22),

$$\frac{dK}{dN} = \frac{K}{N} + \frac{(1-n)F_N}{NF_{NK}} \tag{A4.3.25}$$

The left-hand side of Eq. (A4.3.25) is the slope of the locus of points along which the marginal productivity of labour is constant, or, $F_N = w/p$ as given by Eq. (A4.3.21). Now, $K/N$ is the slope of a ray through the origin. Hence, if $n < 1$, which is a decreasing returns to scale production function, then, the slope of the locus $pF_N = w$ is greater than the slope of the ray through the origin implying a diminishing rate of investment. Analogously, if the production function is a constant returns to scale one, $n = 1$, and the locus $pF_N = w$ is a ray through the origin with the firm maintaining a constant rate of investment.

# Appendix 4.4 Value of the Firm's Equity

The asset market arbitrage condition is given by:

$$\rho V_t = D_t + (V_{t+1} - V_t) - V_t^N$$

This is a difference equation that we solve forwards, by first writing the terms in $V_t$ on the left-hand side:

$$(1 + \rho)V_t = D_t + V_{t+1} - V_t^N \tag{A4.4.1}$$

In period $t + 1$ the expression for the total value of the firm is

$$(1 + \rho)V_{t+1} = D_{t+1} + V_{t+2} - V_{t+1}^N$$

or

$$V_{t+1} = \frac{D_{t+1}}{1+\rho} + \frac{V_{t+2}}{1+\rho} - \frac{V_{t+1}^N}{1+\rho} \tag{A4.4.2}$$

Similarly,

$$V_{t+2} = \frac{D_{t+2}}{1+\rho} + \frac{V_{t+3}}{1+\rho} - \frac{V_{t+2}^N}{1+\rho} \tag{A4.4.3}$$

and generally,

$$V_{t+i} = \frac{D_{t+i}}{1+\rho} + \frac{V_{t+i+1}}{1+\rho} - \frac{V_{t+i}^N}{1+\rho} \tag{A4.4.4}$$

Substituting Eq. (A4.4.3) in Eq. (A4.4.2),

$$\begin{aligned} V_{t+1} &= \frac{D_{t+1}}{1+\rho} + \frac{1}{1+\rho}\left[\frac{D_{t+2}}{1+\rho} + \frac{V_{t+3}}{1+\rho} - \frac{V_{t+2}^N}{1+\rho}\right] - \frac{V_{t+1}^N}{1+\rho} \\ &= \frac{D_{t+1}}{1+\rho} + \frac{D_{t+2}}{(1+\rho)^2} + \frac{V_{t+3}}{(1+\rho)^2} - \frac{V_{t+2}^N}{(1+\rho)^2} - \frac{V_{t+1}^N}{1+\rho} \end{aligned}$$

We can replace $V_{t+3}$ on the right-hand side by using the general expression for $V_{t+i}$ in Eq. (A4.4.4) with $i = 3$, and continue in this way replacing the $V_{t+i}$ terms on the right-hand side so that we can write

$$V_{t+1} = \frac{D_{t+1}}{1+\rho} + \frac{D_{t+2}}{(1+\rho)^2} + \cdots - \frac{V_{t+2}^N}{(1+\rho)^2} - \frac{V_{t+1}^N}{1+\rho}$$

Substituting this into Eq. (A4.4.1)

$$(1+\rho)V_t = D_t + \left(\frac{D_{t+1}}{1+\rho} + \frac{D_{t+2}}{(1+\rho)^2} + \cdots - \frac{V_{t+2}^N}{(1+\rho)^2} - \frac{V_{t+1}^N}{1+\rho}\right) - V_t^N$$

or

$$V_t = \frac{D_t}{1+\rho} + \frac{D_{t+1}}{(1+\rho)^2} + \frac{D_{t+2}}{(1+\rho)^3} + \cdots - \frac{V_{t+2}^N}{(1+\rho)^3} - \frac{V_{t+1}^N}{(1+\rho)^2} - \frac{V_t^N}{1+\rho}$$

or

$$V_t = \sum_{i=0}^{\infty} \frac{1}{(1+\rho)^{i+1}} (D_{t+i} - V_{t+i}^N) \tag{A4.4.5}$$

Hence, the total value of the firm is the present value of the dividend stream adjusted for the present value of new issued shares that would have to be bought by current equity holders to maintain their proportional claim on the firm.

# Appendix 4.5 Tobin's $Q$ and the Rate of Interest

Tobin's $Q$ is given by the ratio of the market value of a business unit to the replacement cost of its assets. The market value is the present value of the future stream of earnings by the firm which is the difference between the revenues generated by the firm in an operating period ($R_t$) and the expenses incurred during the same period ($E_t$). The earnings capitalized at the rate $r$ yields the market value of the firm as

$$\text{Market Value} = \sum_{t=1}^{n} \frac{R_t - E_t}{(1+r)^t}$$

The replacement cost of the assets of the firm ($RC_0$)is the initial expenditure on investment that yields a stream of future earnings to the firm given by $\sum_{t=1}^{n}(R_t - E_t)$. Hence, the internal rate of return ($r_N$) is given by:

$$-RC_0 + \sum_{t=1}^{n} \frac{R_t - E_t}{(1+r_N)^t} = 0$$

or

$$RC_0 = \sum_{t=1}^{n} \frac{R_t - E_t}{(1+r_N)^t}$$

If the earnings is constant over time, then, we may write the $Q$ relationship as

$$Q = \frac{\text{Market Value}}{\text{Replacement Cost}} = \frac{\sum_{t=1}^{n}(R_t - E_t)/(1+r)^t}{\sum_{t=1}^{N}(R_t - E_t)/(1+r_N)^t}$$

$$= \frac{\sum[1/(1+r)^t]}{\sum[1/(1+r_N)^t]}$$

Now, the geometric series $1/x + 1/x^2 + 1/x^3 + \cdots + 1/x^n$ sums to $1/x$ as $n \to \infty$. We may then write Tobin's $Q$ as:

$$Q = \frac{1/r}{1/r_N} = \frac{r_N}{r}$$

This states that a project should be undertaken when $Q > 1$, or, where the internal rate of return is greater than the rate of interest. Alternatively, investment is inversely related to the rate of interest.

# 5

# The TradeB alance and Exchange Rates

CRITICAL QUESTIONS

» *What is a real exchange rate?*

» *How is a country's trade balance influenced by the competitiveness of its goods and services?*

» *How does a devaluation affect the short-run and the long-run trade balance?*

» *How do changes in the income and expenditure of citizens affect the trade balance?*

» *What is uncovered and covered interest rate parity, and how is the exchange rate determined in asset markets?*

» *What is the law of one price? How does it determine the purchasing power parity in the foreign-exchange market?*

» *What are the various exchange-rate regimes between fixed and floating exchange-rate systems, and what effect do they have on macroeconomic management?*

» *What sort of exchange-rate regime is implied by (a) a currency board, (b) dollarization, and (c) monetary union?*

# 5.1 The Real Exchange Rate

Up to now, we have been dealing with a closed economy. However, economies are open to trade and world capital markets and are increasingly integrating with the global economy. In order to export or import goods, any agent must be able to access the foreign exchange that will be accepted in the market. In what follows, we simplify this concept by lumping the rest of the world into one category so that we can speak of "the" exchange rate.[1] In Chapter 2, we had defined the dimension of the exchange rate as the rupee price of the dollar

$$\text{Dimension } [E] = \frac{\text{Rupee}}{\text{Dollar}}$$

As we had discussed earlier, there are two polar systems for setting exchange rates: fixed and flexible. Under a fixed exchange-rate system, the central bank posts the price at which it is willing to buy or sell foreign currencies and then tries to maintain those rates administratively. Under a flexible exchange-rate system, rates are established in private markets for foreign currencies which may be influenced, but not controlled, by central bank's actions. When the exchange rate rises, the rupee is said to depreciate under a system of flexible exchange rates or to be devalued under a system of fixed exchange rates. When the exchange rate falls, the rupee appreciates under flexible rates or is revalued under fixed rates.

The exchange rate is a nominal price for exchanging one currency for another. Adjusting for differences in the price levels between two countries gives us the real exchange rate, or

$$\text{RER} = \frac{EP^*}{P} \tag{5.1}$$

where $P^*$ is the average foreign-price level (dollars per unit of US goods).

The nominal exchange rate is the price of the domestic currency in terms of the foreign currency. The real exchange rate is the price of the average foreign good or service in terms of the average domestic good or service. When we enquire whether on average the goods and services produced by a particular country are expensive relative to the goods and services produced by other countries, we are asking a question about the real exchange rate. A country's REAL EXCHANGE RATE is the price of the average foreign good or service relative to the price of the average domestic good or service, when prices are expressed in terms of a common currency.

› Acou ntry's real exchange rate is the price of the average foreign good or service relative to the price of the average domestic good or service, when the prices are expressed in terms of a common currency.

$$\text{RER} = \frac{EP^*}{P} = \frac{\text{Rupee}}{\text{Dollar}} \times \frac{\text{Dollar/US goods}}{\text{Rupee/Domestic goods}}$$

$$= \frac{\text{Domestic goods}}{\text{US goods}}$$

$$= \frac{\text{Rupee/US goods}}{\text{Rupee/Domestic goods}}$$

We cannot express the real exchange rate as the ratio $P^*/P$ because the two price levels are expressed in different currencies. When $P^*/P$, which is the price of foreign goods relative to the price of domestic goods, is measured in a common currency, we get the real exchange rate. The real exchange rate incorporates both the nominal exchange rate, $E$, and the relative prices of goods and services across countries.

As you would have noticed, economists focus on real variables. When we were examining how to measure income, for instance, we wanted a measure that corrects for changes in the general level of prices—real incomes rather than nominal ones. In a similar vein, a distinction is made between the nominal exchange rate and the real exchange rate. It is the change in the real exchange rate rather than in the nominal exchange rate that affects the exports and imports of a nation—a distinction that is often overlooked, mistakenly.

Let us take an example. Suppose you are planning to purchase furniture for your home, and one option is to import it from Italy. The furniture you have chosen for purchase costs 1,000 euros for being delivered at your home, according to the price list of the Italian furniture brand Friuli. If you were to purchase an Indian brand such as Style Spa, it would cost you INR 60,000. At the time you are contemplating your purchase, the current exchange rate for the euro is INR 57 per euro. In order to purchase the Italian furniture, for 1,000 euros you would need to exchange INR 57,000. When you compare the price of the Italian furniture with the Indian furniture, you find that the latter is costlier by INR 3,000, or 5.3 per cent. To come to this conclusion, you required three pieces of information—the euro price of the Italian furniture, the rupee price of the Indian furniture, and the nominal exchange rate.

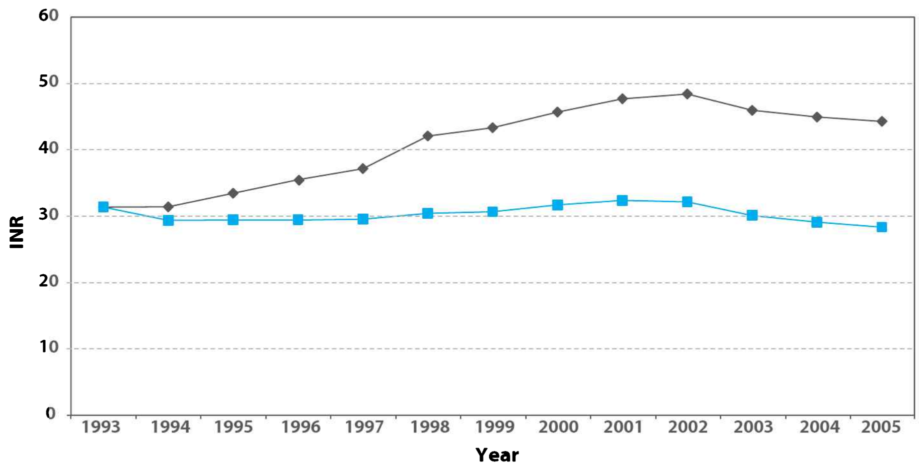
However, if prices rise by 3 per cent a year in Italy, and in India it is 5 per cent a year, then the price of Italian furniture would increase to 1,030 euros, whereas that of Indian furniture would rise to INR 63,000, in a year's time. If the nominal exchange rate does not change—INR 57 is equal to 1 euro—the rupee price of the Italian furniture would be INR 58,710. The Indian furniture is now costlier by INR 4,290, or 7.3 per cent. Even without a change in the nominal exchange rate, Indian furniture is relatively more expensive than Italian furniture, and Indian purchasers of furniture would purchase more Italian furniture—a situation of more imports.

Suppose, on the other hand, the rupee depreciates against the euro so that INR 58.11 equals 1 euro. Then, the rupee price of Italian furniture would rise to INR 59,853.30 (1,030 $\times$ 58.11), which is INR 3,146.70 lower than the cost of Indian furniture (INR 63,000). Indian furniture would now cost 5.3 per cent more than Italian furniture, just like in the situation before the price change. Indian furniture being relatively less costly (5.3 per cent costlier instead of 7.3 per cent), Indians would tend to purchase less from abroad—imports would be less than the case where the nominal exchange rate was constant.

This example illustrates that it is not only changes in the nominal exchange rate but also changes in the price levels in two countries that affect imports and exports. That is why the real exchange rate, which considers the relative price of domestic and foreign goods rather than just the nominal exchange rate, affects the decision to export or import.

### *5.1.1* Real Exchange Rates and Net Exports

When the real exchange rate is high, foreign goods are, on an average, more expensive than domestic goods when they are priced in the same currency. A high real exchange rate means that domestic producers will find it easy to export to other countries while foreign goods will be expensive relative to goods produced at home. Thus, a high real exchange rate will tend to increase exports and reduce imports. Whether the nominal exchange rate is fixed or whether it floats, the real exchange rate always floats. When the nominal exchange rate is pegged, the price level at home must adjust vis-à-vis prices abroad to bring the real exchange rate towards equilibrium. The real exchange rate of the




› **Figure5. 1**
Real and Nominal Exchange Rates. The nominal exchange rate of the INR per US dollar has been depreciating till 2002. After that it has appreciated. The real exchange rate has been more or less constant over the period and has appreciated since 2002.

Hong Kong dollar is thus a floating rate even though the nominal rate is fixed by Hong Kong's currency board.

The graph of the real exchange rate[2] and the nominal exchange rate of the rupee vis-à-vis the US dollar is depicted in Figure 5.1. The real exchange rate for the Indian economy has mildly fluctuated around a constant trend from 1993 onwards. It depreciated till 2002 and has been appreciating since then. However, the nominal exchange rate has been depreciating at a trend growth rate of 3.2 per cent from INR 31.4 per dollar in 1993 to INR 45.3 per dollar in 2006. Since then it has been appreciating rapidly and was INR 39.6 per US dollar on 24 October 2007. It appears that as the inflation differential between the two economies widened the RBI intervened in the foreign exchange market by allowing the nominal exchange rate to depreciate up to 2002, and then appreciate subsequently so as to stabilize the real exchange rate.[3]

The real exchange rate is also important for the definition of the *trade balance*, or the difference between exports ($X$) and imports ($M$). Exports and imports are incommensurable as the former are domestic goods and the latter are foreign goods. To measure imports in the same units as exports—volume of domestic goods—we need to multiply the volume of imports by the real exchange rate.

Let a country produce only one good, which is used for both export and domestic consumption. Let $Q_M$ denote real imports. Then, $P^*Q_M$ is the nominal value of imports in the foreign currency. The nominal value of imports in terms of domestic currency will then be $EP^*Q_M$, where $E$ is the nominal exchange rate. The trade balance, or net exports—exports minus imports—in nominal terms will then be given by

$$\text{Nominal net exports} = PQ_X - EP^*Q_M$$

In real terms—in terms of the price of domestic goods—net exports will be

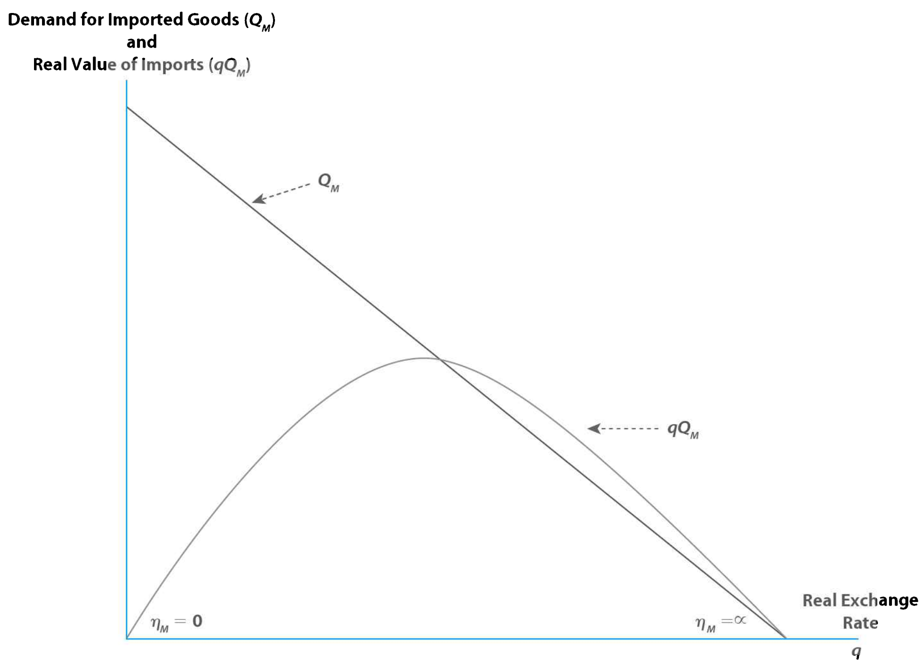
$$\begin{aligned} \text{NX} &= Q_X - \frac{EP^*}{P} Q_M \\ &= Q_X - qQ_M \end{aligned}$$

where $q = (EP^*)/P$ is the real exchange rate.

Hence, net exports (NX) is defined as[4]

$$\begin{aligned} \text{NX} &= Q_X - qQ_M \\ &= X - M \end{aligned} \tag{5.2}$$

› **Figure5.2**
The Demand for Imported Goods and Imports. **As the real exchange rate depreciates, the demand for imported goods ($Q_M$) declines. But if the demand is measured in terms of domestic goods ($qQ_M$), imports initially increase and then eventually decline as the demand for imported goods becomes elastic.**




### *5.1.2* The Marshall–Lerner Condition

› The real exchange rate, which is the nominal exchange rate adjusted for the relative price of foreign and domestic goods, is a measure of the competitiveness of a country's goods and services.

The net exports of a country—as measured by the real exchange rate—are influenced by its COMPETITIVENESS in terms of the amount of another country's goods and services that a unit of our own nation's goods and services can purchase. We discussed earlier how the demand for imported goods declines as the real exchange rate depreciates. In Figure 5.2, the demand for imported goods, $Q_M$, has been shown to be inversely related to the real exchange rate: At low values of the real exchange rate $q$, the demand for imported goods is high, and this demand declines as the real exchange rate increases.

The real value of imported goods in terms of domestic goods is $qQ_M$. Beginning at the origin 0, the initial increase in the real exchange rate ($q$) is not completely offset by a decline in the demand for imported goods ($Q_M$), and the real value of imports ($qQ_M$) increases. At some point, an increase in the real exchange rate is offset by a decline in the demand for imported goods, and the real value of imports decreases. This is because of a well-known property from the theory of demand that at low prices—in this case a low real exchange rate—the quantity demanded is high and the demand is inelastic. When this happens, raising prices results in increased expenditures. Conversely, when the real exchange rate is high and the demand is elastic, raising prices or the real exchange rate results in reduced expenditures on imported goods.

The measure of elasticity of demand is the percentage change in quantity divided by the corresponding percentage change in price. If the absolute value of this quotient is less than 1—an inelastic demand—we know that the percentage change in the quantity demanded is less than the percentage change in the price. When this happens, the price increase results in an increase in the expenditure on imported goods as measured in terms of domestic goods.

The price elasticity of demand for imported goods is defined as

$$\eta_M = -\frac{\Delta Q_M / Q_M}{\Delta q / q}$$

where $\Delta Q_M$ is the change in the demand for imported goods, and $\Delta Q_M/Q_M$ is the percentage change in the demand for imported goods.

If the demand for imported goods is inelastic, then

$$\frac{\Delta Q_M/Q_M}{\Delta q/q} < 1$$

or

$$\frac{\Delta Q_M}{Q_M} < \frac{\Delta q}{q}$$

Multiplying throughout by $qQ_M$, we get

$$qQ_M\left(\frac{\Delta Q_M}{Q_M}\right) < qQ_M\frac{\Delta q}{q}$$

where $qQ_M$ is the prevalent demand for imported goods valued in terms of domestic goods before the change in the real exchange rate. Thus, $qQ_M(\Delta Q_M/Q_M)$ is the extent of the reduction in the demand for imported goods in terms of domestic goods. Similarly, $qQ_M\,(\Delta q/q)$ *is* the increase in the expenditure on imported goods in terms of domestic goods due to the increase in the price of foreign goods in terms of domestic goods—the increase in the real exchange rate. The increased expenditure on imported goods, due to the rise in the real exchange rate being larger than the decline in the demand for imported goods, implies that imports have increased.

Essentially imports have risen because the decline in the demand for imported goods is less than the increase in the expenditure on imported goods at higher real exchange rates. For example, suppose the demand for imported goods is given by $Q_M = 1 - \frac{1}{2}(q)$and the real exchange rate is $q = 0.25$, then $Q_M = 1 - \frac{1}{2}(0.25) = 0.875$. If the real exchange rate rises to $q = 0.30$, then $Q_M = 1 - \frac{1}{2}(0.30) = 0.85$. However, the real value of imports when $q = 0.25$ is $qQ_M = 0.25 \times 0.875 = 0.22$. The real value of imports when $q = 0.30$ is $qQ_M = 0.30 \times 0.85 = 0.26$. The rise in the real exchange rate increases the real value of imports. The increase in the real exchange rate is $\Delta q = 0.05$, and the decline in the demand for imported goods is $\Delta Q_M = 0.025$. The elasticity of demand for imported goods is $\eta_M = (\Delta Q_M/Q_M)/(\Delta q/q) = (0.025/0.85)/(0.05/0.3) = 0.18$. Generally, when the demand for imported goods is inelastic ($\eta_M < 1$) a rise in the real exchange rate increases imports.

Now, suppose the real exchange rate is much higher, $q = 1.75$. Then, the demand for imported goods is $Q_M = 1 - \frac{1}{2}(1.75) = 0.125$. The real value of imports at this real exchange rate is given by $qQ_M = 1.75 \times 0.125 = 0.22$. If now the real exchange rate rises to $q = 1.80$, the demand for imported goods is $Q_M = 1 - \frac{1}{2}(1.80) = 0.10$ and the real value of imports is $qQ_M = 1.80 \times 0.10 = 0.18$. The increase in the real exchange rate has resulted in a decline in imports. The increase in the real exchange rate in this case is $\Delta q = 0.05$ and the decline in the demand for imported goods is $\eta Q_M = 0.025$. However, now the elasticity of demand for imported goods is $\eta_M = (\Delta Q_M/Q_M)/(\Delta q/q) = (0.025/0.10)/(0.05/1.8) = 9$. Generally, when the demand for imported goods is elastic ($\eta_M > 1$) a rise in the real exchange rate reduces imports.

When the REAL EXCHANGE RATE takes small values, the elasticity of the demand for imported goods is less than unity ($\eta_M < 1$) and a rise in the real exchange rate increases imports.

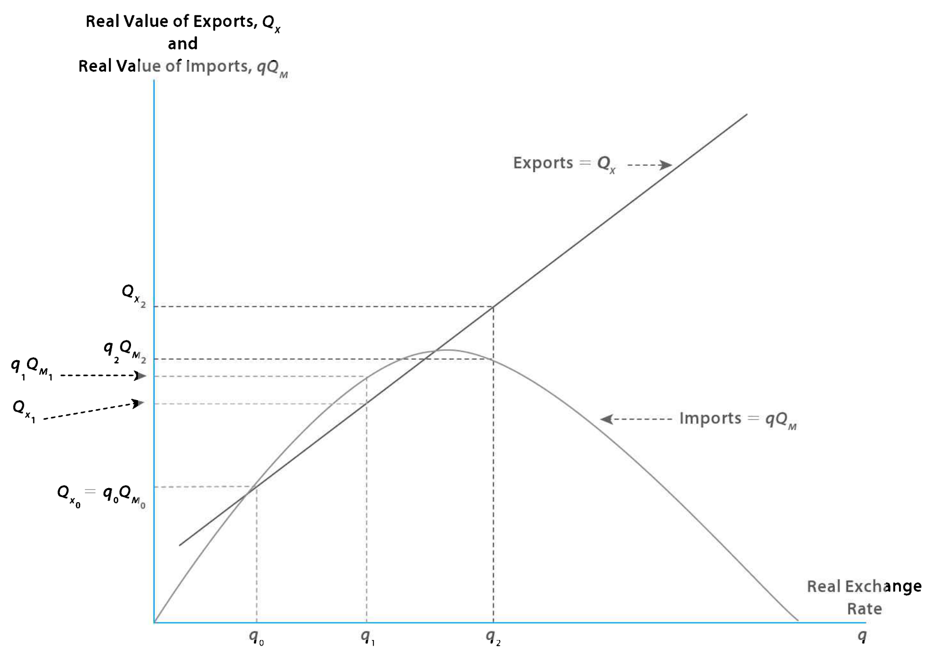
› If the demand for imported goods is inelastic then the increase in the real exchange rate increases imports. If the demand is elastic, an increase in the real exchange rate results in a decline in imports.

Conversely, when the demand for imported goods is elastic, ($\eta_M > 1$), an increase in the real exchange rate reduces imports. At the origin, $q = 0$, and the elasticity of demand for imported goods is $\eta_M = -[(\Delta Q_M/\Delta q)(q/Q_M)] = 0$. In a similar manner, at the point where the demand for imported goods is zero ($Q_M = 0$), the curve representing the demand for imported goods intersects the horizontal axis and the elasticity of demand is perfectly elastic, or $\eta_M = -[(\Delta Q_M/\Delta q)(q/Q_M)] = \infty$. The elasticity of demand for imported goods is, therefore, inelastic when $q$ takes values close to the origin, and it is elastic when the value of $q$ is so high that it makes the demand for imported goods very small. In fact, imports which are measured in terms of domestic goods ($qQ_M$) increases initially as $q$ increases when $\eta_M < 1$, and then decreases as the elasticity of demand for imported goods becomes greater than unity. The graph of imports ($qQ_M$) with respect to the real exchange rate is, therefore, an inverted U-shaped curve, as depicted in FIGURE 5.2. The graph attains its maximum value at the point where $\eta_M = 1$.

› See **Appendix 5.1** for the derivation.

In Figure 5.3 the exports of an economy are depicted as increasing with the real exchange rate and the imports of the economy are depicted as first increasing with the real exchange rate and then decreasing—an inverted U-shaped curve. Suppose the real exchange rate that prevails is $q_0$. Then, exports equal imports, that is $Q_{X0} = q_0 Q_{M0}$, and there is balanced trade. When the real exchange rate depreciates to $q_1$, imports are larger than exports, that is $q_1 Q_{M1} > Q_{X1}$, and the trade balance is in deficit. At an even higher depreciated value of the real exchange rate, $q_2$, exports exceed imports, $Q_{X2} > q_2 Q_{M2}$, and the trade balance is in surplus. As the real exchange rate keeps depreciating beyond $q_2$ Figure 5.3 depicts an increase in the trade balance surplus.

At low values of $q$ then, an increase in the real exchange rate that increases the competitiveness of the domestic economy can actually result in the emergence of trade deficits. Trade surpluses are only guaranteed at substantially larger values of the real exchange rate. At what point then does a real depreciation or devaluation of the exchange rate have beneficial consequences in terms of improving the trade balance of the economy? We had arrived at the conclusion earlier that a real depreciation actually decreases imports when $\eta_M > 1$ or when the demand for




**› Figure5. 3**
The Real Exchange Rate and the Balance of Trade. **An increase in the real exchange rate from $q_0$ to $q_1$ worsens the trade balance with imports increasing more than exports. This is because the increased expenditure on imported goods at higher real exchange rates more than offsets the reduction in the demand for imported goods. At sufficiently high values of the real exchange rate, such as $q_2$, the trade balance improves when the real exchange rate increases.**

imported goods is elastic. It turns out that the more elastic imports and exports are, the more likely is it that a real exchange rate depreciation or devaluation will improve the trade balance. The condition for a real depreciation or devaluation to improve the trade balance is known as the MARSHALL–LERNER CONDITION.[5] This condition states that if the trade balance is initially zero,[6] a real exchange-rate depreciation causes a trade balance surplus if the sum of the relative price elasticities of export and import demand exceeds unity, that is

› A devaluation or depreciation of the real exchange rate improves a country's trade balance when the Marshall–Lerner condition holds true, that is the sum of the relative price elasticities of export and import demand exceeds unity.

$$\eta_X + \eta_M > 1 \tag{5.3}$$

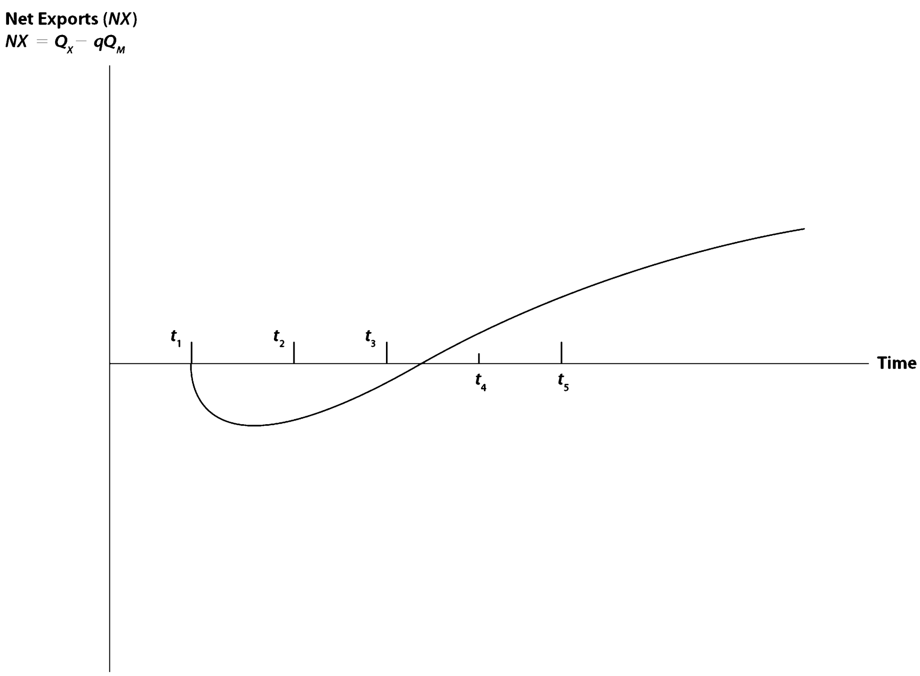
where $\eta_X$ is the relative price elasticity of exports with respect to the real exchange rate and $\eta_M$ is the relative price elasticity of demand for foreign goods with respect to the real exchange rate.

Suppose we consider the level of domestic prices ($P$) and the foreign currency price of the demand for foreign goods ($P^*$) to be given. Then, a change in the real exchange rate is identical to a change in the nominal exchange rate, $\Delta q = \Delta E$. An increase in the relative price $q$ is a pure reflection of a nominal depreciation of the exchange rate. Hence, if prices are fixed or slow to adjust, Eq. (5.3) is the condition for a nominal devaluation by the monetary authority to result in an improvement in the trade balance.

### *5.1.3* The J-Curve

Goldstein and Kahn[7] measured the response of trade flows to changes in the real exchange rate over different adjustment periods. They found that the medium- to long-run elasticities are about twice as large as the short-run elasticities. The short-run elasticities measure the response of trade flows in the first six months after the change in the real exchange rate. The short-run elasticities almost always fail to sum to unity, while the long-run elasticities almost always sum to greater than unity. The possibility that in the short-run the Marshall–Lerner condition may not be fulfilled, although it generally holds true in the longer run, is known as the J-CURVE effect. Following a real devaluation

› The J-curve is the phenomenon where a depreciation of the currency causes the trade balance to initially worsen before it eventually improves.




**› Figure5. 4**
The J-curve. **A real devaluation at time $t_1$ results initially in a deterioration of the trade balance. After some time, $t_3$, the trade balance improves as exports and the demand for imported goods respond to the devaluation.**

› The J-curve effect arises due to adjustment lags. This is because quantity adjustments take long and faster pass through of the exchange rate into prices worsens the trade balance in the short run.

at time $t_1$ (see Figure 5.4), the trade balance deteriorates, and after some time, say, after more than two quarters of a year, it improves.

Magee[8] explains this pattern in terms of ADJUSTMENT LAGS resulting from (a) currency contracts signed prior to devaluation; (b) newer currency contracts signed after devaluation—the period of pass-through; and (c) sluggish quantity adjustments. Let us consider each case.

(a) *Currency contracts signed prior to devaluation:* Prior to signing a contract, economic agents take into account their expectations concerning the future devaluation or appreciation of currencies in order to make a capital gain or avoid a capital loss. An exporter would prefer receiving payments in a currency that is expected to strengthen or appreciate, as after converting receipts into domestic currency units, their takeaway improves. On the other hand, importers prefer to make payments in currencies that are expected to weaken.

Suppose the exchange rate prior to devaluation is INR 35 equals US$ 1. An importer enters into a contract for 100 units of imports denominated in foreign currency of US$ 1 per unit. During the period between the time the contract is entered into and the time the final payment is made, the home currency is devalued to INR 40 equals US$ 1. Importers must now pay INR 4,000 after the devaluation and sustain a capital loss of INR 500. If, on the other hand, the contract was denominated in INR, importers pay INR 3,500 and foreign exporters get only US$87.5, thus absorbing a capital loss of US$ 12.5. Thus, in order to get a capital gain, importers would wish to pay in the currency that depreciates.

Given the similar expectations of both parties, the currency in which contracts are denominated depends on the relative bargaining power of the involved parties. Countries that tend to specialize more in exports than in imports will tend to have more market power in their export markets. Countries that are primary producers tend to have a larger share of import contracts relative to export contracts denominated in foreign currencies. With import contracts in foreign currency, their trade balance measured in their domestic currency deteriorates in the immediate period following devaluation.

› A pass-through effect arises when a change in the exchange rate results in a change in the domestic prices of imported goods.

(b) *Pass-through*: A PASS-THROUGH EFFECT occurs when a change in the exchange rate results in a change in domestic prices of imported goods. Thus, when the domestic currency depreciates, a pass-through effect occurs that results in higher domestic prices of imported goods. This pass-through of the devaluation in prices implies a worsening of the trade balance. Quantity adjustments take longer time in the short run due to supply bottlenecks—supply might be perfectly inelastic for a while because exporters cannot instantly alter their output or sales abroad. Demand may also be perfectly inelastic as importers require time to substitute among commodities and change their order books. When both export and import supplies are inelastic in the short run, the trade balance deteriorates during the fixed pass-through period.

(c) *Quantity-adjustment period*: In the long run, elasticities improve as quantities adjust with the easing of bottlenecks. As the volume effects begin to dominate the price effects, the trade balance improves. Hence, initially contracts already in force in specified currencies dominate the determinants of the trade balance. Over time, new contracts made after devaluation begin to dominate. During the period of pass-through, the trade balance deteriorates due to supply lags. However, as the quantity adjusts, and buying patterns and orders change with price changes, the trade balance improves.

# 5.2 Other Approaches to the Trade Balance

The elasticities approach emphasized the price variables as a determinant of the trade balance. The ABSORPTION APPROACH was developed in the 1950s by Sidney Alexander, an economist at the IMF. It highlights the fact that changes in domestic income are important in understanding trade balances. Absorption ($A$) is nothing but a nation's total expenditures on final goods and services—consumption, investment, government expenditure, and imports. Thus,

› The absorption approach to trade balance determination emphasizes changes in domestic income and expenditure.

$$A = C + I + G + M$$

As you can see, exports are excluded from absorption as they are part of another country's expenditure on final goods and services. The income of the citizens of a nation ($Y^c$), on the other hand, is the receipt from the expenditures on its final goods and services—consumption expenditures, investment, government, and export expenditure. Thus,

$$Y^c = C + I + G + X$$

The absorption approach states that the trade balance of a country is the difference between the income of its citizens and the absorption.

$$\begin{aligned} Y^c - A &= (C + I + G + X) - (C + I + G + M) \\ &= X - M \end{aligned} \tag{5.4}$$

Suppose now there is a devaluation of the domestic currency, which causes foreign goods and services to become relatively more expensive. The absorption approach argues that the entire reduction in the demand for imports due to the devaluation is presumed in the elasticities approach to be switched to the domestic market where goods are relatively cheaper. If the switch of expenditures from imports to domestically produced goods is such that it leaves the overall absorption to be the same, and if the current income of citizens remains the same as well, there would be no change in the trade balance. The trade balance changes only if there is an increase in income relative to absorption.

Alexander[9] argued that the EXPENDITURE SWITCH from the import market to the domestic market, following a devaluation, increases the utilization of domestic resources. Prices rise as producers respond by moving up their supply curve as full employment is approached. The value of an individual's real cash balances—the money stock deflated by the price levels—declines. With a given stock of money, the rise in prices will change the relation of money income to money stocks, and will lead individuals to attempt to build up money stocks. They achieve this through a cutback in spending, and money hoarding at any given level of real incomes is increased. Cutbacks in expenditure by residents implies a decline in absorption, which means that the term ($Y^c - A$) increases. This is equivalent to an increase in the trade balance. According to the absorption approach, expenditure switching due to a devaluation thus results in an improvement of the trade balance. Reduced domestic expenditure on domestic output is absorbed by export demand, and net exports rise.

› An expenditure switch occurs when private expenditures alter or switch among sectors (domestic versus foreign) due to a change in the relative cost of goods. A devaluation, for instance, raises foreign goods prices and induces a switch from imported to domestically produced goods.

This conclusion was contested by Laursen and Metzler[10] who argued that the devaluation by raising the prices of foreign goods relative to domestic goods would lead to substitution towards domestic goods. The composition of expenditure, of course, changes as the relatively more expensive imports are substituted for home goods to some extent, as Alexander pointed out. However, Laursen and Metzler argued that apart from a change in the composition,

› The terms of trade is a relative price that is given by the price of exports ($P$) divided by the price of imports ($EP^*$).

a change will occur in the total expenditure schedule of the economy as well. With given domestic prices, a rise in import prices corresponds to a decline in the country's TERMS OF TRADE. Following a devaluation, any level of income given in domestic terms translates into less income when measured in terms of foreign goods. Consumption expenditure is on a basket of domestic and foreign goods, and the income is now less when measured in terms of foreign goods. In the face of worsened terms of trade, individuals reduce their savings to maintain their living standards. This results in an increase in expenditure that has a negative effect on the trade balance as some of the increase in expenditure is on imports. The elasticities approach stresses the price effect of the devaluation, whereas this approach gives significance to the income effect of the devaluation.

This implies that as import prices rise and the real income corresponding to a given money income decreases, the amount spent on goods and services out of a given money income will rise. The expenditure schedule of a country rises when import prices rise, and it falls when import prices fall. A rise in import prices relative to domestic prices is equivalent in its overall effects to an overall decline in the productivity of the economy. So "just as a decline in physical productivity normally leads to an increase in employment, if not of real income, so also, we believe, will an increase of import prices".[11] Thus, if the total expenditure is highly sensitive to movements in the relative price of foreign and domestic goods, the level of expenditure by residents ($A$) will tend to rise along with a rise in import prices induced by a devaluation. In such a situation, absorption increases unlike what was argued by Alexander. However, if the expenditure depends only to an insignificant degree on the terms of trade, the level of output of the country will be largely independent of the devaluations and revaluations.

### *5.2.1* Devaluation in Developing Countries

Until the late 1970s, it was widely believed that devaluation would improve the trade balance and expand output and employment. An alternative approach

MACROFOCUS 5.1

**Does Devaluation Raise the Price of Imported Goods?**

The dollar depreciated by about 15 per cent against an import-weighted index of currencies from 2002 to 2005. The decline of the dollar should have helped correct the US trade deficit by altering the relative price of US and foreign goods. As the cost of foreign goods rose, the United States demand for imports ought to have weakened. Similarly, US exports ought to have increased as those goods became less costly abroad. Yet, the US trade imbalance has not shrunk in response to the depreciation of the currency.

One explanation for this puzzle is that pass-through—the degree to which a change in the value of a country's currency induces a change in the price of the country's imports and exports—can vary across industries because foreign firms may have different levels of market power relative to domestic firms.* For instance, if a foreign firm exports goods in an industry in which domestic firms have considerable market power, it may be reluctant to raise prices when the domestic currency depreciates. The foreign firm will adjust its markup downward to maintain a market share, and exchange rate pass-through will be lower. If, on the other hand, a foreign firm exports goods in an industry in which domestic firms have less market power, it may have less incentive to keep prices low.

Many other countries also experienced episodes of large depreciations of real exchange rate in the 1990s, which did not result in significant increases in the prices of imported goods. Renowned examples of this include Sweden and the UK after 1992 and Brazil after 1999. Another reason for this phenomenon is the increased emphasis on inflation stabilization by central banks.** When a central bank acts aggressively to stabilize inflation, it tightens policy to offset any inflationary impetus from a rise in import prices. Agents anticipating central bank's actions are less likely to pass through cost increases arising from exchange rate depreciations.

An important implication of import prices becoming less responsive to changes in currency values is that larger devaluations of the currency would be required to narrow trade imbalances. Also, a depreciation of the currency would not result in rising import prices and to that extent would not have an inflationary impact on the economy.

*K. Froot and P. Klemperer, "Exchange Rate Pass-Through When Market Share Matters," *American Economic Review* 79 (September 1989): 637–654.

**J. E. Gagnon and J. Ihrig, "Monetary Policy and Exchange Rate Pass-Through," *International Journal of Finance and Economics* 9, 4 (October 2004): 315–338.

that emerged at the end of the 1970s raised the possibility that devaluation would be CONTRACTIONARY, especially in developing countries. This approach is sometimes referred to as *structuralist* as it considers the economic problem of developing countries to be structural.[12] One of the channels identified by Krugman and Taylor[13] through which a devaluation might cause a reduction in the national output is the income-distribution channel. This income-distribution effect arises because there are two different classes of consumers in society: wage earners, and profit earners that includes rent earners. The marginal propensity to consume for wage earners is higher than that for profit earners as the income level of workers is lower and saving possibilities are limited for them. A devaluation[14] increases the decline in price of traded goods and real wages ($w/P$) since nominal wages are the same but prices are now higher. While the workers' share in national income falls, profit earners' share increases. As the marginal propensity to consume for wage earners is higher than profit earners, there is an eventual decline in aggregate demand.

› A devaluation can be **contractionary** in a developing country if it decreases the share of workers in national income or the increase in the import price of crucial inputs causes a contraction in supply.

Yet another channel through which devaluation works is the imported input cost channel. In many developing countries, the production process is highly dependent on imported inputs in the form of raw materials, or intermediate or capital goods. Devaluation via imported inputs increases the cost of imports in particular, and the cost of domestic production in general. Decreasing imports results in insufficient inputs that are necessary for production. Eventually, due to the lack of inputs and increasing costs, production slows down, leading to a contraction in total supply. This occurs especially in oil-importing, developing countries. A rise in the price of oil increases the cost of production. As the cost of production increases relative to the price of the product, firms tend to produce less, which leads to a reduction in aggregate supply and an increase in unemployment. Devaluation creates the same negative effect as do oil-price shocks on output and employment through increased production costs in developing countries.

### *5.2.2* Evidence on Devaluation

In India the evidence on devaluation is that the effect of changes in the exchange rate on the balance of trade does not yield a J-curve.[15] Singh used quarterly data from 1975 to 1996 for the exchange rate in terms of the trade-weighted effective real exchange rate and found that the first and sixth quarter lags of this variable affect the trade balance with a negative coefficient, while the other lags (up to the eighth lag) are positive. However, only the coefficient of the first lag is statistically significant, suggesting no J-curve effect. While the real exchange rate and domestic income significantly explained the behaviour of the balance of trade in India, Singh found that the world income had no effect on the trade balance. As the lack of a significant relation between the trade balance and the real exchange rate could be due to the use of aggregate trade data, it is advisable to study the short-run and long-run effects of the real depreciation of the rupee on the Indian trade balance using bilateral trade and real exchange rate data. Arora, Bahmani-Oskooee, and Goswami[16] do such a study for the seven largest trading partners of India, who together count for more than 50 per cent of Indian trade, during the quarters of 1977 to 1998. Again, there is no J-curve effect, but long-run real depreciation of the rupee against the currencies of Australia, Germany, Italy, and Japan has a positive impact on India's trade balance with each country. However, the UK, the United States, and France do not show any significant impact of the bilateral real exchange rate on the bilateral trade balance.

# 5.3 Exchange Rates and Assets

As the nominal exchange rate is the price of one country's money in terms of another's, it is also an asset price. The exchange rate, therefore, affects not only the flow of goods and services on the trade account but also the purchase and sale of assets such as bonds, equities, and real estate on the capital account. We should thus expect that the principles governing the behaviour of other asset prices also govern the behaviour of the exchange rate. If foreign currency assets are taken to be a foreign bond or a foreign currency deposit, what are the determinants of the demand for these assets? The demand for a foreign currency deposit will be influenced by the same considerations that apply to the demand for other assets. The expected returns on holding the asset influence the demand for an asset. This is associated with the asset's future value, which in the case of a foreign currency deposit depends on two factors: the interest rate offered abroad and the expected change in the country's exchange rate against other currencies.

Of course, assets are valued by savers for attributes other than the expected rate of return. Savers care about the risk or the variability that the asset contributes to a savers' wealth and the liquidity or the ease with which the asset can be sold. We will assume here that the foreign and domestic financial assets are perfect substitutes so that they have equivalent risk and liquidity attributes. Also, we assume that international capital markets are free from capital controls, and there are no restrictions on the purchase and sale of foreign currency denominated assets. These assumptions allow us to unravel, with reference to asset demand, the determination of the exchange rate.

### 5.3.1 Uncovered and Covered Interest Parity

We now consider the demand for an asset in terms of the return on the asset as compared to competing assets. Under these conditions, consider an agent who must choose between investing an amount INR $X$ domestically or abroad. If they invest domestically, they will receive over a given period (such as a year) a return on the domestic financial instrument $E_D$ measured in the domestic currency, where

$$E_D = (1 + i)X \tag{5.5}$$

and $i$ is the domestic interest rate on domestic deposits.

› When two parties agree to an exchange of financial instruments and execute the deal immediately, the exchange rates governing such "on-the-spot" trading are called spot exchange rates.

If, on the other hand, the money is invested abroad, the amount INR $X$ must first be converted into foreign currency units at the SPOT EXCHANGE RATE, $E$. With INR $E$ equal to \$1, INR $X$ will give \$$(X/E)$. The earnings in foreign currency on this amount will be \$$(1 + i^*)(X/E)$, where $i^*$ is the interest rate on the foreign currency deposit.

This earning must be converted back into domestic currency at the end of the investment period. However, when the investment decision is taken, it is not known what the exchange rate will be at the end of the investment period, and so an expectation is formed concerning the future exchange rate, $E^e$. The expected domestic currency return on the foreign financial instrument will then be

$$E_F^e = (1 + i^*)\frac{X}{E}E^e \tag{5.6}$$

The domestic financial instrument will be preferred as long as the return obtained on it exceeds the return on the foreign financial instrument.

$$E_D > E_F^e,$$

or
$$(1+i)X > (1+i^*)\frac{X}{E}E^e$$

or
$$\frac{1+i}{1+i^*} > \frac{E^e}{E},$$

or
$$\frac{1+i}{1+i^*} - 1 > \frac{E^e}{E} - 1,$$

or
$$\frac{i-i^*}{1+i^*} > \frac{E^e - E}{E},$$

Multiplying throughout by $(1 + i^*)$, we get

$$i - i^* > \frac{E^e - E}{E} + i^*\left(\frac{E^e - E}{E}\right)$$

The product $i^*[(E^e - E)/E]$ is usually a small number and is thus ignored. Hence, the savers will continue to prefer the domestic financial instrument as long as

$$i - i^* > \frac{E^e - E}{E} \tag{5.7}$$

or as long as the interest rate differential exceeds the expected rate of depreciation regarding the domestic currency. The inequality Eq. (5.7) will eventually tend towards an equalization of the domestic and foreign rates of return, $E_D = E_F^e$, through the operation of the following three mechanisms:

(1) A capital inflow as savers abroad seek higher returns domestically, which creates an additional supply of capital and cause a fall in domestic interest rates;
(2) A tightening of the supply of capital abroad due to the savers there preferring a higher return on financial instruments outside their country, which leads to a rise in foreign interest rates;
(3) The capital inflow increasing the supply of foreign currency, which causes the spot exchange rate to appreciate (the $S_X + M_K$ curve in Figure 2.2, Chapter 2, shifts to the right, causing the exchange rate to appreciate) and leads to an increase in depreciation expectations, or an increase in $(E^e - E)/E$.

Hence, equalization of the return on financial instruments domestically and abroad results in

$$i = i^* + \frac{E^e - E}{E} \tag{5.8}$$

This is referred to as the UNCOVERED INTEREST PARITY condition; "uncovered" because the investment involves an exchange risk, which could be eliminated by selling forward the foreign currency amount as soon as the funds are invested abroad. If agents hedged the exchange rate risk via a FORWARD TRANSACTION, they could earn a return $E_F^f$ by selling forward the foreign currency amount as soon as the funds are invested abroad, where $E_F^f = (1 + i^*)(X/E)F$.

Then, the equilibrium condition is as follows:

$$i = i^* + \frac{F - E}{E} \tag{5.9}$$

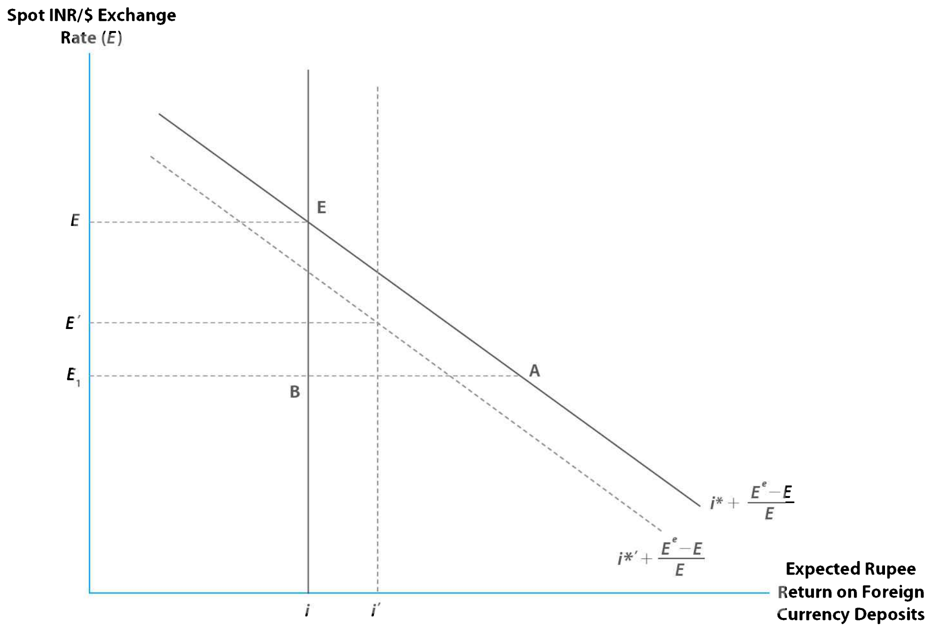
which is referred to as the COVERED INTEREST PARITY condition. The interest parity conditions are an example of the LAW OF ONE PRICE for asset markets. The reason we expect the law to prevail is arbitrage—the process of buying or selling an asset in order to exploit a price differential and to make a riskless profit.

› Uncovered interest parity relates interest differentials of two countries $i - i^*$ to the expected change in the spot exchange rate between those countries.

› The exchange rate quoted for a date on which the parties actually receive the funds they have purchased is a forward exchange rate. Forward exchange rates are typically for 30 days, 90 days, 180 days, and 365 days. The forward exchange rate $F$ in the text is a one-year forward rate.

› Covered interest parity relates the difference between the interest rate domestically and abroad to the forward premium or discount.

› The Law of One Price applied to assets states that if the international arbitrage is unhindered (the market is efficient) then identical assets (with the same risk and liquidity characteristics) must earn the same rate of return.




› **Figure5. 5**
Equilibrium Exchange Rate. The demand for foreign currency deposits increases with an appreciation of the currency. With a domestic interest rate *i*, interest parity determines the exchange rate at *E*. If the exchange rate is instead $E_1$, individuals will purchase the higher return foreign currency deposits. The increased supply of capital abroad decreases the foreign interest rate to *i*′* and the reduced holding of domestic deposits raises the domestic interest rate to *i′*. The exchange rate then settles at *E′*.

The interest parity conditions define the foreign-exchange market to be in equilibrium when the deposits of all currencies offer the same expected rate of return, when measured in the same currency.

### *5.3.2* Exchange Rate Determination

To understand how the foreign-exchange market finds its equilibrium, consider how the current exchange rate affects the expected return on a foreign currency deposit. From Eq. (5.6), a depreciation of the country's currency today (an increase in $E$) reduces the expected domestic currency return on foreign currency deposits, $E^e_F$. Similarly, appreciation of the domestic currency today (a decrease in $E$) raises the domestic currency return expected of foreign currency deposits. For a fixed value of the expected future rupee/dollar exchange rate, $E^e$, and the foreign currency interest rate, $i*$, the expected return on foreign currency deposits is, therefore, a downward-sloping schedule as in Figure 5.5. The reason for this is that a current depreciation of the domestic currency that affects neither the exchange-rate expectations ($E^e$) nor the interest rates ($i$, $i*$) leaves the expected future pay-off of a foreign-currency deposit in terms of rupee as the same, but raises the foreign currency deposit's current rupee cost.[17]

The vertical schedule in Figure 5.5 represents the given level of the return on domestic currency deposits $i$. The intersection of the two schedules at point $E$ gives the equilibrium exchange rate, $E$. Any exchange rate not equal to $E$ will tend to gravitate towards $E$. To know the reason, consider an appreciated exchange rate $E_1$. At this exchange rate, the rate of return on foreign currency deposits (distance $E_1$A) is greater than the rate of return on domestic deposits (distance $E_1$B). Individuals would prefer selling their domestic deposits for more lucrative foreign currency deposits. As domestic residents purchase foreign exchange while seeking higher returns on deposits abroad, they bid up the exchange rate—the curve $D_M + X_K$ shifts to the right in Figure 2.2 of Chapter 2—till the return on domestic and foreign currency deposits are equal at point E and there is no longer an incentive to get out of domestic deposits.

Of course, as discussed earlier, as savers withdraw their holdings of domestic deposits, there is a tightening of the supply of capital which raises the domestic interest rate—the point $i$ shifts to $i'$, where, $i' > i$. Similarly, the capital inflow abroad creates an additional supply of capital that results in a decline in the foreign interest rate—$i^*$ declines to $i^{*\prime} < i^*$. Hence, the final equilibrium exchange rate[18] will be at $E'$ where the dashed lines intersect.

As long as domestic and foreign currency assets are perfect substitutes, the foreign-exchange market is in equilibrium when the interest parity condition given by Eq. (5.8) holds. When imperfect asset substitutability exists, it is possible for the expected returns of the assets to differ in equilibrium. The main factors behind *imperfect asset substitutability* are risk (when the return on an asset fluctuates widely) and liquidity (the high cost and speed at which savers can dispose of an asset). If assets denominated in different currencies have different degrees of risk, investors who are risk averse may be willing to earn a lower expected return on assets that are less risky. This means that they will hold a risky asset only if the expected returns it offers is relatively high. Equilibrium in the foreign-exchange market now requires that the domestic interest rate equals the expected domestic currency returns on the foreign asset, plus a risk premium ($\rho$) that reflects the difference between the riskiness of domestic and foreign currency assets:

$$i = i^* + \frac{E^e - E}{E} + \rho \qquad \textbf{(5.10)}$$

An empirical test procedure of this approach to exchange-rate determination is obtained by combining the uncovered interest parity condition, Eq. (5.8), and the covered interest parity condition, Eq. (5.9). Then,

$$i^* + \frac{E^e - E}{E} = i^* + \frac{F - E}{E},$$

or

$$E^e - E = F - E,$$

or

$$\text{Expectation at time } t \text{ of } E_{t+1} = F_t \qquad \textbf{(5.11)}$$

The forward rate $F_t$ at time $t$ is the market's rational expectation at time $t$ of the future spot rate, $E_{t+1}$. This relationship is tested empirically by estimating an equation of the following type:

$$E_{t+1} = \alpha + \beta F_t + \epsilon_{t+1} \qquad \textbf{(5.12)}$$

where $\epsilon_{t+1}$ is a random error term. If $\alpha$ is close to 0 and $\beta$ is close to 1, the interest parity approach to exchange rates is valid. The estimates for most currencies lead to results that tend to support this approach.[19] Bhoi and Dhall[20] examine the data for India and find that covered interest parity holds for the Indian economy in the post-reform period.

# 5.4 Purchasing Power Parity

Interest rate parity is the application of the law of one price for assets. Applying this law to goods results in PURCHASING POWER PARITY (PPP). It is an old theory of exchange-rate determination.[21] It is based on the law of one price, which states that apart from tariffs and transaction costs (inclusive of transportation costs), international commodity arbitrage will equalize the prices for tradable goods in all locations when quoted in the same currency. The principle here is that when trade is open and costless, identical goods should trade at the same relative prices regardless of where they are sold. If a good $i$ sells in a market abroad at a foreign currency price of US\$ $P_i^*$ and the same good fetches a price of INR $P_i$ in the domestic market, ignoring tariffs and

› Purchasing power parity states that if arbitrage across national borders is unhindered, the price of a good or service in one nation should be the same as the exchange rate adjusted price of the same good in another nation.

transactions costs, the domestic price should be the foreign price times the exchange rate, or

$$P_i = EP_i^* \tag{5.13}$$

If $P_i > EP_i^*$, then an arbitrageur would have an incentive to buy the good abroad where it is cheap and ship it into the domestic market where it is dear. In the process, the arbitrageur will push up the price abroad and push down the domestic price until prices are equal in the two locations. If we assume that the domestic price index $P = f(P_{1 \ldots}, P_i, \ldots, P_n)$ and the foreign-price index $P^* = g(P_1^* \ldots, P_i^*, \ldots, P_n^*)$ are made up of the same goods with the same weights—the reference commodity baskets are identical in the two countries[22]—the law of one price applied to individual goods becomes a law of price levels, which is the ABSOLUTE VERSION OF THE PPP:

› Absolute PPP states that the domestic price level should equal the foreign-price level times the spot exchange rate.

$$P = EP^* \tag{5.14}$$

The left-hand side is the rupee price of the reference commodity basket domestically, and the right-hand side is the rupee price of the reference basket when purchased abroad. These two prices are the same when PPP holds. PPP is thus the proposition that the price levels of all countries are equal when measured in terms of the same currency.

$$E = \frac{P}{P^*} = \frac{\text{Domestic price of the reference basket of goods}}{\text{Foreign price of the same reference basket}} \tag{5.15}$$

Though the law of one price applies to individual commodities, the PPP applies to the general price level, which is a composite of the prices of all commodities contained in the reference basket. The PPP holds as long as the reference basket used to calculate different countries' price levels is the same.[23] When goods become more expensive in one country than in others, the demand for its currency and products declines, and this depreciates its currency and pushes domestic price down in line with PPP.

Absolute PPP makes no sense if the reference baskets whose prices are compared in Eq. (5.14) are not the same. Price indices differ across countries, and they include many goods and services that are not tradable. In addition, even for perfectly homogeneous commodities, there are transactions costs, tariffs, and indirect taxes. It thus makes sense to translate absolute PPP from a statement about price and exchange rate levels into one about price and exchange rate changes. This formulation is referred to as RELATIVE PURCHASING POWER PARITY.[24] If we write the absolute version of PPP, $P = EP^*$, in percentage form, then

› RelativeP PP states that the change in an exchange rate is equal to the difference in the inflation rates between two economies.

$$\frac{\Delta E}{E} = \frac{\Delta P}{P} - \frac{\Delta P^*}{P^*},$$

or

$$\frac{E_t - E_{t-1}}{E_{t-1}} = \frac{\Delta E}{E} = \pi_t - \pi_t^* \tag{5.16}$$

where the inflation rate in a country, defined as the percentage change in the price level between dates $t$ and $(t - 1)$, is

$$\pi_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P}{P}$$

When countries calculate the absolute price level estimates using reference baskets of commodities that differ in coverage and composition, and when factors causing deviations from absolute PPP are by and large stable over time, the percentage changes in relative price levels (or differences in inflation rates) can still approximate the percentage changes in exchange rates.

In monetary policy, the relative version of PPP plays an important role. Many countries adopt a fixed exchange rate vis-à-vis the currency of a country with low inflation so that they can import that country's price stability. When the exchange rate is fixed, $\Delta E/E = 0$, and $\pi_t = \pi_t^*$ so that the country's inflation rate is aligned to the inflation rate of the country to which it pegs its exchange rate.

Alternatively, proponents of a system of flexible exchange rates argue that countries can select their preferred national inflation rate $\bar{\pi}_t$ and that the exchange rate would compensate passively for the differential in inflation, given that $\Delta E/E = \bar{\pi}_t - \pi_t^*$.

The empirical backing for the PPP theory is weak, both for its absolute and relative versions.[25] There are, of course, the transaction costs and restrictions to trade factors. Then, the price indices are based on different commodity baskets. The price indices also include services and non-tradable goods for which the law of one price does not hold. Many goods are not traded internationally because of prohibitive transportation costs. For example, if India were to export haircuts to the United States, it would require transporting an Indian barber to, say, New York every time a New York resident desired a haircut. As transport costs prevent haircuts from being traded internationally, the law of one price does not apply to this service.

Finally, over the last decade, capital flows have become so large that rather than passively accommodating inflation differentials, exchange rates have become an autonomous source for deviations from PPP. The principal weakness of absolute and relative PPP is that they focus solely on the international exchange of goods and services and do not consider the effect of financial flows and money stock on the exchange rate.

Though relative PPP does not hold in general, it is used often to forecast what causes interest rates to differ. From the definition of the real exchange rate, we can write the following identity:[26]

$$\frac{\Delta(RER)}{RER} = \frac{\Delta E}{E} + \pi^* - \pi$$

MACROFOCUS 5.2

## Big Macs, iPods, and Professors' Salaries

Since 1986, the *Economist* magazine has published the Big Mac index which is based on the theory of PPP. This theory states that exchange rates should move towards levels that equalize the price of an identical basket of goods and services in any two countries. The Big Mac PPP rate then is the exchange rate that would leave burgers costing the same in America as elsewhere. In mid-2006, the Big Mac cost 10.5 yuan in China against an average price of $3.10 in the United States. The two prices would be equal at an exchange rate of (10.5/3.10) or 3.39 yuan to a US dollar. The market exchange rate, however, at that time was 8.03 yuan per dollar. When the market-determined value of the yuan is higher than that implied by PPP, it indicates that the local currency is "undervalued" against the dollar. We would then expect the yuan to appreciate relative to the dollar. Similarly, if the local currency exchange rate implied by PPP is lower than the market exchange rate, the local currency is "overvalued" relative to the dollar, and we would expect it to depreciate.

A limitation of the Big Mac index is that burgers cannot easily be traded across borders and prices are distorted due to differences in local taxes, labour laws, and the cost of non-tradable inputs such as rents. In response to these limitations, economists at Commonwealth Securities (CommSec), a brokerage firm in Australia, have developed an iPod index. The key difference between the Big Mac and the iPod index is that Big Macs are made in a host of countries whereas iPods are predominantly made in China. iPods can also be purchased over the Internet so that if there were substantial price differences across locations, customers could switch their purchases to other countries. The CommSec iPod (for a 2 GB nano iPod) indeed reveals that it costs $149 to purchase an iPod in the United States compared to $179.84 in China. Unlike the Big Mac index which would cause us to expect the yuan to appreciate, the iPod index indicates that the yuan should depreciate against the dollar. Again, if Apple sells iPods at different prices in different parts of the world, for example, it gives volume discounts to retailers in the United States, or if freight costs vary between countries, the PPP values of this index are distorted.

In recent times, professors have asked for higher salaries as it is not just higher commercial salaries but also higher salaries overseas that are a barrier to attracting talent into academics. To test whether this is true, Ong and Mitchell* argue that if a Big Mac costs US$2.42 in the United States and A$2.50 in Australia, then an academic would require a salary of A$2.50 in Australia for every US$2.42 earned in the United States in order to have a similar lifestyle. In that case, they find that real academic salaries are highest in Hong Kong and Singapore, followed by Australia and the United States. Canada and New Zealand are unattractive destinations by this measure for academics.

*L. L. Ong and J. D. Mitchell, "Professors and Hamburgers: An International Comparison of Real Academic Salaries," *Applied Economics* 32 (2000): 869–876.

The relationship between the expected change in the real exchange rate, the expected change in the nominal exchange rate, and the expected inflation can then be written as

$$\frac{\Delta(RER^e)}{RER} = \frac{\Delta E^e}{E} + \pi^{*e} - \pi^e$$

The interest parity condition Eq. (5.8) states that

$$i - i^* = \frac{\Delta E^e}{E}$$

Substituting this into the previous identity we obtain

$$\frac{\Delta(RER^e)}{RER} = i - i^* + \pi^{*e} - \pi^e,$$

or

$$i - i^* = \frac{\Delta(RER^e)}{RER} + (\pi^e - \pi^{*e}) \quad \textbf{(5.17)}$$

Hence, the domestic–foreign interest rate differential comprises two components: (1) the expected rate of real rupee depreciation and (2) the expected difference between inflation in India and abroad.

## 5.5 Choice of Exchange-Rate Regimes

We have discussed the role that exchange rates play in affecting the trade balance and competitiveness of the asset markets. Exchange rates are also an area of concern because of their effect on macroeconomic management, particularly inflation, and due to the link between the exchange-rate regime and vulnerability to currency crises. In fact, as Fischer[27] points out, each of the major crises in the 1990s—Mexico in 1994; Thailand, Indonesia, and Korea in 1997; Russia and Brazil in 1998; and Argentina and Turkey in 2000—has involved a fixed exchange-rate system. The advantage of a fixed exchange rate is that countries with a high propensity to inflation are able to import low inflation[28] (see also the discussion in Section 5.4). Fixed exchange rates also reduce devaluation risk and thus lead to lower interest rates that stimulate investment. The fix in the exchange rate is maintained by means of changes in the country's foreign-exchange reserves—buying and selling foreign currency in order to balance the supply and demand in the currency market. This is done to preserve the price of domestic. This is done currency compared with the foreign currency.

To do this successfully, a country generally requires foreign currency reserves equivalent to three to four months' worth of imports. When economic agents anticipate that reserves are inadequate and that the monetary authority will resort to devaluation, agents may launch speculative attacks based on a one-way bet that the peg will be abandoned leading to a currency crisis.

Due to the vulnerability of fixed exchange rates to speculative attacks, floating exchange rates are often advocated as an appropriate regime. However, floating exchange-rate regimes, where the authorities do not intervene in the foreign-exchange market and do not set interest rates for the purpose of effecting the level or path of the exchange rate, do not insulate an economy from the adverse effects of international capital flows. Short-term capital flows seeking ARBITRAGE opportunities and quick return in anticipation of changes in the exchange rate under a currency float can affect the payment system especially when a country's financial sector is poorly regulated or lacks DEPTH. Volatility in the exchange rate increases the currency risk which, in turn, increases the differential in interest rates domestically and abroad. The corresponding higher domestic interest

› Arbitrage means "buy low – sell high" and indicates an activity through which individuals seek immediate profits based on price differentials. If exchange rate differences prevail in different markets forex traders purchase cheap and sell where the price is high to make a profit.

› A financial market is considered to have depth when participants can rapidly execute large volume transactions. The depth is the maximum amount or limit that a dealer stands ready to buy or sell at posted prices.

rate attracts hot money and also increases the likelihood of ORIGINAL SIN. When a national currency cannot be used by local firms or the government to borrow abroad, it suffers from original sin.[29] In such a situation, investors who find the country economically promising must borrow in a foreign currency such as the dollar or borrow short term because the domestic capital market lacks sufficient depth. When a firm borrows in dollars to finance a project that generates INR, a subsequent depreciation can lead to bankruptcy. Investments suffer from a currency mismatch, and when firms see the domestic currency begin to depreciate, they may attempt to buy foreign currency to cover their exposure. This increase in demand for foreign currency causes the domestic currency to depreciate even further, which increases the volatility of exchange rates.

› When a country must borrow abroad in a foreign currency, it suffers from original sin. In such a situation a depreciation of the currency could cause agents to purchase foreign currency to cover their repayments which, in turn, leads to further depreciation and volatility of the exchange rate.

Thus, the appropriate exchange-rate regime for a country depends on its exposure to international capital and the level to which its domestic financial sector has matured. Countries that are heavily exposed to global capital flows and which have a mature and well-regulated financial sector would find floating to be appropriate. Otherwise, countries would operate a hard peg in the form of a currency board or dollarization.[30,31] Under a CURRENCY BOARD the stock of domestic money in circulation is backed by foreign currency reserves, and the monetary authority commits to buying or selling the domestic currency in unlimited quantities at some pre-announced exchange rate. The exchange rate is set by public law, and the commitment to convert currency is limited to the domestic monetary base or reserve money, $M_0$, not the domestic money supply. Recall from Chapter 2 that the money supply is the monetary base (minus currency with the banks and deposits held by them at the central bank) plus demand deposits. Hence,

› Undera currency board the exchange rate between the local currency and, say, the dollar is fixed by law, and the monetary authority commits to buying or selling the domestic currency without any limit at this pre-announced exchange rate.

$$M_0 = \lambda E Q_F$$

where $\lambda$ is the degree to which domestic currency is backed by foreign currency,
$E$ is the pre-announced exchange rate between the domestic currency and the foreign currency to which it is linked, and
$Q_F$ is the monetary authority's holdings of foreign currency.

Occasionally, currency boards have less than 100 per cent backing unlike the textbook story where the domestic base is fully backed by foreign reserves and $\lambda = 1$. The Convertibility Law governing Argentina's currency board, for instance, only required 80 per cent backing of the base.[32] A number of East European countries have adopted currency boards by linking their currencies to the Deutsche mark/Euro—Estonia, Bulgaria, Lithuania, and Bosnia-Herzegovina, to name a few. With the local currency completely equivalent to the foreign anchor currency at the fixed exchange rate, the currency board has no discretionary monetary powers and cannot extend credit if the domestic base is fully backed by foreign currency reserves.

Under DOLLARIZATION, or euroization, a country officially adopts a foreign currency as its legal tender. To dollarize, two exchanges must be accomplished. First, the monetary base has to be exchanged for the adopted foreign currency at a predetermined conversion rate. Second, all contracts in the local currency must be transformed into contracts denominated in the adopted currency. Thus, local currency contracts would be re-denominated and payable in, say, US dollars. Unlike a currency board arrangement, the local currency is completely replaced by the adopted foreign currency rather than being limited in quantity by the number of units of anchor currency held by the currency board. Again there is no scope of independent monetary policy as the monetary policy would be determined by the US Federal Reserve Board, which is desirable if a country's monetary authorities are irresponsible. But it is undesirable if a country needs a more expansionary monetary policy. The dollarizing country

› Under dollarization, a country officially adopts a foreign currency as its legal tender. Dollarization removes a country's control over its monetary policy.

is also faced with the expense of buying dollars to swap with its national currency. It pays for these dollars either with its foreign reserves or with the money from a large dollar-denominated loan. In either case, the cost in terms of forgone interest payments on its reserves or new interest payments on its borrowings could be significant. Argentina, for example, would be required to spend $15 billion to swap its peso currency notes for US dollars.[33] Countries that are dollarized include Marshall Islands, Micronesia, Palau, and Panama, all small economies except Panama, where most goods are traded internationally and most prices are set in dollars.

› In a monetary union a group of countries share a common currency and jointly manage the currency through a common central bank.

Dollarization is an asymmetric MONETARY UNION as one country adopts another country's currency as its own, but does not have any say in the management of that currency. In a symmetric monetary union, a group of countries share a common currency and jointly manage the currency through a common central bank. Under monetary union there is a regionalization, as in the case of the euro, rather than a dollarization of national currencies. With the responsibility for the exchange-rate policy and the balance of payments with the rest of the world assigned to a common central bank, there is a gain in terms of the reduction in transaction costs of doing business within the region. Also, the monetary policy is removed from the domain of populist national politics with a supranational central bank run by designated representatives from each of the participating countries. There are also savings unlike under dollarization. The countries who are party to the monetary union could print the common currency at a low cost and swap it for their outstanding currencies in circulation.[34] If they were to dollarize, they would sell interest-earning dollar reserves or borrow new dollars at high interest rates to swap dollars for the existing currencies.

Finally, monetary union is beneficial when countries have approximately the same economic structure, are subject to the same international shocks, and trade a lot with one another. By adopting a common currency, they fix their exchange rates with one another but remain flexible in relation to other major currencies such as the US dollar. If a country that is part of a monetary union is subject to an economic shock, such as a natural disaster and abrupt shifts in capital flows, which is not common to the union, then as it is deprived of exchange rate flexibility, it needs to adjust according to wage and price flexibility, the mobility of labour and capital, and fiscal compensation. However, there are many political and cultural obstacles to the regionalization of national currencies. If that were not true, Argentina and Brazil would have established one by now.

In between free-floating and hard pegs (currency boards, dollarization, and monetary union) lie *soft pegs* and *managed floats*. Under a soft peg, the authorities adopt a particular target exchange rate and use monetary policy to prevent the market from straying too far from the target. "Too far" is operationalized by the announcement of a target rate and the width of a band around it.

Under an *adjustable peg*, the bands tend to be narrow—less than or equal to ±2.25 per cent—and the target rate is adjusted if the authorities perceive a discrepancy between the target exchange rate and the equilibrium rate. The target rate is adjusted infrequently and by large amounts.[35] The Bretton-Woods system was an adjustable peg system.

Under a *crawling peg*, the authorities target a path for the exchange rate instead of a level. The bands tend to be somewhat wider than the bands around adjustable pegs as they must accommodate the crawling of the pegs. The target rate is altered frequently and by relatively small amounts.[36] Under a *passive crawling peg*, the exchange rate parity is adjusted for past inflation differentials. An active crawling peg pre-announces the rate of the crawl with the aim

MACROFOCUS 5.3

## TheE uro

The European Monetary Union (EMU) is the latest step in a long movement towards integration of Western Europe that began during the mid-1980s, with member countries agreeing to the abolition of internal barriers to the free movement of goods, persons, services, and capital by 1992. In 1989, Jacques Delors, president of the European Union (EU) Commission (the executive body of the EU), proposed a plan for the creation of a single currency to be issued by a European Central Bank (ECB). The proposal was adopted in the 1991 Treaty on EU, often called the Maastricht Treaty. Under the terms of the treaty, countries wishing to join the EMU must meet certain convergence criteria. There are four primary criteria:

1. Prices: For admission to the EMU, the consumer price inflation rate of a country must be no more than 1.5 per cent above the average of those of the three EU member states with the lowest inflation rates.
2. Deficits: Actual and planned budget deficits should not exceed 3 per cent of the GDP, and
   the government debt should not exceed 60 per cent of the GDP. The budget should also be balanced in the medium run in the sense that deficits should be counterbalanced by surpluses in subsequent periods.
3. Interest rates: The long-term interest rate must, on average, not be more than 2 per cent above the average of the three EU members with the lowest inflation rate.
4. Exchange rates: The country should have observed normal fluctuations in the exchange rate, without devaluing on its own initiative against the currency of another EU member.

The motivation for these criteria was to ensure that the single currency will be a hard currency. When an individual country runs into financial problems, it faces the consequence in terms of high interest rates, rising inflation, or even a run on its currency. When countries share a currency, risks are spread, and this can cause a country to take less care in tackling economic problems, such as a fiscal problem, as it escapes the full consequences of lack of action on its part. A profligate government with a loose fiscal policy also pushes up borrowing costs for all countries. The convergence criteria are meant to ensure that countries are sincere and committed to stable policies and are not free riders.

The euro was launched on 1 January 1999 with 11 member countries fulfilling the criteria—Belgium, Germany, Ireland, Spain, France, Italy, Luxembourg, the Netherlands, Austria, Portugal, and Finland. Greece joined a couple of years later, and Slovenia became the 13th member of the euro area in January 2007.

On 1 January 1999, the national central banks of these countries became members of the ECB, and they operate in line with the ECB's guidelines and instructions. They carry out monetary policy operations, such as providing the central bank's money to credit institutions and foreign reserve management operations, as the agents for the ECB.

On 1 January 1999, the nominal exchange rates at which member countries' currencies were to be transferred into euro were fixed. The euro began as an electronic currency in which consumers held bank accounts and credit cards. National currencies were different expressions of what is economically the same currency, and the ECB was obliged to exchange national banknotes against the euro at par. Cash appeared three years later on 1 January 2002. Banks and their ATMs after this date were to dispense only euros, and retailers were to give change only in the single currency. After that, within two months, national notes and coins were to be withdrawn and ceased to be legal tender. Banknotes prior to the euro were adorned with images of Antoine de Saint-Exupéry (France), Vasco da Gama (Portugal), and René Margritte (Belgium). The euro depicts the architectural style of seven periods in Europe's cultural history—Classical, Romanesque, Gothic, Renaissance, Baroque, and Rococo—the age of iron and glass architecture, and modern twentieth-century architecture. They also show three main architectural elements: windows, gateways, and bridges.

The nominal exchange rate of national currencies and the last date of their existence as legal tender are given in Table 1.

The formation of the EMU is the first time that countries of this number, size, or global economic weight have collaborated to share a currency without a single government behind it. It is the first time since the fall of the Roman Empire that much of Western Europe has a single currency. Some have even called it the most momentous currency innovation since the establishment of the US dollar in 1792.

Even though the nominal exchange rates between the members of the euro area have been permanently fixed, their real exchange rates vary depending on differences in their rates of inflation. Between 1999 and 2006, for instance, Italy's real exchange rate has appreciated by more than 20 per cent relative to Germany's. From early 2002, the euro appreciated against the dollar and that was believed to be bad news for the competitiveness of exports from Europe. However, German firms in that period improved productivity and clamped down on wage costs, and this resulted in low inflation and a depreciated real exchange rate that allowed German exports to hold up strongly in world markets.

| Country | National Currency Units per Euro | End of Legal Tender for National Currency in 2002 |
|---|---|---|
| Austria | 13.7603 | February 28 |
| Belgium | 40.3399 | February 28 |
| Finland | 5.94573 | February 28 |
| France | 6.55957 | February 17 |
| Germany | 1.95583 | December 31, 2001* |
| Greece | 340.750 | February 28 |
| Ireland | 0.787564 | February 9 |
| Italy | 1,936.27 | February 28 |
| Luxembourg | 40.3390 | February 28 |
| The Netherlands | 2.20371 | February 28 |
| Portugal | 200.482 | February 28 |
| Spain | 166.386 | February 28 |

* Retailers will accept D-marks till 28 February 2002.

*Source*: The ECB Web site, http://www.ecb.int/home/html/index.en.html of influencing price expectations. Hungary, Israel, and Sri Lanka have crawling band exchange-rate regimes. In a *managed float*, there is no public commitment to a target exchange rate, but the authorities may intervene in the foreign-exchange market and/or set interest rates to influence the exchange rate. The authorities aim may be to smooth short-term volatility or to reverse sustained departure of the exchange rate from what they perceive as its equilibrium value, referred to as misalignment. The system operates in Britain, Japan, the Czech Republic, and Taiwan. When Japan intervened in the currency market in the 1980s to keep down the value of the yen, thereby boosting Japanese exports and restricting its imports, it was described as a dirty float.

Many countries attempt to run a monetary policy aimed at domestic economic conditions while intervening in the foreign-exchange markets to influence exchange rates. Often these goals are mutually incompatible, especially when a country has an open-capital account and is a destination for capital flows. For instance, a country could attempt to correct a recession by an expansionary monetary policy that reduces domestic interest rates relative to international interest rates. With a pegged exchange rate, the lower interest rate reduces the relative return to foreign investors who will pull out their assets. The corresponding outflow of foreign currency will put pressure on the domestic currency and will require the use of the country's foreign-exchange reserves to defend the fixed exchange rate. Ultimately, as reserves get depleted, the country will have to choose between abandoning the fixed exchange rate and its monetary policy by raising its domestic interest rate. This dilemma is referred to as the IMPOSSIBLE TRINITY—it is not possible for a country to have control over a fixed exchange rate, an independent monetary policy (monetary autonomy), and a fully open-capital account simultaneously. Thus, the necessary condition for a government to have both an exchange-rate policy and an independent monetary policy in the short run is that it is imperfectly integrated with world capital markets.

› The impossible trinity refers to the dilemma that a country cannot simultaneously have control over a fixed exchange rate and a monetary policy, and maintain an open-capital account.

The debate over fixed versus floating exchange rates is still a wide open one. The IMF has also shifted its recommendation according to circumstances. After the 1997 Asian Crisis, the IMF accused countries with soft pegs, not for playing a part in the Asian meltdown but rather for amplifying the cost of crisis. Pegged exchange rates encouraged growth of unhedged foreign currency debt and currency mismatch of balance sheets. This pushed up the costs of devaluation for borrowers, triggering claims of bankruptcy and bank failures. These events, along with huge losses incurred by the monetary authorities as they sought to defend their exchange rates from speculative attack, resulted in an even higher cost for the resolution of the crisis. This played a role in the ensuing IMF's doctrinal shift. From considering fixed exchange rates as a tool against inflation, the IMF turned to "corner" solutions based on hard pegs—dollarization or currency board or pure floats—in the late 1990s.[37] However, there was another shift of doctrine, after the Argentine crisis in 2001–2002. Since that time, the IMF has stopped recommending currency boards as a credible solution and has switched to its current doctrine of floating arrangements with inflation targeting.[38]

In the mid-1990s, Obstfeld and Rogoff[39] wrote an articles, *The Mirage of Fixed Exchange Rates*, warning against fixed regimes. They argued that such systems last for a couple of years on an average and are regularly followed by a collapse in the exchange rate and a currency crisis. A fixed exchange rate often causes the real exchange rate to become overvalued, which makes the regime vulnerable to a speculative attack. By the end of the decade, Reinhart[40] wrote an equally famous articles, *The Mirage of Floating Exchange Rates*, where she argued that floating rates are even more of a delusion than fixed ones for the simple reason that they do not exist. Looking at a large sample of countries, she demonstrated that no emerging country actually allows its exchange rate to float because the governments of those countries suffer from what Calvo and Reinhart[41] dubbed the FEAR OF FLOATING. Fear of floating exists when a country claims to be pursuing domestic policy goals, such as price stability and full employment, that are independent of the exchange rate and yet that country repeatedly intervenes, directly or indirectly, to keep the exchange rate in check. This is distinguished from simple dirty floating by the predominant use of domestic monetary policy instruments—mainly the domestic interest rate and reserves—to "indirectly" influence the exchange rate.[42] The fear of floating derives from the actual or perceived cost of exchange rate volatility. For instance, currency fluctuations may adversely affect competitiveness, cause a ratcheting of inflation (exchange-rate pass-through), and adversely affect balance sheets and debt-servicing burdens by raising the domestic currency value of foreign currency–denominated debt. Due to these costs, policymakers in emerging markets believe that the room to pursue an independent monetary policy and to be subject to exchange-rate flexibility is, in practice, limited at best. As Calvo and Reinhart note, when it comes to exchange-rate policy, discretion rules the day. However, it is worth noting that even though no exchange-rate system is best for all countries or for all times, no regime can act as a substitute for good policies or good institutions. The exchange-rate regime should be viewed as part of a coherent monetary order that is itself a part of a sound macroeconomic framework.

› A country exhibits fear of floating when it claims to let its exchange rate be determined in a float but it intervenes directly or indirectly to keep the exchange rate in check.

## SUMMARY

» A country's real exchange rate is the price of the average foreign good or service, relative to the price of the average domestic good or service, when this price is expressed in terms of a common currency.

» The real exchange rate—the nominal exchange rate adjusted for the relative price of foreign and domestic goods—is a measure of the competitiveness of a country's goods and services.

» A devaluation or depreciation of the real exchange rate improves a country's trade balance only when the Marshall–Lerner condition—the sum of the relative price elasticities of export and import demand exceeds unity—holds.

» In the short run, the sum of the relative price elasticity of exports and imports almost always fail to sum to unity, while in the long run these elasticities almost always sum to greater than unity. This gives rise to a J-curve effect.

» The J-curve effect arises due to adjustment lags as quantity adjustments take long and the faster pass-through of the change in exchange rate into prices worsens the trade balance in the short run.

» The absorption approach to the trade-balance determination emphasizes changes in domestic income and expenditure rather than the real exchange rate.

» A currency devaluation switches expenditures to the domestic market from imports and this expenditure could decline (Alexander) or rise (Laursen–Metzler), thereby causing the trade balance to improve or deteriorate.

» A devaluation can be contractionary in a developing country if it decreases the share of workers in the national income or the increase in the import price of crucial inputs causes a contraction in supply.

» When foreign and domestic financial assets are perfect substitutes and there are no capital controls, equalization of the return on financial instruments domestically and abroad results in two types of interest parity conditions:

(a) Uncovered interest parity relates the interest differentials of the two countries to the expected change in the spot exchange rate between those countries.

(b) Covered interest parity relates the difference between the interest rate domestically and abroad to the forward premium or discount.

» PPP states that if arbitrage across national borders is unhindered, the price of a good or service in one nation should be the same as the exchange rate-adjusted price of the same good in another nation.

» Absolute PPP states that the domestic-price level should equal the foreign-price level times the spot exchange rate.

» Relative PPP states that the change in an exchange rate is equal to the difference in inflation rates between two economies.

» Fixed exchange rates enable countries to import low inflation, but anticipation of devaluation due to dwindling reserves can lead to speculative attacks.

- » Floating exchange rates, when a country's financial sector is poorly regulated or lacks depth, can result in greater volatility of exchange rates due to "original sin".
- » When a country must borrow abroad in a foreign currency, it suffers from original sin. In such a situation, a depreciation of the currency could cause agents to purchase foreign currency to cover their repayments, which leads to further depreciation and volatility of the exchange rate.
- » Under a currency board, the exchange rate between the local currency and, say, the US dollar is fixed by law, and the monetary authority commits to buying or selling the domestic currency without any limit at this pre-announced exchange rate.
- » Under dollarization, or euroization, a country officially adopts a foreign currency as its legal tender. Dollarization removes monetary policy from a country's control.
- » In a monetary union, a group of countries share a common currency and jointly manage the currency through a common central bank.
- » Under a soft peg, the monetary authorities adopt a particular exchange rate target and use monetary policy to prevent the market from straying too far from the target.
- » In a managed float, there is no public commitment to an exchange rate target but the authorities may intervene in the foreign-exchange market or the money market to influence the exchange rate.
- » The impossible trinity refers to the dilemma that a country cannot have control over a fixed exchange rate or a monetary policy, and maintain an open-capital account simultaneously.
- » A country exhibits fear of floating when it claims to let its exchange rate be determined in a float, yet it intervenes directly or indirectly to keep the exchange rate in check.

## NOTES

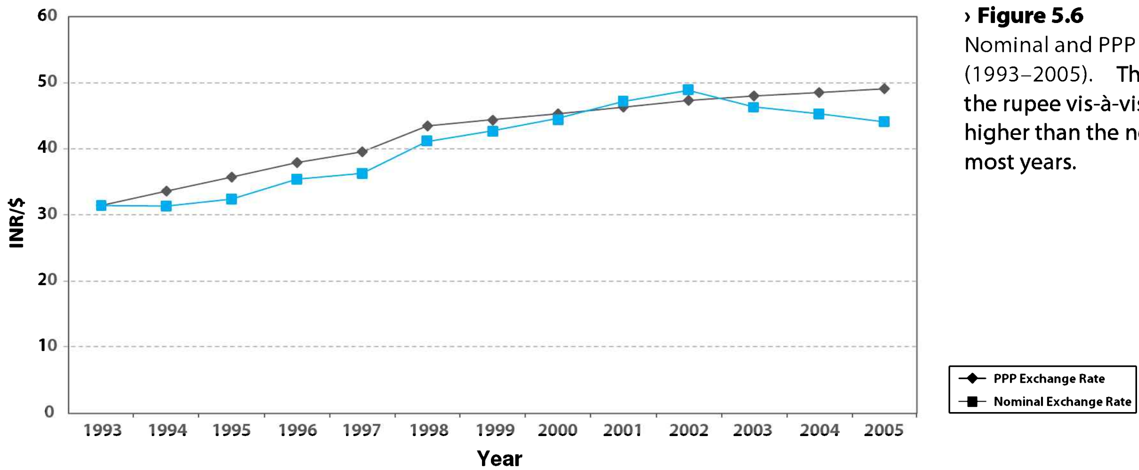
1. This is achieved statistically by calculating a weighted average of a country's bilateral exchange rates, using trade shares as weights. This trade weighted exchange rate is called the nominal effective exchange rate (NEER).
2. $P*$ is the consumer price index (CPI) for urban consumers in the United States, and $P$ is the CPI for urban non-manual employees in India. The real exchange rate is indexed with the base year 1993 value set equal to the nominal exchange rate that year for comparison purposes. The trade-weighted real exchange rate (REER) and nominal exchange rates (NEER) for India depict a similar pattern as seen in the Reserve Bank of India Handbook of Statistics on the Indian Economy.
3. The correlation between the percentage change in the exchange rate and the inflation differential between India and the United States for this time period is 0.72.
4. In the national income accounts, net export is $NX = X - M$. In Eq. (5.2), $X = Q_X$ and $M = (EP^*/P)Q_M$. Multiplying by the real exchange rate converts $Q_M$ which has the dimension as US goods per year to domestic goods per year.
5. See Appendix 5.1, Section B, for the derivation of the Marshall–Lerner condition.
6. The condition is much more complex if the trade balance is not zero initially.
7. Morris Goldstein and Mohsin S. Khan, "Income and Substitute Effects in Foreign Trade," in Ronald W. Jones and Peter B. Kenen (eds.) *Handbook of International Economics* 2 (Elsevier, 1985).
8. Stephen P. Magee, "Currency Contracts, Pass Through and Devaluation," Brookings Papers on Economic Activity 1, 1973, pp. 303–325.
9. Sidney S. Alexander, "Devaluation Versus Import Restriction as an Instrument for Improving the Foreign Trade Balance," International Monetary Fund Staff Papers No. 1, 1951, pp. 379–392; "Effects of a Devaluation on a Trade Balance," International Monetary Fund Staff Papers No. 2, 1952, pp. 263–278.
10. Svend Laursen and Llyod A. Metzler, "Flexible Exchange Rates and the Theory of Employment," *Review of Economics and Statistics* 32 (November 1950): 281–299.
11. Ibid.
12. Krugman and Taylor argue that "In the short run the balance of payments deficit is 'structural', that is, both imports and exports are not very sensitive to price changes for a given level of domestic output."
13. Paul Krugman and Lance Taylor, "Contractionary Effects of Devaluation," *Journal of International Economies* (8 August 1978): 445–456.
14. Krugman and Taylor are discussing a nominal devaluation, which due to the sluggishness with which prices change, results in a real devaluation as well.
15. Tarlok Singh, "Testing J-Curve Hypothesis and Analyzing the Effect of Exchange Rate Volatility on the Balance of trade in India," *Empirical Economics* 29 (2004): 227–245.
16. Swarnjit M. Arora, Mohsen Bahmani-Oskooee, and Gour Goswami, "Bilateral J-Curve between India and Her Trading partners," *Applied Economics* 35 (2003): 1037–1041.
17. Of course, current exchange rates affect future expected exchange rates, but here the focus is on a temporary change, which is so transitory that it has no effect on the expected exchange rate.
18. The expected future exchange rate, $E^e$, will of course also be affected by a change in interest rates. We ignore that for now, but note that a rise in the expected future price of foreign currency in terms of domestic currency shifts the downward slopping schedule in Figure 2.2 to the right and causes the spot rate to depreciate as foreign currency deposits offer a higher expected rate of return and

individuals seek to cash in on this by exchanging domestic deposit for foreign currency deposits.

19. The estimates are OLS estimates, which are problematic as the data are non-stationary. Thus, researchers prefer to estimate $E^e - E = F - E$, or $E_t(\Delta E_t) = F_t - E_t$, which leads to estimating $\Delta E_{t+1} = \alpha + \beta(F_t - E_t) + \epsilon_{t+1}$. This yields bad results for the theory with a negative sign for $\beta$. This negative $\beta$ is referred to as the "forward premium puzzle." For more on this see Ken Froot and Richard Thaler, " Anomalies: Foreign Exchange," *Journal of Economic Perspectives* 4, 3 (1990): 179–192.
20. B. K. Bhoi, and Sarat C. Dhal, "Integration of Financial Markets in India: An Empirical Evaluation," RBI Occasional Papers No. 19 (4), Mumbai 1998, pp. 345–380.
21. For a history of purchasing power parity, see Paul Einzig, *The History of Foreign Exchange* (London and Basingtoke: Macmillan St. Martin's Press, 1990).
22. In that case, the homogeneous of degree one $g(.)$ and $f(.)$ functions are the same.
23. Of course, this is ignoring tariffs and transport costs.
24. This can be straightforwardly written taking the log of Eq. (5.14), which is $\log P = \log P^* + \log E$ and differentiating to obtain $\Delta E/E = (\Delta P/P) - (\Delta P^*/P^*)$, which is Eq. (5.16).
25. See Paul Krugman, "Purchasing Power Parity and Exchange Rates: Another Look at the Evidence," *Journal of International Economics* 8 (August 1978): 397–407.
26. Taking the logarithm of $RER = EP^*/P$, we get $\log(RER) = \log E + \log P^* - \log P$. Totally differentiating gives $\Delta(RER)/RER = (\Delta E/E) + (\Delta P^*/P) - (\Delta P/P) = (\Delta E/E) + \pi^* - \pi$.
27. Stanley Fischer, "Exchange Rate Regimes: Is the Bipolar View Correct?" *Journal of Economic Perspectives* 15, no. 2 (Spring 2001): 3–24.
28. Rudi Dornbusch, "Fewer Monies, Better Monies," NBER Working Paper No. 8324, Cambridge, MA: NBER, 2001.
29. Ricardo Hausmann, "Should There be Five Currencies or One Hundred and Five?" *Foreign Policy* 116 (Fall 1999): 65–79.
30. Ibid.
31. Maurice Obstfeld and Kenneth Rogoff, "Mirage of fixed Exchange Rates," *Journal of Economic Perspectives* 9 (1995): 73–95.
32. Less than complete backing of the currency gives the central bank some leeway in case financial crises occur. The central bank can then act as a lender of last resort to stem the crisis.
33. Jeffrey Sachs and Felipe Larrain, "Why Dollarisation is More Straitjacket than Salvation," *Foreign Policy* 116 (Fall1999): 80–92.
34. As we will see later, this is a saving achieved by not forgoing seigniorage—the income a monetary authority receives when the value of its currency exceeds the cost of producing the currency.
35. Peter B. Kenen, "The International Financial Architecture: What's New? What's Missing?" *Institute for International Economics* (Washington, D.C.: IIE, 2001).
36. Ibid.
37. Stanley Fischer, op. cit.
38. Kenneth S. Rogoff, Aasim M. Husain, Ashoka Mody, Robin Brookes, and Nienke Oomes, "Evolution and Performance of Exchange Rate Regimes," IMF Working Paper No. WP/2003/243, 2003.
39. Maurice Obstfeld and Kenneth Rogoff, op. cit.
40. Carmen M. Reinhart, "The Mirage of Floating Exchange Rates," *American Economic Review*, May 2002.
41. Guillermo A. Calvo and Carmen M. Reinhart, "Fear of Floating," *Quarterly Journal of Economics* 117 (2002): 379–408.
42. Countries can use domestic open-market operations and thus the interest rate to stabilize the exchange rate, leading to the expectation that interest rate volatility will be higher under a fixed exchange rate regime than under a floating one. Similarly, foreign exchange reserve variability would be higher under a fixed exchange rate regime than under a floating one. Hence, countries that claim to be floaters under IMF classification but which exhibit high variability of interest rates and/or reserves are said to exhibit fear of floating.

## TEST YOURSELF

1. What is the real exchange rate, and why do we focus on it rather than on the nominal exchange rate in macroeconomics?
2. How does the real exchange rate affect the exports and imports of a nation? If the real exchange rate is devalued, will there be an unambiguous impact on the trade balance?
3. What is the J-curve effect, and how can one explain this phenomenon?
4. How does the absorption approach explain the impact of devaluation on the trade balance?
5. Explain what is uncovered and covered interest parity and in what way are they different from each other.
6. How does the exchange rate attain equilibrium under uncovered interest parity?
7. What is absolute and relative purchasing power parity and to what extent are they good explanations for exchange-rate determination?
8. Can a country run an independent monetary policy to achieve domestic economic policy goals, if it is committed to the maintenance of a fixed exchange rate?
9. When is an exchange rate considered to be overvalued or undervalued? Suppose the exchange rate is overvalued and the policymakers wish to correct this by changing the domestic interest rate. What should they do?
10. Describe the various types of "corner" solutions advocated by the IMF for exchange-rate regimes in the 1990s.




› **Figure 5.6**
Nominal and PPP Exchange Rates (1993–2005). The PPP value of the rupee vis-à-vis the US dollar is higher than the nominal value for most years.

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
   (c) Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy".
2. Click on "Table 40: Consumer price index—annual average" and create a spreadsheet of data on the consumer price index for urban non-manual employees (UNME)—column 4 of Table 40—from 1993–1994 to 2005–2006. This column of data is the Indian consumer price index or $CPI_{India}$.
3. Then click on "Table 150: Exchange rate of the Indian rupee …" and create a data column in your spreadsheet on the exchange rate of the rupee corresponding to the US dollar from 1993 to 2005 (column 3 of Table 153). This is annual data on the nominal exchange rate.
4. Now go to the home page of the Federal Reserve Bank of Dallas (URL: http://dallasfed.org/index.cfm).
5. Click on the "Economic Data" icon.
6. Under US Economic Date, click on "Prices."
7. Under "Prices" click on "Consumer Price Index."
8. Click on "View the Data Table."
9. Create a data column in your spreadsheet on the annual average consumer price index for all urban consumers (fourth column of the data table you are viewing) from 1993 onwards up to 2005. This column of data is the US consumer price index or $CPI_{US}$.
10. Divide the Indian CPI in a year by the United States CPI for that year to get the ratio $CPI_{India}/CPI_{US}$.
11. We want to create a relative price index. Hence, set the relative consumer prices in the two countries for the year 1993 at $CPI_{India}/CPI_{US} = 216/144.48 = 31.44$, where 31.44 is the rupee-dollar nominal exchange rate in the year 1993.
12. For 1994, we index the relative prices accordingly. Thus, with $CPI_{India}/CPI_{US} = 237/148.23$ for 1994 we have an index value of 33.62 for the year.
13. Similarly, the index value for 1995 is $CPI_{India}/CPI_{US} = 259/152.38 = 35.74$.
14. Calculate for the remaining years. The index of the ratio of CPIs between the two countries with the base 1993 ratio set equal to the actual nominal rate can be considered to be the PPP exchange rate between the two countries.
15. Now plot a graph of your calculated PPP exchange rate of the rupee versus the US dollar with the actual nominal exchange rate for the time period. Your graph should resemble the one in Figure 5.6.
16. You would observe that the PPP exchange rate is consistently below the nominal exchange rate up to 2000. Thereafter, the PPP exchange rate is marginally below the nominal exchange rate for a couple of years before reverting back above it. Is the exchange rate overvalued relative to PPP?
17. Is it possible that the RBI intervened in the foreign-exchange market to nominally devalue the currency in order to keep the exchange rate from straying away from its long-run fundamental PPP value?

# Appendix 5.1 The Trade Balance and the Real Exchange Rate

## Section A: Imports and the Real Exchange Rate

Suppose that the demand for imported goods is inversely related to the real exchange rate and this relationship is linear, or

$$Q_M = \mu_0 - \mu_2 q \tag{A5.1.1}$$

where $\mu_0$ and $\mu_2$ are positive constants. Then, the demand for imported goods measured in terms of domestic goods or imports is equal to

$$qQ_M = \mu_0 q - \mu_2 q^2 \tag{A5.1.2}$$

This is a quadratic equation in the real exchange rate which results in an inverted U-shaped curve as depicted in Figure 5.2.

The maximum value of $qQ_M$ is attained where

$$\frac{\partial(qQ_M)}{\partial q} = \mu_0 - 2\mu_2 q = 0,$$

or

$$q = \frac{\mu_0}{2\mu_2}$$

Substituting this into Eq. (A5.1.1),

$$Q_M = \mu_0 - \mu_2\left(\frac{\mu_0}{2\mu_2}\right)$$

$$= \frac{\mu_0}{2} \tag{A5.1.3}$$

The quadratic equation, Eq. (A5.1.2), intersects the horizontal axis or takes a value zero when

$$\mu_0 q - \mu_2 q^2 = 0,$$

or

$$q = \frac{-\mu_0 \pm \sqrt{\mu_0^2}}{-2\mu_2}$$

Hence,

$$q = 0,$$

or

$$q = \frac{\mu_0}{\mu_2}$$

The graph of $qQ_M$, therefore, starts at the origin ($q = 0$), attains a maximum at $q = \mu_0/2\mu_2$, and then declines till the demand for imported goods is driven to zero at $q = \mu_0/\mu_2$.

Now,

$$\eta_M = \frac{\Delta Q_M}{\Delta q}\frac{q}{Q_M}$$

From Eq. (A5.1.1),

$$\frac{\Delta Q_M}{\Delta q} = -\mu_2$$

At the maximum value of the graph of $qQ_M$, $q = \mu_0/2\mu_2$ and $Q_M = \mu_0/2$. Substituting these into the definition of the elasticity of demand for imports,

$$\eta_M = -(-\mu_2)\frac{\mu_0/2\mu_2}{\mu_0/2}$$

$$= \mu_2\frac{\mu_0}{2\mu_2}\frac{2}{\mu_0}$$

$$= 1$$

Hence, the graph of imports with respect to the real exchange rate attains its maximum value when the elasticity of demand for imports is unity.

Section B: The Marshall–Lerner Condition

A nation's exports depend on two factors: the price competitiveness of its goods and the ability of buyers abroad to pay-the income of the rest of the world, $Y^*$. The real exchange rate is a measure of the competitiveness of a country's goods. Thus, export volumes are directly affected by increases in foreign income and the real exchange rate. We may write,

$$Q_X = Q_X\left(\underset{(+)}{Y^*}, \underset{(+)}{q}\right)$$

In the case where the export demand function is linear,

$$Q_X = \delta_0 + \delta_1 Y^* + \delta_2 q \qquad \textbf{(A5.1.4)}$$

The export function reflects the expectation that a depreciation or devaluation, which increases $E$, reduces the real relative price of exports abroad ($q$ rises), and results in a rise in the volume of exports.

Similarly, a nation's imports depend on the real exchange rate and the ability of its citizens to pay, as given by domestic income, $Y$. Thus,

$$Q_M = Q_M\left(\underset{(+)}{Y}, \underset{(-)}{q}\right)$$

where the signs below the arguments in the import function represent the direction in which imports are affected when there is an increase in the argument. A rise in domestic incomes raises the demand for goods from abroad, and a rise in the real exchange rate makes imports more expensive and reduces the volume of imports. In the linear case,

$$Q_M = \mu_0 + \mu_1 Y^* - \mu_2 q \qquad \textbf{(A5.1.5)}$$

Again the import function reflects the expectation that a depreciation or devaluation makes foreign goods more expensive at home and discourages imports. An appreciation or revaluation of the currency has the opposite effect.

When does a devaluation of the currency lead to an improvement in the nation's trade balance? A devaluation makes a given level of imports more expensive in terms of the domestic output that must be exchanged for it. At the same time, the devaluation results in the output of the economy becoming cheaper for foreigners and increases their demand for exports. There is thus a critical level of price sensitivity for exports and imports that ensures that devaluation leads to a higher level of net exports. This condition for a devaluation to improve a trade balance is known as the *Marshall–Lerner condition*.

This condition concentrates on demand conditions and assumes that supply elasticities for exports and imports are perfectly elastic. To derive the Marshall–Lerner condition, we first identify the factors that change the net export position of a country. The change in net exports is given by writing the differential of Eq. (5.2)

$$\Delta NX = \Delta Q_X - q\Delta Q_M - Q_M \Delta q$$

Now, the change in the quantity of demand for exports ($\Delta Q_X$), and for imported goods ($\Delta Q_M$), is given from Eqs. (A5.1.4) and (A5.1.5) as

$$\Delta Q_X = \delta_1 \Delta Y^* + \delta_2 \Delta q$$

$$\Delta Q_M = \mu_1 \Delta Y^* - \mu_2 \Delta q$$

Substituting this into the change in $NX$,

$$\Delta NX = \Delta Q_X - q\Delta Q_M - Q_M \Delta q$$

$$\Delta NX = \delta_1 \Delta Y^* + \delta_2 \Delta q - q(\mu_1 \Delta Y - \mu_2 \Delta q) - Q_M \Delta q$$
$$= (\delta_2 + q\mu_2 - Q_M)\Delta q + \delta_1 \Delta Y^* - q\mu_1 \Delta Y$$

Define the relative price elasticity of demand for exports and imported goods as

$$\eta_X = \frac{\Delta Q_X / Q_X}{\Delta q / q} = \frac{\Delta Q_X}{\Delta q} \times \frac{q}{Q_X} = \delta_2 \frac{q}{Q_X}$$

$$\eta_M = -\frac{\Delta Q_M / Q_M}{\Delta q / q} = -\frac{\Delta Q_M}{\Delta q} \times \frac{q}{Q_M} = \mu_2 \frac{q}{Q_M}$$

as $\Delta Q_X / \Delta q = \delta_2$ and $-(\Delta Q_M / \Delta q) = \mu_2$ from the demand for exports and imported goods functions (A5.1.4) and (A5.1.5). Hence,

$$\delta_2 = \eta_X \frac{Q_X}{q}$$

and,

$$q\mu_2 = \eta_M Q_M$$

We can then write the above expression for $\Delta NX$ as

$$\Delta NX = \left( \eta_X \frac{Q_X}{q} + \eta_M Q_M - Q_M \right) \Delta q + \delta_1 \Delta Y^* - q\mu_1 \Delta Y$$

Suppose that trade is balanced initially, that is $Q_X = qQ_M$. Then,

$$\Delta NX \,|_{NX=0} = (\eta_X + \eta_M - 1) Q_M \Delta q + \delta_1 \Delta Y^* - q\mu_1 \Delta Y$$

The impact on net exports due to a change in the real exchange rate when trade is initially balanced may be written as

$$\frac{\Delta NX}{\Delta q} = \underbrace{(\eta_X + \eta_M - 1)}_{\text{Direct Effect}} Q_M + \underbrace{\delta_1 \frac{\Delta Y^*}{\Delta q} - q\mu_1 \frac{\Delta Y}{\Delta q}}_{\text{Indirect Effect on Foreign and Domestic Incomes}} \qquad \textbf{(A5.1.6)}$$

The Marshall–Lerner condition restricts itself to the direct effect of a change in the real exchange rate. The condition states that if the trade balance is initially zero, a real exchange rate depreciation causes a trade balance surplus if the sum of the relative price elasticities of export and import demand exceeds unity

$$\frac{\partial NX}{\partial q} > 0, \text{ if } \eta_X + \eta_M - 1 > 0,$$

or

$$\eta_X + \eta_M > 1 \qquad \textbf{(A5.1.7)}$$

With given levels of domestic prices $(P)$ and the foreign currency price of imports $(P^*)$, $\Delta q = \Delta E$, and an increase in the relative price $q$ is a pure reflection of a nominal depreciation. Hence, if prices are fixed or slow to adjust, the above condition is the condition for a nominal devaluation to result in an improvement in the trade balance. The Marshall–Lerner condition, Eq. (A5.1.7), for a real devaluation to improve the trade balance is only a necessary and not sufficient condition as it ignores the induced income effects and assumes that income is constant when the real exchange rate changes.

# 6

# The Demand for Money

CRITICAL QUESTIONS

- *What are the major functions of money?*
- *How do we determine the price of a bond?*
- *Why does macroeconomics consider money and bonds as the only two components of private sector wealth?*
- *What is the difference between nominal and real interest rates, and how are they related?*
- *How is an individual's budget constraint affected by the holding of financial assets?*
- *How does a person decide the amount of real money balances they should hold?*
- *How do we decide the allocation of a financial portfolio consisting of money and bonds?*
- *How does the government gain revenue (seigniorage) from the printing of money?*
- *What are the other means of achieving seigniorage apart from printing money?*

# 6.1 Money, Bonds, and Private Wealth

Macroeconomists get around the problem of specifying what money is by describing instead what properties something must have to be called money. The three major functions of money are (1) as a medium of exchange, (2) a unit of accounting, and (3) a store of value.

In a barter economy with no money, trading takes the form of the direct exchange of goods. However, finding someone who has the good you want and is moreover willing to exchange it for something you have can be very time consuming. Money makes searching for a trading partner unnecessary. By operating as a MEDIUM OF EXCHANGE—a device for making transactions—money allows people to trade with less time and effort.

› Money's most important function is as a medium of exchange, when it is used to purchase goods and services.

Besides, money is also a unit of accounting—the basic unit that is used to assess economic value. In India, for instance, all prices, wages, asset values, debts, and financial contracts are expressed in INR. A single and uniform measure of value makes it straightforward to compare different goods and assets. As goods and services are exchanged for money—the medium of exchange function—it is also natural to express economic value in terms of money—as a UNIT OF ACCOUNTING function.[1]

› As a unit of accounting, money is the yardstick for measuring economic value. It measures the value of goods, services and assets, and thereby enables a comparison of these based on their values.

The third function of money is as a STORE OF VALUE. Money can be used to buy goods not only today but also tomorrow and so it embodies and stands for future purchasing power. Money can, thus, be used as a store of value or a way of holding wealth. However, there are other assets—stocks, bonds, gold, real estate—that can be used as a store of value and they typically outperform money in this role as they normally yield a higher rate of return than money. The distinguishing feature of money is, therefore, its role as the medium of exchange—the holding of non-interest-bearing currency and non-(or low) interest-bearing checking deposits for use as a medium of exchange in planned transactions. Any other commodity can serve as a medium of accounting without being a medium of exchange at the same time, and there are various non-money assets that are capable of serving as stores of value.

› Asa store of value, money is one way of holding wealth.

### 6.1.1 Bond Pricing

Money is needed to make purchases—it facilitates transactions. Without money you cannot buy groceries, or pay rent and utility bills. With money—liquidity—you can obtain goods and services that you consume. Money is a financial asset held by households and firms. The other asset that the private sector holds are bonds. Macroeconomic models work either with PERPETUAL BONDS OR FIXED-PRICE BONDS. In a fixed-price bond, the interest rate is variable. The present value of such a bond depends on the sequence of spot or instantaneous interest rates where the interest rate varies over time.[2] A *perpetuity* is a non-maturing bond or a bond that has no final date of maturity. Perpetuities pay a fixed annual amount, or coupon return of INR $C$, per year forever. A perpetuity that is purchased today will pay INR $C$ next year, the year after that, and every other year into the future. That means the discounted present value of this bond, given a discount rate $i$ , is the sum of the discounted present values of INR $C$, or the infinite sum.

› A bond is a legal promise to repay a debt. The principal amount or the amount originally lent is paid at some fixed date in the future called the maturation date. A bondholder receives regular interest, or coupon payments, until the bond's maturation date.

$$\frac{C}{(1+i)} + \frac{C}{(1+i)^2} + \frac{C}{(1+i)^3} + \cdots \tag{6.1}$$

This discounted sum is today's value of the coupon returns. The sum we ought to be willing to pay for the bond today should equal this discounted current value of coupon returns. Hence, the price of the perpetual bond, $P_B$, will equal

$$P_B = \frac{C}{1+i} + \frac{C}{(1+i)^2} + \frac{C}{(1+i)^3} + \cdots \tag{6.2}$$

Multiplying throughout by (1 + i)

$$(1+i)P_B = C + \frac{C}{1+i} + \frac{C}{(1+i)^2} + \cdots \tag{6.3}$$

Subtracting Eq. (6.2) from (6.3),

$$iP_B = C$$

› Bond **prices** and interest rates are inversely related

Hence, the PRICE OF A PERPETUITY is

$$P_B = \frac{C}{i} \tag{6.4}$$

This implies that if the nominal interest rates rise, bond prices will fall and people holding bonds will incur a nominal capital loss.

### *6.1.2* Private Sector Net Wealth

It may seem surprising that apart from money and bonds, households and firms in the private sector apparently hold no other assets and liabilities like real estate, houses, and equity. This is because the net wealth of the private sector consists of the sum of its assets minus the sum of its liabilities. Aggregating over individuals, the assets of a person or a firm are at the same time liabilities of another person or business unit, and, thus, cancel out. The net wealth of the private sector, thus, consists of its net claims on agents outside that sector. It is only the outside financial wealth that we are concerned with and there are two sources of such wealth—the government and the overseas sector. Here we restrict ourselves to an increase in the government debt brought about via a budget deficit that increases the financial asset holding of the private sector.

The real net wealth of households is the stock of physical capital assets and financial assets they hold (see Table 6.1)—corporate bonds and equity, holdings of money issued by the central bank, and holdings of government bonds. Similarly, the net wealth of firms in the economy constitutes their assets—capital stock, money, and government bonds—minus their liabilities, corporate bonds, and equity. In the aggregate, if we consolidate the balance sheet of households and firms into a single balance sheet for the private sector, then, as shown in Table 6.2, the net wealth of the private sector constitutes money issued by the

› **Table 6.1**
The Balance Sheet of Households and Firms

| Households | |
|---|---|
| **Assets** | **Liabilities** |
| Corporate Bonds | |
| Corporate Equity | |
| Money | |
| Government Bonds | |
| Other Physical Capital Assets | Net Worth |
| **Firms[3]** | |
| **Assets** | **Liabilities** |
| Capital Stock | |
| Government Bonds | Corporate Bonds |
| Money | Corporate Equity |

| Assets | Liabilities |
|---|---|
| Money | |
| Government Bonds | |
| Capital Stock | Net Worth |

› **Table 6.2**
The Balance Sheet of the Private Sector

central bank, the government bonds, and the real quantity of physical capital used by households or firms in production, or its physical assets.

Here, we will adopt the simplification that the capital stock is constant when we define net worth or financial wealth, even though when discussing investment—change in capital stock—we allow for its full-flow effects.[4] Then, the composition of WEALTH OF THE PRIVATE SECTOR is given by[5]

› Then et wealth of the private sector consists of its net claims on agents outside that sector—the stock of money and the stock of government bonds.

$$W_t = M_{t-1} + \frac{B_{t-1}}{i_t} \tag{6.5}$$

where, $W_t$ is the nominal wealth variable determining the private sector's decision at time $t$ for time $t+1$;

$M_{t-1}$ is the stock of money held by the private sector at time $t$; and
$B_{t-1}$ is the stock of bonds held at time $t$.

Without loss of generality we are assuming that the coupon on the bond is INR 1 or $C = 1$. Then, with bonds priced at $P_B = 1/i$, $B_{t-1}/i_t$ is the nominal value of bonds held at time $t$. Note that we can write:

$$\frac{B_{t-1}}{i_t} = \frac{B_{t-1}}{i_{t-1}} + \left(\frac{B_{t-1}}{i_t} - \frac{B_{t-1}}{i_{t-1}}\right),$$

where the first term is the wealth in the form of bonds at the beginning of the period, and the second term in brackets is the capital gains or losses incurred during that period.[6] Capital gains/losses are rightly treated as revaluations of wealth rather than as a component of income. Note also that as the value of the bond at time $(t-1)$ is $B_{t-1}/i_t$, the amount of interest paid during period $t$ by the government to the private sector is $i_t(B_{t-1}/i_t) = B_{t-1}$.

The notion of time being advanced here is of agents with a unit planning horizon who make plans at time $t$ for time $t+1$. This means that the concern here is the end-of-period equilibrium since agents wish to achieve their goals at the end of the planning period.

There are a couple of points worth noting:

1. In such a set-up, the current stocks of financial assets must be the end result of decisions, and not a determinant of them, requiring that such variables appear with a one-period lag. The appropriate periodization of variables is often overlooked in macroeconomic analysis.
2. The interest rate being referred to is the nominal interest rate. Money differs from bonds in that the nominal price of money is always equal to one unit of money, for example, US$ 1, INR 1, and 1 yen. In contrast, the nominal price of a bond can change over time as $P_B = C/i$. An individual who owns a bond earns a capital gain if the nominal price of a bond increases over an interval of time, or a capital loss if the nominal price of a bond falls. In contrast, a note of INR 1 has the same nominal value over an interval of time and agents cannot earn nominal capital gains or incur nominal capital losses if they hold all their financial assets as currency. As financial wealth is allocated between holdings of money and of bonds, $W = M + (B/i)$, any change in holdings of bonds for a given stock of financial wealth must be offset by an equal change in money holdings in the opposite direction.[7] As a

change in bond holdings requires a change in money holdings and because bonds earn a nominal interest return, it is the nominal not the real interest rate that influences the desired holding of money.

## 6.2 Nominal and Real Interest Rates

When analysing the demand for money as an asset we refer to the nominal rate of interest, $i$. This is because in most economies contracts for borrowing and lending are denominated in the unit of account, which is the currency of the economy. Assets such as money and bonds also represent future purchasing power as they enable goods which are to be produced in the future to be traded for goods that are currently produced. Underlying the nominal rate of interest, then, is the real rate of interest. Let us illustrate this with an example.

Suppose a farmer lent five sacks of rice to their poor neighbour under a stipulated contract that the neighbour is to repay with six sacks of rice a year later. This would translate into a real rate of interest of 20 per cent per annum. Instead of lending the commodity directly, if the farmer was to lend money, then assuming that a sack of rice cost INR 300, the farmer could have equivalently lent INR 1,500 with a contract to repay INR 1,800 a year later. This would amount to a nominal rate of interest of 20 per cent per annum. As long as the price of rice stays the same in each year, a loan denominated in units of the commodity is equivalent to a loan denominated in INR. But when the price of commodities changes from one time period to the next, the real rate of interest differs from the nominal rate of interest.

Let us assume now that the price of a sack of rice goes up to INR 350. To get back six sacks of rice the next year, the farmer would have to receive INR $(350 \times 6)$ = INR 2,100 from the person to whom he has loaned the money. He would have to charge a nominal rate of interest

$$i = \frac{2{,}100 - 1{,}500}{1{,}500} \times 100 = \frac{600}{1{,}500} \times 100 = 40 \text{ per cent}$$

Thus, it is only on charging a 40 per cent nominal rate of interest that the farmer receives the INR 2,100, which represents the purchasing power of six sacks of rice. In commodity (real) terms, the real rate of interest, however, is still $r = [(6 - 5)/5] \times 100 = 20$ per cent.

We can generalize this example. If at time $t$, the nominal rate of interest is $i_t$ and the price is $P_t$, then the lender receives $P_t(1 + i_t)$ in the next period. The number of units of the commodity that can be purchased with this when the price of the commodity is INR $P_{t+1}$ is $[P_t(1 + i_t)]/P_{t+1}$. Hence, the real rate of interest is given by:

$$1 + r_t = \frac{P_t(1 + i_t)}{P_{t+1}}$$

$$= \frac{1 + i_t}{P_{t+i}/P_t}$$

$$= \frac{1 + i_t}{1 + \pi_t}$$

where $\pi_t = (P_{t+1}/P_t) - 1$ is the rate of inflation.

Hence,
$$r_t = \frac{i_t - \pi_t}{1 + \pi_t} \tag{6.6}$$

Writing the expression in terms of $i_t$,

$$i_t = r_t + \pi_t + r_t\pi_t$$

The term $r_t\pi_t$ is negligibly small and usually ignored so that we may write

$$i_t = r_t + \pi_t \tag{6.7}$$

In other words, we can say that the NOMINAL INTEREST RATE is the real rate of interest plus the rate of inflation.

› The nominal interest rate is the real rate of interest plus the rate of inflation.

Alternatively we may write,

$$r_t = i_t - \pi_t \tag{6.8}$$

This states that the real rate of interest is the nominal rate adjusted for a change in the purchasing power of money.

Equation (6.7) implies that nominal interest rates tend to be high when inflation is high, and low when inflation is low. This tendency of nominal interest rates is referred to as the FISHER EFFECT after the economist Irving Fisher. This can be understood as the response of borrowers and lenders to the inflation they anticipate will continue in the future. Lenders can be expected to raise their nominal interest rate so that their real rate of return will not be affected. Borrowers will accept this raising of the nominal interest rate on the understanding that a higher nominal rate is a compensation to the lender for the fact that the loan will be repaid in currency of reduced real value. The cost of borrowing does not changed when the nominal interest rate and the inflation rate increase by the same amount. Thus, there will be a tendency for nominal interest rates to follow the pattern of the inflation rates in the economy.

› The Fisher effect describes the full adjustment of the nominal interest rate to a change in the inflation rate.

The Fisher effect may not hold if money and capital are substitutes of each other. If money and capital are substitutes, a rise in the inflation rate will encourage people to reduce their holding of money and increase their holdings of capital. This is referred to as the Tobin effect.[8] Recall that the productivity of capital increases at a diminishing rate—the diminishing marginal productivity of capital. An increase in the capital due to the Tobin effect occurs along with a reduction in the marginal product of capital. This represents a reduction in the real rate of return on capital, and a reduction in the real rate of interest as assets will tend to have an equalization of their rates of return. The increase in inflation still leads to a rise in the nominal interest rate but now it is associated with a simultaneous decline in the real interest rate. In this case, the nominal rate will not rise by the full amount of the rise in inflation.

When the individual saves and allocates their savings to financial assets that earn nominal rates of return, it is the nominal rate of interest that becomes useful in describing the intertemporal budget constraint of the individual. We now examine how to construct this budget constraint.

## 6.3 Financial Assets and the Budget Constraint

In Chapter 2, Section 2.1, we have mentioned the expression for private disposable income as:

**Private Disposable Income = GDP + Net Factor Incomes Abroad**
**+ Transfers from Abroad**
**+ Transfers Received from Government**
**− Taxes**

For simplicity, we ignore the presence of net factor incomes from abroad, transfers from abroad, transfers received from government, and taxes. Then, private disposable income is equivalent to GDP ($Y$).

Saving is the part of the disposable income that is not consumed. Saving is nothing but an increase in the wealth of the private sector. Hence,

$$Y - C = S = \frac{\Delta W}{P} 2$$

In nominal terms we may write this as:

$$PY - PC = PS = \Delta W$$

or

$$PY = PC + \Delta W$$

The nominal wealth of the private sector constitutes its holdings of money and bonds—Eq. (6.5)—and a change in nominal wealth is the addition to the holdings of these assets, or

$$\Delta W = (M_t - M_{t-1}) + \left(\frac{B_t}{i_t} - \frac{B_{t-i}}{i_{t-1}}\right)$$

In a two-period set-up—such as when we looked at the intertemporal aspects of consumption—we may write for the first period

$$P_1Y_1 = \overbrace{P_1C_1}^{\substack{\text{Nominal}\\\text{Consumption}}} + \overbrace{(M_1 - 0) + \left(\frac{B_1}{i_1} - 0\right)}^{\substack{\text{Nominal}\\\text{Saving}}}$$

as the individual is taken in a two-period set-up to have no initial endowment of assets ($M_0 = 0, B_0/i_0 = 0$), and to end the second period with no holding of assets either ($M_2 = 0, B_2/i_2 = 0$).

Thus, the first-period budget constraint of the individual in nominal terms is

$$P_1Y_1 = P_1C_1 + M_1 + \frac{B_1}{i_1} \tag{6.9}$$

In the second period, the individual earns an interest on the holdings of bonds given by $i_1(B_1/i_1) = B_1$ and the budget constraint in the second period in nominal terms is

$$P_2Y_2 + B_1 = P_2C_2 + (M_2 - M_1) + \left(\frac{B_2}{i_2} - \frac{B_1}{i_1}\right)$$

or since this is a terminal period and the individual will plan not to hold any assets beyond this period,

$$P_2Y_2 + B_1 = P_2C_2 + (0 - M_1) + \left(0 - \frac{B_1}{i_1}\right)$$

$$P_2Y_2 + B_1 = P_2C_2 - M_1 - \frac{B_1}{i_1} \tag{6.10}$$

To get the lifetime budget constraint we add together the first-period budget constraint—Eq. (6.9)—to the present discounted value of the second-period budget constraint—Eq. (6.10).

As the budget constraints are in nominal terms the present discounted value of the second-period budget constraint is achieved by discounting with the nominal interest rate. Hence, the present discounted value of the second-period budget constraint is

$$\frac{P_2Y_2}{(1+i_1)} + \frac{B_1}{(1+i_1)} = \frac{P_2C_2}{(1+i_1)} - \frac{M_1}{(1+i_1)} - \frac{B_1}{(1+i_1)i_1} \quad \textbf{(6.11)}$$

Adding together Eqs. (6.9) and (6.11) gives us:

$$(P_1Y_1) + \frac{P_2Y_2}{(1+i_1)} + \frac{B_1}{(1+i_1)} = \left(P_1C_1 + M_1 + \frac{B_1}{i_1}\right) + \frac{P_2C_2}{(1+i_1)} - \frac{M_1}{(1+i_1)} - \frac{B_1}{(1+i_1)i_1}$$

where on substituting the relationship between the nominal and real interest rates we obtain (see the derivation in Appendix 6.2):

$$Y_1 + \frac{Y_2}{(1+r_1)} - \frac{i_1}{(1+i_1)}\frac{M_1}{P_1} = C_1 + \frac{C_2}{(1+r_1)} \quad \textbf{(6.12)}$$

This is the TWO-PERIOD BUDGET CONSTRAINT when we explicitly allow for the accumulation of assets between periods. This equation is the same budget constraint as written in the chapter on consumption [see Eq. (3.8) in Chapter 3]. The significant difference is that the term $i_1M_1/(1+i_1)P_1$ appears in Eq. (6.12). This is because holding money implies losing the interest that would have been earned if bonds were held in its place instead. The cost of holding money is the interest lost in the second period, which would have been the amount $i_1$ $(M_1/P_1)$. This loss in the present value terms is obtained by discounting this by $1 + i_1$ to obtain $i_1M_1/(1+i_1)P_1$. Holding money implies a cost in terms of the consumption that is foregone and this cost increases with a rise in the nominal interest rate $i_1$ as $i_1/(1+i_1)$ increases with an increase in $i_1$. There is, thus, a trade-off between holding money and enjoying the consumption that is possible from the purchasing power of this money. This trade-off is depicted in Figure 6.1. We can rewrite Eq. (6.12) as

› The two-period budget constraint with asset accumulation is the present value of the income in the two periods less the discounted value of interest lost on account of holding money balances rather than bonds.

$$C_1 = \left[Y_1 + \frac{Y_2}{(1+r_1)} - \frac{C_2}{(1+r_1)}\right] - \frac{i_1}{(1+r_1)}\frac{M_1}{P_1}$$

When the individual holds no money balances, then, with $M_1/P_1 = 0$, the consumption possibility in period 1 is given by

$$C_1 = \left[Y_1 + \frac{Y_2}{(1+r_1)} - \frac{C_2}{(1+r_1)}\right]$$

This is depicted by point A in Figure 6.1.

The other case is where the individual does not consume in period 1 and saves the entire income in the form of money balances, $C_1 = 0$ and the lifetime budget constraint is

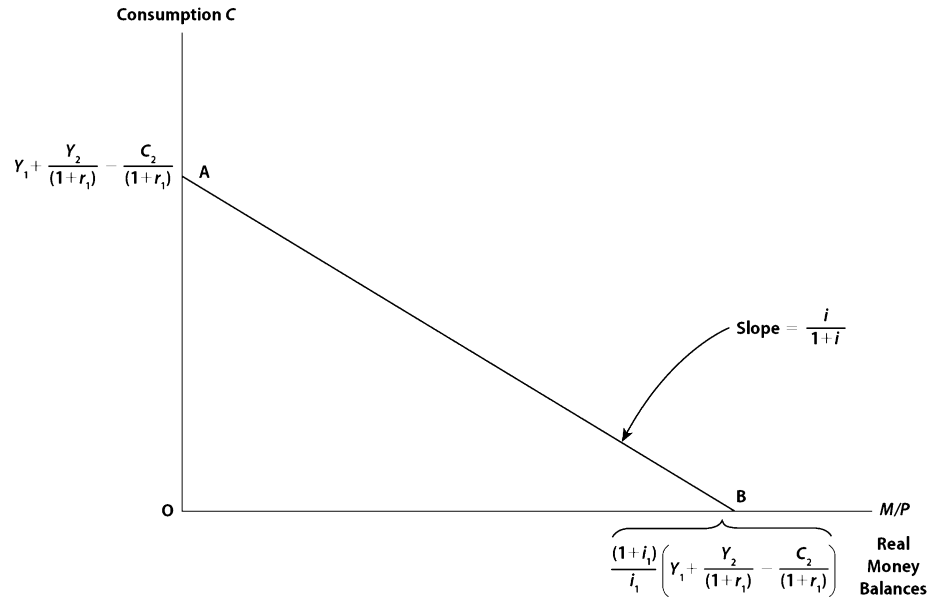
$$\left[Y_1 + \frac{Y_2}{(1+r_1)} - \frac{C_2}{(1+r_1)}\right] - \frac{i_1}{(1+i_1)}\frac{M_1}{P_1} = 0$$

$$\frac{M_1}{P_1} = \frac{(1+i_1)}{i_1}\left[Y_1 + \frac{Y_2}{(1+r_1)} - \frac{C_2}{(1+r_1)}\right].$$

This is given by point B in Figure 6.1.

Joining points A and B gives us the budget line representing the various combinations of present consumption and holdings of real money balances that are possible over the life cycle.

**› Figure 6.1**
Income Constraint when Allocating Savings to Money Balances. **As an individual allocates their savings to money balances they lose the implied interest income from bonds given in present value terms by $i_1M_1/[(1 + i_1)P_1]$ As the individual transfers savings to bonds they increase their consumption possibility, but at the same time decrease their liquidity in terms of money balances as depicted by line AB.**




The slope of the budget line is the ratio of the vertical distance OA to the horizontal distance OB and is given by

$$\text{Slope of budget line} = \frac{\{Y_1 + [Y_2/(1+r_1)] - [C_2/(1+r_1)]\}}{[(1+i_1)/i_1]\,\{Y_1 + [Y_2/(1+r_1)] - [C_2/(1+r_1)]\}} = \frac{i_1}{1+i_1}$$

### *6.3.1* The Determinants of the Real Money Demand

The private sector opts to forgo an income from interest earnings by holding money. The benefit to the household and firms of holding money, however, is that it provides liquidity without which transactions minus the inconvenience of barter would not be possible. This trade-off of constraints on the actions of the private sector can be represented by the constraint line in Figure 6.1.

Which point would an agent choose on the constraint AB in Figure 6.1? It depends on preferences, and preferences, in turn, are depicted by the indifference curves derived from a utility function. The two arguments in the utility function are consumption[9] and money. The reason for utility depending on consumption is that in the final analysis we are interested in the consumption made possible by income. Since money has value because it is used to exchange goods and services, we should measure money in units of output. For instance, if INR *P* is the general price level that enables us to purchase one unit of output, INR *M* will enable us to buy *M/P* units of output. Hence, the value of money measured in terms of the output is *M/P,* and is referred to as real money balances. When people decide how much money to carry, what matters is the purchasing power of the money they hold.

Suppose, for instance, you decide to carry INR 30 to cover your intended purchases for lunch and tea during the day. Imagine, however, what happens if the price level doubles for some reason after you leave home. The cost of lunch and tea is now INR 60. You will not be able to cover the cost of the lunch you planned to eat and you may have to forgo your cup of tea. The purchasing power of your cash of INR 30 is now half of its previous value. To purchase the same meal and your favourite cup of tea, you would have to double your nominal money balances to INR 60. A doubling of prices requires a doubling of nominal money balances to maintain the required purchasing power for the day's expenses. The utility function can accordingly be written as:

$$U = U\left(C, \frac{M}{P}\right). \tag{6.13}$$

This utility function generates standard indifference curves that are downward sloping. The aim of an individual agent is to maximize utility subject to the income constraint[10] given in Figure 6.1.

$$\left.\begin{aligned} &\underset{\{C_1,(M_1/P_1),(B_1/i_1P_1)\}}{\text{Max}} U\left(C_1, \frac{M_1}{P_1}\right) \\ &\text{such that } Y_1 + \frac{Y_2}{(1+r_1)} - \frac{i_1}{(1+i_1)}\frac{M_1}{P_1} = C_1 + \frac{C_2}{(1+r_1)} \end{aligned}\right\} \tag{6.14}$$

The agent chooses consumption, nominal money balances (M), and nominal government bond holdings (B), to maximize utility subject to the budget constraint. The agent takes the nominal interest rate *i*, and the price level *P* as given.[11]

### *6.3.2* Parametric Changes on the Money Demand

At this point we are only interested in the agent's choice of money balances. The graphical solution is depicted in Figure 6.2 as the point E where the highest indifference curve is tangential to the income constraint. The demand for real money balances at this point is given by $M^{d*}/P$. At point E, the slope of the income constraint is equal to the slope of the indifference curve.

Recall that the slope of the income constraint is the ratio of the vertical distance OA to the horizontal distance OB, or

$$\text{Slope of income constraint} = \frac{\text{Distance OA}}{\text{Distance OB}} = \frac{i_1}{1+i_1}$$

What happens to the demand for money when there is a rise in the interest rate? An increase in the interest rate to $i'$ makes the income constraint steeper and the increased interest rate will result in an income constraint such as AB′ (Figure 6.3).

**MACROFOCUS 6.1**

### Stock Markets, Exchange Rates, and the Money Demand

Two Nobel laureates, Milton Friedman and Robert Mundell, have argued that the money demand, in addition to the standard variables—income and nominal interest rates—is also affected by stock market prices* and by the exchange rate.** Friedman observed a correlation between stock prices and money demand in the US economy, and suggested a number of explanations for this phenomenon: (1) Wealth effect—a rise in the stock prices is a rise in nominal wealth and an increase in the wealth-to-income ratio will be reflected in a higher money-to-income ratio; (2) Substitution effect—the higher the real stock prices, the more attractive are the equities as a component of a portfolio that includes equities and money; (3) Risk considerations—a rise in stock prices is an increase in return from risky assets relative to liquid assets, which may be counterbalanced by increasing the weight of money in an aggregate portfolio. Friedman found that for the United States, the relationship between money and the stock market is time varying with both wealth and substitution effects showing up.

The Mundell conjecture is rationalized as follows. Assume wealth holders evaluate their portfolio in terms of domestic currency. A depreciation means an increase in the value of the foreign securities held by domestic residents. If the increased domestic currency value of foreign assets held by domestic residents is perceived as an increase in wealth, the demand for money could increase. The effect of the exchange rate depreciation could also lead to a decline in the demand for money. This is because the public expects further depreciation and, therefore, demands more of foreign currency and less of domestic currency.

Dua and Subramaniam† find the presence of a strong wealth effect of stock prices on broad money (M3) in India but find the exchange rate has no effect on the money demand for the period of 1980–2003. An increase in stock prices, thus, increases the demand for money in India. However, Bahmani-Oskooee and Tanku†† argue that in developing economies, where there is a black market for foreign exchange due to foreign exchange controls, it is the black market exchange rate rather than the official exchange rate which is a determinant of the demand for money. They find that the black market exchange rate, as well as the official exchange rate, affects the demand for money in countries like Argentina, Egypt, India, Indonesia, Mexico, and Philippines. Thus, money demand is a function of not just income and the nominal interest rate, but also stock market prices and the exchange rate.

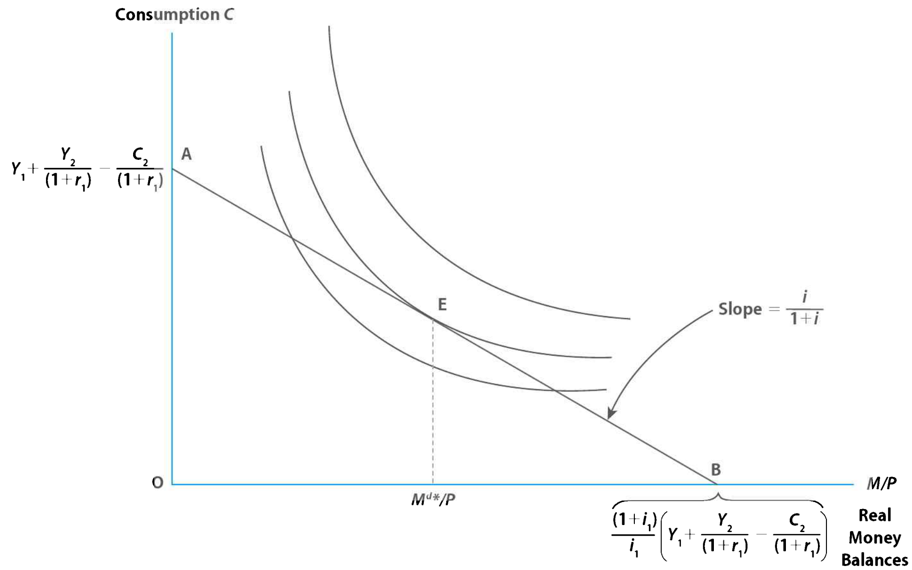
* M. Friedman, "Money and the Stock Market," *Journal of Political Economy* 96 (1988): 221–245.

** R. A. Mundell, "Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates," *Canadian Journal of Economics and Political Science* 29 (1963): 475–485.

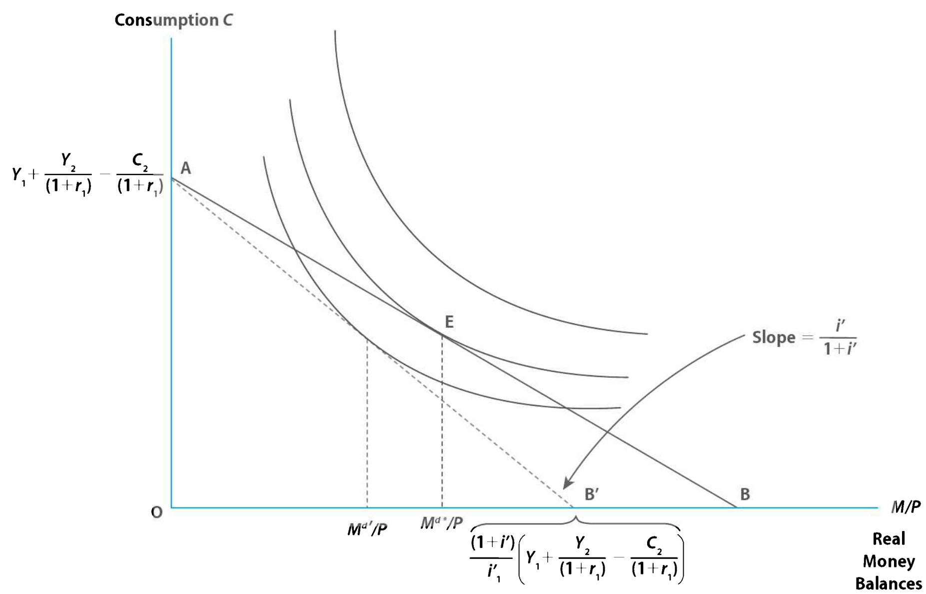
† P. Dua and S. Subramaniam, "Money Demand in the Indian Economy: A Time Varying Parameter Approach," in B. B. Bhattacharya and A. Mitra (eds.) *Studies in Macroeconomics and Welfare* (New Delhi: Academic Foundation, 2005), pp. 117–141.

†† M. Bahmani-Oskooee and A. Tanku, "Black Market Exchange Rate, Currency Substitution and the Demand for Money in LDCs," *Economic Systems* 30 (2006): 249–263.

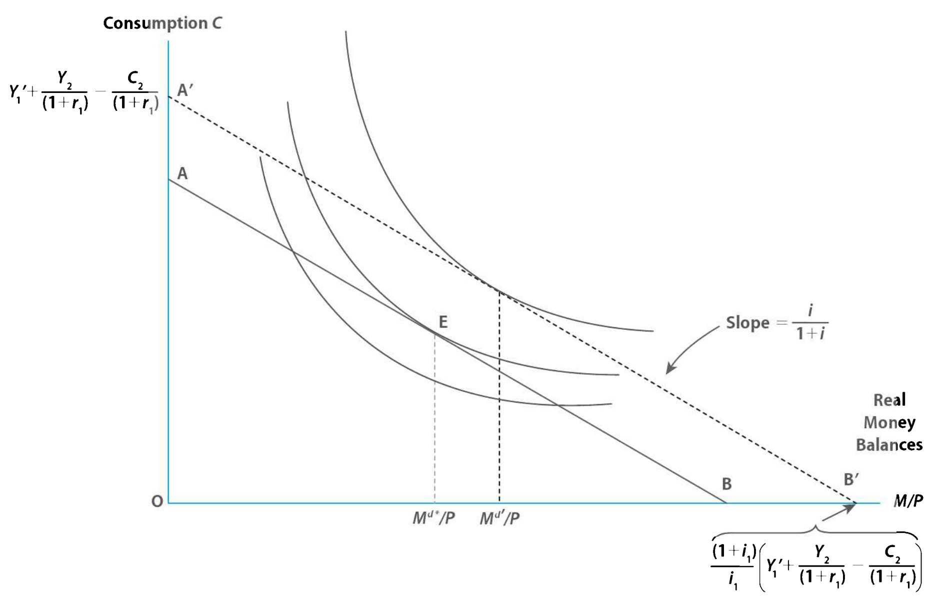
**› Figure 6.2**
Equilibrium Money Balances Held. The demand for real money balances—the value of money measured in units of commodities—is found by equating the preference for liquidity versus consumption as given by the slope of an indifference curve, with the slope of the income constraint line.




**› Figure 6.3**
Effect of a Rise in Interest Rate. A rise in the interest rate causes an individual to reallocate savings from money to bonds thereby reducing the demand for real money balances.




**› Figure 6.4**
Effect of a Rise in Income. A rise in income increases the requirement of liquidity for transactions and causes the individual to increase the demand for real money balances.



As depicted in Figure 6.3, the rise in the interest rate induces a decline to $M^{d'}/P$ in the demand for money balances. The opportunity cost of holding money is the forgone interest. A rise in the interest rate induces agents to reallocate their portfolio of financial assets from money to bonds.

What happens if the real income rises? A rise in real income shifts the budget line outwards—parallel to itself. In Figure 6.4 this is depicted by the new income constraint A′B′. The rise in real income results in an increase in the demand for money to $M^{d'}/P$. A rise in income increases the volume of planned transactions in the economy, which requires additional money balances.

In general, we can write the demand for money balances as a function of the nominal interest rate and income in the economy. The MONEY DEMAND decreases with a rise in the interest rate and increases with a rise in income. Thus,

› The demand for real money balances rises with the income and declines with a rise in the nominal interest rate.

$$\frac{M^d}{P} = f\Big(\underset{(-)}{i}, \underset{(+)}{Y}\Big) \tag{6.15}$$

# 6.4 Money as a Store of Value

So far, the return on money and bonds is known by the agent. As the return on bonds ($i$) is higher than that on the money (0), only bonds are held as a store of value. However, once we allow for *uncertainty*—the agent faces uncertainty regarding the yield on their holding of a stock of financial assets—by allowing the yield on bonds to vary, agents are subject to capital risk. We continue to allow the yield on money to be known with certainty so that money constitutes a "safe" asset from the point of view of the agent.[12] The yield on money is zero compared to the variable yield on bonds. The return on bonds consists of two components. One is the interest or a coupon of a known amount received at regular intervals. The other is the prospect of a capital gain or loss. The anticipated capital gain or loss is the difference between the price expected at the end of the holding period and the current price ($P_B^e - P_B$). With a coupon of $C$ = INR 1, for instance, the total expected gain from holding a bond is given as a ratio of the purchase price by

$$g_B = \frac{1 + (P_B^e - P_B)}{P_B} \tag{6.16}$$

By contrast the gain from holding money $g_M$ is nil—the nominal price of a unit of money is always equal to a unit of money, no matter when we are considering its nominal value. Let the portfolio share of money in the aggregate financial asset holding be

$$\theta = \frac{M}{W}$$

The expected value of aggregate financial assets at the end of the holding period, $W^e$, will be

$$W^e = [(1+g_M)\theta + (1+g_B)(1-\theta)]W$$

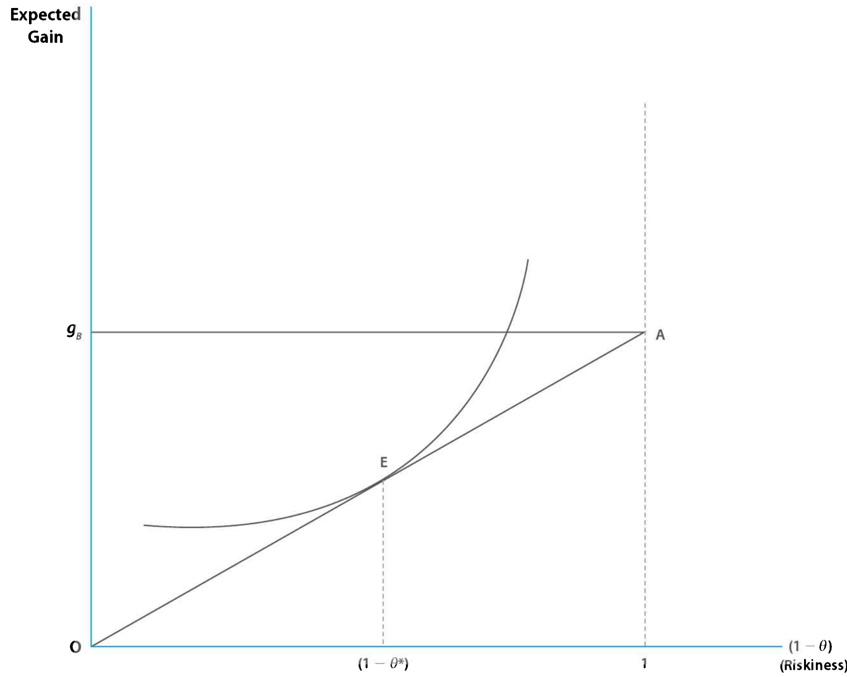
or,

$$W^e = [\theta + (1 + g_B)(1 - \theta)]W$$

$$= [1 + (1 - \theta)g_B]W \tag{6.17}$$

If the agent invests all their financial wealth in money then with $\theta = 1$, there is no change in the value of the stock of financial wealth and $W^e = W$.

If the agent invests all their financial wealth in bonds then $\theta = 0$ and $W = B/i$, and we may write $W^e = (1 + g_B)W$.

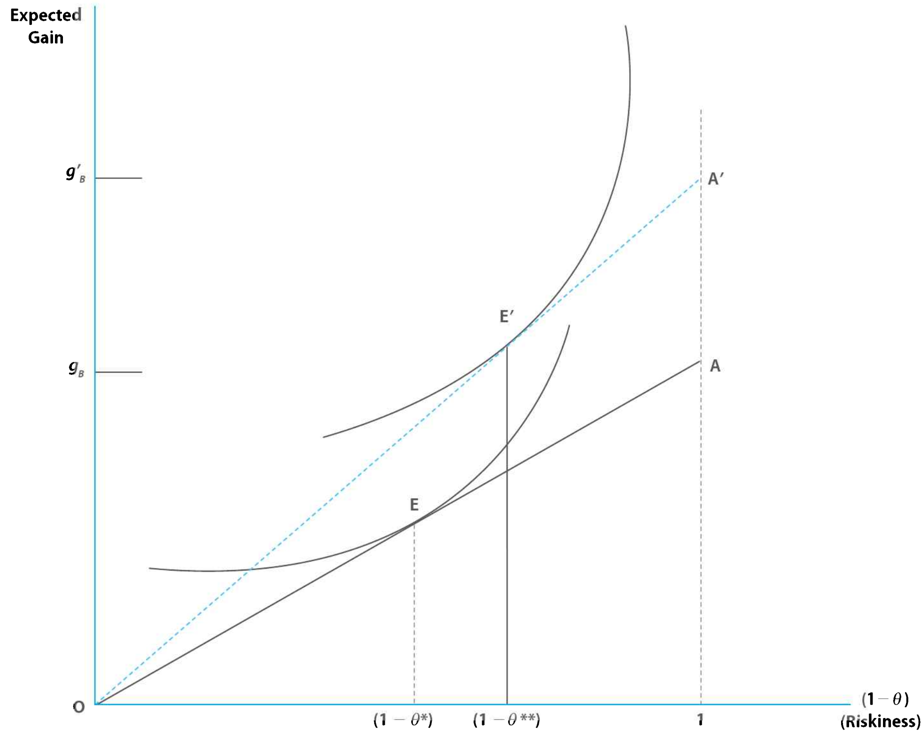



› **Figure 6.5**
Proportion of Wealth in Bonds = $(1-\theta)$. **As the savings portfolio is allocated to bonds increasingly, with $(1-\theta)$ the proportion of the portfolio allocated to bonds increasingly taking values away from 0 and towards 1, the capital risk of the portfolio also increases. An individual with a preference for risk versus return, as depicted by an indifference curve, will allocate a proportion $(1-\theta^*)$ of the portfolio to bonds.**

In Figure 6.5 we measure the proportion of the portfolio in bonds, $1-\theta$, on the $X$-axis and the expected gain from the portfolio on the $Y$-axis. If the agent holds the entire portfolio in money, $\theta = 1$, the expected gain is 0 as $g_M = 0$ and the agent is at the origin, O. On the other hand, if the entire portfolio is in bonds, $\theta = 0$, the agent expects a gain given by $g_B$—point A in Figure 6.5. The agent can hold any proportion $\theta$ of the portfolio in money. Joining points O and A give the expected gain as the agent varies the proportion of the portfolio that is held in money. This line represents the individual's portfolio constraint. Which point would an individual select on this portfolio line?

To answer this question note that the expected price of bonds at the end of the holding period, $P_B^e$, may not be the actual price. Bond prices could rise or fall by more or less than the expected price. Bonds are risky assets whereas money is riskless with a known nominal return of zero. If the individual is at point O in Figure 6.5, the entire portfolio is in money and there is no exposure to risk, but the flipside is that the return is zero. As the individual moves towards point 1 on the $X$-axis their holdings of bonds increase and since bonds are risky, the riskiness of their portfolio increases as well. The $X$-axis, thus, measures the proportion of the portfolio in bonds as well as the riskiness of the portfolio simultaneously, whereas the $Y$-axis measures the expected return from the portfolio. The individual is trading off the risk of the portfolio with the return on the portfolio. A risk-averse individual would have a preference for a portfolio with a higher risk only if that higher risk is compensated by a higher return. The indifference curve of an individual, which depicts the equal satisfaction that an individual receives for various combinations of risk and return, is accordingly upward sloping.

The individual seeks to achieve the highest level of satisfaction possible, subject to the portfolio constraint. This occurs where an indifference curve is tangential to the portfolio constraint—point E in Figure 6.5. At this point the agent holds $(1-\theta^*)$ of the portfolio in bonds and $\theta^*$ of the portfolio in money. The individual is holding a diversified portfolio of money and bonds. If there was no risk associated with bonds and $P_B^e > P_B$, the individual would hold only bonds rather than




**› Figure 6.6**
Effect of the Increase in Expected Gain on Bonds. **As the expected capital gain on bonds increases, due to an increase in the expected price of the bond or a rise in the interest rate, the individual allocates a greater proportion of the portfolio to bonds.**

earn the zero return on money. However, $P_B^e$ is not known with certainty and the individual hedges their bets by holding a combination of money and bonds.

Suppose now there is some news that causes the individual to revise their expectation of the price of the bond upwards to $P_B^{e\prime}$, where, $P_B^{e\prime} > P_B^e$. From Eq. (6.15), $g_B$ increases to $g'_B$ and point $A$ shifts to $A'$ in Figure 6.6.

The new equilibrium is at point $E'$ where the proportion of the portfolio in bonds has increased to $(1 - \theta^{**})$. The portfolio constraint could also swivel out to OA′ for a given expectation if the interest rate increases. Note that A is the point where

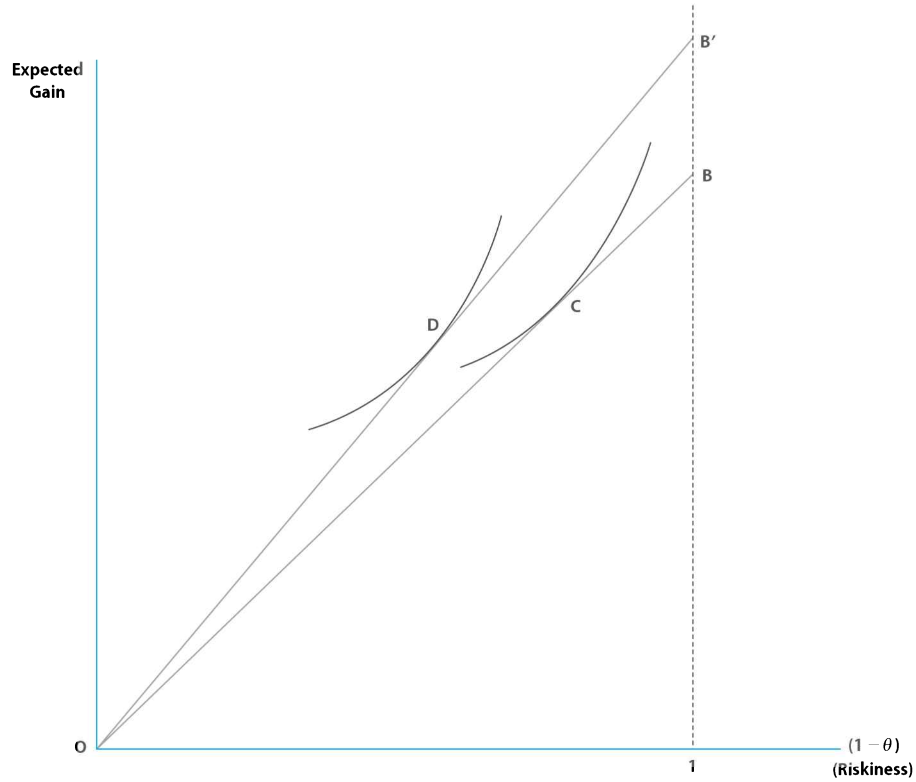
$$g_B = \frac{1 + P_B^e}{P_B} - 1$$

If $P_B$ declines to $P'_B$, where $P'_B < P_B$, the denominator in the equation for the expected gain on bonds declines and there is an increase in the expression for the expected return on bonds to $g'_B > g_B$. Recall that the price of a bond is inversely related to the interest rate so that a decline in the price is simultaneously a rise in the interest rate on bonds.

We can, therefore, conceive of the portfolio constraint shifting out to OA′ as a result of an increase in the interest rate. A rise in the interest rate is, thus, associated with an increase in the demand for bonds and a decrease in the demand for money. The demand for money as a store of value is, therefore, a portfolio or speculative demand for money and it varies inversely with the rate of interest. Therefore, whether money is a demand for transaction purposes as a medium of exchange, or a store of value for portfolio/speculative purposes, it is inversely related to the interest rate. The two functions of money together imply that the demand for money is positively related to the level of income and negatively related to the interest rate, as depicted in Eq. (6.15).

› The agent decreases their allocation to money in their **portfolio** as the interest rate increases up to an acceptable level of risk. As the risk keeps on increasing, however, at some point the individual seeks to compensate for the risk by increasing the allocation of money even though interest rates increase.

As the interest rate rises and the portfolio constraint swivels out, the share of bonds in the portfolio increases. However, the RISKINESS OF THE PORTFOLIO also increases along with the rise in the interest rate. It is conceivable that eventually,




**› Figure 6.7**
The Demand for Money Increasing with Interest Rates. **At very high rates of interest the individual can face significant capital losses and will then reduce their holding of bonds despite the high current return from bonds.**

the rise in the interest rate will increase the riskiness of the portfolio to the point where it triggers a response for a reduction in risk. This is done by reducing the proportion of bonds in the portfolio as interest rates rise. This is depicted in Figure 6.7 where the interest rate rise swivels the portfolio constraint from OB to OB′ and the equilibrium portfolio composition shifts from point C to D where the share of bonds in the portfolio has declined. This is due to the standard substitution and wealth effects that operate when budget constraints change. In Figure 6.6 the substitution effect dominates the wealth effect and the demand for money declines with a rise in the interest rate. In Figure 6.7 the wealth effect dominates the substitution effect and the individual trades off the increased risk by changing the composition of the portfolio and increasing the money holding so as to reduce the riskiness of the portfolio.

The demand for real money balances being negatively related to the nominal rate of interest as in Eq. (6.11), thus, represents a relation where the increase in the inflation rate means that there is a decline in the demand for real money balances. In developing countries since the inflation rates are high, a substitution is induced not just between money and other financial assets such as bonds, but also between money and real assets such as gold and real estate. In such cases, it is standard to include expected inflation as well as an explanatory variable in the demand for money function. Finally, as the demand for money is the demand for a stock of an asset, there is a period of adjustment before individual agents can achieve their desired stock of real money balances. It is, therefore, usual to have a one-period lag in real money stock as an explanatory variable to incorporate the partial stock adjustment in money balances. For the Indian economy, M. J. Manohar Rao[13] estimated the following money demand function:

$$\log\left(\frac{M_t}{P_t}\right) = 0.48 + 0.73\log\left(\frac{M_{t-1}}{P_{t-1}}\right) + 0.38\log Y - 0.23i - 0.0043\log\pi^e \qquad \textbf{(6.18)}$$

If we take a longer time period such as from 1953 to 2003[14] it still turns out that the money demand is positively related to income and negatively to the nominal interest rate—as predicted by the theory.

## 6.5 Seigniorage

That the demand for money balances declines with increasing inflation implies that there is a limit to the discretionary power of governments to produce fiat money. As real money yields direct utility, inflation affects the demand for money by reducing the purchasing power of money. The issuing authority, therefore, finds that there is an optimal rate at which it may issue money. In the monetary and macroeconomics literature there are three theories to provide a framework for understanding what the government does when it expands the stock of (base) money in the economy. Each of the theories focuses on a different goal as the central objective of the central bank. These theories are

1. Seigniorage: The central bank pursues the fiscal interest of the government and generates revenue by printing money.
2. Political business cycles: The central bank objective is the re-election of the incumbent in power through its influence on macroeconomic conditions.
3. Time inconsistency: The central bank uses its discretionary policy to manipulate the economy to achieve a desirable combination of inflation and unemployment. However, when the public understands the incentive structure by which it is achieving this, the options open to it to continue with its discretionary policy get foreclosed (please refer to Chapter 15 for an explanation).

› The use of money creation as a revenue device by the government is called seigniorage.

The basic idea behind SEIGNIORAGE is that a government reaps profit when it produces new money at an expense less than the value of the money produced. The government finances additional expenditure by spending the new units of money. As new money is interchangeable with old money, the expansion of the money stock dilutes the value of existing money balances and implicitly taxes moneyholders. Seigniorage is the net revenue from producing money, or the difference between the exchange value of the produced money and the cost of producing and maintaining it in circulation.

The expression seigniorage dates from the time of the gold standard when every citizen was entitled to have gold bars minted in the form of gold coin. A duty was payable on this service known as seigniorage since it was due to the ruler—the French seigneur or lord—who enjoyed the prerogative of money issue. If the coins were produced in a mint then the mint operations were subject to the following accounting identity:

$$M = PQ + C + S \tag{6.19}$$

where, $M$ is the nominal value (for example, INR 100) assigned to a batch of coins, $P$ is the nominal price paid by the mint per gram of a precious metal, $Q$ is the number of grams of the precious metal embodied in the batch of coins, $C$ is the remaining average cost of operating the mint (called "brassage"), and $S$ is the nominal seigniorage.

If, out of every INR $M$ worth of coins minted, $PQ$ is paid for the silver, brass or gold, and $C$ covers other mint expenses, then $S$ is retained as the seigniorage profit available for spending by the king. The total seigniorage revenue per annum just depends on how many batches of coins are produced per annum.[15]

In today's world of fiat money standards, the bullion content of the base money is zero, and production costs are almost zero.[16] Setting $Q = 0$ and $C = 0$ it follows that $M = S$. A government's nominal seigniorage per annum is just the change in the stock of base money per year

$$S = \Delta M_0 \tag{6.20}$$

where $M_0$ is the stock of base money (see Chapter 2). Real seigniorage is

$$s = \frac{\Delta M_0}{P} \tag{6.21}$$

The budget constraint of a government that issues fiat money is

$$G + B_{t-1} = T + \frac{\Delta B}{iP} + \frac{\Delta M_0}{P} \tag{6.22}$$

where, $G$ is government expenditure, $B_{t-1}$ is interest payments on the debt, $T$ is tax revenue,

$\Delta B/i$ is the change in interest-bearing debt held by the non-government public, and $\Delta M_0$ is nominal seigniorage or the change in fiat base money held by the public.

The budget constraint makes the process of seigniorage transparent—it helps to finance the government spending with the central bank printing new notes to pay for the bills the government runs up. Another way in which seigniorage is obtained is when the central bank expands the monetary base through an open-market operation. By purchasing $\Delta M_0$ worth of government securities in the market, the central bank retires that much debt[17] and makes it possible for the government and the finance minister to finance a stream of new spending whose present value equals $\Delta M_0$. Essentially the central bank's open-market purchase increases $M_0$ and reduces $B/i$. The finance ministry issue of new debt brings $B/i$ back up to its original level and finances a rise in $G$. The net impact is an increase in $G$ financed by $\Delta M_0$.

› As the government increases the growth **rate** of money in circulation (the tax rate), individuals respond by reducing their holding of real money balances (the tax base) so that seigniorage does not increase without bound.

Seigniorage has a revenue implication in that the RATE OF MONETARY EXPANSION is not infinite. This is equivalent to the problem of profit maximization for a monopolist who has to take into account the decline in sales as they raise the price. In this case the government is a monopoly producer of base money. The inflation rate is the "price" of holding money. As the price is increased, the real quantity of the demanded base money shrinks. Hence, the revenue gained from monetary expansion is finite. Seigniorage can equivalently be thought of as the application of a tax rate on a tax base.

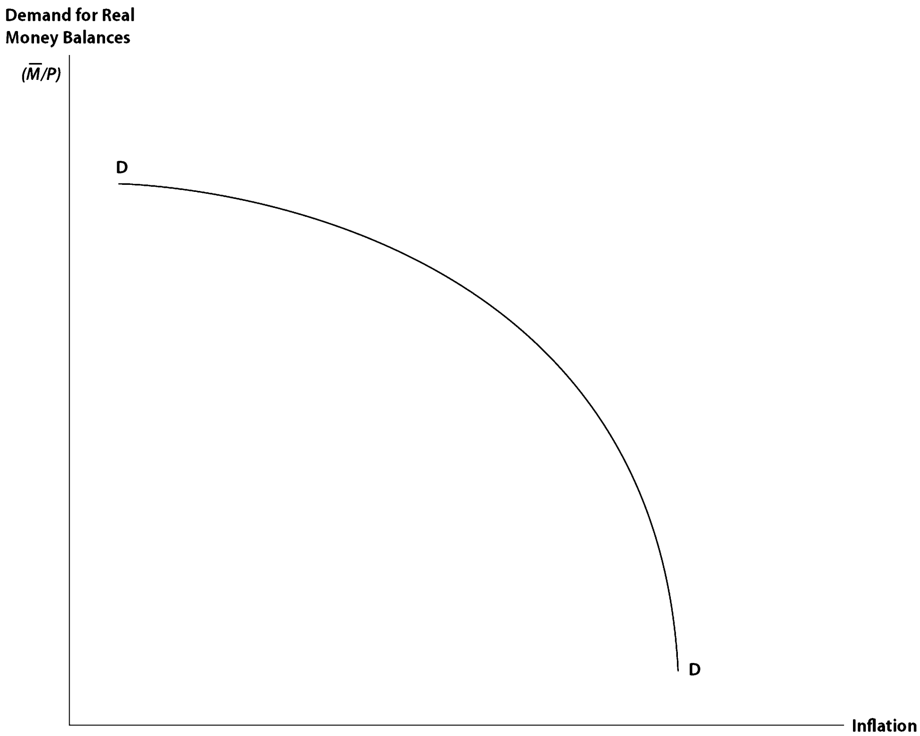
Write seigniorage as:

$$s = \frac{\Delta M_0}{P} = \frac{\Delta M_0}{M_0}\frac{M_0}{P} \tag{6.23}$$

$\Delta M_0/M_0$ is the growth rate of the base money, and is taken to represent a "tax rate" whereas the real base money stock, $M_0/P$, represents the "tax base". Although the faster growth in base money (an increase in the tax rate) increases seigniorage, it reduces the real money supply held in an economy (a narrower tax base), as the resulting increase in inflation and nominal interest rates causes a decline in the holding of the money balances. The process by which this occurs is taken up in the next section.

### **6.5.1 The Optimal Level of Seigniorage

Recall that the demand for money is inversely related to the nominal interest rate $i$ which, in turn, is the sum of the real rate of interest and inflation, $i = r + \pi$. The demand for money is then inversely related to the rate of inflation.




**› Figure 6.8**
The Demand for Steady-State Real Money Balances. **In a steady state the demand for money is inversely related to the rate of inflation. As the rate of inflation increases, the demand for money declines at an increasing rate, as depicted by the slope of DD.**

It is convenient to concentrate on the situation where the money supply grows at a constant rate and ask what happens to the demand for money balances. It turns out (see Appendix 6.3) that the demand for money balances may, depending on initial conditions (the initial holding of real money balances), either decrease over time, increase over time, or, continue to stay at the same level indefinitely. We focus on this third possibility where the demand for money is in a STEADY STATE such that $M_{t+1}/P_{t+1} = M_t/P_t = \bar{M}/P$.

› In a steady state the demand for real money balances is no longer changing and is, therefore, a constant, $\bar{M}/P$.

In a steady state with individuals pursuing consumption smoothing ($C_{t+1} = C_t = \bar{C}$), the increased money supply has no medium of exchange function and individuals have no incentive to hold the increase in the stock of money. As the money supply increases, individuals extinguish their holdings of the increased money stock and this results in a bidding up of prices such that the percentage increase in prices equals the percentage increase in money supply, $\Delta P/P = \Delta M_0/M_0$, and as a result $M_0/P$ is constant at the level required to meet the transactions associated with constant consumption. If we concentrate on STEADY STATES, then, an increase in the growth in money supply results in an equivalent increase in inflation and a reduction in the holding of real money balances because the demand for money is inversely related to the inflation rate. In fact, the relationship between inflation and real money demand is as depicted in Figure 6.8 (see Appendix 6.3 for a proof).

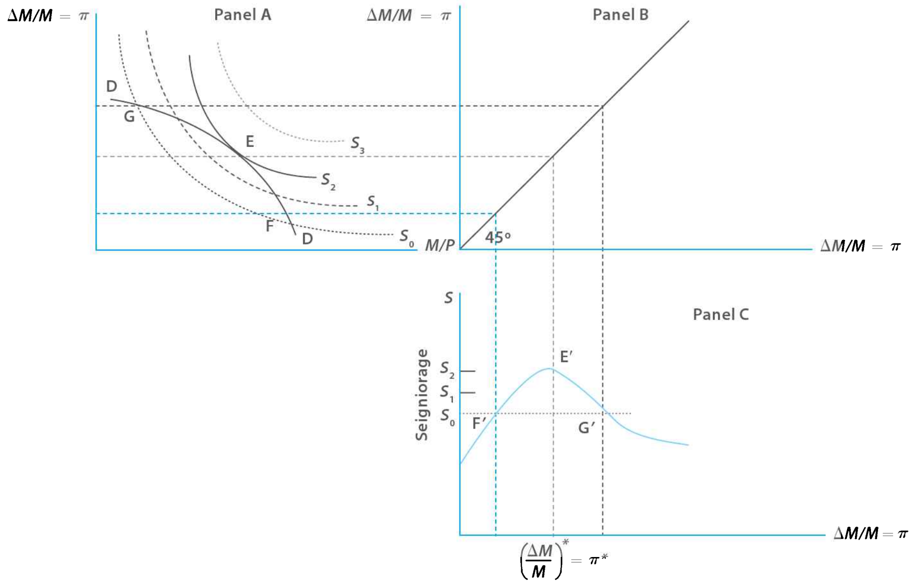
› An economy is in a steady state when various quantities grow at constant rates. If $M/P$ is constant then the rate of growth of the numerator ($M$) must equal the growth rate of the denominator ($P$) that amounts to the condition that the growth of money supply equals the inflation rate.

The demand for real money balances—DD, as depicted in Figure 6.8, decreases as inflation increases. At high rates of inflation, as shown in the Figure 6.8, the demand for real money balances declines rapidly.

Abstracting from a fractional reserve banking system,[18] high-powered money or base money is equivalent to an increase in money stock, that is $\Delta M_0 = \Delta M^s$. Real seigniorage is $s = (\Delta M^s/M^s)(M^s/P)$ from Eq. (6.23)—the product of the growth rate in the nominal money supply and the real money supply.

For any given level of seigniorage the product of $\Delta M^s/M^s$ and $M^s/P$ is a constant—any given level ($s_0$) of seigniorage can be achieved both as a combination of a higher real money supply and a lower growth rate of money

› **Figure 6.9**
Optimal Seigniorage and Inflation. An increase in the growth rate of nominal money supply increases the rate of inflation and results in a reduction in the demand for real money balances. At some point, depicted by point E in panel A and E′ in panel C, the reduction in the demand for money offsets the percentage increase in the money supply and reduces seigniorage.




(point F in Figure 6.9 for $s = s_0$), and as a combination of a lower real money supply and a higher growth rate of money (point G in Figure 6.9 for $s = s_0$). Accordingly, the graph for seigniorage in Figure 6.9(A) is downward sloping.[19]

A higher level of seigniorage, $s > s_0$, is represented in this graph by combinations of real money supply and growth rate of money to the right of $s_0$. In this way various levels of seigniorage $s_0$, $s_1$, $s_2$, ... can be represented by downward-sloping curves, as depicted in Figure 6.9(A). The real money demand function is depicted by the curve DD. For a given level of seigniorage, $s_0$, money demand equals money supply at points F and G. Point F represents a lower growth of the money supply compared to G, but both yield the same seigniorage revenue $s_0$. Using panel B this information can be plotted in panel C of the diagram where points F′ and G′ correspond to points F and G in panel A.

MACROFOCUS 6.2

## Saddam Dinars and Swiss Dinars

The most common reason advocated for people holding money is that it has the backing of a government and is a legal tender for the settlement of payments—the medium of exchange function. Yet, there have been currencies in existence, such as cowrie shells in India, which have had no government backing. Social convention rather than monopoly issue by a government often makes people willing to attribute a medium of exchange property to an asset. Mervyn King, the Governor of the Bank of England, cites an interesting example of this.*

After the Gulf War, southern Iraq was under Saddam's control and northern Iraq became a Kurdish protectorate. UN sanctions made Saddam resort to printing money to finance increasing fiscal deficits. In May 1993, the central bank of Iraq announced a three-week period during which citizens would have to exchange their old 25 dinar notes for new "Saddam dinars" that bore his portrait. Saddam dinars were printed in abundance with the face value of the cash in circulation rising from 22 billion dinars towards the end of 1991, to 584 billion four years later, accompanied by an inflation averaging 250 per cent a year.

The Kurds had no opportunity to exchange their Swiss dinars—named this because the notes had been printed by the company De La Rue on plates manufactured in Switzerland. The Swiss dinar was, therefore, in fixed supply, unlike the Saddam dinar, and it became more valuable. Though there was no trade between North and South Iraq, both the currencies were traded against the dollar and the cross-exchange rate reveals that by 2003 it took 300 Saddam dinars to buy one Swiss dinar.

King observes that by the fall of 2002, as it became more likely that the United States would invade, the Swiss dinar became more valuable and began to appreciate. This was driven by expectations that the Swiss dinar would be honoured by future governments. That expectation turned out to be correct when on 7 July 2003, the head of the Coalition Provisional Authority, Paul Bremer, announced the creation of a new Iraqi dinar that would be exchanged for the two existing currencies at a rate that implied that one Swiss dinar would be worth 150 Saddam dinars. Thus, expectations of changes in a policy regime can have a major impact on the demand for a currency, quite independently of policies pursued by a current government.

* M. King (2004) "The Institutions of Monetary Policy—The Ely Lecture", http://www.bankofengland.co.uk/publications/speeches/2004/speech208.pdf

As depicted, higher levels of seigniorage are possible as we move along the money demand curve from point F towards E as the demand curve intersects successively higher seigniorage curves. After point E as we move along the demand curve towards point F we reverse this trend as the demand curve intersects successively lower seigniorage curves. This gives us the inverted U-shaped seigniorage curve in panel C. The maximum seigniorage is the point of tangency of the demand curve for money and the seigniorage curve $s_2$ at point E in panel A, E′ in panel C. A higher rate of monetary expansion has the direct effect of generating more seigniorage initially but it also increases inflation and reduces the demand for real money balances, which results in an offsetting effect. This offsetting effect grows larger and eventually exceeds the direct effect at points E and E′ as the rate of monetary base expansion is pushed up higher. The shape of the seigniorage curve in panel C suggests that a government trying to maximize its revenues should set a money supply expansion rate and a corresponding inflation rate that is optimal for it at $(\Delta M/M)^* = \pi^*$. A government financing its expenditure via seigniorage would then achieve a constant rate of inflation given by $\pi^*$.

Fischer, Sahay, and Végh[20] find that for 25 countries that experienced high inflationary episodes defined by inflation in excess of 100 per cent per annum, there is a nonlinear relationship between seigniorage and inflation, as depicted in Figure 6.9. As government deficits are financed by money creation, seigniorage rises but eventually declines due to the reduction in the tax base (real money balances) as the monetary base expands.

The first paper to empirically study seigniorage was by Phillip Cagan in 1956 where he restricted himself to hyperinflations during the period 1920–1946. A hyperinflation was defined by Cagan as inflation in excess of 50 per cent per month for at least a year. Cagan found that during these hyperinflationary episodes in countries such as Russia, Germany, Poland, Hungary, and Greece, the actual rate of monetary growth was in excess of the seigniorage maximizing rate of money growth—to the right of point E′ in Figure 6.9. Between 1947 and 1984, there were no hyperinflations in the world economy. Since 1984 there have been at least seven countries that witnessed hyperinflation episodes. The Nicaraguan hyperinflation that occurred over a period of 58 months between June 1986 and March 1991 was the worst among these seven countries[21] with a cumulative inflation of 11,895 billion.

Such phenomena whereby the government tries to finance a deficit by increasing the money supply at a rate that is higher than the seigniorage-maximizing rate arises because the public adjusts by reducing its holdings of real money balances. Sargent[22] shows how after World War I, Austria printed money at extremely high rates to finance its government deficit. Austrian notes in circulation increased by over 70 per cent from July to August 1922. This rapid increase in money creation led to annual inflation rates that approached 10,000 per cent per year, and people responded by reducing their holdings of real money balances drastically. Real money balances fell as the government resorted to increasing the notes in circulation since individuals expected inflation to rise, and they sought to avoid capital losses on the holding of money balances.

### *6.5.2* Alternative Means of Obtaining Seigniorage

Governments enhance seigniorage revenues through various policies. One such policy is the imposition of reserve requirements on banks, which compels banks to hold more base money than they otherwise would, and, thus, increases the real demand for base money. The impact of such a policy can be understood by assuming that the currency held by the public is zero and the public holds all money only in the form of deposits.[23] Thus, all base money is held by the banks. Base money,[24] in this case, is given by $xM^s$ with a required reserve ratio of $x$. Assume that banks intermediate with zero operating costs and that perfect competition prevails. The interest on deposits in banks $i_D$ must be given by

$$i_D = i_L(1 - x) \tag{6.24}$$

This is because a fraction, $x$, of deposits must be held by banks as the required reserves of non-interest-bearing base money, and so banks can invest and competitively pass earnings back to customers from only the remaining $(1 - x)$ share of their deposits.

Now consider a bank that has a deposit base of INR 100 and let the bank voluntarily hold INR 10 in reserves for meeting the liquidity needs of depositors and INR 90 in interest-earning assets. Let the nominal interest rate be 10 per cent. The bank earns INR $(90 \times 0.1)$ = INR 9 in interest, which it passes through to depositors. The difference between the lending and deposit rates is 1 percentage or 100 basis points. If the government imposes an additional 10 per cent reserve requirement, the bank holds INR 20 in reserves and INR 80 in interest-earning assets that earn INR $(80 \times 0.1)$ = INR 8 in interest, which is passed on to depositors. The difference between the lending and deposit rates has increased from 1 to 2 percentage points.

Now suppose the government increases the rate of monetary expansion by 10 percentage points. Correspondingly, inflation will also increase by 10 percentage points as $\Delta M/M = \pi$ in a steady state, and the rise in inflation raises the nominal interest rate[25] from 10 to 20 per cent. If the additional government-mandated reserve requirement were not imposed, the bank would earn INR $(90 \times 0.2)$ = INR 18 in interest, which is passed on to the depositors. However, with the mandated reserve requirement, the bank earns INR $(80 \times 0.2)$ = INR 16 in interest, which is the interest that is paid on deposits. The increase in inflation in conjunction with the required reserves mandated by the government increases the difference between lending and deposit rate to 4 percentage points—up from 2 percentage points.

Increasing required reserve ratios at a given rate of monetary expansion forces banks to hold more base money and enhance their real demand for base money. This means that any given level of seigniorage can now be obtained at a reduced expansion rate of money supply and, hence, a reduced inflation rate. A given amount of seigniorage is obtained via a higher reserve requirement and lower inflation rate. However, at any given inflation rate seigniorage is now extracted by a lower interest paid to depositors rather than by a higher inflation rate.[26]

› **Seigniorage** revenues can be enhanced by imposing reserve requirements on banks and putting restrictions on substitutes for base money.

Governments also adopt other methods for increasing SEIGNIORAGE—putting restrictions on substitutes for base money, for instance, so as to increase the real demand for base money. By reducing the availability of close substitutes for base money, the government increases the real base money demand by lessening the inflation sensitivity of the demand for base money and thereby increases its real seigniorage. A close substitute for domestic base money can be foreign currency, and governments often restrict citizens from holding or doing business in foreign currencies. This imposition of restrictions on financial markets would be prevalent whenever there are seigniorage-motivated governments.

These are governments who would benefit from preventing large substitution from domestic to foreign currency since that shrinks the demand for domestic currency and reduces seigniorage. Economies that are dollarizing such as in Latin America would be good examples of this. A government may also restrict interest rates on demand deposits of various kinds. This again reduces the appeal of substitutes for the currency of the government and increases seigniorage.[27]

Click's[28] study of 90 countries for the period 1971–1990 found that only 10 depended on seigniorage for more than 20 per cent of government spending—countries such as Paraguay, Uganda, Burma, Israel, Argentina, and Yugoslavia. Most industrialized nations, including the United States, financed less than 3 per cent of government spending with seigniorage. The median country in the sample collected siegniorage equal to 1.7 per cent of the GDP and 7.8 per cent of government spending.

India's position was slightly above the median situation in that it collected seigniorage equal to 1.72 per cent of the GDP and 11.8 per cent of the government spending. Moreover, inflation rates in most countries seem to be far too low for us to consider seigniorage maximization as a reliable approach towards understanding the objectives of central banks and governments with regard to money supply. Governments might want to keep in reserve the ability to raise seigniorage in a time period when there is a peak need for revenue. This could happen during a war or famine. The government encourages larger money holding and secures a larger tax base, which can be taxed when there is a short-run need for revenues. For a seigniorage-motivated government there are limits to monetary expansion as the real base money stock accommodates itself to the real quantity of base money demanded through the adjustment of the price level. In a fiat-money system, according to the seigniorage approach, the equilibrium real-money stock is determined entirely on the demand side where the real quantity of base money demanded falls as the rate of monetary expansion rises.

## SUMMARY

- Money serves three functions—as a medium of exchange, a unit of account, and a store of value.
- The most important function of money is that of a medium of exchange—when it is used to purchase goods and services.
- Money, as a unit of account, is the yardstick for measuring the economic value of goods, services, and assets, and thereby enables a comparison of these based on their values.
- As a store of value, money is one way of holding wealth, which can be compared with other ways of holding wealth such as stocks, bonds, or real estate.
- A bond is a legal promise to repay a debt. The principal amount or the amount originally lent is paid at some fixed date in the future called the maturation date. A bondholder receives regular interest or coupon payments until the bond's maturation date.
- Bond prices and interest rates are inversely related.
- The net wealth of the private sector consists of its net claims on agents outside the sector—the stock of money and the stock of government bonds ($W_t = M_{t-1} + B_{t-1}/i_t$).
- The real interest rate is the percentage increase in the purchasing power of a financial asset.
- The nominal interest rate is the real rate of interest plus the rate of inflation.
- The Fisher effect is the tendency for nominal interest rates to be high when inflation is high, and for them to be low when inflation is low.
- The two-period budget constraint—asset accumulation is allowed—implies that the lifetime income is the present value of income in the two periods less the discounted value of interest lost on account of holding money rather than bonds.

$$Y_1 + \frac{Y_2}{1+r} - \frac{i_1}{1+i_1}\frac{M_1}{P_1} = C_1 + \frac{C_2}{1+r}$$

- An individual decides on the amount of real money balances to hold by maximizing their preference for consumption and real money balances, subject to the intertemporal budget constraint.
- The demand for real money balances rises with income and declines with a rise in the interest rate.
- When trading off risk versus returns, while deciding on the composition of the assets in their portfolio, the agent decreases the allocation to money as the interest rate rises up to an acceptable level of risk. As the risk keeps on increasing, at some point the individual seeks to reduce their exposure to risk by increasing the allocation to money even though the interest rates increase.
- The use of money creation as a revenue device by the government is called seigniorage ($s = \Delta M_0 / P$).
- As the government increases the growth rate of money in circulation (the tax rate), individuals respond by reducing their holding of real money balances (the tax base) so that the rate of increase in seigniorage diminishes eventually.
- Seigniorage revenues can be enhanced by imposing reserve requirements on banks and putting restrictions on substitutes for base money.

## NOTES

1. If all goods are expressed in terms of money and money is the unit of account, then only *n* different absolute prices for the different goods need to be recorded. If $P_i (i = 1, \dots, n)$ denotes absolute price for good *i,* then the implied relatives prices are $P_i / P_j$. If there was no unit of account, in an economy with *n* goods we would need to record $n(n-1)/2$ distinct relative prices which can be a fairly large number.
2. The present value of the amount INR 1 that is to be received at time *t* is given by $\exp\{-\int_0^t r(s)ds\}$.
3. We are assuming the market value of the firms' debt and equity just equal the market value of their assets so that the net worth of the firm is zero. Conventional corporate finance theory often invokes this assumption.
4. Similarly, whether government bonds constitute private net wealth is a debated issue that is unresolved. Here we treat government bonds as private wealth and later discuss this issue again while discussing Ricardian equivalence.
5. The accumulation of financial wealth takes place by agents deciding to save rather than consume. Hence, $\Delta W/P = S = Y - C$. This is akin to stating that agents accumulate wealth by saving and adding to their stocks of money and bonds.
6. Recall that $P_{B,t} = B_{t-1}/i_t$. Hence, $(B_{t-1}/i_t) - (B_{t-1}/i_{t-1}) = P_{B,t} - P_{B,t-1}$ is the change in the price of the bond or the capital gains.
7. For *W* constant, $\Delta M + \Delta(B/i) = 0$ or, $\Delta(B/i) = -\Delta M$.
8. Scott Freeman , "Transactions Costs and the Optimal Quantity of Money," *Journal of Political Economy* 93 (February 1985), 146–157.
9. In a general version of the transactions demand for money such as shopping cost models, utility is a function of consumption and leisure in each time period. Money reduces transaction cost associated with exchange and thus increases leisure. See Bennett McCallum, *Monetary Economics: Theory and Policy* (New York: Macmillan, 1989). Robert Feenstra demonstrated that there exists a functional equivalence between models with money in the utility function and models where money affects "liquidity costs" which in turn show up in the budget restriction. For more, see Robert Feenstra, "Functional Equivalence Between Liquidity Costs and the Utility of Money," *Journal of Monetary Economics* 17 (1986): 271–291.
10. Saving as pointed out earlier is the change in wealth given by $S = \Delta B/iP + \Delta M/P$
11. See Appendix 6.1 for the formal solution to this problem.
12. In an inflationary environment the yield on money is $1 + r^M = 1/(1+\pi)$, where $\pi$ is the inflation rate. This can be seen by writing the nominal budget constraint as $P_tY_t + M_{t-1} + B_{t-1} + (B_{t-1}/i_{t-1}) = P_tC_t + M_t + (B_t/i_t)$. Dividing throughout by $P_t$, and defining inflation by $\pi_t = (P_t/P_{t-1}) - 1$ we get, $Y_t + [b_{t-1}(1+\pi_t)] = C_t + \{m_t - [m_{t-1}/(1+\pi_t)]\} + \{(b_t/i_t) - (b_{t-1}/[i_{t-1}(1+\pi_t)])\}$ where $m_t = M_t/P_t$ and $b_t = B_t/P_t$.
13. M. J. Manohar Rao, "The Demand for Money in India," in K. L. Krishna (ed.) *Econometric Applications in India,* Oxford India Themes in Economics (New Delhi: Oxford University Press, 1997).
14. B. Bhaskara Rao and Rup Singh, "Demand for Money in India: 1953–2003," *Applied Economics* 38, no. 11 (2006): 1319–1326.
15. Note that if it were a perfectly competitive minting industry competition would ensure price equals marginal cost or $M = PQ + C$, and $S = 0$. A royal mint can only earn seigniorage if it has a legally protected monopoly.
16. Well, not quite. Printing costs are typically 2 to 3 per cent of smaller denominations of paper currency.
17. Interest on that debt now goes to the central bank which essentially rebates it back to the Ministry of Finance.
18. Otherwise, the money supply is the money multiplier times base money.
19. More technically, it is a rectangular hyperbola.
20. Stanley Fischer, Ratna Sahay and Carlos A. Végh, "Modern Hyper- and High Inflations," *Journal of Economic Literature* (September 2002): 837–880.
21. Ibid.
22. Thomas J. Sargent, "The Ends of Four Big Inflations," in Robert E. Hall (ed.) *Rational Expectations and Inflation* (New York: Harper & Row, 1986).
23. For the case where the public holds both currency and deposits see Dwyer and Saving (1986). Gerald P. Dwyer Jr. and Thomas R. Saving, "Government Revenue from Money Creation with Government and Private Money," *Journal of Monetary Economics* 17 (1986): 239–249.
24. See the section on the money multiplier in Chapter 2. There, with no currency holdings by the public $\theta = 0$, and money supply *is* $M^s = (1/x)M_0$.
25. Recall that $i = r + \pi$.
26. Essentially the higher tax rate on deposits in the form of a reduction in interest paid on deposits makes the demand for deposits (money) more sensitive to the inflation rate.
27. See Donald A. Nichols, "Some Principles of Inflationary Finance," *Journal of Political Economy* 82 (1974): 423–430.
28. Reid W. Click, "Seigniorage in a Cross-Section of Countries," *Journal of Money, Credit and Banking* 30 (May 1998), 154–171.

## TEST YOURSELF

1. Prices were quoted in cowrie shells in medieval India for payments related to trade in food, clothing, and other services. How did cowrie shells fulfil the three functions of money?
2. You own a perpetuity that pays you INR 5,000 as a coupon. You plan to sell the bond to finance a holiday. Can you expect to receive INR 5,000 in the bond market for the bond? Explain why.
3. How does the demand for money depend on the nominal interest rate and income?
4. The cash that you hold earns you no interest. What is the real rate of return to holding cash in your pocket?
5. Derive the two-period budget constraint for an individual whose wealth can be held in money or perpetual bonds.
6. Why is the demand for money stated in real terms ($M/P$) rather than in nominal terms?
7. How does an individual decide to allocate their portfolio between bonds and money when the latter yields a risky return?
8. Why does the government expand the stock of base money in the economy?
9. Explain the response of an individual's demand for real money balances when the government increases the rate of monetary expansion in the economy.
10. What are the other means available to the government for raising seigniorage revenues besides resorting to increasing the notes in circulation?

## ONLINE APPLICATION

1. Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
2. Click on the "Database" icon.
3. Under the classification of Annual data, click on the "Handbook of Statistics on Indian Economy".
4. Click on "Table191: Yield of SGL Transactions in Treasury Bills for Residual Maturities". Look at the data for treasury bills with a term to maturity of 15–91 days. Create a spreadsheet of monthly data on yields for this maturity of treasury bills for the period April 1996–1997 to March 2005–2006. This data give you the nominal interest rates for the period.
5. Click on "Table 251: Wholesale Price Index—Annual Variation" and create a data column with the data for all commodities for the period April 1996–1997 to March 2005–2006. These data are the inflation rate for the period.
6. Create a graph depicting the nominal interest rate available from the data in Step 4 and the inflation rate available from the data in Step 5. Do the nominal interest rate and the rate of inflation graphs have the same pattern? (You could find the correlation coefficient between the nominal interest rate and inflation as well to get a number that tells you the association between inflation and the nominal interest rate.)
7. Are inflation and the nominal interest rate closely related? Why or why not?

# Appendix 6.1 Preference Maximizing Choice of Consumption and Real Money Balances

More generally, the agent maximizes an intertemporal utility function subject to the constraint that consumption plus saving is less than or equal to income (inclusive of interest payments, $B_{t-1}/P_t$), or,

$$\left.\begin{aligned} &\text{Max} \sum_{t=0}^{\infty} \frac{1}{(1+\delta)^t} U\left(C_t, \frac{M_t}{P_t}\right) \\ &\text{subject to } C_t + \frac{M_t}{P_t} + \frac{B_t}{i_t P_t} \leq Y_t + \frac{B_{t-1}}{P_t} + \frac{M_{t-1}}{P_t} + \frac{B_{t-1}}{i_{t-1} P_t} \end{aligned}\right\} \quad \textbf{(A6.1.1)}$$

The Langrangean for this is

$$L = \frac{1}{(1+\delta)^t} U\left(C_t, \frac{M_t}{P_t}\right) + \lambda_t \left(Y_t + \frac{B_{t-1}}{P_t} + \frac{M_{t-1}}{P_t} + \frac{B_{t-1}}{i_{t-1} P_t} - C_t - \frac{M_t}{P_t} - \frac{B_t}{i_t P_t}\right)$$

The first-order conditions are

$$\frac{\partial L}{\partial (M_t/P_t)} = \left(\frac{1}{1+\delta}\right)^t U_M - \lambda_t + \lambda_{t+1} \frac{P_t}{P_{t+1}} = 0$$

$$\frac{\partial L}{\partial C_t} = \left(\frac{1}{1+\delta}\right)^t U_C - \lambda_t = 0$$

$$\frac{\partial L}{\partial (B_t/i_t P_t)} = -\lambda_t + \lambda_{t+1} \left[\frac{i_t P_t}{P_{t+1}} + \frac{P_t}{P_{t+1}}\right] = 0$$

These can be rewritten as:

$$\left(\frac{1}{1+\delta}\right)^t U_M = \lambda_t - \lambda_{t+1} \frac{P_t}{P_{t+1}}$$

$$\left(\frac{1}{1+\delta}\right)^t U_c = \lambda_t$$

$$\lambda_t = \lambda_{t+1}(1+i_t) \frac{P_t}{P_{t+1}}$$

Substituting the third first-order condition into the first,

$$\left(\frac{1}{1+\delta}\right)^t U_M = \lambda_t - \lambda_{t+1} \frac{P_t}{P_{t+1}} = \lambda_t - \frac{\lambda_t}{1+i_t} = \lambda_t \left(\frac{i_t}{1+i_t}\right)$$

Dividing this expression by the second first-order condition,

$$\frac{U_M}{U_C} = \frac{i_t}{1+i_t}$$

the slope of the indifference curve equals the slope of the budget line as depicted in Figure 6.1. If we now lead the second first-order condition by one period:

$$\left(\frac{1}{1+\delta}\right)^{t+1} U_{C_{t+1}} = \lambda_{t+1}$$

Hence,

$$\frac{(U_{C_{t+1}})/(1+\delta)^{t+1}}{(U_{C_t})/(1+\delta)^{t}} = \frac{\lambda_{t+1}}{\lambda_t}$$

or,

$$\frac{1}{(1+\delta)}\frac{U_{C_{t+1}}}{U_{C_t}} = \frac{\lambda_{t+1}}{(1+i_t)(P_t/P_{t+1})\lambda_{t+1}}$$

However, we know that

$$1 + r_t = \frac{P_t(1+i_t)}{P_{t+1}}$$

This allows us to rewrite the previous expression as:

$$\frac{1}{(1+\delta)}\frac{U_{C_{t+1}}}{U_t} = \frac{1}{(1+r_t)}$$

or,

$$\frac{U_{C_{t+1}}}{U_t} = \frac{1+\delta}{(1+r_t)}$$

This is the intertemporal Euler equation for consumption and is equivalent to Eq. (3.16) derived for the optimal consumption in Chapter 3. Thus, including money in the utility function retains the optimal consumption decision—the marginal rate of substitution between consumption across periods equals the rate of time preference divided by the rate of interest. Now, however, the individual in addition has to ensure that the marginal rate of substitution between consumption and real balances of money is equal to its opportunity cost, which is the present discounted value of the interest rate on bonds, $i_t/(1+i_t)$.

# *Appendix 6.2* The Two-Period Budget Constraint

Derivation of Eq. (6.12):
Adding the first- and second-period budget constraint we obtained

$$(P_1Y_1) + \frac{P_2Y_2}{(1+i_1)} = P_1C_1 + \frac{P_2C_2}{(1+i_1)} + \left[M_1 - \frac{M_1}{(1+i_1)}\right] + \left[\frac{B_1}{i_1} - \frac{B_1}{(1+i_1)} - \frac{B_1}{(1+i_1)i_1}\right]$$

or,

$$(P_1Y_1) + \frac{P_2Y_2}{(1+i_1)} = P_1C_1 + \frac{P_2C_2}{(1+i_1)} + \left[\frac{(1+i_1)M_1 - M_1}{(1+i_1)}\right] + \left[\frac{(1+i_1)B_1 - i_1 B_1 - B_1}{(1+i_1)i_1}\right]$$

or,

$$(P_1Y_1) + \frac{P_2Y_2}{(1+i_1)} = P_1C_1 + \frac{P_2C_2}{(1+i_1)} + \frac{i_1M_1}{(1+i_1)}$$

Dividing throughout by $P_1$, we have

$$Y_1 + \frac{P_2Y_2}{P_1(1+i_1)} = C_1 + \frac{P_2C_2}{P_1(1+i_1)} + \frac{i_1}{(1+i_1)}\frac{M_1}{P_1}$$

Now, recall that the real interest rate is given by

$$1 + r_1 = \frac{P_1(1+i_1)}{P_2}$$

This implies that $P_2/[P_1(1+i_1)] = 1/(1+r_1)$, which we substitute into the above budget constraint to obtain

$$Y_1 + \frac{Y_2}{(1+r_1)} = C_1 + \frac{C_2}{(1+r_1)} + \frac{i_1}{(1+i_1)}\frac{M_1}{P_1}$$

or,

$$Y_1 + \frac{Y_2}{(1+r_1)} - \frac{i_1}{(1+i_1)}\frac{M_1}{P_1} = C_1 + \frac{C_2}{(1+r_1)} \qquad \textbf{(A6.2.1)}$$

# Appendix 6.3 Seigniorage and the Real Money Demand

We can write the expression for seigniorage as

$$s = \frac{M_{t+1} - M_t}{P_{t+1}}$$

Subtracting and adding $M_t / P_t$ on the right-hand side,

$$s = \left(\frac{M_{t+1}}{P_{t+1}}\right) - \frac{M_t}{P_t} + \frac{M_t}{P_t} - \left(\frac{P_t}{P_{t+1}}\frac{M_t}{P_t}\right)$$

$$= \Delta\left(\frac{M}{P}\right) + \left(1 - \frac{P_t}{P_{t+1}}\right)\frac{M_t}{P_t}$$

where, the increase in the real stock of money is

$$\Delta\left(\frac{M}{P}\right) = \frac{M_{t+1}}{P_{t+1}} - \frac{M_t}{P_t}$$

Now, the rate of inflation is given by

$$\pi = \frac{P_{t+1} - P_t}{P_t} = \frac{P_{t+1}}{P_t} - 1$$

Thus,

$$1 - \frac{P_t}{P_{t+1}} = \frac{\pi}{1 + \pi}$$

allowing us to write

$$s = \Delta\left(\frac{M}{P}\right) + \left(\frac{\pi}{1 + \pi}\right)\frac{M_t}{P_t} \qquad \textbf{(A6.3.1)}$$

This equation states that seigniorage is the sum of the increase in the real stock of money plus the change in the real money stock. This change would have occurred with a constant nominal money stock due to inflation $(\pi/ 1 + \pi)$ $M_t / P_t$. This second term is referred to as the inflation tax with $(\pi/ 1 + \pi)$ denoting the tax rate and $M_t/ P_t$ the tax base. Seigniorage may also be written as:

$$s = \frac{M_{t+1} - M_t}{M_t}\frac{M_t}{P_{t+1}} = \mu\frac{M_t}{P_{t+1}} \qquad \textbf{(A6.3.2)}$$

where, $\mu = (M_{t+1} - M_t)/M_t$ is the growth rate of the money supply.

We analyse seigniorage by abstracting from a fractional reserve system so that the stock of money in the economy is the monetary base, $M = M_0$.

Recall that the individual at time $t$ maximizes the utility from consumption and the real stock of money $U[C_t, (M_t/P_t)]$ subject to their intertemporal budget constraint—Eq. (6.14). The maximum utility is attained where the slope of the indifference curve equals the slope of the budget line, or

$$\frac{U_{M/P}[C_t, (M_t/P_t)]}{U_c C_t, (M_t/P_t)} = \frac{i_t}{1+i_t} \qquad \textbf{(A6.3.3)}$$

where, $U_{M/P}$ is the incremental (or marginal) utility of a unit increase in real money holding, $M/P$; and

$U_c$ is the marginal utility of consumption.

This may be rewritten (see the derivation at the end of this appendix) on the basis of the relationship between the real and nominal interest rates as

$$\frac{U_{M/P}[C_t, (M_t/P_t)]}{U_c C_t, (M_t/P_t)} = 1 - \frac{(P_t/P_{t+1})}{1+r} \qquad \textbf{(A6.3.4)}$$

where, as before, the right-hand side is the cost for a period in terms of forgoing a unit of consumption good and holding a unit of real money balances. To see this, consider what happens when an individual forgoes a unit of consumption on date $t$. This gives the individual $P_t$ units of money which when carried over to the next period $t + 1$ is worth $P_t/P_{t+1}$ due to inflation. The present value of $P_t/P_{t+1}$ units of currency is $(P_t/P_{t+1})/(1 + r)$. Hence, forgoing a unit of consumption on date $t$ gives $(P_t/P_{t+1})/(1 + r)$ units of consumption on date $t + 1$ in present value terms. The net cost to forgoing a unit of consumption and holding real money balances instead is $1 - [(P_t/P_{t+1})/(1 + r)]$.

So far, we have focused on the demand for real money balances. The individual in an intertemporal world is also interested in deciding on how much of the income to consume today and how much to transfer to another period's consumption. In Chapter 3, we saw that the condition for optimal consumption [Eq. (3.17)] is

$$\frac{\partial U/\partial C_t}{\partial U/\partial C_{t+1}} = \frac{1+r}{1+\delta}$$

or,

$$\frac{C_{t+1}}{C_t} = \frac{1+r}{1+\delta}$$

where, $\partial U/\partial C_{t+1} = 1/C_{t+1}$ and $\partial U/\partial C_t = 1/C_t$. We focus on the case of consumption smoothing where $r = \delta$. Recall that the slope of an indifference curve was given by

$$-\frac{dC_{t+1}}{dC_t} = (1+\delta)\frac{\partial U/\partial C_t}{\partial U/\partial C_{t+1}} = (1+\delta)\frac{C_{t+1}}{C_t}$$

Hence, when $r = \delta$,

$$\frac{\partial U}{\partial C_{t+1}} = \frac{\partial U}{\partial C_t}$$

With consumption the same across the two-time periods ($C_{t+1} = C_t$) under consumption smoothing, the marginal utility of consumption must also be equal. Let

$$\frac{\partial U}{\partial C_{t+1}} = \frac{1}{C_{t+1}} = \frac{\partial U}{\partial C_t} = \frac{1}{C_t} = \frac{1}{\bar{C}}$$

Then,
$$U_C = \frac{1}{\bar{C}}$$

Accordingly, Eq. (A 6.3.4) can be written as

$$\frac{M_{t+1}}{P_{t+1}} \frac{M_t}{M_{t+1}} \frac{1}{1+r} = \frac{M_t}{P_t}\left(1 - \bar{C}U_{M/P}\right) \tag{A6.3.5}$$

Suppose the government increases the money supply at the constant growth rate $\mu$ [Eq. (6.24)], where, $1 + \mu = M_{t+1}/M_t$. Then, when the utility function is additive and given by

$$U\left(C_t, \frac{M_t}{P_t}\right) = \log C + \log \frac{M}{P}$$

we get the following expression for the marginal utility of money:

$$U_{M/P} = \frac{1}{M/P} \tag{A6.3.6}$$

Substituting Eq. (A6.3.6) into Eq. (A6.3.5) (see the derivation at the end of this appendix) gives us

or,
$$\frac{M_{t+1}}{P_{t+1}} = (1 + \mu)(1 + r)\left(\frac{M_t}{P_t} - \bar{C}\right) \tag{A6.3.7}$$

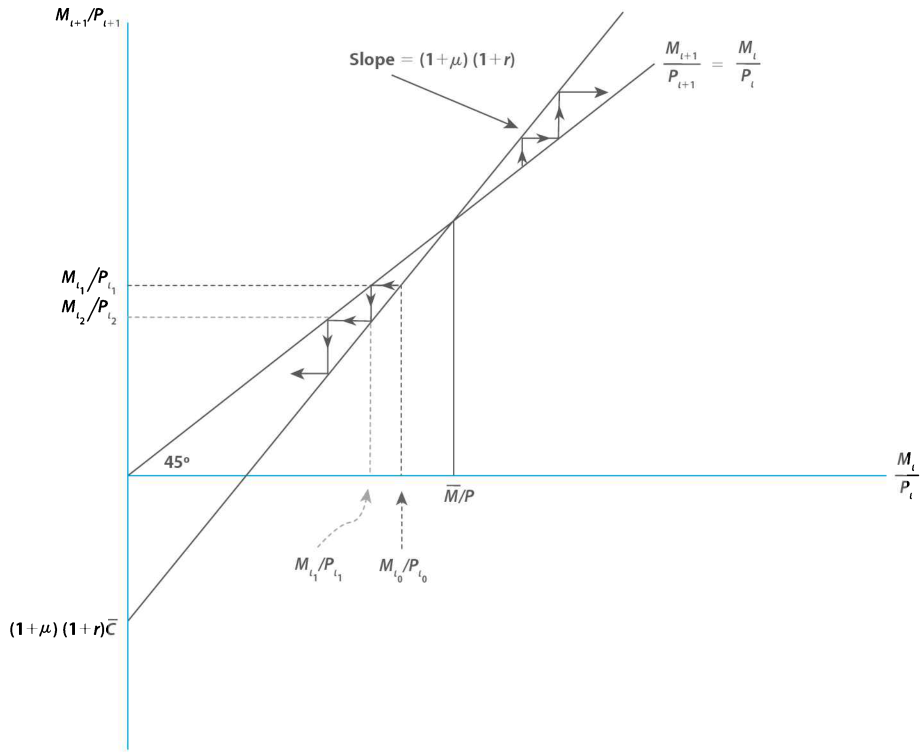
This reveals how the holding of real money balances evolves over time. The graph of this equation is given in Figure A6.3.1.

If the individual were to allocate their real money balances such that it is stationary over time (just as $C_{t+1} = C_t = \bar{C}$ is stationary), this occurs (see Figure 6.8) where

$$\frac{M_{t+1}}{P_{t+1}} = \frac{M_t}{P_t} = \frac{\overline{M}}{P}$$

or, solving for real money balance holdings,

$$\frac{\overline{M}}{P} = \bar{C}\left[\frac{(1 + r)(1 + \mu)}{(1 + r)(1 + \mu) - 1}\right] \tag{A6.3.8}$$




**› Figure A6.3.1**
Evolution of Real Money Balance Holding. The stock of real money balances may decline over time to the left of point A if there are capital losses from the holding of money balances. The stock of money balances increases with time to the right of point A as capital gains are associated with those holdings. At point A there is no change in the holding of money balances with time and this characterizes the steady state holding of money balances.

Suppose the economy was not at point A where the real money balances are given by $\overline{M}/P$. For instance, suppose at time $t_0$, real money balances were $M_{t_0}/P_{t_0}$. In the next period, $t_1$, real money balances will be $M_{t1}/P_{t1}$. Use the 45-degree line to project this period $t+1$ balance back on to the horizontal axis. We can see that

$$\frac{M_{t_1}}{P_{t_1}} < \frac{M_{t_0}}{P_{t_0}}$$

The real money balances held have, therefore, declined. In fact, at time $t_2$ the real money balances that are held are even lower, as indicated by the arrows in the figure. Similarly, if the stock of real money balances held was greater than $\overline{M}/P$ at time $t_0$, real money balances holding would increase without bounds as indicated in Figure 6.3.1. Hence, if the economy is not at $\overline{M}/P$, the holding of real money balances diverges further and further away from $\overline{M}/P$ as time goes by.

When the economy is characterized by a point to the right of A in the diagram, the demand for real money balances keeps increasing with time $M_{t+1}/P_{t+1} > M_t/P$. This occurs when individuals expect to receive capital gains on the holding of this money, which implies that there is an expectation of a reduction in the rate of inflation. However, any reduction in inflation adds to the holding of real balances. This would increase the expectation of capital gain and for this expectation not to be contradicted it would require individuals to expect a more rapid decline in inflation the following period. Thus, there is an unstable upward path for the holding of real money balances to the right and left of point A. We, thus, have a large number of unstable solutions possible to Eq. (A6.3.5).

We concentrate on the solution at point A—assume the economy always goes to the stationary state where real money balances are given by $\overline{M}/P$ and consumption expenditure by $\overline{C}$. Recall the demand for money is given by Eq. (A6.3.4) that we rewrite here

$$\frac{U_{M/P}\left[C_t, (M_t/P_t)\right]}{U_C\left[C_t, (M_t/P_t)\right]} = 1 - \frac{P_t/P_{t+1}}{1+r} \tag{A6.3.9}$$

In a steady state where both real balances and consumption are constant, $U_{M/P} = 1(\overline{M}/P)$ and $U_C = 1/\overline{C}$, so that

$$\frac{1/(\overline{M}/P)}{1/\overline{C}} = 1 - \frac{P_t/P_{t+1}}{1+r}$$

or,

$$\frac{\overline{C}}{\overline{M}/P} = 1 - \frac{P_t/P_{t+1}}{1+r}$$

However, $P_t/P_{t+1} = 1/(1+\pi)$. Hence,

$$\frac{\overline{M}}{P} = \overline{C}\left[\frac{(1+r)(1+\pi)}{(1+r)(1+\pi)-1}\right] \tag{A6.3.10}$$

Comparing Eq. (A6.3.10) with Eq. (A6.3.8) we see that

$$\mu = \pi$$

or,

$$\frac{\Delta M}{M} = \pi$$

In a steady state, then, the rate of growth of the money supply is equal to the rate of inflation. This conclusion is also reached by considering that in a steady state the holding of real money balances is a constant, $k$ , or,

$$\frac{M}{P} = k$$

Hence,

$$\Delta\left(\frac{M}{P}\right) = \Delta k = 0$$

or,

$$\frac{M}{P} = \left(\frac{\Delta M}{M} - \frac{\Delta P}{P}\right) = 0$$

The left-hand side of the above is zero provided that

$$\frac{\Delta M}{M} = \frac{\Delta P}{P} = \pi$$

A steady state then results in the condition that the rate of monetary expansion equals the inflation rate.

The response of real money balance holdings to an increase in inflation is given by

$$\frac{\partial(M/P)}{\partial\pi} = -\frac{\overline{C}(1+r)}{[(1+r)(1+\pi)-1]^2} < 0 \tag{A6.3.11}$$

Money demand, therefore, declines with a rise in inflation. The rate of decline is given by

$$\frac{\partial^2(M/P)}{\partial \pi^2} = -\frac{2\bar{C}(1+r)}{[(1+r)(1+\pi)-1]} < 0 \tag{A6.3.12}$$

This implies a relationship between the demand for real money balances and inflation as depicted in Figure 6.8.

Derivation of Eq. (A6.3.4):

We had earlier (Section 6.2) mentioned the relationship between the real and nominal interest rates as

$$1 + r = \frac{P_t(1+i_t)}{P_{t+1}}$$

or,

$$1 + i_t = (1+r)\frac{P_{t+1}}{P_t}$$

Hence,

$$\frac{i_t}{1+i_t} = \frac{(1+r)[(P_{t+1}/P_t) - 1]}{(1+r)(P_{t+1}/P_t)}$$

$$= 1 - \frac{P_t/P_{t+1}}{1+r}$$

Then, we may write

$$\frac{U_{M/P}[C_t, (M_t/P_t)]}{U_C[C_t, (M_t/P_t)]} = 1 - \frac{P_t/P_{t+1}}{1+r} \tag{A6.3.4}$$

Derivation of Eq. (A6.3.5):

Under consumption smoothing Eq. (A6.3.4) is

$$\frac{U_{M/P}[C_t, (M_t/P_t)]}{1/\bar{C}} = 1 - \frac{P_t/P_{t+1}}{1+r}$$

or,

$$\bar{C}U_{M/P} = 1 - \frac{P_t/P_{t+1}}{1+r}$$

or,

$$\frac{P_t/P_{t+1}}{1+r} = 1 - \bar{C}U_{M/P}$$

Multiply throughout by $M_t/P_t$

$$\frac{M_t}{P_t}\frac{P_t}{P_{t+1}(1+r)} = \frac{M_t}{P_t}(1 - \bar{C}U_{M/P})$$

or,
$$\frac{M_{t+1}}{P_{t+1}}\frac{M_t}{M_{t+1}}\frac{1}{1+r}=\frac{M_t}{P_t}(1-\bar{C}U_{M/P}) \tag{A6.3.5}$$

Derivation of Eq. (A6.3.7):

Substitute $1+\mu = M_{t+1}/M_t$ and $U_{M/P} = 1/M/P$ in Eq. (A6.3.5)

$$\frac{M_{t+1}}{P_{t+1}}\frac{1}{(1+\mu)(1+r)}=\frac{M_t}{P_t}-\bar{C}$$

or,
$$\frac{M_{t+1}}{P_{t+1}}=(1+\mu)(1+r)\left(\frac{M_t}{P_t}-\bar{C}\right) \tag{A6.3.7}$$

Derivation of Eq. (A6.3.8):

Set $M_{t+1}/P_{t+1} = M_t/P_t = \overline{M}/P$ in Eq. (A6.3.7)

$$\frac{\overline{M}}{P}=(1+\mu)(1+r)\left(\frac{\overline{M}}{P}-\bar{C}\right)$$

or,
$$\frac{\overline{M}}{P}\left\{\frac{[(1+\mu)(1+r)]-1}{(1+\mu)(1+r)}\right\}=\bar{C}$$

or,
$$\frac{\overline{M}}{P}=\bar{C}\left\{\frac{(1+r)(1+\mu)}{[(1+r)(1+\mu)]-1}\right\} \tag{A6.3.8}$$

# *Appendix 6.4* Currency Crises and Speculative Attacks

A *speculative attack* is a process by which investors change the composition of their portfolios—reducing the proportion of domestic currency and raising the proportion of foreign currency. The change in composition is justified by a change in relative yields. The change in yields occurs as the government is no longer able to defend the exchange rate and the currency begins depreciating. A model for the balance-of-payments crisis must have two characteristics:

1. The demand for domestic currency depends on the exchange rate,
2. The exchange rate that clears the domestic-money market changes over time.

We will look at a simplified model of the currency crisis when the exchange rate is fixed or predetermined. We make four assumptions:

1. The demand for money is a demand for transactions purposes as asset markets are relatively underdeveloped. Alternatively, the interest rate is that given by the interest rate prevailing in the rest of the world which is then fixed via the interest parity condition. Hence,

$$\frac{M^d}{P} = f(Y)$$

or,

$$M^d = kPY$$

where, $k$ is a constant.
2. In the long run, output and employment tend towards their full employment or natural levels. Thus, for the purpose of analysis, $Y$ can be treated as fixed or known.
3. The authorities cannot neutralize the monetary impact of the balance-of-payments deficit on the money supply in the long run.
4. After due allowance for tariffs and transport costs, arbitrage will ensure that the prices of similarly traded goods will tend to be equalized in the long run. Thus, with $P$, the domestic-price level, $P^*$ the foreign level, and $E$ the spot exchange rate, $P = P^*E$.

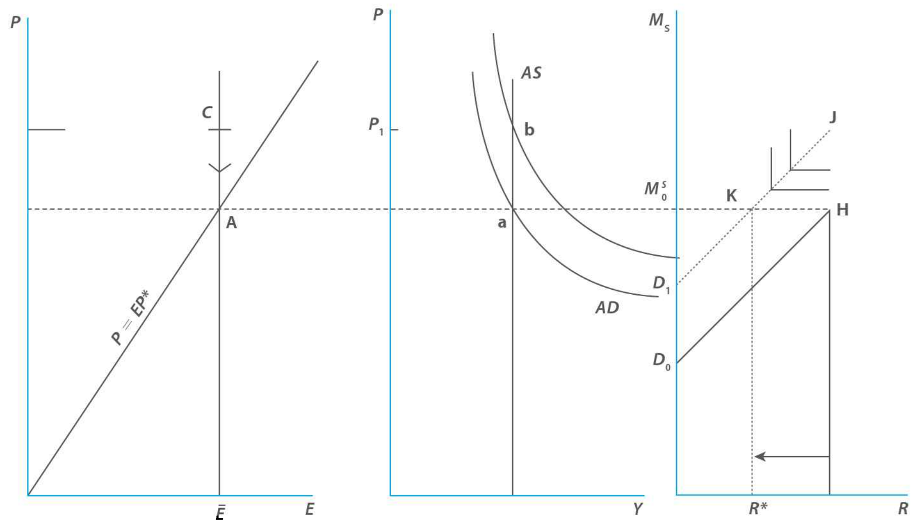
With given foreign prices and the exchange rate fixed or pegged by the government, the domestic-price level is also constant. If the exchange rate is allowed to be market determined, any movement in the exchange rate is automatically reflected in domestic prices.

The supply of money in the absence of a commercial banking system is domestic credit, $D$, plus international reserves, $R$.

$$M^S = D + R$$

In a fixed exchange rate regime, money supply is not an exogenous variable. Instead, the policy variable is domestic credit so that the burden of adjustment to changes in exogenous variables falls on foreign exchange reserves, $R$.

We look at the effect of an expansionary monetary policy, undertaken, say, to finance the budgetary deficit. The third section of the Figure A6.4.1 plots the money stock ($M^s$) against levels of the foreign exchange reserves, $R$. Suppose the initial level of reserves is $R_0$ and domestic credit, $D_0$. Measure $D_0$ on the vertical axis and from that point, extend a 45-degree line. That ray shows how the money stock increases with additions to reserves, given the initial quantity of domestic credit, $D_0$. The money stock is initially $M_0^s$, at the starting point in the figure, H.




› **Figure A6.4.1**
Currency Crises. In a fixed exchange rate regime, an increase in domestic credit creation is exactly offset by a decline in foreign exchange reserves. Individuals with perfect foresight will launch a speculative attack on the currency to avoid windfall capital losses when the reserves run out and the exchange rate is allowed to float.

Now, in money market equilibrium, $M^s = M^d$, or, for $M^s = M_0^s$,

$$M_0^S = kPY$$

or,

$$PY = \frac{M_0^S}{k}$$

For a given value of the money stock, nominal income, $PY$, is a constant. The graph of this is a rectangular hyperbola. Any particular aggregate demand schedule such as $AD_0(M_0^s)$ whose equation is $Y = M_0^s / Pk$, is drawn for a given money stock. An increase in money stock shifts the demand curve out to the right in such a way that the increase in prices is in the same proportion as the increase in money supply.

The aggregate supply curve is vertical by virtue of that being the natural level of output—the level of output can change due to productivity increases and thrift.

Now, set the initial value of the money supply and price level at unity. $M_0^s = P_0 = 1$. This has the advantage that it allows us to translate the money stock value on the vertical axis of the third panel in the figure directly into the implied price level in the first and second panels.

The first panel plots the PPP line. The exchange rate, however, is pegged at $\bar{E}$ and so the economy must at all times be restricted to points that lie along the vertical line.

The system starts off from points A, a, and H in the three panels. Suppose there is a monetary expansion to $D_1$. The 45-degree line in the third panel shifts upward by the amount of the expansion and the money stock is now $M_1^s$ at point J.

In the second panel, the increase in money supply shifts the aggregate demand curve upwards to a price level $P_1$. From the PPP line it is clear that the higher price level has made the domestic economy uncompetitive as with unchanged foreign prices and a fixed exchange rate, $\bar{E}P^*/P$ has fallen and a real appreciation has occurred. At point C, there is an incentive for importing from abroad which must result in a deficit in the balance of trade. This deficit is financed by foreign exchange reserves as the authorities are not permitting the exchange rate to depreciate but are committed instead to selling foreign currency for the home currency which results in a loss of international reserves. Each successive reduction in the reserves tends to reduce the money supply, pushing the economy along the arrowed paths so that the system moves from point J to K, from b back to a, and from C down to A. The reinstatement of the original equilibrium is possible because the money supply has contracted to the point where it is once more at its initial level $M_0^s$ due to a loss of foreign exchange reserves to the extent KH.

The important thing to note is that as long as the government has reserves left, the domestic money supply is determined by the portfolio preferences of domestic residents. With the exchange rate fixed and output given, and perfect foresight, the demand for holding money is

$$M^d = kPY = k(\bar{E}P^*)\bar{Y}$$

$$\therefore \quad M^s = M^d \Rightarrow D + R = k(\bar{E}P^*)\bar{Y}$$

or,

$$D = k(\bar{E}P^*)\bar{Y} - R$$

and a rise in domestic credit creation is exactly offset by a fall in reserves. When the government issues more domestic money than the private sector is willing to hold, private investors can always withdraw the excess money from circulation by trading it for foreign money at the exchange window. Thus, a commitment to pegging the exchange rate is tantamount to having no control on the money supply and accepting the amount of domestic money supply that residents want. With a lower bound on reserves and domestic credit expanding, the fixed exchange rate regime cannot survive forever. When reserves run out, the exchange rate must be allowed to float.

With a positive rate of domestic credit growth, rational agents will anticipate that without speculation reserves will eventually fall to the lower bound and will, therefore, foresee the ultimate collapse of the system. In equilibrium, under perfect foresight, agents can never expect a *discrete jump* in the level of the exchange rate, since a jump would provide them with profitable arbitrage opportunities. Thus, arbitrage in the foreign exchange market requires that the exchange rate that prevails immediately after the attack is equal to the fixed rate prevailing at the time of the attack.

Formally, the time of collapse is where the "shadow floating rate" that reflects market fundamentals is equal to the prevailing fixed rate. The shadow floating rate is the exchange rate that would prevail with the current money stock if reserves had fallen to the minimum level and the exchange rate were allowed to float freely. The shadow floating rate is given by the PPP line in the first panel. When the fixed exchange rate is depreciated more than the shadow floating rate, the fixed-rate regime is viable—beyond that point the fixed rate is not sustainable. This is because if the shadow floating rate falls below the prevailing fixed rate, speculators would not profit from driving the government's stock of reserves to its lower bound and precipitating the adoption of a floating rate regime as that would require them to bear an instantaneous *capital loss* on their purchases of foreign currency. To the right of $\bar{E}$ the exchange rate is more depreciated, $E = E_+ > \bar{E}$ and so $M^s/E_+ < M^s/\bar{E}$ which implies a capital loss. (Note that we are writing $M^s/E$ instead of $M^s/EP^*$ as $P^*$ is fixed.)

On the other hand, if the shadow exchange rate is above the fixed rate, $\bar{E} = E_- < E$, speculators would experience an instantaneous capital gain as $M^s/E_- > M^s/\bar{E}$. Neither anticipated capital gains nor losses at an infinite rate are compatible with a perfect foresight equilibrium as speculators would compete with each other to eliminate such opportunities. This implies an *equilibrium attack* or a *crisis* where the arbitrage condition that the pre-attack fixed rate equals the post-attack floating rate holds. If investors correctly anticipate events, the reserves of the government must be eliminated by a speculative attack that enables investors to avoid windfall capital losses. For there to be no windfall capital loss, the speculative attack must not lead to a discrete change in the price level and so we must have $S = \bar{\bar{E}}$ at the instant of the collapse. The time of the collapse is determined by the intersection of the PPP and $\bar{\bar{E}}$ curves and the crisis occurs before the government would have run out of reserves in the absence of speculation. If the government announced that it would allow the currency to be market determined when reserves reach $R^*$ in the figure, the moment the domestic credit expansion results in a 45-degree line passing through point K such as $D_1J$, the speculative attack occurs. Prior to this attack, the money stock is constant, but its composition varies since domestic credit rises but reserves decline. In the post-collapse regime, the money stock equals the domestic credit plus the minimum level of reserves $R^*$ the government has, but does not offer foreign exchange at the discount window.

An economy with a balance-of-payments problem will, thus, pass through three stages:

1. A period of gradually declining reserves,
2. A sudden speculative attack, and
3. A post-crisis period during which the currency gradually depreciates.

There has long been an argument that since a flexible exchange rate may be subject to arbitrary speculative fluctuations, the exchange rate should be fixed. The currency crisis literature analysis shows that arbitrary speculative behaviour, identical in nature to that which may manifest itself under floating rates, can also render a fixed exchange rate collapse. Hence, "speculative behaviour is an economic force which is masked, not purged, by fixing of exchange rates".*

* Robert P. Flood and Peter M. Garber, "Collapsing Exchange Rate Regimes: Some Linear Examples," *Journal of International Economics* 17 (August 1984): 1–13.

# Appendix 6.5 The Inventory Demand for Money (Baumol–Tobin)

The demand for money for the purposes of transactions—people holding money because they want to use it for payments—was developed by Baumol* and Tobin.** It derives the optimal money holdings for an individual who receives an annual income in instalments which is then spent for the purchase of goods and services. Some individuals receive their income monthly—12 instalments—and some every fortnight—26 instalments. In what follows, we presume (for simplicity) that the income is received annually. Typically individuals receive their income in the form of a cheque which they then deposit into a bank in order that the amount may be credited to their bank account. The money stock in the bank account of an individual can command $PY$ worth of goods and services in nominal terms. The choice problem the individual faces is that while all receipts and payments are made in money, money bears no interest. However, there is an alternative asset—bonds—that generates an interest income. The individual can convert their money into bonds and earn interest income but there are transactions costs involved in exchanging money for bonds and vice versa. So the individual must choose the character and number of transactions involving the purchase and sale of bonds so as to maximize their interest receipts less the transactions costs to be paid. We introduce the following notation:

$x$ is the number of transactions involving bonds, B—purchases or sales—per payment period

$C$ is the fixed transaction cost per transaction in the purchase or sale of bonds.

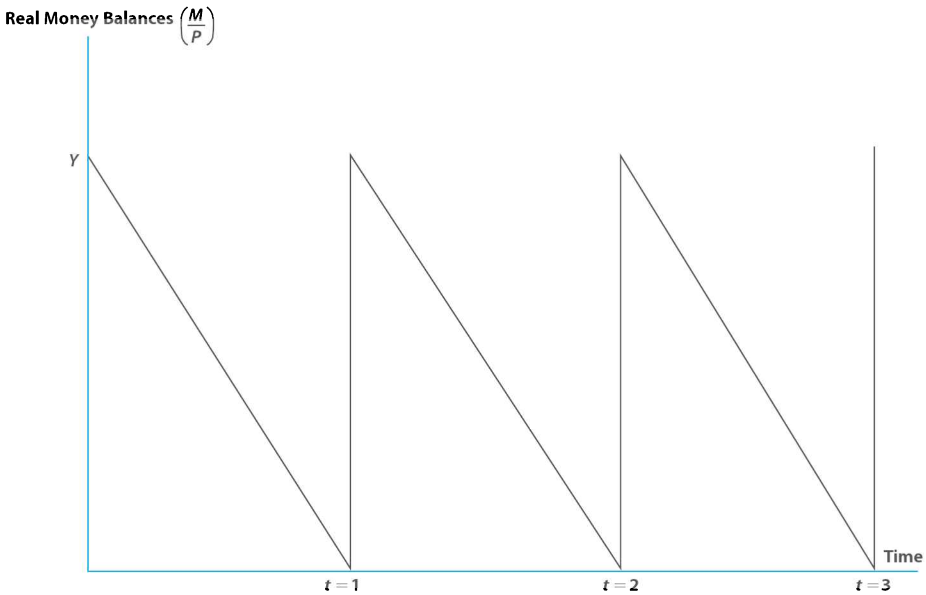
Transaction costs can be thought of as the costs of time spent in going to deposit and withdraw money from the bank, transport costs, and the various restrictions and penalties that may be imposed on withdrawal. Some of the transaction costs could also be psychic—nuisance value of dealing with brokers who purchase/sell bonds on your behalf—rather than explicitly pecuniary.

Suppose $x = 0$. Then, the individual holds only money balances and no bonds. At regular time intervals the individual receives $Y$ of income which is then used to spend on goods and services. As in each time period the individual will get paid their regular income and as we are assuming all income is for the purpose of expenditure it makes perfect sense for the individual to deplete their cash balance to zero the moment before income is received. Thus, the cash balance is $Y$ at the beginning of a time period and by the end of a time period it is depleted to zero. With expenditures being even and continuous (a simplifying assumption made) the cash balances must be depleted in a smooth manner throughout any unit time interval. This results in the profile of money balances, $M/P$, as given in Figure A6.5.1. The average cash balance the individual holds in a time interval, where they begin with real balances equal to $Y$ and ends with real balances of zero, is $(1/2)(Y + 0) = Y/2$.

Now, suppose the individual puts half of their income receipt into bonds as soon as they receive, the income; and is then left with half of their cash balances to spend, which as expenditure is even and continuous, will deplete half way through the time interval. Requiring cash balances half way through the income payment interval they would sell their bonds purchased earlier to get the money

*William J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* 66, no. 4 (November 1952): 545–556.

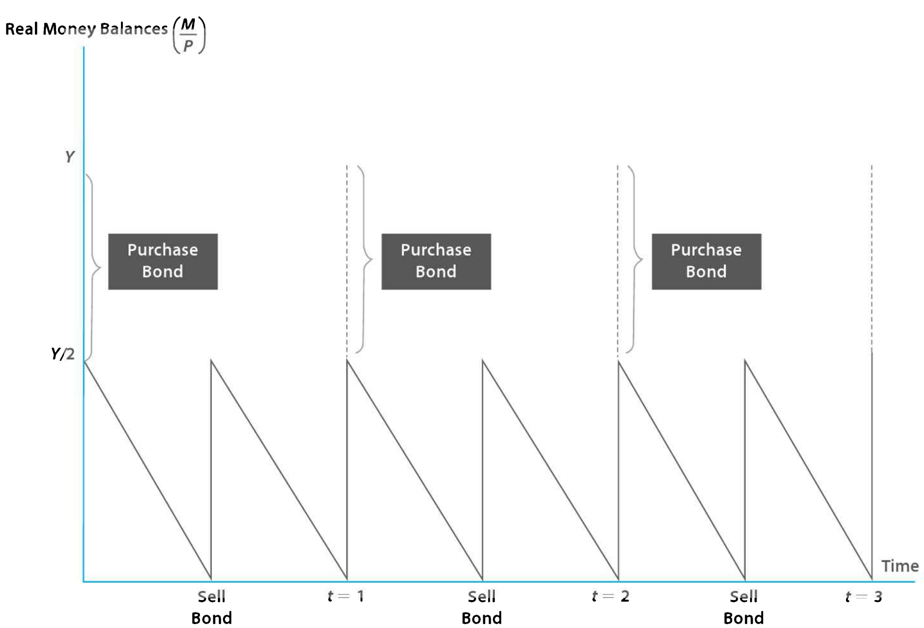
**James Tobin, "The Interest Elasticity of Transactions Demand for Cash," *Review of Economics & Statistics* 38 (August 1956): 241–247.



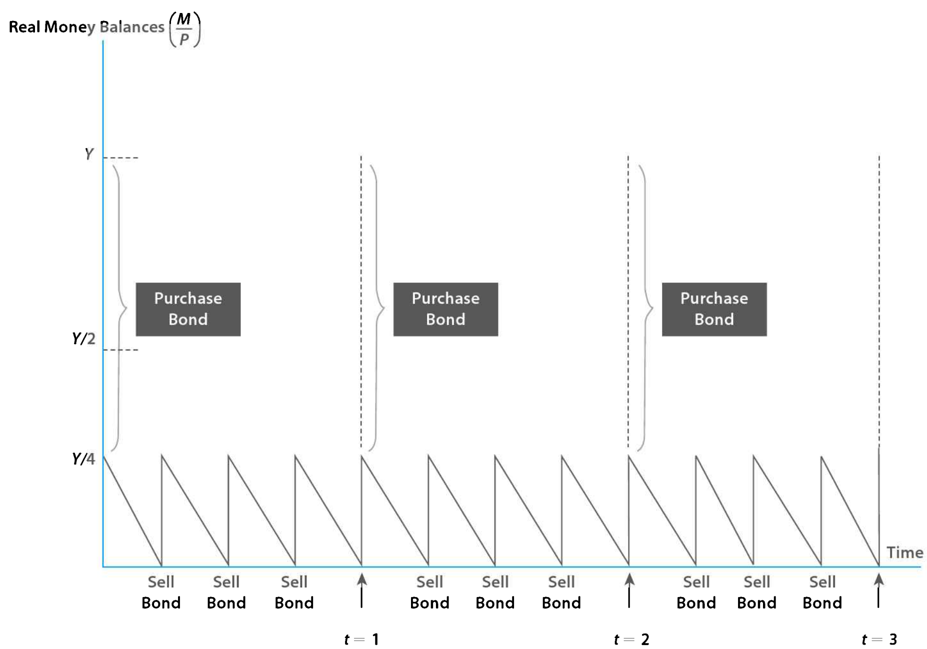

**› Figure A6.5.1**
Time Profile of Money Balances when $x = 0$. **The individual receives an income held in the form of money balances at discrete regular intervals. With even and continuous expenditures within a time period the money balances deplete at a uniform rate during each time interval.**

balances that can be used for purchases of goods and services in the remainder of the time period. The number of transactions in bonds is then $x = 2$. At the beginning of the next income receipt period they again set half of the income aside to purchase bonds and continue with the same behaviour as in the previous period. The time profile of money balances is now that depicted in Figure A6.5.2. The average cash balance held by the individual when $x = 2$ is then given by $Y/4$ .

If now the individual puts three-fourths of their cash balances into bonds, then, they are left with real cash balances equal to $Y/4$. With even expenditures on goods they will run out of cash balances by the time one-fourth of the time interval has elapsed. They then sell each quarter (1/3) of the bonds purchased earlier to finance their expenditure over each quarter period. In each time interval they purchase bonds at the beginning of the interval and sell bonds thrice at the beginning of each second, third, and fourth quarters of the interval. Hence, $x = 4$. The time profile of money balances is that given in Figure A6.5.3. The average money balances held by the individual when $x = 4$ is then $Y/8$.




**› Figure A6.5.2**
Time Profile of Money Balances when $x = 2$. **At the moment the individual receives their income they put half of it into bonds. Half way through the interval 1, as money balances have depleted, they sell the bonds to gain liquidity for the purposes of expenditures.**




**› Figure A6.5.3**
Time Profile of Money Balances when $x = 4$. On receiving their income at the beginning of a time period the individual puts three-fourth of it into bonds. Then each quarter they sell one-third of the bonds purchased in the beginning of the period to replenish their liquid money balances that have been depleted.

Generalizing, when $x$ is an integer and $x \geq 2$, the average holding of money balances will be given by $Y/2x$. What about the average bond holding? Bonds are held whenever money is not required for transactions. Money required for transactions on average is $Y/2x$. Money held on average when bonds are neither purchased nor sold ($x = 0$) is $Y/2$. Thus,

$$\text{average bond holding} = \frac{Y}{2} - \frac{Y}{2x}$$
$$= \frac{Y}{2}\left(1 - \frac{1}{x}\right)$$

With an interest rate $i$ on bonds, the interest receipt on holding bonds on average is the average bond holding multiplied by the interest rate, or,

$$\text{interest receipts on average bond holding} = \frac{iY}{2}\left(1 - \frac{1}{x}\right).$$

Each transaction in a bond—purchase or sale—is subject to a transaction cost of $C$ per transaction. With $x$ number of transactions we have,

$$\text{transactions costs} = Cx$$

The individual's objective is to select $x$, the number of transactions in bonds, to maximize $\phi(x)$, the difference between the interest receipts from holding bonds and the transactions costs of those bond holdings

$$\underset{\{x\}}{\text{Max}}\ \phi(x) = \frac{iY}{2}\left(1 - \frac{1}{x}\right) - Cx$$

Hence, $$\frac{d\phi}{dx} = \frac{iY}{2}x^{-2} - C = 0$$

or, $$x^{-2} = \frac{2C}{iY}$$

or, $$x^2 = \frac{iY}{2C}$$

or,
$$x^* = \sqrt{\frac{iY}{2C}}$$

For example, if we take the estimated per capita national income at factor cost in 2005–2006 at 1999–2000 prices, it is INR 21,005. Let the transaction cost of buying bonds be INR 20 and the bond yield be 6 per cent per year. Then,

$$x^* = \sqrt{\frac{(0.06)(21{,}005)}{40}} = 5.6$$

With the number of transactions an integer,

$$x^* \approx 6$$

An average income person then hardly transacts in bonds during a year.

As the second-order condition is $\phi'' = -(iY/2x^3)$ the negative root can be disregarded.

The average holding of money balances as we saw is $M/P = Y/2x$. Substituting the value of $x = x^*$ we obtain

$$(M/P)^* = \frac{Y}{(2\sqrt{iY/2C})}$$

$$= \sqrt{\frac{CY}{2i}}$$

From this equation we can see the main results of the Baumol–Tobin approach to the demand for money. The demand for money depends

(i) positively on real income (Note that the relationship is not proportional. In fact, real money demand grows less than real income which means that there are economies of scale in the holding of money balances.)

(ii) negatively on the interest rate on bonds

(iii) positively on transactions costs. In fact, when transactions costs fall drastically as the efficiency of financial transactions improves the demand for money drops significantly.

7

# The Labour Market

CRITICAL QUESTIONS

» *What are the determinants of the demand and supply of labour?*

» *How is the aggregate supply* (AS) *curve of output derived?*

» *What form does the* AS *curve take when individuals are subject to money illusion?*

» *How does unemployment result when the* AS *curve is Keynesian?*

» *How does Friedman use the hypothesis of mistaken price expectations to generate an* AS *of output function?*

» *What is the natural rate of unemployment and cyclical unemployment?*

» *What is the Phillips curve and what does it depict?*

» *How is the slope of the Phillips curve affected by inflation expectations?*

» *What are rational expectations and how do they influence the trade-off between inflation and unemployment?*

# 7.1 Profit Maximization and Labour Demand

The supply of the economy deals with the production of output. To produce output, firms use factors of production such as capital and labour. The production function summarizes this relationship:

$$Y = F(K, N) \tag{7.1}$$

where $Y$ is the output produced, and the two inputs are $K$ and $N$. $K$ is a measure of the aggregate stock of land, machines, and other physical inputs. $N$ is the number of employee hours or the man-days hired by the firm. The number of employee hours is the product of the number of workers hired times the average number of hours worked per person. By writing the product $N$, rather than its two separate components, we are assuming that the firm gets the same output when it hires 10 workers for an 8-hour day as when it hires 20 workers for a 4-hours shift. In the interests of simplicity, we ignore the distinction between the number of workers hired and the number of hours worked, and refer to labour input $N$ as the number of workers hired by the firm.

The firm produces that level of output where profits ($\Pi$) are maximum. Profits are defined as total revenue less total costs, or

$$\Pi = pY - wN - p^K(r + \delta)K$$

where $p$ is the price at which the firm sells its output, $w$ is the wage rate or the cost of hiring an additional worker, and $p^K(r + \delta)$ is the rental on capital. The total revenue is $pY$ and the total cost is $wN + p^K(r + \delta)K$. We assume that the firm operates in a competitive environment so that the price of the output is unaffected by how much the firm produces and sells, and input prices are unaffected by how much labour and capital the firm hires.

Consider the situation faced by a profit maximizing firm in the short run—a time span that is sufficiently short so that the firm cannot increase or reduce the size of its plant, or purchase and sell machines and physical equipment. With capital stock fixed at some level $K$, short-run profits are given by

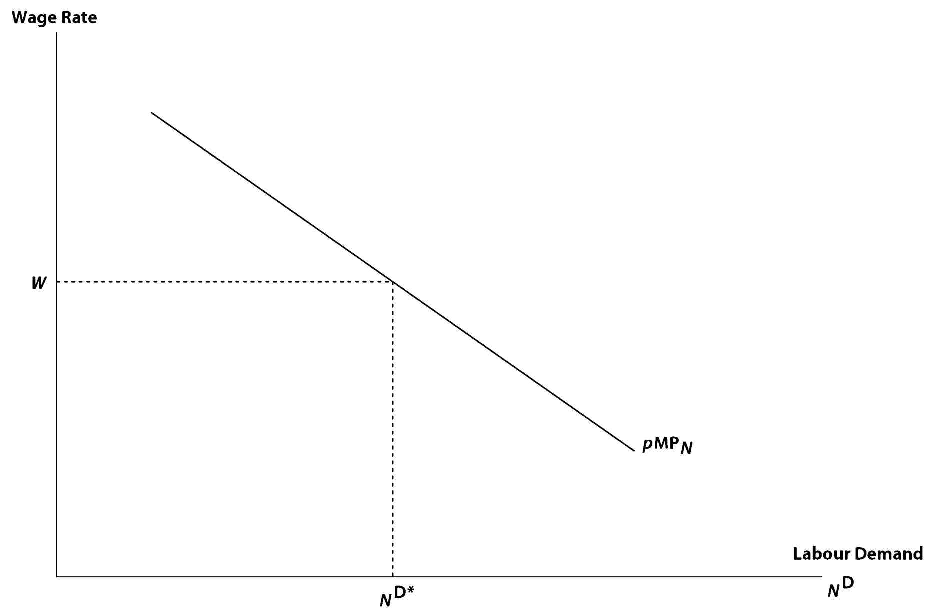
$$\Pi = pF(\overline{K}, N) - wN - p^K(r + \delta)\overline{K} \tag{7.2}$$

Since the firm is competitive and input prices are fixed, $p^K(r + \delta)K$ is akin to a fixed cost, and $wN$ is the variable cost of the firm. The firm can, in the short run, affect profits by the number of workers it hires. It selects this number as given by the profit-maximization condition:[1]

$$p\text{MP}_N = w \tag{7.3}$$

› Profit-maximizing firms hire workers up to the point where the wages equal the value of the marginal product of labour.

MAXIMIZING PROFITS requires that the marginal cost of hiring an additional worker ($w$) equals the marginal gain from hiring that additional worker, which is the value of the marginal product of labour ($p\text{MP}_N$). The marginal product of labour declines as more labour is hired because more and more workers added to a fixed capital stock reduces the gains from specialization. This means that the profit-maximization condition in Eq. (7.3) implicitly gives the number of workers to be hired[2] by $N^{D*}$, as depicted in Figure 7.1. $p\text{MP}_N$ is the short-run demand curve for labour and depicts how much labour the firm will hire at different wage rates.




**› Figure 7.1**
The Demand for Labour. **The demand for labour is downward sloping, reflecting the diminishing marginal productivity associated with hiring additional labour. Firms hire labour up to the point where the revenue from hiring an additional unit of labour equals the nominal wage.**

# 7.2 Utility and Labour Supply

The supply of labour is a decision made by individuals and that decision is a choice between labour and leisure. Of course, a person does spend some time on household activities. In our simple set-up though we abstract from this and assume that a person decides to allocate time between the labour market and leisure activities. Moreover, the person receives satisfaction from the income received as it can be used for consumption of goods and services, and the person also receives satisfaction from the consumption of leisure ($l$). The idea that the individual gets satisfaction from consuming goods and leisure is summarized by the utility function[3]

$$U = U(C, l)$$

The consumption of goods and leisure is constrained both by the individual's time and by income. Let $H$ be the number of hours after rest, sleep, and household activities, which the individual can allocate to the labour market. The time constraint then with $N^S$ hours allocated to labour is

$$H = N^S + l$$

In addition to a time constraint, an individual also faces a budget constraint. Part of a person's income is independent of how many hours they work and can be through other sources such as receipts from lotteries, dividends, and other asset incomes. This can be denoted as the non-labour income $V$. The individual also gets an income by working at a real wage rate of $w/p$ per hour for $N^S$ hours. Hence the consumption expenditure that is feasible for the individual is determined by the sum of the labour earnings and non-labour income:[4]

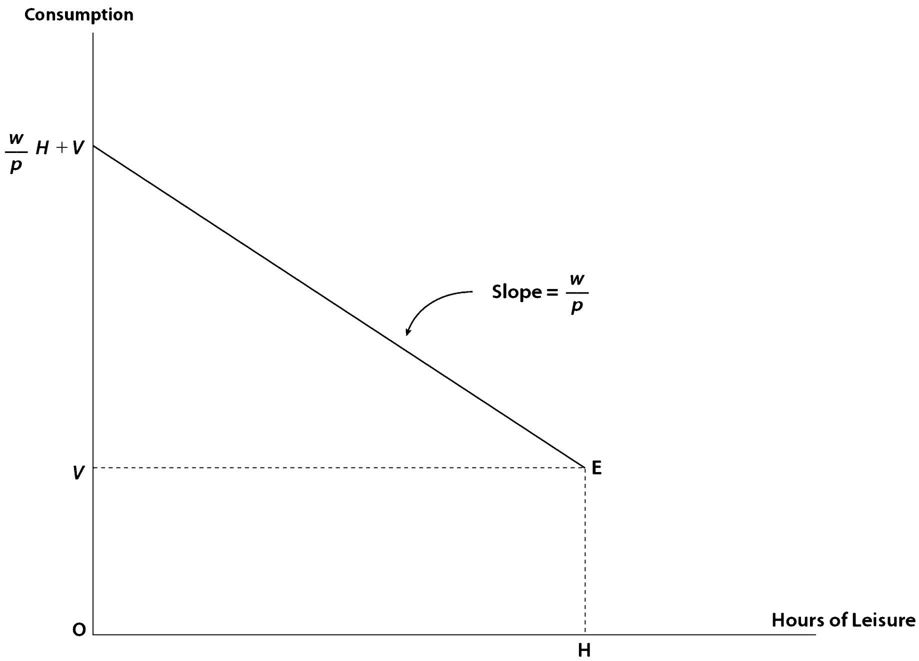
$$C = \frac{w}{p} N^S + V$$

We will combine the budget constraint and the time constraint by writing $N^S = H - l$ and plugging this into the income constraint to give

$$C = \frac{w}{p}(H - l) + V$$

or

$$\frac{w}{p} H + V = C + \frac{w}{p} l$$




**› Figure 7.2**
The Budget Line of an Individual. **The budget line depicts the consumption possibilities inclusive of the accounting for non-labour income *V* as the individual gives up leisure and works in the labour market at a real wage *w/p*.**

The left-hand side is the income received if the individual spends all the time in the labour market working $H$ hours. The non-labour income that is received is $V$. This income of $(w/p)H + V$ can be spent on INR $C$ worth of consumption goods and the rest $(\{w/p\}l)$ on leisure activities. Thus, each hour of leisure has a price and that is the wage rate $w/p$ forgone.

The budget line of the individual is given in Figure 7.2. Point E indicates that if the person decides not to work at all (all $H$ hours devoted to leisure), their income is $V$. If the individual gives up all leisure activities, their income is $(w/p)H + V$. From point E, as the individual moves up the budget line for every hour of leisure that is given up, they receive a real wage of $w/p$ , which is the slope of the budget line.

What point on this budget line will the individual select? That will be determined by the individual's preferences for the consumption of goods versus leisure, as given by the utility function.

Formally, the individual maximizes utility subject to the budget constraint, or

$$\max_{\{N^S\}} U(C, l)$$

such that

$$C = \frac{w}{p}N^S + V \tag{7.4}$$

The graphical solution to this is depicted in Figure 7.3 where the individual decides to supply $N^{S*}$ hours of labour and take $(H - N^{S*})$ hours of leisure. The EQUILIBRIUM is where the slope of the budget line equals the slope of an indifference curve (the marginal rate of substitution of consumption for leisure) or $\text{MRS}_{C,l}$

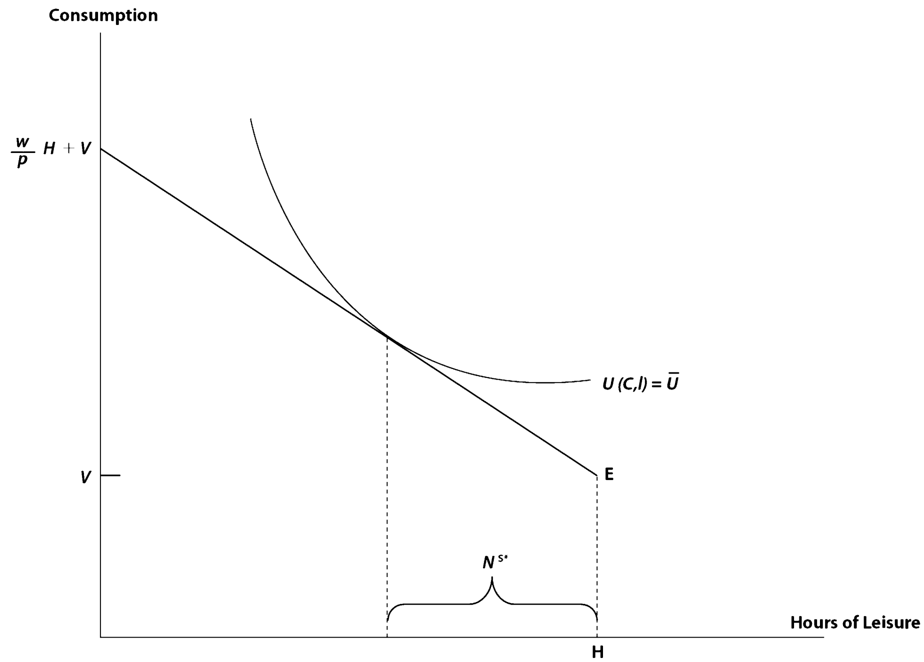
> › Workers maximize utility so that the last rupee spent on leisure activities yields the **same utility** as the last rupee spent on consumption.

$$\overbrace{\text{MRS}_{C,l}}^{\substack{\text{slope of} \\ \text{indifference} \\ \text{curve}}} = \underbrace{\frac{w}{p}}_{\substack{\text{slope of} \\ \text{budget} \\ \text{line}}}$$

or,

$$\frac{\text{MU}_l}{\text{MU}_C} = \frac{w}{p} \tag{7.5}$$

› **Figure 7.3**
The Optimal Labour Supply. The optimal supply of labour is determined from the point where the trade-off between consumption and leisure, as given by the slope of an indifference curve, equals the slope of the budget line.




The slope of an indifference curve[5] is the ratio of the marginal utility of leisure to the marginal utility of consumption. The above condition can be written as:

$$\frac{MU_l}{w/p} = MU_C$$

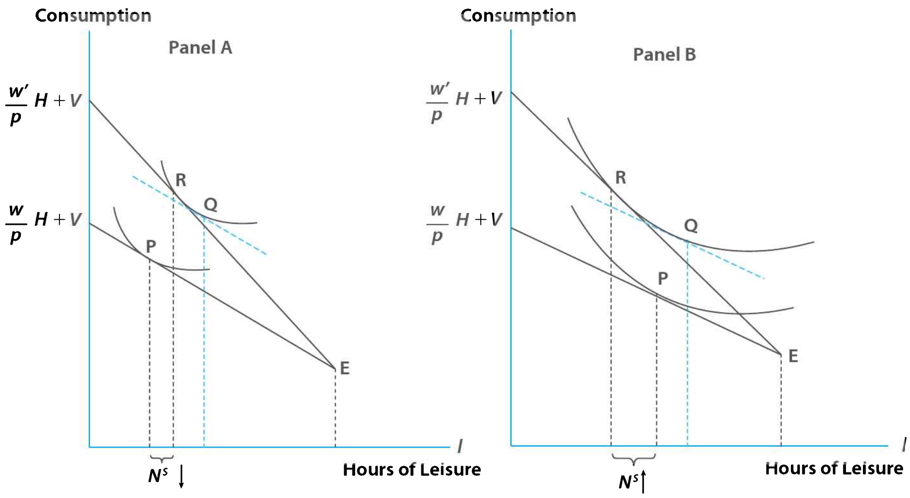
Recollect that an hour of leisure costs $w/p$ rupees. The extra utility from that hour of leisure is $MU_l$. Hence $MU_l/(w/p)$ is the utility received by spending an additional rupee on leisure. The tangency solution thus implies that the last rupee spent on leisure activities yields the same utility as the last rupee spent on consumption goods.

What happens to the hours of work as the wage changes? The wage increase to $w'$ rotates the budget line around the point E and makes it steeper. Panel A in Figure 7.4 depicts a shift from point P to R where the individual consumes more leisure and the hours of work decline. Panel B depicts the opposite result—the wage increase reduces leisure and the individual puts in more hours at work. This is because of well-known opposing tendencies in economics—at high wages individuals want to enjoy the benefits of high incomes and would want to consume more leisure, the income effect of a move from point P to Q in the diagram. The move from point P to Q is associated with a change in income, holding the wage rate constant, which is why the dashed budget line is parallel to the original budget line.

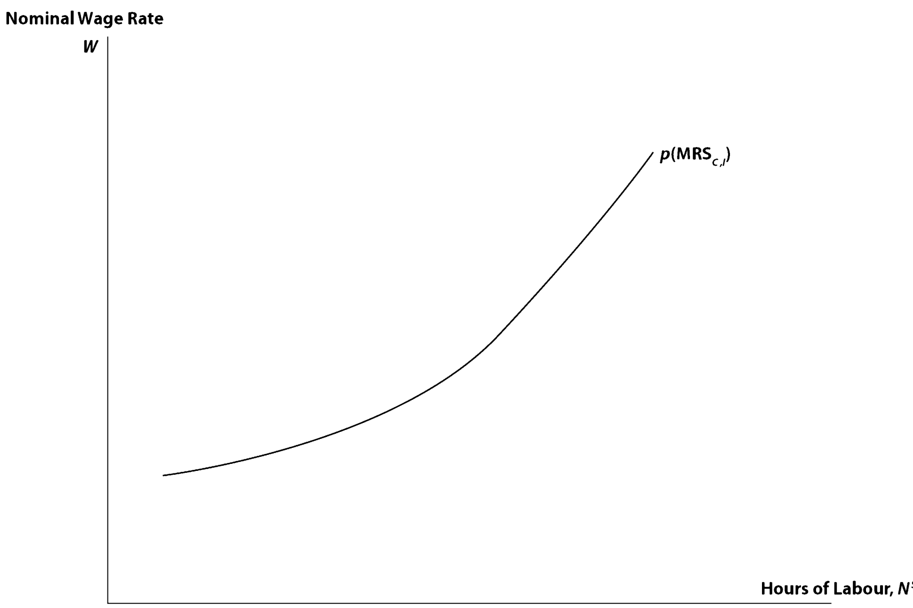
At high wages, leisure time is also very expensive and the individual finds it difficult to afford to take time off from work, the substitution effect of a move from point Q to R in the diagram. The substitution effect illustrates how consumption changes as the wage increases, holding utility constant along an indifference curve.

› An increase in the wage results in both an income and substitution effect. The labour supply curve is upward sloping when the substitution effect dominates.

We will assume that the SUBSTITUTION EFFECT dominates the INCOME EFFECT—panel B of Figure 7.4—so that as wages rise, individuals increase their supply of labour. Thus, the supply of labour schedule as given by Eq. (7.5), $w = p\text{MRS}_{C,l}$, is an upward-sloping schedule as illustrated in Figure 7.5.




**› Figure 7.4**
Effects of a Wage Increase. An increase in the wage causes the budget line to rotate around the point E that depicts the non-wage income. This increases the demand for leisure (income effect). Also as leisure is now relatively more expensive than before, the individual is induced to substitute leisure for labour (substitution effect). In panel A the income effect dominates and the supply of labour decreases. In panel B the substitution effect dominates and the supply of labour increases.

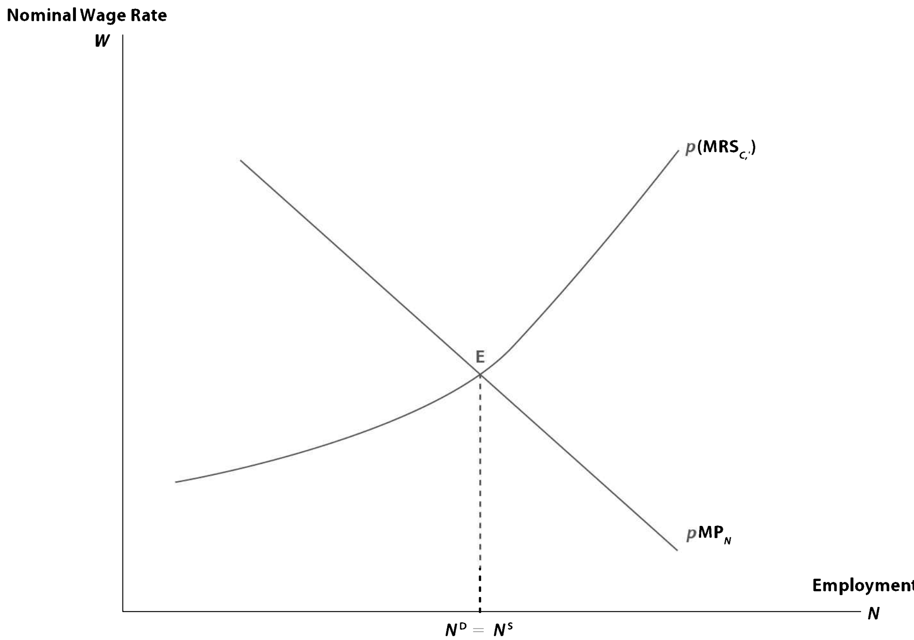



**› Figure 7.5**
The Labour Supply Schedule. The supply of labour increases with the nominal wage rate as the individual responds to the increase in the wage by offering to work more and giving up leisure.

# 7.3 Aggregate Supply with/without Money Illusion

By the familiar economic logic of demand equals to supply in the labour market, $N^D$ equals $N^S$. Putting the demand and supply curves from Figures 7.1 and 7.5 together in Figure 7.6, we get the situation as depicted by point E. Recall that the reason for our exploration into the workings of the labour market is that we are interested in how firms produce the output that is to be sold in the market. That output is produced in the short run by varying the labour hired, the variable factor of production.

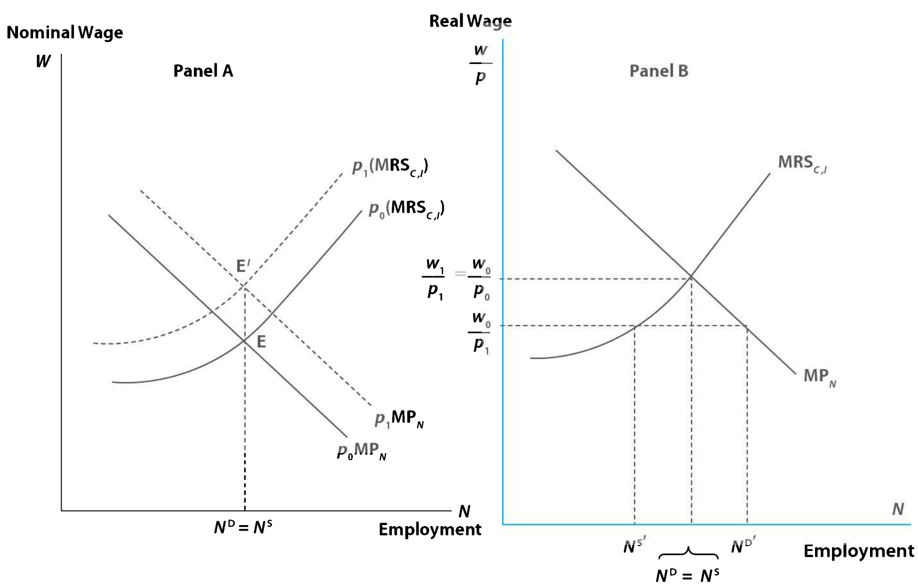
To generate the supply of output curve by the firm, we now inquire into how the output produced by labour changes as the price of the good to be sold varies. Suppose the price of the product goes up from $p_0$ to $p_1$, where $p_1 > p_0$.




**› Figure 7.6**
Labour Market Equilibrium. The demand for labour by firms that are profit maximizers equals the supply of labour by individuals who maximize their preferences for consumption versus leisure, at point E.

Then, the demand and supply curves shift upwards from $p_0\text{MP}_N$ and $p_0\text{MRS}_{C,l}$ to $p_1\text{MP}_N$ and $p_1\text{MRS}_{C,l}$, as depicted in panel A of Figure 7.7. As drawn, there is no change in the labour market outcome with exactly the same amount of labour being demanded and supplied as prior to the price change. Why? Let us employ Figure 7.7 (panel B) to understand this outcome where the vertical axis now graphs the real wage $w/p$ instead of the nominal wage $w$ as before.

Essentially, the increase in the price reduces the real wage cost from $w_0/p_0$ to $w_0/p_1$. With a decline in the real wage, the cost of labour has become cheaper and the firm responds by going down its demand curve $\text{MP}_N$ and increasing the demand for labour to $N^{D\prime}$. The reduction in the real wage reduces the return to labour supply, and individuals react by reducing their supply of labour to $N^{S\prime}$.




**› Figure 7.7**
Effect of Price Increase. An increase in the price results in an equal increase in the nominal wage (panel A) that results in a constant real wage (panel B). The change in price then has no effect on the employment in the labour market.

At $w_0/p_1$ there is an excess demand for labour to the extent of $N^{D'} - N^{S'}$. As the firm faces a shortage of labour, the price of labour must be bid up to induce the required labour supply. Hence, the nominal wage rises up to $w_1$, the point where demand again equals supply, or where $w_1/p_1 = w_0/p_0$, and the equilibrium in the labour market is restored.

This implies that the percentage increase in the price exactly equals the percentage INCREASE IN THE WAGE, or

$$\frac{p_1 - p_0}{p_0} = \frac{\Delta p}{p} = \frac{\Delta w}{w} = \frac{w_1 - w_0}{w_0}$$

and the real wage is a constant as a result.

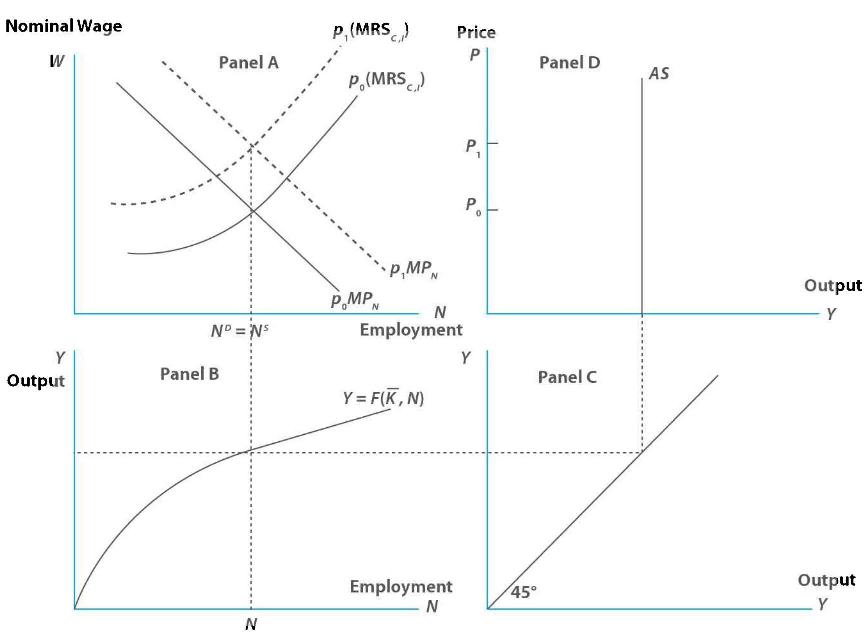
› If a change in price results in an equiproportionate change in the money wage, employment and output are unaffected. This is referred to as a situation of no money illusion—a situation where individuals do not ordinarily make mistakes in assessing nominal versus real changes.

The effect on output is depicted in Figure 7.8. Panel A of this diagram is the same as panel A of Figure 7.7. In panel B, the short-run production function $Y = F(\bar{K}, N)$ is graphed—the decline in the slope of this production function as $N$ increases reflects the diminishing marginal productivity of labour. Panel C is just a 45-degree line in output that allows us to mirror the output changes from panel B to panel D, which traces the relationship between output produced and prices—the *AS* relationship.

As the change in the price leads to no change in labour input, there is no corresponding change in output. The *AS* curve is vertical. In effect, when both sides of the market—firms and workers—think in real terms, any increase in price is offset by an equal increase in money wages. With no change in real income to individuals or real cost to firms the real demand ($MP_N$) and supply ($MRS_{C,l}$) curves for labour are unchanged and there is no change in employment and output. When individuals think in real terms, they are said to be subject to no money illusion and the vertical *AS* curve is called a CLASSICAL SUPPLY CURVE.

› The classical *AS* curve is vertical, and the quantity of output produced depends on real factors such as technology (productivity and labour demand), and preferences (labour supply), with no effect of nominal variables on output.

Now, suppose that individuals suffer from some degree of money illusion and do not formulate their labour supply decision on the basis of the real wage. Suppose that rather than the real value of the good that the money wage can buy ($w/p$), they are interested only in their money wages ($w$) because they are more interested in comparing their earnings position vis-à-vis the money earnings of




**› Figure 7.8**
Aggregate Supply (Real Wages). A rise in the price raises money wages by an identical amount when there is no money illusion (panel A). With no change in employment, output does not change (panel B). Output is, therefore, constant and the *AS* curve is vertical (panel D).

others in the labour market. In that case, workers bargain with their employers taking into account the nominal wages they receive compared to a reference group of workers with whom they make comparisons in terms of similarities in working conditions and job description. Employers, on the other hand, continue to make decisions about labour demand in terms of the real cost to them of any nominal wage they pay. An individual firm can observe its own price as well as the wage paid to its workers so that it can make decisions with respect to the real wage. In this case, the demand and supply curves for labour are given, respectively, by

$$w = p\text{MP}_N$$

and

$$w = \text{MRS}_{C,l}$$

Figure 7.9 depicts this situation and shows that when the price rises from $p_0$ to $p_1$, the *AS* curve is upward sloping like a standard supply curve. Hence, if workers suffer from some MONEY ILLUSION and make decisions about labour supply on the basis of nominal wages, the *AS* curve is upward sloping.

› When individuals formulate labour supply decisions on the basis of the money wage, they are subject to money illusion. The *AS* curve is then upward sloping.

## 7.4 Introducing Unemployment

By now you would have noticed that in both Figures 7.8 and 7.9 we have proceeded by equating the demand curve with the supply curve of labour. With $N^D = N^S$, there can never be any unemployment, which is a characteristic of all economies. The early attempts to introduce unemployment into a

**MACROFOCUS 7.1**

### Why Has Wage Inequality Increased?

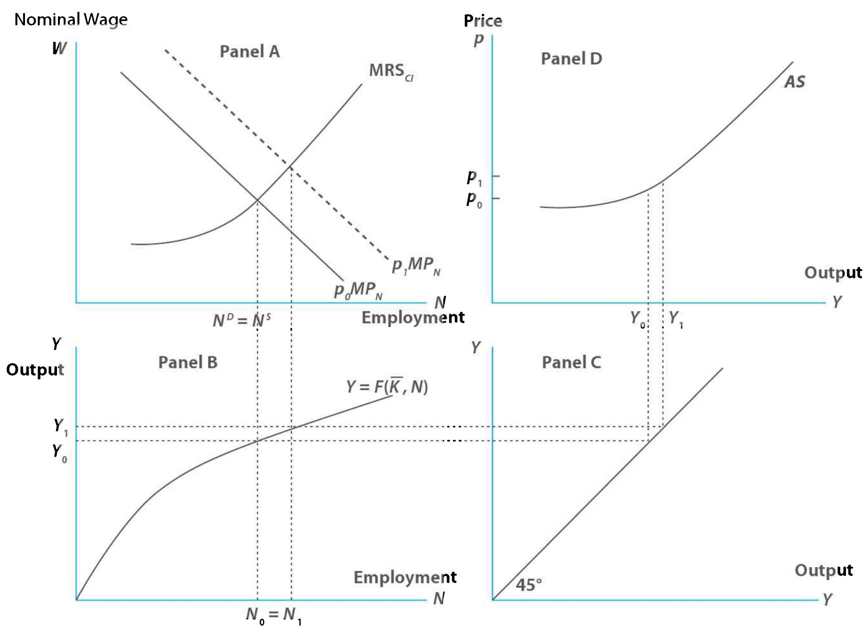
The distribution of wages has become more unequal since the 1980s in most countries across the world. In the United States, for instance, during the 1980s, most income groups faced a decline in real wages*. However, this widening of the wage distribution in the United States is not a story of the rich getting richer and the poor poorer since the wage of more skilled workers remained roughly constant while that of unskilled workers declined rapidly. In India, by contrast, during the 1980s and 1990s, even the workers on the lower end of the wage distribution gained in terms of real wage increases. The accelerating wage inequality showed mainly on the upper half of the wage distribution.

What causes wage inequality? One explanation is in terms of the simple supply–demand framework with the added recognition that there are two types of workers in the labour market—skilled and unskilled. Then the wages of skilled workers relative to the wage of unskilled workers would rise over time, if the supply of labour curve for skilled workers did not shift out as much as the supply curve for unskilled workers, or, if the demand curve for skilled workers shifted out by more than the demand curve for unskilled workers.

If the supply curve of college graduates shifts outwards, it would tend to reduce the wage for skilled workers relative to less-educated workers. If, however, the demand curves for skilled workers shifts out by more than the demand curve for unskilled workers, an economy can witness increases in returns to skilled workers. In India after the mid-1990s, there was increasing deregulation of job security legislation that required employers to seek government permission before dismissing an employee. Since then, about 15 per cent of workers in organized manufacturing lost their jobs but the employment of supervisors nevertheless increased steadily. Deregulation of labour law is thus one reason for an increase in the relative demand for skilled labour.

As tariff and non-tariff barriers to trade reduced, many cheap imported goods also became available in India. These are goods that are typically produced in India in the small-scale sector by unskilled workers. However, small-scale industries are subject to a reservation policy that allows certain, mainly consumer, goods to be produced only if production is of limited scale. With restrictions on their scale, such industries have been unable to increase the demand for unskilled labour. Kijima† thus found that the relative demand for skilled workers grew faster than the supply, which increased inequality in India. Policies to decrease wage inequality would then require to focus on increasing the supply of skills through enabling access to secondary and tertiary (college-level) education. Moreover, labour-intensive manufactured export growth is also adversely affected to the extent that such exports are from the small-scale sector.

*K. M. Murphy and F. Welch, "The Structure of Wages," *Quarterly Journal of Economics* 107 (*February* 1992): 285–326.
†Y. Kijima, "Why Did Wage Inequality Increase? Evidence from Urban India 1983–99," *Journal of Development Economics* 81 (2006): 97–117.




› **Figure 7.9**
Aggregate Supply (Money Wages). Individuals making labour supply decisions on the basis of money wages increase their supply of labour when firms increase their demand for labour due to an increase in the price of output (panel A). The increase in employment increases output (panel B). The price rise is then accompanied by an increase in the supply of output (panel D).

macroeconomic framework introduced a friction, or a rigidity, into the main model. We now look at two such manipulations to the above framework.[6]

The first is essentially the Keynesian approach in which money wages are treated as constant until the point where those who want to be employed at this constant wage find employment, after which the money wages are flexible upwards.[7]

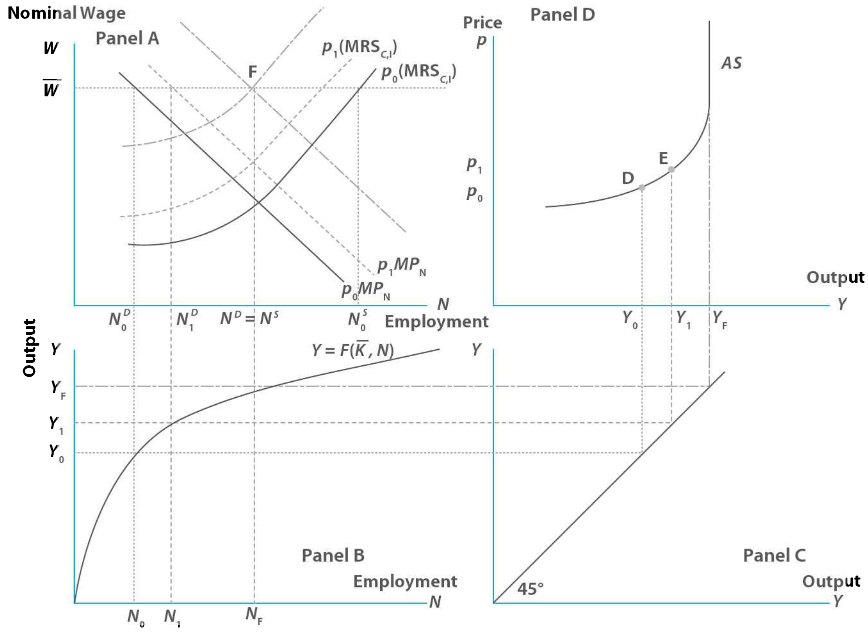
The second approach is due to Friedman who argues that there is an asymmetry of information between workers and firms. Essentially, firms observe the actual value of their own price and the wage rate. Workers can observe their nominal wage but it is inherently more difficult a task for them to estimate the prices of the whole basket of goods that they actually consume with their wage. Workers, therefore, form an estimate of the price level they will face for the goods in their consumption basket and the aggregate of these prices will be denoted by $p^e$, the price level they expect. Workers in Friedman's view supply labour on the basis of expected prices, whereas firms demand labour on the basis of actual prices. In Friedman's analysis, it is because workers have mistaken expectations that unemployment occurs. Let us look at each view starting first with the Keynesian *AS* curve.

### *7.4.1* Keynesian Aggregate Supply

Figure 7.10 derives the Keynesian aggregate supply curve. Let the initial price be $p_0$ and the fixed ruling money wage rate be $w$. This results in a demand for labour of $N_0^D$ and a supply of $N_0^S$ with unemployment given by $N_0^S - N_0^D$. Even though more workers are willing to work, firms are not willing to offer them jobs as the wage cost is set high and it is not profitable to employ more than $N_0^D$ workers. Actual employment is thus given by

$$N = \min[N^D, N^S]$$

Given the level of employment determined in the labour market, $N_0 = N_0^D$, we can determine the output $Y_0$ that is produced by these $N_0$ units of




**› Figure 7.10**
Keynesian Aggregate Supply.
With the money wage fixed in the short run, firms are able to hire additional labour at fixed nominal costs when the price of the output increases up to the point of full employment F in panel A. The increase in employment raises output (panel B). The increased output that results when the price rises is the upward-sloping *AS* curve in panel D. The *AS* curve is vertical at the point of full employment where the demand for labour equals the supply of labour.

labour from the production function (panel B). Using the 45-degree line (panel C), this level of output $Y_0$ can be combined with the price associated with that output of $p_0$ as specified in panel A to give one point D on the *AS* curve.

› When money wages are fixed and employment is determined by the demand for labour, we obtain a Keynesianag gregate supply curve that is upward sloping till all those who offer to supply labour at the fixed money wage have been employed.

To derive the rest of the SUPPLY CURVE, let the price at which output is sold rise to $p_1 > p_0$. This causes the demand and supply curves to shift up resulting in increased employment ($N_1 = N_1^D$). This is because the real cost of labour has declined from $\bar{w}/p_0$ to $\bar{w}/p_1$, making it profitable to employ more labour. The resulting new output produced of $Y_1$ can be traced out from panels B and C to panel D where the new price output point of $(p_1, Y_1)$ on the supply curve is given by point E. As price rises, the demand for labour rises and more labour is hired, resulting in more output produced. The *AS* curve is, therefore, an upward-sloping curve as long as $N^D < N^S$. Once $N^D = N^S$, as occurs at point F in panel A, all workers who are willing to work at that wage $\bar{w}$ have been fully employed, and even if the firm advertises more vacancies it will find no takers for those jobs. The level of employment at point F, $N_F$, is characterized by full employment and the output produced corresponding to this level of employment is the full employment level of output, $Y_F$. The Keynesian supply curve thus has an upward-sloping portion that becomes vertical at full employment.[8]

### *7.4.2* Friedman on Mistaken Expectations

Another explanation of the labour market is that provided by Friedman who argues that workers on the supply side of the labour market are imperfectly informed about the aggregate price level of the typical basket of goods they consume. They thus form an estimate of the aggregate price level and on the basis of this the supply of labour schedule is not given by $w = p\text{MRS}_{C,l}$ but is defined in terms of the price level that they expect, $p^e$, so that,

$$w = p^e\text{MRS}_{C,l} \tag{7.6}$$

The demand for labour is still given by

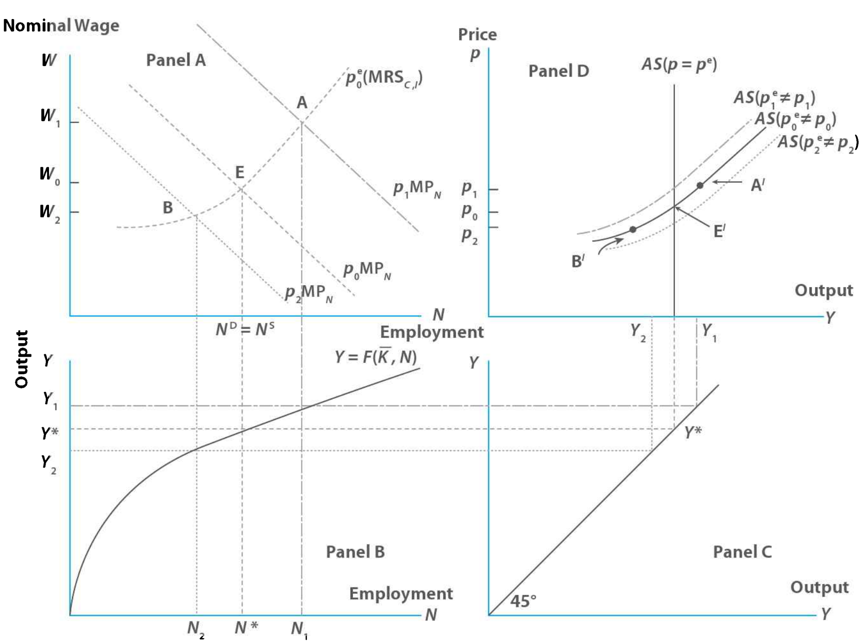
$$w = p\text{MP}_N \quad \textbf{(7.7)}$$

If $p^e = p$, there are no expectational errors and both sides of the market are making decisions in real terms, and the outcome will be a vertical supply curve as in Figure 7.8. The level of employment and the output produced when expectations are realized, $p^e = p$, is called the NATURAL RATE OF EMPLOYMENT ($N^*$) AND OUTPUT ($Y^*$).

› When expected prices equal actual prices, expectations are realized and the economy is at the natural rate of employment and output, with a vertical *AS* curve.

In Figure 7.11, we begin with an initial price of $p_0$. Let us assume that the expected price is also equal to this, $p_0^e = p_0$. With workers making no expectational error, we can trace out the natural rate of employment and output and the *AS* curve is the vertical curve given by $AS$ ($p = p^e$). Now, suppose the price increases to $p_1 > p_0$. The demand curve for labour shifts up. The expected price, however, is still $p_0^e$. The labour supply curve is, therefore, unchanged. The demand and supply curves for labour intersect at point A, where the nominal wage is $w_1$, and employment and output are given by $N_1 > N^*$ and $Y_1 > Y^*$. Tracing out through the labour input-product output, and 45-degree line panels, we get the point A′ which denotes the point ($p_1$, $Y_1$) in the price versus output produced panel.

Employment and output have increased at point A because the actual real wage has declined making it profitable for firms to employ more workers. More workers accept work because in their mental map the expected real wage is the same. Workers have underestimated the price level $p_0^e < p_1$ and as a consequence overestimated the real wage earnings, $w_1/p_0^e > w_1/p_1$. If the actual price declines to $p_2 < p_0 < p_1$, the demand curve for labour shifts down to $p_2\text{MP}_N$. With unchanged price expectations, the new demand and supply curves intersect at point B where the nominal wage is $w_2$ and employment and output have declined to $N_2 < N^*$ and $Y_2 < Y^*$. Again tracing out through the two panels in the bottom-half of the diagram, we get the point B′, which denotes the point ($p_2$,$Y_2$) in the price-supply of output space. With a decline in the price, the real cost of employing labour has increased and firms downsize. Workers again have overestimated prices $p_0^e > p_2$




› **Figure 7.11**
Mistaken Expectations Aggregate Supply. When workers supply labour on the basis of expected prices, firms are able to employ more labour as the price of the output increases, and employ less labour when the price falls (panel A). An increase in employment is associated with a rise in the output produced (panel B). The increase in output associated with the rise in the price results in an *AS* curve when price expectations are not realized: *AS* ($p^e \neq p$), which is upward sloping. When price expectations are realized, the *AS* curve, *AS* ($p^e = p$), is vertical.

› When workers have mistaken expectations regarding the price of output, the *AS* curve for a given price expectation is **upward sloping.**

and underestimated their real wage earnings, $w_2/p_0^e < w_2/p_2$. Joining all points such as B′ and A′ we get the UPWARD-SLOPING *AS* CURVE that traces out various price-supply of output combinations for constant price expectations, $p_0^e$.

For higher price expectations, $p_1^e > p_0^e$, it can be traced out that the *AS* curve will be higher everywhere or $AS\,(p_1^e) > AS\,(p_0^e)$. Similarly, for lower price expectations, $p_2^e < p_0^e$, the corresponding *AS* curve will be lower down as given by $AS\,(p_2^e) < AS\,(p_0^e)$. There is thus a family of *AS* curves, each one corresponding to a given level of price expectation, and resulting due to mistaken expectations on the part of workers $AS\,(p^e \neq p)$. If expectations were always realized, $p^e = p$, the relevant *AS* curve would be the vertical $AS\,(p^e = p)$ curve that is depicted in the diagram. Friedman's *AS* curve can accordingly be labelled as the mistaken expectation *AS* curve and the equation of this[9] curve can be written as:

$$Y = Y^* \left(\frac{p}{p^e}\right)^{\phi},\ \phi > 0 \qquad \textbf{(7.8)}$$

When $p = p^e$ in Eq. (7.8), output is at its natural level, or $Y = Y^*$. Similarly, when $p \neq p^e$, Eq. (7.8) gives the $AS\,(p \neq p^e)$ curve for each price expectation. Thus, if actual prices are greater than expected, $p > p^e$, Eq. (7.8) implies that $Y > Y^*$, as was derived in Figure 7.11 for point A′. If $p < p^e$, then, $Y < Y^*$ as derived in Figure 7.11 for point B′.

› Under adaptive expectations, individuals revise their price expectations in every period to incorporate a fraction of the errors in the expectation ($p - p^e$) from the previous period.

Friedman is also of the view that price expectations change as individual workers realize that their expectations were mistaken. Specifically, Friedman advocated the ADAPTIVE EXPECTATIONS HYPOTHESIS where the expected price is given in the short run but moves slowly to correct for past expectational errors. Thus,

$$p_{t+1}^e = p_t^e + \lambda(p_t - p_t^e),\ 0 < \lambda < 1 \qquad \textbf{(7.9)}$$

where $\lambda$ represents the speed with which households update their price expectation. The lower is the value of $\lambda$, the more sluggish is the adjustment of expectations. Equation (7.9) states that households expect the price in the future period $t + 1$ to be the actual price in the current period $t$, $p_{t+1}^e = p_t$ if their expectations proved to be correct in the current period, $p_t^e = p_t$. If, however, they mistake the price level in the current period ($p_t^e \neq p_t$), then, they incorporate a fraction $\lambda$ of the expectation error ($p_t - p_t^e$) in the revision of their expectation in the next period. In terms of Figure 7.11, at point A′, $p_1 > p_0^e$, and so $\lambda\,(p_{t,1} - p_{t,0}^e)$ is the amount by which price expectations are revised upwards as given by $p_{t+1}^e - p_t^e = \Delta p_t^e = \lambda(p_{t,1} - p_{t,0}^e)$. The next time period will then see workers on a higher $AS\,(p \neq p^e)$ curve than the current one. Similarly, at point B′, $p_2 < p_0^e$ and workers revise their expectations downwards and move to a lower $AS\,(P \neq P^e)$ curve.

# 7.5 Cyclical Unemployment and the Output Gap

Economists think of someone as unemployed if they are not holding a job but are looking for one. The unemployed are broadly classified as being of four types—those unemployed due to seasonal unemployment, frictional unemployment, structural unemployment, or cyclical unemployment.

› Seasonal unemployment varies predictably with the seasons.

Unemployment that varies predictably with the seasons is SEASONAL UNEMPLOYMENT. Workers in agriculture and tourism are regularly subject to seasonal unemployment. In the seaside state of Goa, for instance, tourism increases in the winter months, and correspondingly the number of jobs that cater to tourists increase in these months. Agriculture in India is known to have peak demand for labour during the sowing season.

Seasonal unemployment follows a predictable pattern and is unlike other types of unemployment that affect workers as a part of the normal transition from one job to another. FRICTIONAL UNEMPLOYMENT, for instance, results when people are in transition between one job and another, or when they first enter the labour market to search for work. Though the labour market serves to match available jobs with available workers, this can be a time-consuming process because both jobs and workers are heterogeneous. Jobs, for instance, differ in the skills they require, in their working conditions, and in their location. Similarly, workers differ in their abilities and experience, in their willingness to travel and relocate as a job requires, and in their career goals. Thus, the process of matching jobs with workers can be time-consuming. The short-term unemployment associated with the process of matching is called frictional unemployment. Frictional unemployment is not considered to be a burden for the economy to the extent that the process leads to a better match between a worker and a job. In that sense, it is actually productive since the turnover of workers on such jobs would be lower and disruptions to output less frequent.

› Frictional unemployment is the short term unemployment associated with matching workers with jobs.

STRUCTURAL UNEMPLOYMENT is a more long-term and chronic form of unemployment as compared to frictional unemployment, and exists independently of the rate of output that is being produced in the economy. This often occurs in periods of economic change when the new jobs being created have requirements different from the old jobs being lost. Since the early 1980s, in India, the textile industry declined for many years as mills closed down, while the computer software industry has been growing rapidly. It would be ideal if textile workers who lose jobs could be gainfully employed in the software sector. However, textile workers lack the education and training that is a prerequisite for working in the computer industry. With textile workers skills no longer being demanded, these workers enter into chronic or long-term unemployment that is structural. Structural unemployment also arises where there are labour market barriers to employment such as minimum-wage laws, job security legislation, discrimination, and language requirements associated with jobs. Structurally unemployed workers, unlike frictionally unemployed ones, are not involved in productive work for long spells, and this represents a substantial economic loss.

› Structural unemployment is a long-term unemployment that does not vary with the pace of economic activity.

CYCLICAL UNEMPLOYMENT is associated with a reduction in real output during periods of recession.[10] Seasonal, frictional, and structural unemployment are always present in the labour market, even when the economy is operating normally. Cyclical unemployment, by contrast, occurs only during recessions.

› The unemployment associated with periods of low levels of economic activity is cyclical unemployment.

The sum of seasonal, frictional, and structural unemployment is called the NATURAL RATE OF UNEMPLOYMENT. Actual total unemployment in the economy is the sum of the natural rate of unemployment and cyclical unemployment. In practice, it is often very difficult to make sharp distinctions between the different categories of unemployment. However, the unemployment concept that is the FOCUS OF MACROECONOMICS is cyclical unemployment—the unemployment that increases when there is a slowdown in economic activity, and that decreases when real output increases.

› The natural rate of unemployment is the sum of the seasonal, frictional, and structural unemployment.

› Macroeconomics focuses on cyclical unemployment.

Emerging markets like India do not have statistics generated on the natural rate of unemployment. In the United States, the Congressional Budget Office provides estimates of the natural rate of unemployment regularly.[11]

The natural rate of unemployment changes over time due to such factors as changes in the age composition of the labour force. In the United States, the natural rate has declined in the last 15 years. This is some times explained as due to the rising average age of US workers.[12] Older workers are not as prone to change jobs, a reduction in frictional unemployment, and also have greater experience and skills (a reduction in structural unemployment).

› Full employment is a state where cyclical unemployment is zero.

› Cyclical unemployment is the difference between total unemployment and the natural rate of unemployment.

In macroeconomics FULL EMPLOYMENT occurs when there is no cyclical unemployment. There will still be the existence of seasonal, frictional, and structural unemployment when full employment prevails. Macroeconomics focuses on CYCLICAL UNEMPLOYMENT, which is the difference between total unemployment and the natural rate of unemployment.

**Cyclical unemployment = Total unemployment − Natural rate of unemployment**

or,

$$\text{Cyclical unemployment} = U - U^*$$

When economic activity decreases, the actual unemployment exceeds the natural rate of unemployment, and cyclical unemployment is positive. When there is an expansion in economic activity and real output, the actual unemployment rate is lower than the natural rate, and cyclical unemployment is negative.

It was Arthur Okun, President Kennedy's chief economic advisor who suggested that there is a relationship between the cyclical unemployment in the economy and the output gap. The output gap is the difference between potential output and actual output. Potential output, $Y^*$, or the full employment output, is what the economy would produce if labour were fully employed at normal levels of overtime, and plant and machinery were used at their normal rates of utilization. POTENTIAL OUTPUT is the maximum *sustainable* amount of real GDP that can be produced. Real GDP can temporarily exceed its potential when individuals temporarily put in extra overtime. However, this is not sustainable as individuals cannot work overtime continuously and even plant and machinery must be shut down for maintenance. Changes in potential output are then part of the reason for the rise or decline in economic activity. When the economy's factors of production such as capital and labour are not fully utilized, the actual output $Y$ is below potential output $Y^*$. This underutilization of factors of production constitutes a recession. When actual output exceeds potential output, factors of production are being utilized more than normal and this results in a boom. The percentage difference between potential and actual output is called the OUTPUT GAP.

› Potential or full employment output is the maximum sustainable amount of output that an economy can produce.

› The output gap is the difference between the potential output and the actual output as a fraction of potential output.

$$\text{Output gap} = \frac{\text{Potential output} - \text{Actual output}}{\text{Potential output}} = \frac{Y^* - Y}{Y^*}$$

Okun[13] pointed out that as an economy gets out of a recession, output increases by a greater percentage than the rise in employment. Similarly, when the economy heads into a recession, output decreases by a greater percentage than the reduction in employment. OKUN'S LAW states that a 1 per cent increase in the unemployment rate results in about a 3 per cent reduction in the output gap.

› Okun's Law states that each percentage point increase of cyclical unemployment is associated with a 3 percentage point increase in the output gap.

$$\frac{Y^* - Y}{Y^*} = 3(u - u^*)$$

where the unemployment rate ($u$) is the ratio of the number unemployed to those in the labour force. Okun's Law allows us to link fluctuations in the unemployment rate with fluctuations in the output gap. The unemployment rate by one measure (current daily status) increased by 1.2 per cent in India between 1994 and 2000, and was 7.3 per cent in that year. Suppose the natural rate of unemployment is 6 per cent. Then, cyclical unemployment is 7.3 minus 6 per cent, or 1.3 per cent. According to Okun's Law that would be associated with a $3 \times 1.3 = 3.9$ per cent negative output gap. Real GDP in the year 2000 was roughly INR 19,600 billion. Hence, a 3.9 per cent output gap means that the economy was producing 3.9 per cent of INR 19,600 billion, or INR 764.4 billion less than its potential. The lost output represents the opportunity[14] cost associated with the higher cyclical unemployment. A negative output gap implies that factors of production are not fully utilized and, therefore, output and employment are below the maximum sustainable levels. A rise in cyclical unemployment thus means that the economy is not matching up to its potential and this cost can have drastic impacts on the popularity of policymakers and politicians. Similarly, a positive output gap is a situation where the utilization of factors of production is stretched beyond sustainable levels. This unsustainable situation is resolved through price increases that reflect the increased scarcity of capacity.

Expansionary gaps thus result in inflation that also impacts adversely on the popularity of policymakers and politicians. Policymakers therefore attempt to eliminate output gaps and stabilize the economy. This process of aligning actual output with potential output is achieved, as we will see, through the usage of fiscal and monetary policies. Such policies eliminate output gaps by influencing total expenditure in the economy. Since a positive output gap (high cyclical unemployment) is associated with inflation, it implies a negative relationship between inflation and unemployment. This negative relationship was empiri-

MACROFOCUS 7.2

### Unemployment and Wages in India

Data on employment and unemployment in India are available from the large sample National Sample Survey (NSS) surveys, the last three of which were carried out in 1987–1988, 1993–1994, and 1999–2000. The sample is large in the sense that a sample of 10 households (12 households in the 1999–2000 round) per selected village or urban block is considered for canvassing the schedules. The NSS has evolved a system of capturing employment and unemployment through the time-use method. A worker is asked if he or she was working (a) last year (usual status), (b) last week (current week status), and (c) yesterday (current daily status). If a worker has not been working, but sought or was available for work during the reference period, he or she is classified as part of the labour force but unemployed. The usual status basis for classification is the one that makes macroeconomic sense as it refers to the period of a year.

*Usual principal status* (*UPS*): A person is counted as being in the labour force on a UPS basis if he or she was engaged in economic activity (work), or was seeking or was available for work for the major part of the preceding 365 days. Those classified as being in the labour force on

**› Table 1**

Employment as per Usual Status as on 1-1-1994 (Numbers in thousands)

| Activity | Rural Areas | | | Urban Areas | | |
|---|---|---|---|---|---|---|
| | Male | Female | Total | Male | Female | Total |
| Employed | 187,744 | 104,739 | 292,483 | 64,456 | 17,209 | 81,665 |
| Unemployed | 2,716 | 958 | 3,674 | 2,727 | 1,110 | 3,837 |
| Unemployment Rate | 1.43 | 0.91 | 1.24 | 4.06 | 6.06 | 4.49 |
| Labour Force | 190,460 | 105,697 | 296,157 | 67,183 | 18,319 | 85,502 |
| Not in Labour Force | 149,041 | 213,629 | 362,670 | 56,771 | 92,818 | 149,588 |
| Total | 339,501 | 319,326 | 658,827 | 123,953 | 111,026 | 234,979 |

Note: Due to rounding off the figures, totals do not tally.

(continued on next page.)

this basis are further classified as employed or unemployed depending on whether the majority of the days in the labour force were spent in economic activity or in seeking/being available for work. The UPS unemployment rate is the proportion of those classified as being unemployed on this basis, and is expressed as a percentage of those classified as being in the labour force. Based on this criterion, people can be counted as being employed even if they were unemployed (or were outside the labour force) for a significant part of the year. Equally, a person can be counted as unemployed even though he or she may have been employed for part of the year.

As can be seen from Tables 1 and 2, the unemployment rates have gone up between 1994 and 2000. It is widely argued that countries like India suffer from widespread underemployment and much less open unemployment as it is hard for a person to remain unemployed without a formal and effective social security system. Unfortunately, there are no statistics on underemployment. However, as a rough measure, researchers take the difference between the weekly and daily unemployment status as a measure of underemployment.

The unemployment rate for India (rural and urban combined) on the currently weekly status was 3.63 per cent in 1994, and 4.41 per cent in 2000, whereas on the current daily status it was 6.03 per cent in 1994 and 7.32 per cent in 2002. An estimate for underemployment on the basis of these figures is that in 1994, 6.03 – 3.63 = 2.40 per cent of the labour force was underemployed, and in 2000, 2.91 per cent of the labour force was unemployed. Thus, an estimate of the unemployment rate, when we account for underemployment, is 1.97 + 2.40 or 4.37 per cent in 1994 and 2.20 + 2.91 or 5.11 per cent in 2000.

What about wage rates? The real hourly wage rates for various categories of workers are given in Table 3. As would be expected, urban wages are higher than rural wages, men get paid more than women, regular workers earn more than casual workers, better-educated workers get paid more than illiterate workers, and wages are higher in the services sector, followed by industry and then agriculture.

› **Table 2**

Employment as per Usual Status as on 1-1-2000 (Numbers in thousands)

| Activity | Rural Areas | | | Rural Areas | | |
|---|---|---|---|---|---|---|
| | Male | Female | Total | Male | Female | Total |
| Employed | 198,610 | 105,687 | 304,297 | 75,382 | 18,202 | 93,584 |
| Unemployed | 3,366 | 1,060 | 4,427 | 3,493 | 1,048 | 4,540 |
| Unemployment Rate | 1.67 | 0.99 | 1.43 | 4.43 | 5.44 | 4.63 |
| Labour Force | 201,976 | 106,747 | 308,724 | 78,875 | 19,249 | 98,124 |
| Not in Labour Force | 172,054 | 246,721 | 418,774 | 66,650 | 111,699 | 178,349 |
| Total | 374,030 | 353,468 | 727,498 | 145,525 | 130,948 | 276,473 |

Note: Due to rounding off the figures, totals do not tally.

› **Table 3**

Average Hourly Wages and Salaries as per Constant 2000 Prices

*Source:* R. Hasan and R. Magsombol, "Labour Markets in India: Some Findings from NSS Data," Economics Research Department, Asian Development Bank, 2005, mimeo.

| | Rural | | Urban | |
|---|---|---|---|---|
| | 1994 | 1994 | 1994 | 1994 |
| **By gender** | | | | |
| Male | 7.83 | 10.06 | 17.07 | |
| Female | 2.80 | 4.22 | 8.33 | 12.39 |
| **By type** | | | | |
| Regular | 8.90 | 14.55 | 14.37 | 19.92 |
| Casual | 3.42 | 5.02 | 5.55 | 7.35 |
| **By education** | | | | |
| Illiterate | 3.20 | 4.62 | 5.99 | 7.28 |
| College | 14.86 | 23.17 | 22.97 | 32.51 |
| **By sector** | | | | |
| Agriculture | 3.23 | 4.67 | 4.74 | 6.09 |
| Industry | 5.70 | 8.63 | 10.83 | 13.78 |
| Services | 9.33 | 14.58 | 13.64 | 19.01 |

cally established by the British economist Alban W. Phillips and is called the Phillips curve. Let us now examine how the analysis of the labour market by Keynes and Friedman can give us an insight into the Phillips curve relationship.

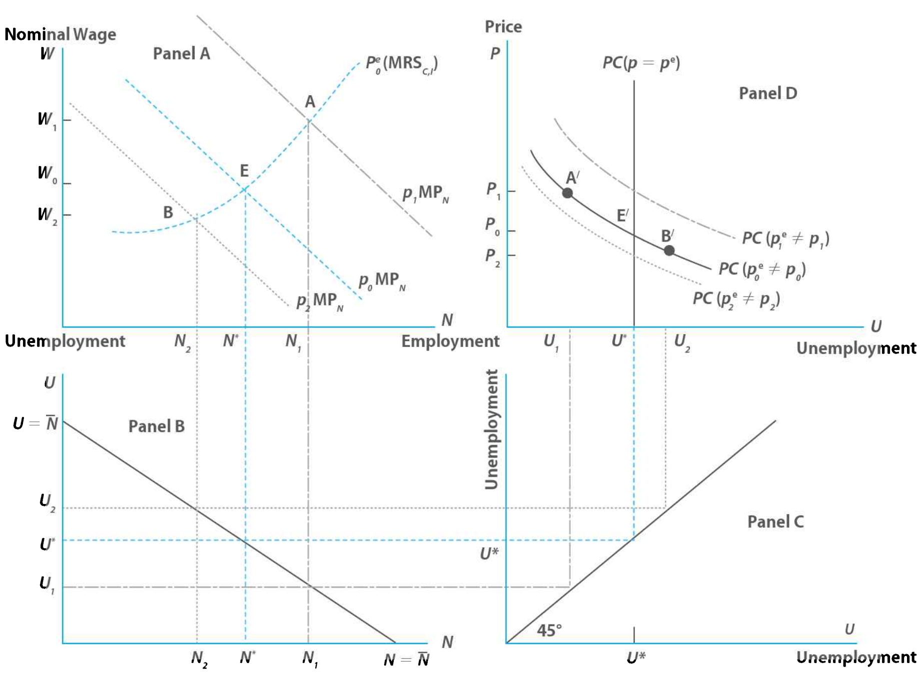
# 7.6 The Static Phillips Curve

It is best to approach the Phillips curve, the trade-off between unemployment and inflation, via its static version that depicts the trade-off between the price level and unemployment. Section 7.7 deals with the standard dynamic version of the Phillips curve.

In Figure 7.12, panel A represents the demand for labour given by the condition that the nominal wage equals the revenue generated from the productivity of the marginal unit of labour hired, or $w = p\text{MP}_N$. The supply curve of labour is upward sloping with the supply of labour increasing with an increase in the expected price, as individuals substitute labour for leisure, or $w = p^e$ $\text{MRS}_{C,l}$. Let the expected price be $p_0^e$ and the demand and supply curves of labour intersect at point E. We know that point E represents an equilibrium in the labour market where with price expectations realized, $p_0^e = p_0$, the natural rate of employment prevails. Let the total labour force of the economy be given by $\overline{N}$. Those who are not employed in the labour market are considered to be unemployed so that

$$\overline{N} = N + U$$

where $U$ is the number of people who are unemployed. In panel B of Figure 7.12, we measure unemployment on the vertical axis and employment on the horizontal axis. At $U = \overline{N}$, everyone in the labour force is unemployed and at $N = \overline{N}$, there is full employment. Joining the points $U = \overline{N}$ and $N = \overline{N}$ gives us a locus of points, each of which informs us as to what level of unemployment is associated with a given employment determined in the labour market.




**› Figure 7.12**
The Static Phillips Curve. When workers supply labour on the basis of the expected price, firms are able to employ more labour if the price of the output increases (panel A). The increase in employment reduces the unemployment in the economy (panel B). A reduction in unemployment is associated then with a rise in the price when workers supply labour on the basis of expected prices. This results in the inverse relation between price and unemployment that is the static Phillips curve $PC(p^e \neq p)$. When price expectations are realized, the static Phillips curve $PC(p^e = p)$ is vertical (panel D).

Panel C again is a 45-degree line that transfers the level of unemployment from the vertical axis of panel B to the horizontal axis of panel D. Associated with the natural rate of employment, $N^*$ when the price level in the economy is $p_0$, is the natural rate of unemployment $U^*$ in panel D. The combination $p_0$ and $U^*$ in panel D at point E′ represents a combination of price level and unemployment, which depicts one point among the possible trade-offs between prices and unemployment.

If the actual price rises to $p_1$, for given price expectation $p_0^e$, the demand and supply of labour equilibrate at point A in panel A. Employment rises to $N_1$ (panels A and B) and unemployment declines to $U_1$ (panels B–D). The lower unemployment $U_1$ is associated with the price level $p_1$ and point A′ in panel D represents another point on the static short-run Phillips curve.

In a similar vein, if the price level declined to $p_2$ for given price expectation $p_0^e$, employment declines to $N_2$ (panels A and B), and unemployment rises to $U_2$ (panels B–D). Unemployment $U_2$ is associated with a price $p_2$ and depicted as point B′ in panel D. Joining the points A′, E′, and B′ gives the inverse relationship between the price level and unemployment, which is known as the STATIC PHILLIPS CURVE $PC(p_0^e \neq p_0)$.

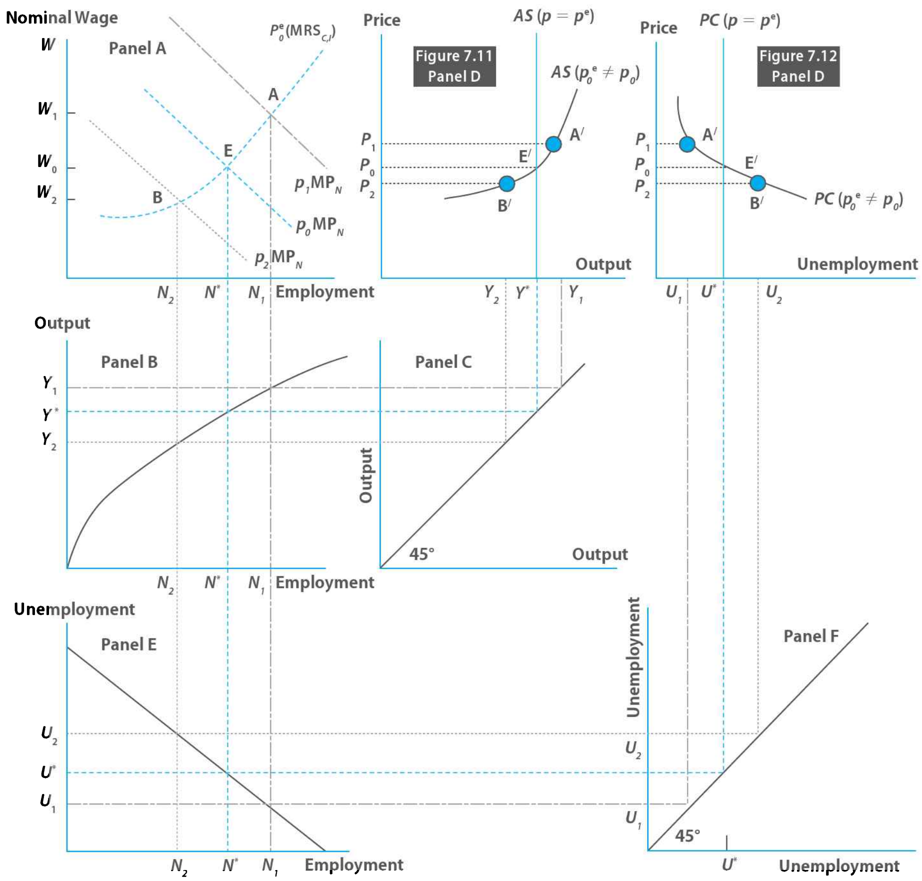
› The static Phillips curve depicts a negative relationship between the price level and unemployment.

What happens at a point such as A′ where prices are greater than expected? Recall that workers change their price expectation according to the adaptive expectations hypothesis. As $p_1 > p_0^e$, workers revise their price expectation in the next period by a fraction $\lambda$ of the error in expectation $(p_1 - p_0^e)$. They then supply labour on the basis of their new expectation about the price level $p_1^e$, where $p_1^e > p_0^e$. This shifts up the supply curve of labour to $p_1^e \text{MRS}_{C,l}$ in panel A (not shown in the diagram) and the new short-run Phillips curve is $PC(p_1^e \neq p_1)$. Similarly, at point B′, price expectations are revised downwards resulting in a new short-run Phillips curve $PC(p_2^e \neq p_2)$. If price expectations are realized, we obtain the vertical Phillips curve $PC(p^e = p)$ at the natural rate of unemployment $U^*$.

Since panel A in Figures 7.11 and 7.12 are identical, it is but obvious that the panel D in Figures 7.11 and 7.12 are graphically stating the same thing in different ways. Figure 7.13 replicates both Figures 7.11 and 7.12 in one diagram. It is obvious that the factors resulting in points A′, B′, and E′ of Figure 7.11 (panel D) result as well in the outcomes depicted by points A′, B′, and E′ in Figure 7.12 (panel D). In fact, if one were to place a mirror along the vertical axis of the farthest north-eastern quadrant of Figure 7.13—along the vertical axis of Figure 7.12 (panel D)—the Phillips curve would appear as a mirror image of the mistaken expectations *AS* curve. The PHILLIPS CURVE and the *AS* curve are, therefore, interchangeable ways of stating the same—that there is a positive trade-off between changes in price (or inflation) and output, or there is a negative trade-off between changes in the price level and unemployment.

› The Phillips curve is the mirror image of the *AS* curve. The former posits a negative trade-off between the price level and the unemployment, and the latter posits a positive relationship between the price level and the output.

The Phillips curve is usually stated in its dynamic form—the inverse relation between inflation (not price levels) and unemployment. This dynamic version of the Phillips curve can be derived in much the same way as the static version. The dynamic *AS* curve, with the output supplied related to the rate of inflation, will then form the basis for deriving the dynamic Phillips curve. Diagrammatically, this implies replacing $p$ and $p^e$ in Figure 7.12 with inflation and expected inflation and proceeding as before. We leave this way of proceeding to the interested reader and in its place derive the dynamic Phillips curve a bit more formally in Section 7.7.




› **Figure 7.13**
The Phillips Curve as an Image of the Aggregate Supply Curve.
A rise in the price raises the demand for labour in panel A. This increase in employment raises the output produced in panel B and reduces unemployment in panel E. The increased output of panel B that accompanies the rise in the price is represented by the upward sloping *AS* curve *AS* ($p_0^e \neq p_0$) that is a redoing of Figure 7.11 panel D. Similarly, the reduced unemployment that accompanies the rise in the price is represented by the downward-sloping *PC* ($p_0^e \neq p_0$) that is a redoing of Figure 7.12 panel D. The same price rise of panel A is then associated with an increase in the output and a decline in the unemployment that are depicted in panel D. The *AS* curve is, therefore, the mirror image of the *PC* curve.

## **7.7 The Dynamic Phillips Curve

Following the discussion of Figure 7.10, we argued that as prices increase, the demand for labour curve shifts upwards and unemployment declines. This gives an inverse relationship between price increases or inflation and unemployment. This result hinges on the assumption that money wages are rigid so that an increase in prices reduces the real wage cost of labour and makes it profitable for the firm to employ more workers. However, to be fair to the Keynesians, like Friedman, they do maintain that there is some adjustment to wages over time—specifically, as unemployment increases, the excess supply of labour results in a pressure that reduces the money wage rate. We now incorporate this adjustment process of nominal wages.[15]

In the Keynesian set-up, employment is determined by the demand for labour as money wages have been set higher than market clearing levels, resulting in unemployment. So whatever output is produced depends strictly on how much labour the firm demands, $N = N^D$. The demand for labour is given by the condition that the wage cost equals the value of the marginal product of labour, $w = pMP_N$. This means that[16] the percentage change in wages ($\Delta w/w = dw/w$) is equal to the percentage change in prices $\Delta p/p$ plus the percentage change in the marginal product of labour, $\Delta MP_N/MP_N$ or

$$\frac{\Delta w}{w} = \frac{\Delta p}{p} + \frac{\Delta MP_N}{MP_N} \tag{7.10}$$

$\Delta MP_N / MP_N$ is the growth in labour productivity which we denote as

$$\rho = \frac{\Delta MP_N}{MP_N}$$

Also, $\Delta p/p$ is the percentage change in prices or the inflation rate that we denote as $\pi$. Then, we may write,

$$\pi = \frac{\Delta w}{w} - \rho \tag{7.11}$$

Now, if there is unemployment that means there is an excess supply of labour, pressure is generated by the unemployed to reduce money wages as they are willing to undercut those who are already employed. We could write

$$\frac{\Delta w}{w} = \alpha_0 - \alpha_1 u \tag{7.12}$$

› The unemployment rate is the ratio of the number of the unemployed to those in the labour force. $u = U/\bar{N}$

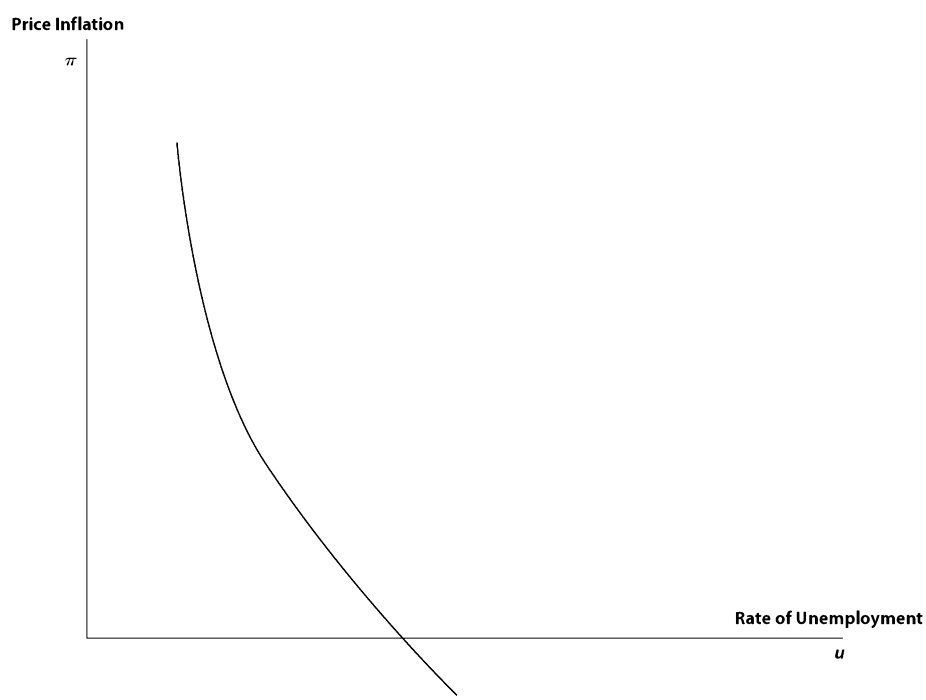
where $u$ is the UNEMPLOYMENT RATE. Substituting this into Eq. (7.11), we obtain

$$\pi = \alpha_0 - \alpha_1 u - \rho \tag{7.13}$$

› The Phillips curve depicts the unemployment rate as being inversely related to the rate of inflation.

The Keynesian set-up thus involves an INVERSE RELATIONSHIP BETWEEN INFLATION AND UNEMPLOYMENT, a relationship empirically observed by the English economist Alban W. Phillips for long periods of time and particularly marked from 1861 to 1913 and again from 1948 to 1957.[17] The Phillips curve is an empirical relationship regarding inflation–unemployment trade-offs, and is not a complete theory of inflation as it involves the unemployment rate, which is itself an endogenous variable that needs to be explained.

The Phillips curve should, therefore, be thought of as describing the trade-offs in the economy between two endogenous variables—price inflation and unemployment. It stresses that a lower level of unemployment can only be achieved at the expense of a higher rate of inflation and vice-versa. The graph of the Phillips curve is depicted in Figure 7.14.[18] It is important to note that the position of the Phillips curve depends on the growth of labour productivity. An increase in the marginal productivity of labour lowers the rate of inflation for a given unemployment rate, $u$ , and shifts the Phillips curve towards the origin.




**› Figure 7.14**
The Phillips Curve. The Phillips curve is named after the economist Alban W. Phillips who found an empirical relationship between inflation and unemployment, as depicted in the diagram.

The Phillips curve, with a negative empirical relationship between inflation and the unemployment rate, started breaking down from the middle of the 1960s. Empirical estimates became unstable and their degree of explanatory power began declining. This was attributed to the fact that over the course of the 1960s, inflation rates increased worldwide. Friedman and Phelps[19] argued that once inflation rates increase, inflation expectations begin to play a major role in wage negotiations, separate from the role played by an excess supply of labour or unemployment. Thus, what matters for workers in wage negotiations is not the increase in nominal wages, which is moderated by the unemployment rate, but the increase in expected real wages. As the actual change in the price level at the time of wage negotiations is not known, it is not actual inflation but the expected inflation $\pi^e = (\Delta p/p)^e$ that is important to wage bargains. Hence,

$$\frac{\Delta w}{w} = \alpha_0 - \alpha_1 u + \alpha_2 \pi^e,\ 0 \leq \alpha_2 \leq 1, \alpha_1 > 0 \tag{7.14}$$

Substituting into the differential of the demand for labour, $\pi = (\Delta w/w) - \rho$, as derived in Eq. (7.11), we get

$$\frac{dp}{p} = \pi = \alpha_0 - \alpha_1 u + \alpha_2 \pi^e - \rho \tag{7.15}$$

Assume the simplest expectations hypothesis—the static hypothesis—where the anticipated rate of inflation for period $t + 1$, $\pi^e_{t+1}$, equals the previous rate of inflation,[20] $\pi_t$ or

$$\pi^e_{t+1} = \pi_t \tag{7.16}$$

Then,

$$\pi_{t+1} = \alpha_0 - \alpha_1 u_{t+1} + \alpha_2 \pi_t - \rho \tag{7.17}$$

Equation (7.15) can be thought of as a short-run trade-off between the current inflation rate and unemployment for a given level of inflationary expectations. As these expectations change, as given by Eq. (7.16), the position of the short-run Philips curve shifts. Substituting Eq. (7.16) into Eq. (7.15) gives us Eq. (7.17), which describes how the Phillips curve shifts over time. In the long-run, inflation will converge to the STATIONARY STATE, where $\pi_{t+1} = \pi_t = \pi$, or

› A stationary or steady state is one where the growth rate of a variable is constant, or, where there is no tendency for the variable to change with time.

$$(1 - \alpha_2)\pi = \alpha_0 - \alpha_1 u - \rho \tag{7.18}$$

or

$$\pi = \frac{\alpha_0 - \rho}{1 - \alpha_2} - \frac{\alpha_1}{1 - \alpha_2} u \tag{7.19}$$

The LONG-RUN PHILLIPS CURVE has a negative slope of $\alpha_1/(1 - \alpha_2)$ as given by the coefficient of $u$ in Eq. (7.19), which is steeper than the slope of the short-run Phillips curve of $\alpha_1$ given in Eq. (7.15).

› The long-run Phillips curve is steeper than the short-run Phillips curve as inflation expectations are updated in wage negotiations.

In the long-run, as price expectations are updated, wage negotiations take this information into account. A given rate of unemployment is transformed into a higher inflation rate as compared to the short-run situation when expectations are unchanged.

Figure 7.15 depicts the long-run and short-run Phillips curves. If $\alpha_2 = 1$, the long-run Phillips curve of Eq. (7.18) is vertical. $\alpha_2 = 1$ is the case when wages are negotiated in expected real wage terms ($w/p^e$). When the Phillips curve is vertical, there is no long-run trade-off between inflation and the rate of unemployment. In fact, at the inflation rate that is achieved after expectations have adjusted in the long-run the economy will always tend towards the same rate of unemployment. This prompted Friedman to term it as the *"natural rate of unemployment".*

To understand this further, consider what happens in the "natural rate" case when wage negotiations are going on in expected real wage terms, $\alpha_2 = 1$. Equation (7.18 ) can then be written as

$$\alpha_0 - \alpha_1 u - \rho = 0$$

or

$$u^* = \frac{\alpha_0 - \rho}{\alpha_1} \tag{7.20}$$

where $u^*$ is the natural rate of unemployment. The impact of the natural rate can be understood in terms of the dynamics of the short-run Phillips curve as given in Eq. (7.17) when $\alpha_2 = 1$, or

$$\pi_{t+1} = \alpha_0 - \alpha_1 u_{t+1} + \pi_t - \rho$$

$$= (\alpha_0 - \rho) - \alpha_1 u_{t+1} + \pi_t$$

The natural rate of unemployment from Eq. (7.20), $\alpha_1 u^* = (\alpha_0 - \rho)$, can be substituted into the above to give

$$\pi_{t+1} = \alpha_1 (u^* - u_{t+1}) + \pi_t \tag{7.21}$$

Let the initial rate of inflation be $\pi_0$, and let the economy be at the natural rate of unemployment, $u^*$. The short-run Phillips curve for period 1 is given by

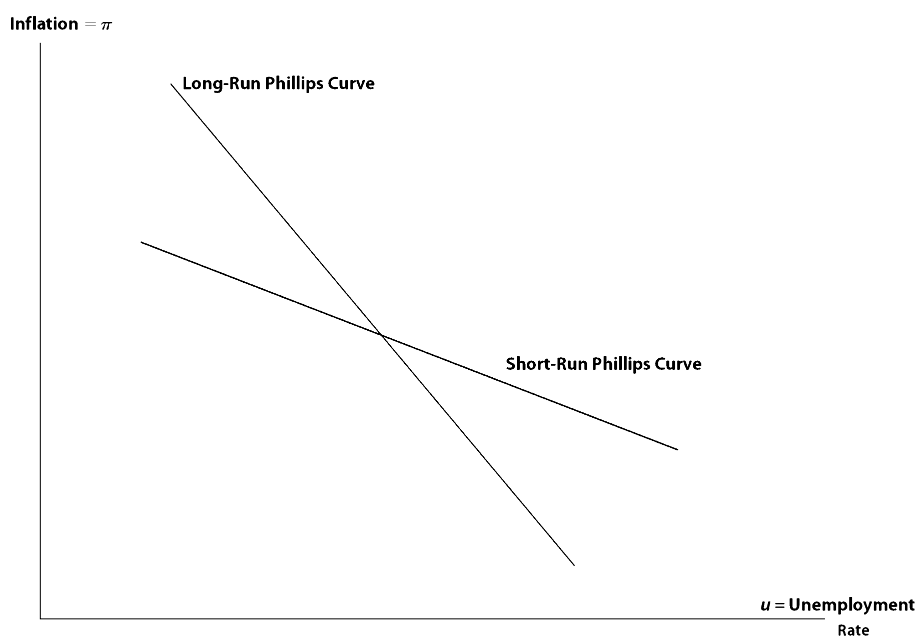
$$\pi_1 = \alpha_1 (u^* - u_1) + \pi_0$$

and represented by $P_1P_1$ in Figure 7.16. With unemployment at its natural rate ($u_1 = u^*$), the economy is at point A and inflation is $\pi_0$.

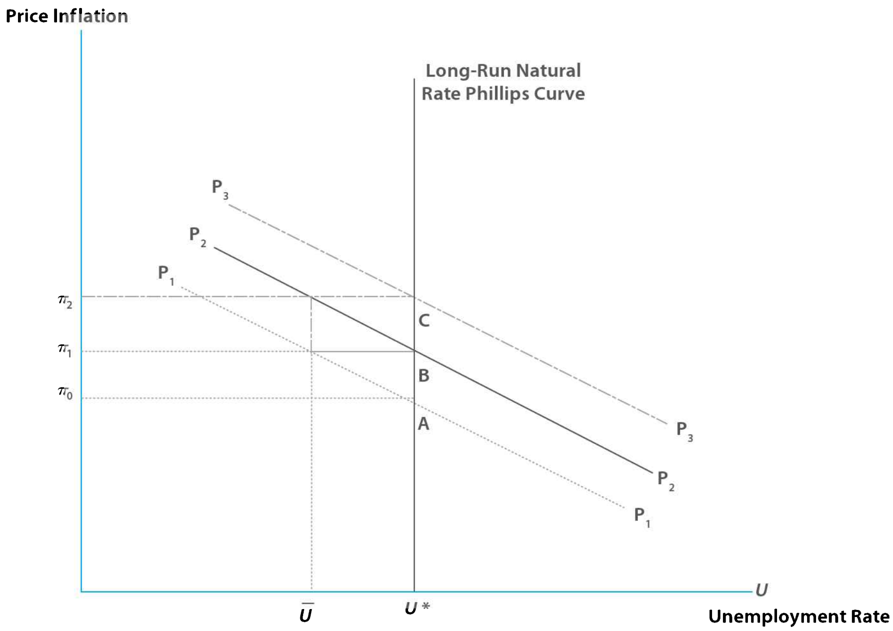
Suppose the government is bothered about unemployment, as elections are round the corner, and it attempts to reduce the level of unemployment to $u = \bar{u}$. In the short run, this immediately raises the rate of inflation to $\pi_1$. This raises the expectations of inflation in the next period to $\pi_2^e = \pi_1$ and the short-run Phillips curve for period 2, $P_2P_2$, is given by

$$\pi_2 = \alpha_1 (u^* - \bar{u}) + \pi_1$$

The curve $P_2P_2$ is parallel to $P_1P_1$ but the inflation associated with $u^*$ on this curve is higher at $\pi_1$. If in period 2 the government gives up its resolve to




**› Figure 7.15**
The Short- and Long-run Phillips Curve. **As inflation rises, individuals revise their estimate of expected inflation. Wage negotiations that result from increases in expected inflation over time increase the slope of the Phillips curve in the long run.**




**› Figure 7.16**
The Natural Rate of Unemployment. **A government that attempts to reduce unemployment to $\bar{U}$ raises the short-run rate of inflation initially to $\pi_1$. As individuals revise their expectations in response to this increase in inflation, the short-run Phillips curve shifts up to $P_2P_2$. If the government persists in maintaining unemployment at $\bar{U}$ the rate of inflation keeps accelerating and the economy moves along successive short-run Phillips curves. It is only at $u^*$ that the current rate of inflation does not change and this point is referred to as the non-accelerating inflation rate of unemployment or NAIRU.**

reduce unemployment because it is sure to obtain another term in office, the economy will return to the natural rate of unemployment at point B. However, now the economy has obtained a higher rate of inflation than earlier. A ONE-TIME REDUCTION IN UNEMPLOYMENT leads to a *permanent* increase in the rate of inflation.

› A government that attempts a one-time reduction in the unemployment rate will increase the rate of inflation permanently.

If the government decides to sustain a lower rate of unemployment, $\bar{u} < u^*$, then in period 3, the new short-run Phillips curve $P_3P_3$ will pass through the point ($\pi_2$, $u^*$). As long as the government continues to maintain, this rate of unemployment given by $\bar{u} < u^*$, the short-run Phillips curves will continue to shift outwards and the rate of inflation will continue to increase. This is referred to as the ACCELERATION HYPOTHESIS, as price increases will continuously accelerate as long as unemployment is held constant at $\bar{u} < u^*$, and the rate of inflation will not settle at any stable value.

› A government that attempts to sustain the unemployment rate below the natural rate will cause accelerating inflation. This is called the accelerator thesis.

Conversely, if unemployment is held above the natural rate, $\bar{u} > u^*$ prices will decelerate continuously. As long as $\bar{u} < u^*$, then, inflation continues to accelerate in an upward spiral and when $\bar{u} > u^*$ inflation continues to decelerate. At $\bar{u} = u^*$, the current rate of inflation, whatever it is, will persist—it will neither accelerate nor decelerate. The point $u^*$ is thus sometimes also called the non-accelerating inflation rate of unemployment or NAIRU.

› The NAIRU, or the non-accelerating inflation rate of unemployment, is that rate of unemployment at which the current rate of inflation will continue, neither accelerating nor decelerating. This is the point at which the Phillips curve cuts the X-axis.

Even though we have illustrated the results for the case of static price expectations, the argument holds in general. The general form of Eq. (7.21) may be written[21] as

$$\pi_{t+1} - \pi^e_{t+1} = \alpha_1 (u^* - u_{t+1}) \quad \textbf{(7.22)}$$

Thus, as long as INFLATION IS NOT FULLY ANTICIPATED in the short-run, the left-hand side of Eq. (7.22) is positive, and there is a temporary trade-off between inflation and unemployment. When inflation is perfectly anticipated as occurs in the long-run, the left-hand side of Eq. (7.22) equals zero, and unemployment is at its natural rate on the vertical long-run Phillips curve.

› Inflation is unanticipated when the actual inflation does not equal the expected inflation. Unanticipated inflation results in a Phillips curve.

The Phillips curve, as derived so far, was based on the adaptive-expectations mechanism. This mechanism assumes that individual agents make systematic errors in forecasting the rate of inflation. This has been questioned by rational expectations theorists who argue that the public does not make systematic errors in forecasting future economic developments. They submit that rational forecasters do make mistakes. However, they do not consistently over or underestimate what is going to happen. If, for example, agents found from experience that on average they had underestimated inflation by 2 per cent, they would correct this systematic error and add 2 per cent to their forecasts towards improving future predictions. The basic assumptions of the RATIONAL EXPECTATIONS approach were first stated by John F. Muth.

› Rational expectations theorists argue that information is scarce and agents do not waste it (efficiency); that expectations are formed in a well-specified way on the structure of the system describing the economy (consistency); and that agents do not make systematic mistakes in their forecasts (unbiasedness).

John F. Muth[22] argued that rational agents must use all the relevant information available at the time at which they form their expectations, a property known as efficiency. Second, when forming expectations at time $t$ over successive time horizons, they must utilize all the information available at $t$ in a logically consistent manner—the consistency property of rational expectations.

For example, if agents know in April that, say, two months down the line they would alter their expectation of inflation for September, then they should use that information in April to alter their view of inflation for September. Consistency demands that only *new* information can lead one to alter one's expectations. Finally, on average, expectations should be correct—the unbiasedness property. This property requires individuals to forecast correctly, they do not systematically under or over estimate the actual future value of a variable.

The rational expectations theorists thus assume that inflation expectations are formed according to

$$\pi_t^{\text{ratex}} = \pi_t + \varepsilon_t \tag{7.23}$$

where $\varepsilon_t$ is the forecast error $\varepsilon_t > 0$ when the rate of inflation is underestimated and $\varepsilon_t < 0$ when the actual inflation falls short of the anticipated rate. When forecasts are rational, the forecast errors $\varepsilon_t$ must average out to zero and be unpredictable rather than systematic—they are purely random.

When workers do not make systematic errors in anticipating inflation, they bargain in expected real wage terms and are not misled by inflation. In that case, $\alpha_2 = 1$ in Eq. (7.15), since workers anticipate inflation on average correctly. Then, Eq. (7.15) may be written as

$$\pi_t = \alpha_0 - \alpha_1 u_t + \pi_t^e - \rho \tag{7.24}$$

However, inflation expectations are formed rationally as given by Eq. (7.23), $\pi_t^e = \pi_t^{\text{ratex}}$, thus allowing us to write

$$\pi_t = \alpha_0 - \alpha_1 u_t + \pi_t + \varepsilon_t - \rho$$

Solving for $u_t$, we obtain

$$u_t = \frac{\alpha_0 + \varepsilon_t - \rho}{\alpha_1} \tag{7.25}$$

Under rational expectations then, the unemployment rate is unaffected by inflation even in the short run unlike other expectations mechanisms such as static expectations [compare Eq. (7.25) with Eq. (7.17)]. Moreover, the gap between the unemployment in any period under rational expectations and the natural rate of unemployment—as given by Eq. (7.20)—is given by

$$u_t - u^* = \frac{\alpha_0 + \varepsilon_t - \rho}{\alpha_1} - \frac{\alpha_0 - \rho}{\alpha_1} = \frac{\varepsilon_t}{\alpha_1} \tag{7.26}$$

This means that when EXPECTATIONS ARE RATIONAL, unemployment deviates randomly from the natural rate of unemployment. Even in the short run, deviations from the long-run Phillips curve or the *AS* function are purely random unlike the case of adaptive expectations.

› In rational expectations theory, the deviation of unemployment from its natural rate is random and not an extrapolation of the past as it is in adaptive expectations.

Rational expectations theorists attribute the departure of the short-run *AS* function from the vertical long-run *AS* function due to random misperceptions about price rather than any systematic errors of expectation. An important implication of rational expectations then is that if governments attempt to reduce unemployment below the natural rate to $\bar{u} < u^*$, and if this policy action is anticipated, there will be no effect on output and employment. Only those policy actions that come as a complete surprise, as represented by the random term $\varepsilon_t$ in Eq. (7.25), will result in an impact on the level of unemployment. If a government persists in attempting to maintain unemployment below the natural rate, the policy is no longer random and is anticipated. In that case, there will be no effect of the policy on the economy. Thus, when expectations are rational, government policy is ineffective in its attempts to stimulate the economy, when prices and wages adjust quickly to equate the demand and supply of labour. This has been labelled the POLICY INEFFECTIVENESS PROPOSITION.[23]

› The policy ineffectiveness proposition states that the government cannot systematically exploit the Phillips curve to guide the economy to a point on it. This is because with rational agents the policy will come to be anticipated and workers will seek to protect their real wages, with no consequent effect on employment and output.

## SUMMARY


» Profit-maximizing firms hire workers up to the point where the wage equals the value of the marginal product of labour.

» Workers maximize their utility over consumption and leisure in such a way that the last rupee spent on leisure activities yields the same utility as the last rupee spent on consumption

$$\frac{MU_l}{w/p} = MU_c.$$

» An increase in the wages results in both an income and a sub-

» stitution effect. The income effect reduces labour supply while the substitution effect increases it. The labour-supply curve is upward sloping when the substitution effect dominates.

» When there is no money illusion, an increase in the price of output results in an equiproportionate increase in the money wage and no change in employment and output. This results in a classical *AS* curve that is vertical.

» When individuals formulate labour supply decisions on the basis of the money wage, they are subject to money illusion. The corresponding *AS* curve is upward sloping.

» The assumption that money wages are fixed and that employment is determined by the short side of the labour market, which is usually the demand for labour side, results in an upward-sloping Keynesian *AS* of output. This is up to the point where all those who offer to supply labour at the fixed money wage have been employed.

» When workers have mistaken expectations regarding the price of output, the supply curve for a given price expectation is upward sloping.

» When expected prices equal actual prices, expectations are realized and the economy is at the natural rate of employment and output, with a vertical *AS* curve.

» Under adaptive expectations, individuals revise their price expectation in any period to incorporate a fraction of the error in expectation ($p - p^e$) from the previous period.

» Actual unemployment in an economy is the sum of the natural rate of unemployment and cyclical unemployment.

» The natural rate of unemployment is the sum of the seasonal, frictional, and structural unemployment that do not vary with the level of economic activity. In theory, the natural rate of unemployment is that rate of unemployment that prevails when price expectations are realized.

» Cyclical unemployment is the unemployment associated with recessions and expansions in economic activity, and equals the unemployment rate less the natural unemployment rate.

» Potential or full employment output is the maximum sustainable amount of output that an economy can produce.

» The output gap is the difference between potential output and actual output measured in relation to potential output.

» Okun's Law relates cyclical unemployment to the output gap. It states that each extra percentage point of cyclical unemployment is associated with a 3 percentage point increase in the output gap.

» The Phillips curve depicts the unemployment rate as being inversely related to the rate of inflation for a given growth in labour productivity.

» The long-run Phillips curve is steeper than the short-run Phillips curve as inflation expectations are updated in wage negotiations.

» A government that attempts to sustain the unemployment rate below the natural rate of unemployment will cause accelerating inflation—the acceleration hypothesis.

» Rational expectations theorists argue that information is scarce and agents do not waste it (efficiency); that expectations are formed in a well-specified way on the structure of the system describing the economy (consistency); and that agents do not make systematic mistakes in their forecasts (unbiasedness).

» With rational expectations, there are no systematic errors of expectations and governments are unable to systematically exploit the Phillips curve—the policy ineffectiveness proposition.

## NOTES

1. Maximizing Eq. (7.2) we obtain $\partial \Pi / \partial N = p\,(\partial F / \partial N) - w = 0$ or $p\,(\partial F / \partial N) = p\text{MP}_N = w$.
2. In the long run, the firm chooses both how many workers to hire and how much plant and equipment to invest in. If the two factors are co-operative in the sense that increasing one factor raises the marginal productivity of the other factor: $F_{NK} = F_{KN} > 0$, the long-run demand curve for labour will be more elastic than the short-run demand curve.
3. Firms and workers are connected in two different ways—through the goods market and through the labour market. Hence, behaviour in one market can affect behaviour in the other market. This creates a problem that demand variables are dependent on supply side variables and can result in circular reasoning. To avoid this, we assume that preferences are separable so that labour market variables do not affect money and goods market variables (see Appendix 7.2). How to deal with this issue remains unresolved in macroeconomics.
4. For simplicity, we are ignoring that some income is saved.
5. Differentiating $U(C, l) = \overline{U}$, we get

$$dU = 0 = \frac{\partial U}{\partial C} dC + \frac{\partial U}{\partial l} dl$$

or

$$-\frac{dC}{dl} = \frac{\partial U / \partial l}{\partial U / \partial C} = \frac{\text{MU}_l}{\text{MU}_c}$$

6. See Franco Modigliani, "Liquidity Preference and the Theory of Interest and Money," *Econometrica* 12 (1944): 45–88.
7. Later we will look at other explanations for unemployment as provided by the efficiency wage and insider–outsider frameworks.
8. Even though Keynesians believe that money wages decline in the face of unemployment and excess supply of labour, $N^S > N^D$, they conjecture that this occurs with a time lag due to the complexities of wage negotiations and so in the short-run money wages can be taken to be rigid.
9. See Appendix 7.1 for the derivation of this equation.
10. Slowdowns in economic growth are called recessions. When the slowdown is severe as occurred in the period following 1929, it is called a depression.
11. This can be accessed online at http://www.cbo.gov.
12. Lawrence Katz and Alan Krueger, "The High-Pressure U.S. Labour Market of the 1990s," *Brookings Papers on Economic Activity* 1 (1999): 1–88.
13. Arthur M. Okun, *The Political Economy of Prosperity* (New York: Norton, 1970).
14. The opportunity cost of an activity is the value of the next-best alternative that must be forgone in order to undertake the activity.
15. If an increase in prices reduces unemployment, which, in turn, raised the money wage because of a reduction in the effective supply of labour, then, for the real wage and consequently unemployment to decline we would require that the percentage change in wages is less than the inflation rate. Keynesians therefore presume that money wages are somewhat sticky.
16. This follows from taking logarithms and totally differentiating the condition for the demand for labour.
17. See Alban William Phillips, "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," *Economica* 25 (1958): 283–299.
18. Even though our analysis suggests a linear relationship, empirical evidence suggests a hyperbolic relationship, where $\Delta p/p = \alpha_o - \alpha_1 U^{-1} - \rho$.
19. See Milton Friedman, "The Role of Monetary Policy," *American Economic Review* 58, no.1 (1968): 1–17; Edmund S. Phelps, "Phillips curves, Expectations of Inflation and Optimal Unemployment Over Time," *Economica* 34 (1967): 254–281.
20. This can be also thought of as the adaptive expectations hypothesis when expectations prove correct in the current period, or $\pi_t^e = \pi_t$.
21. We have taken $\pi_t$ in Eq. (7.24) to the left-hand side and replaced it by its general form, $\pi_{t+1}^e = \pi_t$.
22. John F. Muth, "Rational Expectations and the Theory of Price Movements," *Econometrica* (July1961).
23. Thomas J. Sargent and Neil Wallace, "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics* 2 (1976): 169–183.

## TEST YOURSELF

1. Why is the demand curve for labour downward sloping?
2. How does a worker decide how many hours to allocate to the labour market?
3. What does it mean to say that an individual is subject to money illusion?
4. How is the *AS* of output curve derived if workers are subject to money illusion?
5. Derive the *AS* curve when workers are subject to mistaken expectations about the price of output.
6. Define four types of unemployment and the causes for the same. Which type of unemployment is least costly for the economy? Why?
7. What is Okun's Law? Explain how it is possible for an economy to produce an output greater than potential output.
8. Why is there a trade-off between unemployment and inflation in the short run?
9. If unemployment is pushed below its natural rate what will happen to inflation?
10. What is the long-run trade-off between inflation and unemployment?

## ONLINE APPLICATION

1. The following is the data on the unemployment rate (as measured by the usual principal status) in India for some years:

| *Incidence of Unemployment: Usual Status* | |
|---|---|
| 1972–1973 | 1.6 |
| 1977–1978 | 2.6 |
| 1982–1983 | 1.9 |
| 1987–1988 | 2.7 |
| 1993–1994 | 1.9 |
| 1999–2000 | 2.8 |
| 2004–2005 | 3.1 |

*Source:* NSSO, Central Statistical Organisation, India.

2. Go the RBI Web site (URL: www.rbi.org.in/home.aspx).
3. Click on the "Database" icon
4. Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy".
5. Click on "Table 39: Wholesale Price Index—Annual Average". Calculate the inflation rate for all commodities for the years listed in the table above from this web page.
6. Create a scatter diagram between the inflation rate and the unemployment rate. Your diagram should be as depicted in Figure 7.17.
7. In this scatter diagram note that when unemployment increased between 1982–1983 and 1987–1988, and again between 1999–2000 and 2004–2005, inflation increased rather than decreased contrary to what a Phillips curve would predict.
8. Why would inflation increases have been accompanied by unemployment increases in the recent past? To try and understand the factors responsible for the rise in inflation in 2004–2005, read the relevant sections of the government's Economic Survey, 2005.
   - Go to (URL: http://indiabudget.nic.in).
   - Under "Previous Economic Surveys" click on "2005–2006".
   - Click on "Prices and Food Management".
   - Click on "Inflation".


› **Figure 7.17**
Inflation and Unemployment in India.

# Appendix 7.1 The Mistaken Expectations Aggregate Supply Curve

The demand and supply curves of labour are given by

$$w = p\text{MP}_N \tag{A7.1.1}$$

$$w = p^{e}\text{MRS}_{C,l} \tag{A7.1.2}$$

Taking logarithms of Eq. (A7.1.1) and differentiating,

$$\frac{dw}{w} = \frac{dp}{p} + \frac{d(\text{MP}_N)}{\text{MP}_N} \tag{A7.13}$$

Now,

$$d(\text{MP}_N) = \frac{\partial(\text{MP}_N)}{\partial N^{\text{D}}}\, dN^{\text{D}}$$

Thus,

$$\frac{d(\text{MP}_N)}{\text{MP}_N} = \frac{\partial(\text{MP}_N)/\partial N^{\text{D}}}{\text{MP}_N}\, dN^{\text{D}} = \frac{\partial(\text{MP}_N)/\text{MP}_N}{\partial N^{\text{D}}/N^{\text{D}}}\frac{dN^{\text{D}}}{N^{\text{D}}}$$

The wage elasticity of labour demand is defined as

$$\varepsilon_{\text{D}} = \frac{-\partial N^{\text{D}}/N^{\text{D}}}{\partial(w/p)/(w/p)}$$

or

$$\varepsilon_D = -\frac{\partial N^{\text{D}}/N^{\text{D}}}{\partial(\text{MP}_N)/(\text{MP}_N)}$$

Hence, we may write

$$\frac{d(\text{MP}_N)}{\text{MP}_N} = -\frac{1}{\varepsilon_{\text{D}}}\frac{dN^{\text{D}}}{N^{\text{D}}}$$

Substituting this result into Eq. (A7.1.3),

$$\frac{dw}{w} = \frac{dp}{p} - \frac{1}{\varepsilon_{\text{D}}}\frac{dN^{\text{D}}}{N^{\text{D}}}$$

or

$$\frac{dN^{\text{D}}}{N^{\text{D}}} = -\varepsilon_{\text{D}}\left(\frac{dw}{w} - \frac{dp}{p}\right) \tag{A7.1.4}$$

Similarly, if we take the logs of Eq. (A7.1.2), totally differentiate, and define the wage elasticity of labour supply as

$$\varepsilon_s = \frac{\partial N^{\text{S}}/N^{\text{S}}}{\partial(\text{MRS}_{C,l})/(\text{MRS}_{C,l})}$$

we obtain

$$\frac{dN^{\text{S}}}{N^{\text{S}}} = \varepsilon_{\text{S}}\left(\frac{dw}{w} - \frac{dp^{e}}{p^{e}}\right) \tag{A7.1.5}$$

Equilibrium in the labour market requires that

$$N^{\text{D}} = N^{\text{S}}$$

or

$$\frac{dN^{\text{D}}}{N^{\text{D}}} = \frac{dN^{\text{S}}}{N^{\text{S}}}.$$

Equating Eqs. (A7.1.4) and (A7.1.5),

$$\varepsilon_S\left(\frac{dw}{w}-\frac{dp^e}{p^e}\right)=-\varepsilon_D\left(\frac{dw}{w}-\frac{dp}{p}\right)$$

or

$$(\varepsilon_S+\varepsilon_D)\frac{dw}{w}-\varepsilon_D\frac{dp}{p}=\varepsilon_S\frac{dp^e}{p^e}$$

Adding $-\varepsilon_S(dp/p)$ to both sides,

$$(\varepsilon_S+\varepsilon_D)\frac{dw}{w}-(\varepsilon_D+\varepsilon_S)\frac{dp}{p}=-\varepsilon_S\left(\frac{dp}{p}-\frac{dp^e}{p^e}\right)$$

or

$$\frac{dw}{w}-\frac{dp}{p}=\frac{-\varepsilon_S}{\varepsilon_D+\varepsilon_S}\left(\frac{dp}{p}-\frac{dp^e}{p^e}\right) \tag{A7.1.6}$$

Substituting this result into the labour–demand Eq. (A7.1.4) gives us

$$\frac{dN}{N}=\frac{\varepsilon_D\varepsilon_S}{\varepsilon_D+\varepsilon_S}\left(\frac{dp}{p}-\frac{dp^e}{p^e}\right) \tag{A7.1.7}$$

The amount of labour employed produces the output defined by the production function $Y = F(\bar{\bar{K}}, N)$. Totally differentiating the production function gives us

$$dY=\mathrm{MP}_N\,dN$$

Dividing throughout this expression by $Y$ gives us

$$\frac{dY}{Y}=\frac{\mathrm{MP}_N}{Y}dN=\frac{\mathrm{MP}_N N}{Y}\frac{dN}{N} \tag{A7.1.8}$$

Now, $\mathrm{MP}_N$ is the marginal product of labour which is set equal to the real wage—see Eq. (A7.1.1). Hence, $\mathrm{MP}_N N$ is the total real wage bill and $\mathrm{MP}_N N/Y = wN/pY$ is the share of wages in total output. Define this as

$$\omega=\frac{\mathrm{MP}_N N}{Y}=\frac{wN}{pY} \tag{A7.1.9}$$

Then, Eq. (A7.1.8) can be written as

$$\frac{dY}{Y}=\omega\frac{dN}{N}$$

and on substituting in Eq. (A 7.1.7), we obtain

$$\frac{dY}{Y}=\frac{\omega\varepsilon_D\varepsilon_S}{\varepsilon_D+\varepsilon_S}\left(\frac{dp}{p}-\frac{dp^e}{p^e}\right) \tag{A7.1.10}$$

Define

$$\frac{\partial Y/Y}{\partial p/p}=\phi=\frac{\omega\varepsilon_D\varepsilon_S}{\varepsilon_D+\varepsilon_S}>0 \tag{A7.1.11}$$

where $\phi$ is the inverse of the elasticity of the short-run $AS$ curve in $(p, Y)$ space when $p \neq p^e$. The higher values of $\phi$ represent flatter short-run $AS$ curves. If we integrate Eq. (A7.1.10) we can write it as

$$Y=Y^*\left(\frac{p}{p^e}\right)^{\omega\varepsilon_D\varepsilon_S/(\varepsilon_D+\varepsilon_S)}$$

or

$$Y = Y^*\left(\frac{p}{p^e}\right)^{\phi} \qquad \textbf{(A7.1.12)}$$

where $Y^*$ is the natural rate of output when price expectations are realized or $p^e = p$. The fact that Eq. (A7.1.12) follows from Eqs. (A7.1.10) and (A7.1.11) can be checked by taking the logs of Eq. (A7.1.12) and totally differentiating, which will give us Eq. (A7.1.10). Equation (A7.1.12) is the equation of Friedman's *AS* curve, as noted in the text. Equation (A7.1.12) can also be written as

$$p = p^e\left(\frac{Y}{Y^*}\right)^{1/\phi} \qquad \textbf{(A7.1.13)}$$

where $(Y/Y^*) < 1 \Rightarrow Y < Y^*$ or there is a gap between actual output and the natural rate of output. This is akin to Okun's gap, which is a measure of unemployment.

# Appendix 7.2 Separability Between Aggregate Demand and Supply

A central problem of macroeconomics is that there tend to be *spillovers* between markets. These make it difficult to derive separate supply and demand curves for aggregate output. For instance, if there is unemployment then this is likely to affect the demand for goods in the goods market through the substitution between consumption and leisure. In that case, a supply side curve affects a demand side curve and this can lead to a situation where supply affects demand and demand, in turn, affects supply as the reduced demand for goods makes firms reduce their demand for labour even further. It, therefore, becomes necessary to make the supply and demand for goods independent of the situation in the labour market. This is achieved if we assume that preferences are separable into consumption, real money balance holdings, and labour supply. Then,

$$U\left(C, \frac{M}{P}, l\right) = U\left(C, \frac{M}{P}\right) + U(l)$$

The general problem for a representative individual then becomes

$$\max_{\{C_t, N_t^s, M_t, B_t\}} \sum_{t=0}^{\infty} \frac{1}{(1+\delta)^t} U\left(C_t, \frac{M_t}{P_t}\right) + U(H - N_t^S) \qquad \textbf{(A7.2.1)}$$

subject to

$$C_t + \frac{M_t}{P_t} + \frac{B_t}{i_t P_t} \leq \frac{w_t}{P_t} + V_t + \frac{B_{t-1}}{P_t} + \frac{M_{t-1}}{P_t} + \frac{B_{t-1}}{i_{t-1} P_t} \qquad \textbf{(A7.2.2)}$$

and

$$N_t^s \leq H \qquad \textbf{(A7.2.3)}$$

The individual thus chooses a sequence of $C_t$, labour supply $N_t^S$, nominal money balances $M_t$, and bond holdings $B_t$, to maximize utility [Eq. (A7.2.1)] subject to the budget constraint [Eq. (A7.2.2)] and the time constraint [Eq. (A7.2.3)]. In doing so, the individual takes the sequence of nominal interest rates, $i_t$, prices $P_t$, nominal wages $w_t$, and non-wage income $V_t$ as given. The specification of the firm's maximization problem will determine the output, price, nominal wages, and profits, which will accrue to consumers since firms are ultimately owned by them (we have ignored this in the specification mentioned earlier).

Since the preferences in ($C$, M/P) are separable from those in $N^S$, the substitution between consumption and real money holdings (as reflected in the slope of the indifference curve) is independent of the labour supply. Thus, demand in the goods market is independent of the specification of the labour market and the supply of output. We have followed this approach to avoid circularity. This chapter on labour supply and demand therefore assumes that the firm maximizes profits and the individual maximizes a utility function that does not impinge on the profits of the firm through consumption expenditure that is financed from incomes received from the firm. Even though we write the utility function as $U(C, l) = U(C, H - N^S)$, the consumption in this utility function is serving more the role of a numeraire that allows the individual to value the leisure (and thereby labour supply) that they choose.

# 8

# The *IS–LM* Model

CRITICAL QUESTIONS

- *What is Walras' Law and how is it applied in macroeconomics?*
- *What is the difference between the nominal and the real rate of interest?*
- *What is the* IS *curve and how does it represent the condition for equilibrium in the product market?*
- *What parameters cause the* IS *curve to shift, and in what direction?*
- *What is the* LM *curve and how does it represent the condition for equilibrium in the money market?*
- *When is the* LM *curve elastic and when is it inelastic?*
- *How does monetary policy affect interest rates and output in an* IS–LM *situation?*
- *What is the impact of fiscal policy on interest rates and output?*
- *What is crowding out and to what extent does it depend on the elasticity of the* IS *and* LM *curves?*
- *What is Ricardian Equivalence* (RE)*?*

# 8.1 Walras' Law

By now considerable insight has been gained into the determinants of consumption, investment, net exports, the demand and supply of money, and the demand and supply of labour. Our focus on these aggregate economic magnitudes was informed by the macroeconomic method of aggregating the economy up to four markets:

1. the output market
2. the money market
3. the bond market
4. the labour market

In macroeconomics, Walras' Law is used to eliminate the explicit treatment of one of these markets. Traditionally, the market-rendered implicit in the analysis is usually the bond market.

To intuitively understand Walras' Law, consider the income and expenditure statement of households, firms, and government, as depicted in Table 8.1. The expenditure of an agent is the income left over after deducting necessary payments. For the firm, income is identified as retained profits ($\Pi_r$). After getting receipts from the sale of output, $Y$, the firm pays out[1] taxes ($T_f$) and wages $[(w/p)N]$ followed by distributing to its equity holders (households) some part of post-tax profits ($\Pi_d$). Hence, retained profits that are used to finance investment expenditures are

$$\Pi_r = Y - \frac{w}{p}N - T_f - \Pi_d = I \tag{8.1}$$

Income less payments for households are equal to expenditure on consumption as well:

$$\frac{w}{p}N + \Pi_d - T_h = C \tag{8.2}$$

Finally, the income–expenditure statement for the government is

$$T_h + T_f = G \tag{8.3}$$

Adding Eqs. (8.1) to (8.3),

$$Y = C + I + G \tag{8.4}$$

However, expenditure need not be financed only by current income. Agents also hold stocks of assets that they can liquidate in order to finance current expenditures. The financial assets they hold are money and bonds that can be depleted in order to finance expenditures.[2] Accordingly, agents make plans regarding their desired holdings of money and bonds, which constitute their real demands for money and bonds, keeping in mind their planned

**› Table 8.1**
Income and Expenditure of Agents in the Economy

| Agent | Income | Payments | Expenditure |
|---|---|---|---|
| Households | Wages $(w/p)N$<br>Distributed Profits ($\Pi_d$) | Taxes on Households (net of Transfers) ($T_h$) | Consumption($C$) |
| Firms | Retained Profits ($\Pi_r$) | Wages $(w/p)N$<br>Distributed Profits ($\Pi_d$)<br>Taxes on Firms (net of Transfers)($T_f$) | Investment ($I$) |
| Government | Taxes on Households ($T_h$)<br>Taxes on Firms ($T_f$) | | Government Expenditure ($G$) |

*Note*: Expenditure = Income − Payments.

expenditures.[3] The planned holdings of money and bonds, together with the plans for expenditure on consumption, investment, and government expenditure, are accordingly constrained by the following factors:

1. the value of financial wealth at the beginning of the period under consideration and
2. the income earned over the period.

Thus, in addition to income on the left-hand side of Eq. (8.4), we must add the stock of financial assets held (money, $m$ and bonds, $b$), and to the right-hand side of Eq. (8.4), we must include the planned demand for money ($m^d$) and bonds ($b^d$).[4] Accordingly,[5]

$$m + b + Y = C + I + G + m^d + b^d \tag{8.5}$$

We can write Eq. (8.5) in implicit form as:

$$(m - m^d) + (b - b^d) + (Y - C - I - G) = 0 \tag{8.6}$$

› Walras' Law states that the sum of the excess demand for money, bonds, and current output must equal zero.

This is WALRAS' LAW—the sum of the excess demand for money, bonds and current output must equal zero. We could rearrange this as

$$(b - b^d) = -(Y - C - I - G) - (m - m^d) \tag{8.7}$$

This states that the excess demand for bonds (the left-hand side) must correspond to an excess supply of some other quantity—current output and/or money (the right-hand side).

Hence, once we have specified the workings of the output and the money market, the specification of the bond market becomes superfluous. This is why there is an explicit treatment of the markets for commodities (or output) and money in modern macroeconomics, and the bond market is ignored.

In macroeconomics, the output and money markets are combined to derive the aggregate demand curve of the economy whereas the labour market determines the output or the aggregate supply curve of the economy. Aggregate demand and aggregate supply together determine the output and income as well as the price at which that output is sold and income is spent. We, therefore, proceed by analysing the economics of the commodities, goods and services, and money markets, and then combine these to develop the aggregate demand function of the economy.

At the outset, we need to differentiate between the nominal rate of interest and the real rate of interest.

## 8.2 Nominal Versus Real Rate of Interest

As we discussed in Chapter 6, in the money market, the interest rate we referred to was the nominal interest rate as the opportunity cost of holding money was the forgone interest earnings on bonds which is a nominal return. By contrast, when we were analysing the determinants of real consumption and investment expenditure, the rate of interest that we used as a discount rate was the real rate of interest. As we are eventually going to combine the goods and money markets, we need to ensure that the interest rate referred to when goods market variables are being considered is the same as the interest rate that is referred to when the money market variables are being analysed. As we have already seen in the chapter on money demand, the REAL RATE OF INTEREST is the nominal interest rate minus the inflation rate

› The real rate of interest is the nominal interest rate minus the rate of inflation.

$$r = i - \frac{\Delta P}{P}$$
$$= i - \pi \qquad \textbf{(8.8)}$$

The inflation rate $\pi$ in Eq. (8.8) is the actual inflation rate.

The trouble with an actual inflation rate is that it is known only after the future time period for which decisions are being made today actually commences. For instance, if the interest rate is with reference to a loan, the inflation rate is what occurs over the life of the loan and is calculated as the percentage change in prices between the present, when the loan is being repaid, and the past when the loan was taken. As the inflation rate here is the actual inflation rate, the real interest rate in Eq. (8.8) is known as the *ex post* real rate, where, *ex post* means "after the fact".

In actuality the inflation rate is not known at the time when borrowers and lenders enter into commitments on the various aspects of the loan and agents must form an expectation of the future inflation rate. This leads to a second definition of the real interest rate known as the *ex ante* real interest rate, $r^e$, that is, the difference between the nominal interest rate and the expected inflation rate, $\pi^e$, where, *ex ante* means "before the fact", that is, before inflation is known.

$$r^e = i - \pi^e \qquad \textbf{(8.9)}$$

Comparing Eqs. (8.8) and (8.9), we can see that the *ex post* real interest rate is the same as the *ex ante* real interest rate when agents have perfect foresight and can predict the future price level with certainty so that $\pi = \pi^e$. Hence, the real interest rate that was used to discount the future value of income and consumption, when analysing consumption behaviour or to represent the opportunity cost of finance when analysing investment, is the *ex ante* real interest rate.

# 8.3 The *IS* Curve

With these preliminaries mentioned, we are in a position to analyse the equilibrium in the commodity market for a closed economy where output equals expenditure, or

$$Y = C + I + G$$

Looking at the determinants of consumption, we found that most approaches to consumption argued that consumption expenditure depends on income and the real interest rate

$$C = C(\underset{(+)}{Y}, \underset{(-)}{r})$$

Consumption increases with income and declines with a rise in the real interest rate, which makes future consumption more attractive relative to present consumption.[6]

We now work with a more appropriate definition of income for the consumption function, which is the disposable income, $Y^D$. In reality, consumption can occur only with respect to that income which is in the hands of the individual after taxes (net of transfers), $T$, have been paid to the government. Consumption is out of the disposable income, $Y^D$, where[7]

$$Y^D = Y - T$$

The appropriate consumption function[8] may be written as

$$C = C(\underset{(+)}{Y - T}, \underset{(-)}{i - \pi^e}). \tag{8.10}$$

For simplicity and without much loss of generality, we proceed with a linear version of the CONSUMPTION FUNCTION as

› Consumption is a positive function of the aggregate disposable income and is negatively related to the real interest rate.

$$C = c_0 + c_1(Y - T) - c_2(i - \pi^e) \tag{8.11}$$

where $c_0$ is the autonomous consumption that is independent of income, $c_1$ is the responsiveness of consumption to a change in disposable income, and $c_2$ is the responsiveness to a change in the *ex ante* real rate of interest.

› Investment is positively related to the productivity of capital, and negatively related to the real interest rate.

Similarly, INVESTMENT EXPENDITURE was found to be positively related to the productivity of capital and negatively related to the *ex ante* real rate of interest. In linear terms, we may write this as

$$I = a - b(i - \pi^e) \tag{8.12}$$

where $a$ is the shorthand for business confidence and the productivity of investment and $b$ is the parameter that summarizes by how much investment declines in response to an increment in the *ex ante* real interest rate.

› An exogenous variable is a variable that is taken as given and is not explained within an economic model.

We treat government expenditures to be EXOGENOUSLY determined. Whenever a factor that determines demand is outside the explanatory adopted framework, we treat it as an exogenous or autonomous change in demand. Government expenditures are determined in this fashion by some extraneous factors such as the desire to appease vote banks to increase the likelihood of being elected. Though government expenditures comprise both consumption and investment expenditure, we shall treat government expenditure as of the consumption type and ignore public investment.[9] We write the government expenditure function as being exogenously determined at some level $\bar{G}$, so that:[10]

$$G = \bar{G} \tag{8.13}$$

The condition for equilibrium in the product market for a closed economy is given by

$$Y = C + I + G$$

or

$$Y - C - G = I$$

Adding and subtracting taxes, $T$, on the left-hand side,

$$\underbrace{(Y - T) - C}_{\text{Private Savings}} + \underbrace{T - G}_{\text{Government Savings}} = I \tag{8.14}$$

The term $(Y - T)$ is disposable income, and disposable income less consumption expenditure constitutes private savings. Similarly, the subtraction of government expenditure from tax receipts constitutes government savings. EQUILIBRIUM IN THE PRODUCT MARKET thus requires that savings—private plus government savings—equals investment. Substituting for $C$ and $I$ from Eqs. (8.11) and (8.12)

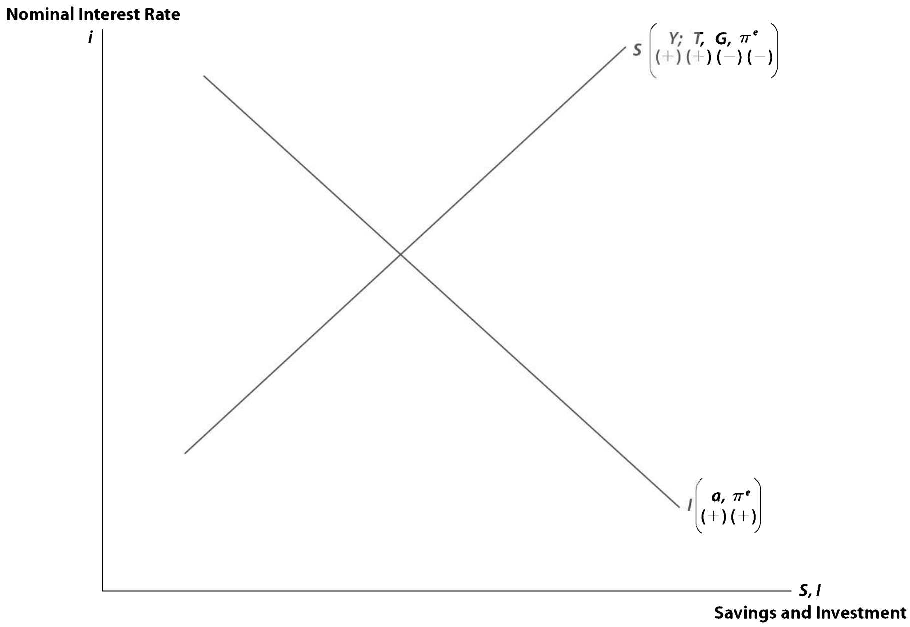
› Equilibrium in the product market is given by the condition that aggregate savings—private and public—equals investment.

$$(Y - T) - c_0 - c_1(Y - T) + c_2(i - \pi^e) + (T - G) = a - b(i - \pi^e)$$

or

$$(1 - c_1)(Y - T) - c_0 + c_2(i - \pi^e) + (T - G) = a - b(i - \pi^e) \tag{8.15}$$

Aggregate savings in the economy, the left-hand side, increases with the nominal interest rate (the coefficient $c_2$ appears with a positive sign). The aggregate investment, the right-hand side, decreases with the nominal interest rate (the coefficient $b$ appears with a negative sign). The graphs of savings, $S$, and




**› Figure 8.1**
Aggregate Investment and Savings. **An increase in the interest rate increases aggregate savings because the relative return to future consumption has risen; and reduces aggregate investment because of the decline in investment opportunities.**

investment, $I$, as a function of the nominal interest rate and for given values of $\{Y; T, G, \pi^e\}$ are depicted in Figure 8.1. In Figure 8.1, the nominal interest rate is measured on the vertical axis, and savings and investment are measured on the horizontal axis. The savings schedule is upward sloping as an increase in the nominal interest rate is associated with an increase in savings. The investment schedule is downward sloping as an increase in the interest rate reduces investment expenditure.

Note that aggregate income $Y$'s value is what we are trying to determine whereas $T$, $G$, $\pi^e$ are exogenous variables. An increase in $Y$ increases disposable income and increases the savings out of disposable income, which shifts the savings curve to the right in Figure 8.1. This is the motivation for the POSITIVE SIGN UNDER THE VARIABLE $Y$ in the savings function.

› An increase in income increases disposable income and aggregate savings.

An increase in taxes ($\Delta T$) reduces the disposable income of the private sector by $\Delta T$ and increases the tax receipts of the government by $\Delta T$. The private sector would have spent $c_1\Delta T$ on consumption and saved $(1 - c_1)\Delta T$ of this income that is now transferred to the government.

For the government, on the other hand, the entire amount $\Delta T$ is an addition to receipts and that for a fixed government expenditure results in an increase in government savings by $\Delta T$. The NET ADDITION TO AGGREGATE SAVINGS is thus $\Delta T - (1 - c_1)\Delta T$, or $c_1\,\Delta T$, justifying the positive sign under the variable $T$ in Figure 8.1.

› An increase in taxes reduces private savings but increases government savings, with the net effect being a rise in the aggregate savings.

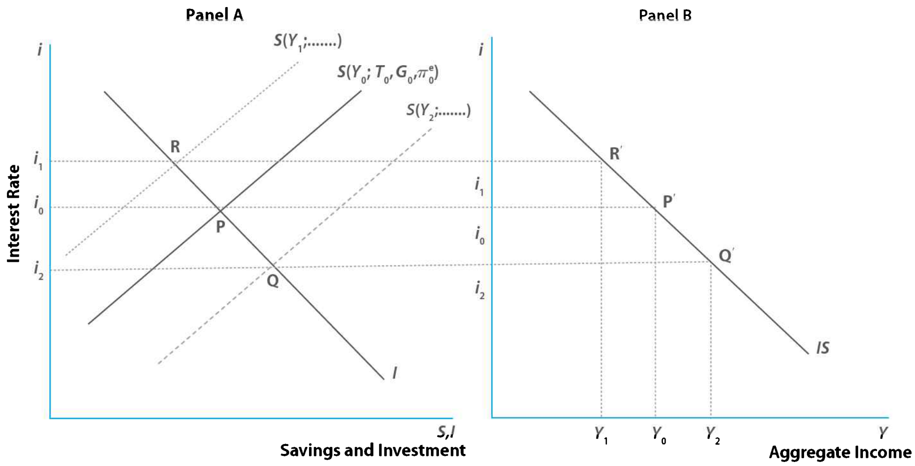
Similarly, an INCREASE IN GOVERNMENT EXPENDITURE corresponds to a reduction in government saving $(T - G)$, and the $S$ schedule shifts to the left when this occurs. an INCREASE IN EXPECTED INFLATION reduces the real interest rate, which makes savings for the purposes of future consumption less attractive and reduces savings, thereby shifting the savings schedule to the left (see Figure 8.1).

› An increase in the government expenditure reduces public savings and aggregate savings.

› An increase in the expected rate of inflation reduces the real interest rate and makes savings for purposes of future consumption less attractive.

The investment schedule in Figure 8.1 shifts to the right with an increase in business confidence or productivity (increase in $a$) that improves investment opportunities. Also, an increase in expected inflation reduces the real interest rate and rental on capital, thereby making it profitable to increase investment, which shifts the investment schedule to the right.

We are now set up to derive the *IS* curve. This is the curve that plots combinations of the interest rate and income along which equilibrium prevails




**› Figure 8.2**
The *IS* Curve. The *IS* curve depicts combinations of the interest rate and aggregate income that are associated with aggregate investment (*I*) equal to aggregate savings (*S*).

in the product market. Equilibrium in the product market is where income equals expenditure, or aggregate savings—private plus government—equals investment as written in Eq. (8.14). Since investment ($I$) equals saving ($S$) at every point on the curve, it is termed the *IS* curve. The *IS* curve is easily derived in Figure 8.2. Suppose the savings schedule in ($i$, $S$) space is drawn for a particular income level $Y_0$ and this savings equals investment at point P in Figure 8.2, panel A. The nominal interest rate, when savings equals investment at P, is thus associated with an interest rate–income combination ($i_0$, $Y_0$). This combination is plotted in ($i$, $Y$) space in panel B as point P′. P′ is an INTEREST RATE–INCOME COMBINATION where savings equals investment.

› The *IS* curve is downward sloping in the nominal interest rate–aggregate income space as a rise in income raises the supply of savings to the capital market and causes the interest rate to decline.

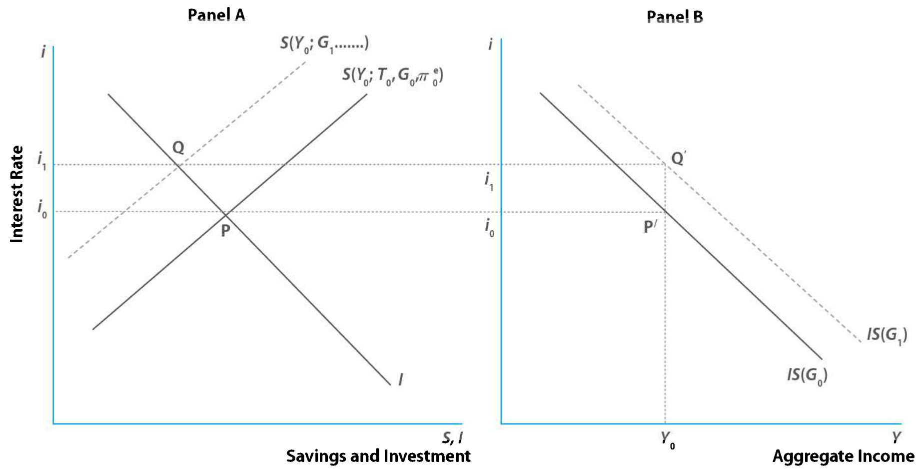
Suppose income increases to ($Y_2 > Y_0$). Then, savings increases with the increase in income, and the savings schedule shifts down to intersect the investment schedule at point Q where the interest rate is $i_2$. The increased saving at higher incomes increases the supply of savings to the capital market, and the increased supply of funds pushes down the interest rate to $i_2$. Again, the interest rate–income pair ($i_2$,$Y_2$), when savings equals investment at point Q, is depicted as point Q′ in panel B.

Analogously, if income declines from $Y_0$ to $Y_1$—the change in income is ($Y_0$ − $Y_1$)—there is an accompanying decline in savings that shifts the savings schedule up. Savings is now equal to investment at point R, and the accompanying interest rate–income pair ($i_1$, $Y_1$) depicted in panel B as point R′.

By joining all the points such as P′, Q′, R′, which represent different combinations of the nominal interest rate and income for which savings equals investment, we obtain the *IS* curve for the economy along which the goods market is in equilibrium. The *IS* curve is downward sloping because an increase in income increases savings, the increased supply of which reduces interest rates. A reduction in interest rates makes investment projects more profitable and investment rises to match the increased savingsx.

### *8.3.1* Parametric Shifts of the *IS* Curve

We now look at how parametric changes affect the *IS* curve. First, we analyse the impact of an increase in government expenditure. An increase in government expenditure to $G_1 > G_0$ corresponds to a decline in government savings ($T - G$). The decline in savings shifts the savings schedule upwards, and savings now equals investment at point Q in Figure 8.3 where the interest rate has risen to $i_1$ from $i_0$.




**› Figure 8.3**
Increased Government Expenditure. **An increase in government expenditure reduces the savings of government (panel A) and shifts the *IS* curve to the right (panel B).**

The INCREASED GOVERNMENT EXPENDITURE in the absence of any change in taxes is financed by a sale of bonds by the government to the private sector. The increased supply of bonds pushes down the price of bonds which, in turn, raises the interest rate (see the price formula for bonds given by Eq. (6.4) in Chapter 6 on the demand for money). Essentially the government competing with the private sector for the funds of savers drives up the interest rate. The interest rate at the original income level $Y_0$, and indeed at all income levels, is now higher and the *IS* curve shifts out to the right.

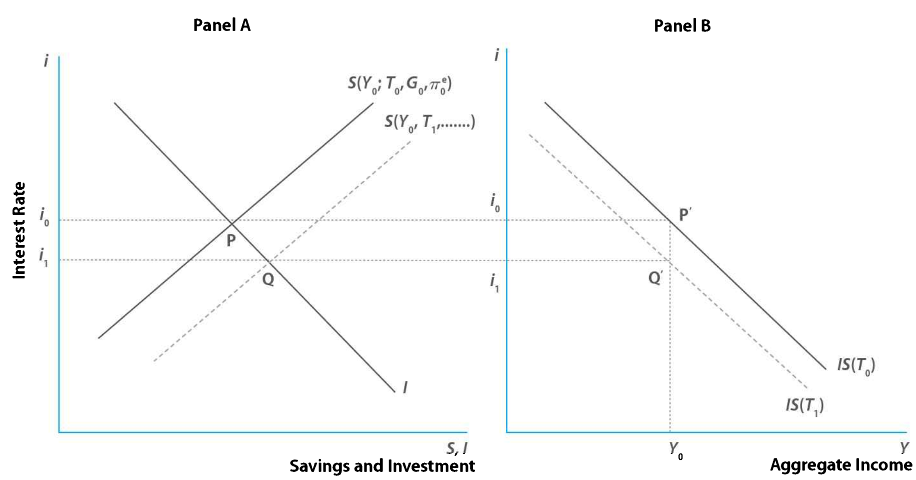
› The *IS* curve shifts to the right with an increase in the government expenditure as the increased supply of bonds causes bond prices to decline and the interest rate to rise.

What if TAXES INCREASE? Then, as we saw earlier, the private sector would have saved $(1 - c_1)$ of the increase in taxes equal to $(1 - c_1)\Delta T$, whereas the government's receipts, which accrues as savings, increase by $\Delta T$. Hence, savings increase by $\Delta T - (1 - c_1)\Delta T = c_1\Delta T$. The savings schedule shifts to the right by the amount $c_1\Delta T$ in Figure 8.4, where $\Delta T = T_1 - T_0$ and the increased supply of savings pushes down the nominal interest rate. In fact, it is now lower at the original and indeed at all levels of income, causing the *IS* curve to shift to the left.

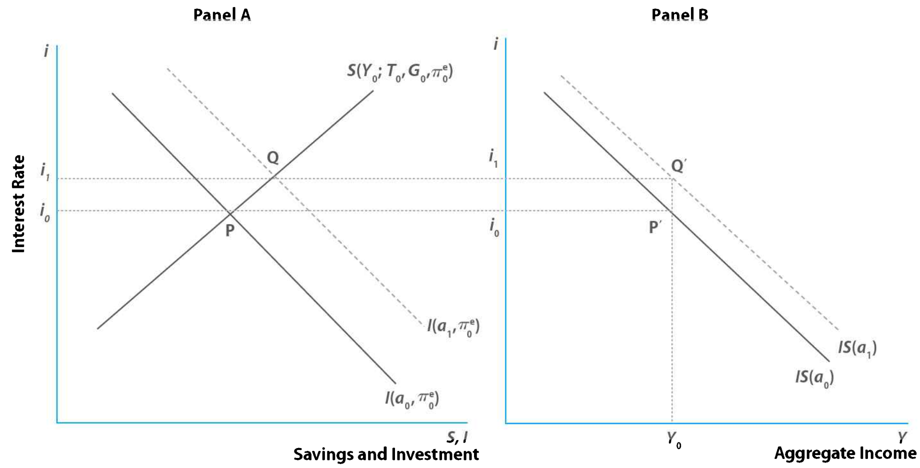
› An increase in taxes increases the aggregate savings and the nominal interest rate declines, thus, shifting the *IS* curve leftwards.

If firms come up with a new technology that INCREASES PRODUCTIVITY, they will undertake investments to exploit this technology. Expectations about the future profitability of a technology shifts the investment schedule to the right—an increase in the parameter $a$ in Eqs. (8.12) and (8.15). Again, the increased demand for the funds supplied by savers bids the interest rate up and the *IS* curve shifts to the right (see Figure 8.5). We could also interpret the parameter $a$ as the expectations of future profitability due to the mass psychology around

› An increase in the productivity of capital increases the demand for investment and the interest rate, thereby shifting the *IS* curve out to the right.




**› Figure 8.4**
Increase in Taxes. **An increase in taxes increases aggregate savings (panel A) and shifts the *IS* curve to the left (panel B).**




**› Figure 8.5**
Increase in Productivity. An increase in productivity increases investment (panel A) and shifts the *IS* curve to the right (panel B).

business conditions that characterizes entrepreneurs and business persons—what Keynes referred to as *animal spirits*.

› An increase in the expected inflation leaves the aggregate savings and investment unchanged but shifts the *IS* curve out to the right as the nominal interest rate rises.

Finally, if there is an increase in the EXPECTED INFLATION, say, because of the current nominal rate seeming low in terms of the surge in production and accompanying constraints on capacity that will occur in the future, the *IS* curve shifts to the right (see Figure 8.6).

As savings and investment are conducted in real terms, the reference interest for decisions regarding these variables is the real interest rate. With no change in the real interest rate, there is no change in savings and investment, which stays at $S_0 = I_0$. However, the nominal interest rate rises by the amount of the increase in expected inflation.

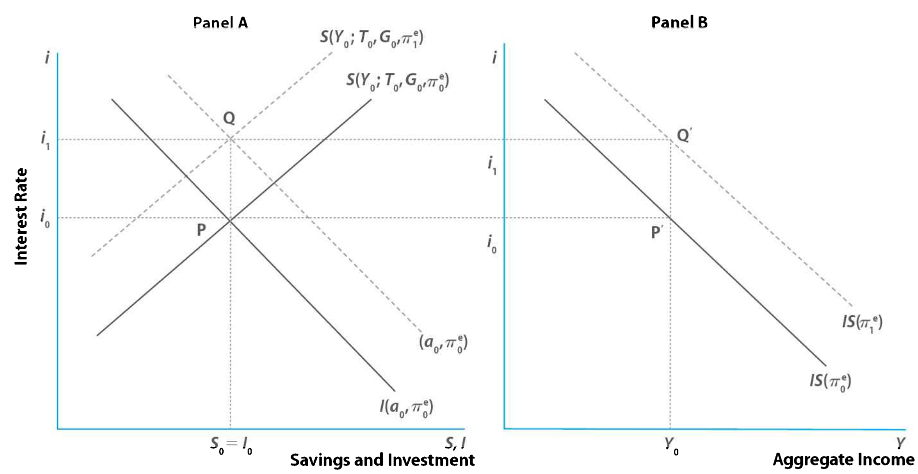
### 8.3.2 Slope of the *IS* Curve

What can we say about the slope of the *IS* schedule? Take the savings equals investment Eq. (8.15), and find the expression for the nominal rate of interest by first taking the terms in $i$ on the left-hand side

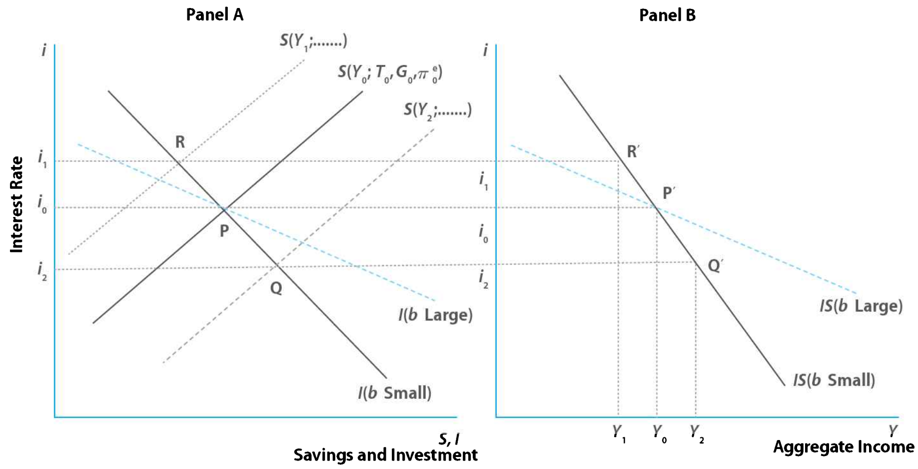
$$(b + c_2)\, i = a + c_0 - (1 - c_1)(Y - T) - (T - G) + (b + c_2)\pi^e$$

Hence,

$$i = \frac{a + c_0 - (1 - c_1)(Y - T) - (T - G)}{b + c_2} + \pi^e \tag{8.16}$$




**› Figure 8.6**
Increase in Expected Inflation. An increase in expected inflation raises the nominal interest rate but does not affect the real rate of interest. Savings and investment in real terms are unaffected (panel A) and the *IS* curve shifts out to the right (panel B).




**› Figure 8.7**
*IS* Curve when Investment Sensitivity to Interest Rate Is Low and High. **If investment is not very sensitive to interest rates (*b* is small), the demand for investment curve is inelastic (steep)—panel A. This results in a steep *IS* curve—panel B. The *IS* curve is flat when investment s sensitive to interest rate changes (*b* is large).**

So, if investment and consumption are insensitive to interest rates ($b + c_2$ small), the *IS* CURVE, algebraically written as Eq. (8.16), is steep. In Figure 8.7, if the interest sensitivity of investment ($b$) is assumed to be very small the investment schedule is close to vertical.[11]

› The steeper the *IS* curve, the lower is the responsiveness of the aggregate income to the interest rate.

If income increases to $Y_2$, savings equals investment at ($i_2$, $Y_2$) and if income declines to $Y_1$, savings equals investment at ($i_1$, $Y_1$). The resulting *IS* curve is fairly steep, reflecting the low-interest sensitivity of investment. If, on the other hand, investment is fairly sensitive to the interest rate and the parameter $b$ is large, the investment schedule—the dashed downward-sloping line in Figure 8.7—is relatively flat and the corresponding *IS* schedule is also relatively flat. The *IS* curve is a description of equilibrium in the goods market. We now describe the *LM* curve or the condition for equilibrium in the money market.

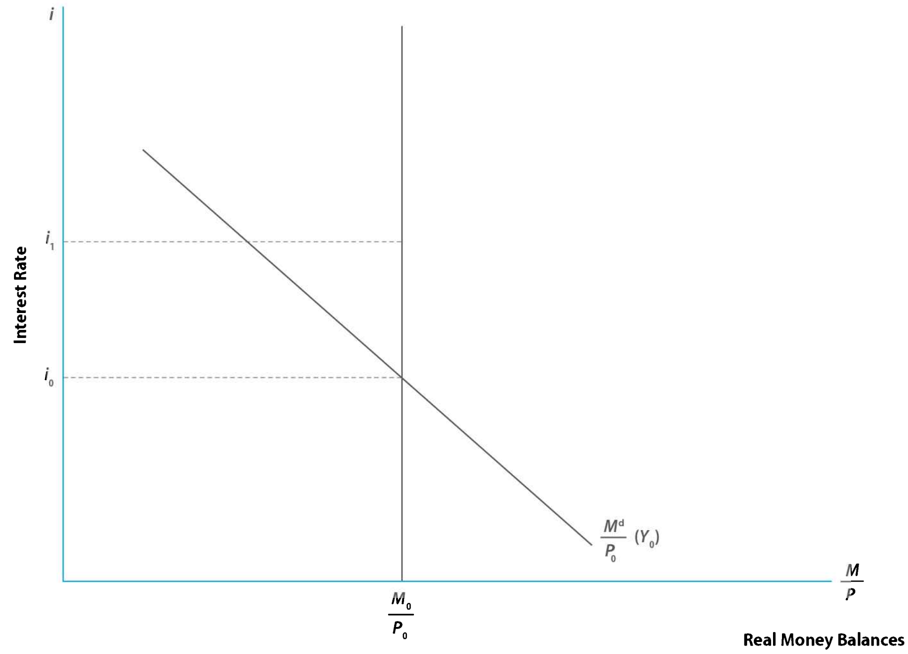
# 8.4 The *LM* Curve

The *LM* curve depicts combinations of the nominal interest rate and income that represent equilibria between the supply of and demand for money. In deriving the *LM* curve it is assumed that the entire money supply is an exogenous variable directly under the control of the central bank. Realistically, the central bank's ability to control the money supply in circulation is imperfect. This is because changes in interest rates induced by money supply changes cause banks to respond by altering the amount of deposits they issue.[12] Nevertheless, it is worth simplifying by assuming that the central bank can supply any quantity of nominal money balances it desires in the form of non-interest bearing cash and deposits. The quantity of supplied money that is independent of the nominal interest rate is then given by

$$M^s = \overline{M} \tag{8.17}$$

The real purchasing power of money is what matters to individuals and, therefore, a central bank would find it desirable to have direct control over the quantity of real money balances. However, central banks are not in a position to set prices that are determined in the markets for goods and services.

At best, central banks can conduct policies that influence the price level. But price levels can change without any policy action being initiated by a central bank. Hence, the supply schedule for real money balances ($M^s/P$) is affected by both the nominal money stock and the price level.




**› Figure 8.8**
Equilibrium Real Money Balances. **The demand for real money balances ($M^d/P_0$) is inversely related to the nominal rate of interest. The supply of money ($M_0/P_0$) is independent of the interest rate and is a vertical curve. In equilibrium demand equals the supply of money and the nominal interest rate is determined.**

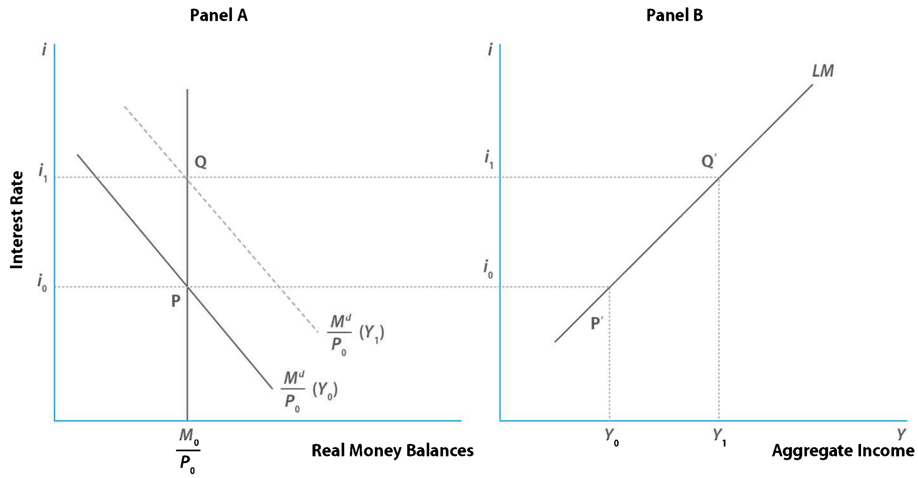
The demand for real money balances has already been derived in Chapter 6 (on the demand for money) as $M^d/P = f(Y, i)$. In a linear form, we write this as

$$\frac{M^d}{P} = h + nY - ei \qquad \textbf{(8.18)}$$

The demand for real money balances is, therefore, inversely related to the nominal interest rate. This relation is because of a rise in the interest rate, which represents income forgone from holding alternative financial assets like bonds that induce people to give up the liquidity associated with money. Alternatively, a rise in interest rates, which is a decline in the price of bonds for given expectations, makes individuals anticipate increased capital gains on bond holdings and allocate more of their portfolio to bonds.

The demand and supply schedules for real money balances is depicted in Figure 8.8 for a given price level, $P_0$. The two schedules intersect at a point where all individuals in the economy are satisfied holding the nominal money stock supplied by the central bank, $M_0$, deflated by the current price level, $P_0$. The nominal interest rate adjusts to achieve equilibrium where real money demand equals real money supply. To understand how equilibrium is achieved, consider what transpires if the interest rate happened to be $i_1$. At the higher interest rate $i_1$, the quantity of real money supplied exceeds the demand for real money balances. People wishing to hold smaller real money balances than that supplied by the central bank will attempt to get out of money and hold more bonds in its place. They already hold the money balances required for transacting in goods and services at $Y_0$ and so instead of spending in the goods market they acquire bonds. The additional demand for bonds cause bond prices to rise, and as there is an inverse relation between the price of bonds and the nominal interest rate, the interest rate falls towards the equilibrium rate, $i_0$.

We can now derive the *LM* curve by asking about the relation between the nominal interest rate and income when the demand for real money balances equals the supply of real money balances. If income increases to $Y_1$, then, there is an increase in the demand for money required for transactions purposes,[13] and the demand for money schedule shifts to the right (see Figure 8.9, panel A).




**› Figure 8.9**
The *LM* Curve. An increase in income shifts the demand for money curve to the right and raises the interest rate in the money market (panel A). An increase in income results in a new money-demand-and-supply equilibrium that is associated with an increase in the interest rate. Tracing out such points where money demand equals money supply at various levels of income gives the *LM* curve (panel B).

How is the demand for money going to be met with a fixed money supply? Essentially, individuals rebalance their portfolio of financial assets by selling bonds in order to acquire the money balance required to finance the higher transactions associated with income $Y_1$. The increase in the supply of bonds, in turn, causes the price of bonds to fall and the nominal interest rate, which is inversely related to the price, to rise till $i_1$, where demand and supply for real money balances are again equal. We plot the combinations of nominal interest rate and income such as $(i_0, Y_0)$ and $(i_1, Y_1)$, where money demand equals money supply in panel B of Figure 8.9. The graph of these combinations of real income and nominal interest rate, where money demand equals money supply, is called THE *LM* CURVE. The name *LM* follows a notation introduced by John Hicks[14] in 1937, where he referred to the demand for money as desired liquidity L. Setting desired liquidity equal to the supply of real money balances M, L = M, results in a schedule where the money market is in equilibrium, called the *LM* schedule. The *LM* curve is upward sloping because an increase in real income increases the demand for real money balances, which is satisfied by selling bonds that causes a decline in bond prices and a corresponding rise in the nominal interest rate.

› The *LM* curve is upward sloping in the nominal interest rate—aggregate income space since an increase in income raises the demand for real money balances, which results in a rise in the interest rate.

### *8.4.1* Impact of Varying the Money Supply

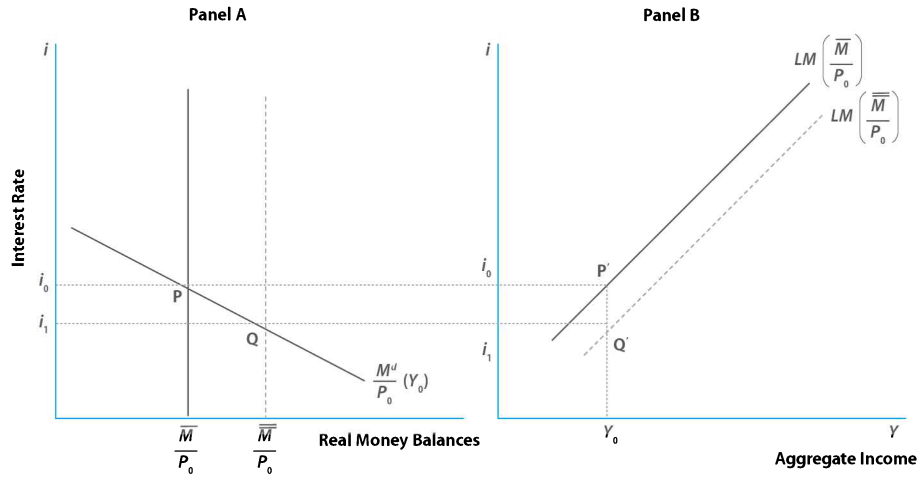
What if the central bank manipulates the money supply? Manipulation of the money supply with a view to influence endogenous variables, such as the interest rate or the level of income, is referred to as *monetary policy*. The increase in nominal money supply from $\overline{M}$ to $\overline{\overline{M}}$ shifts the money supply curve to the right as depicted in Figure 8.10, panel A.

There is now an EXCESS SUPPLY OF MONEY as $\overline{\overline{M}} / P_0$ real money balances are sufficient to finance transactions associated with the real income $Y_0$. Individuals then deplete their excess money balance by purchasing bonds, which drive up bond prices and reduce the nominal interest rate to $i_1$ where desired money demand again equals money supply. The *LM* curve shifts to the right as the nominal interest rate associated with any level of real income is lower, as depicted in Figure 8.10, panel B.

› The *LM* curve shifts to the right if there is an increase in the money supply. This is because interest rates decline in response to individuals' rebalancing their financial portfolio by purchasing bonds when the money supply increases.

### *8.4.2* Slope of the *LM* Curve

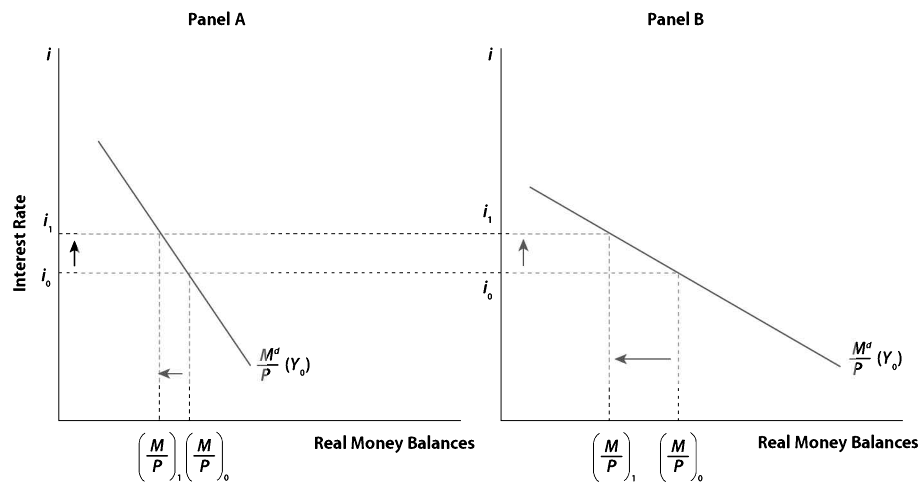
What can be said about the slope of the *LM* schedule? In Figure 8.11, the interest sensitivity of real money demand is low for the curve in panel A—an *interest inelastic* money demand schedule.



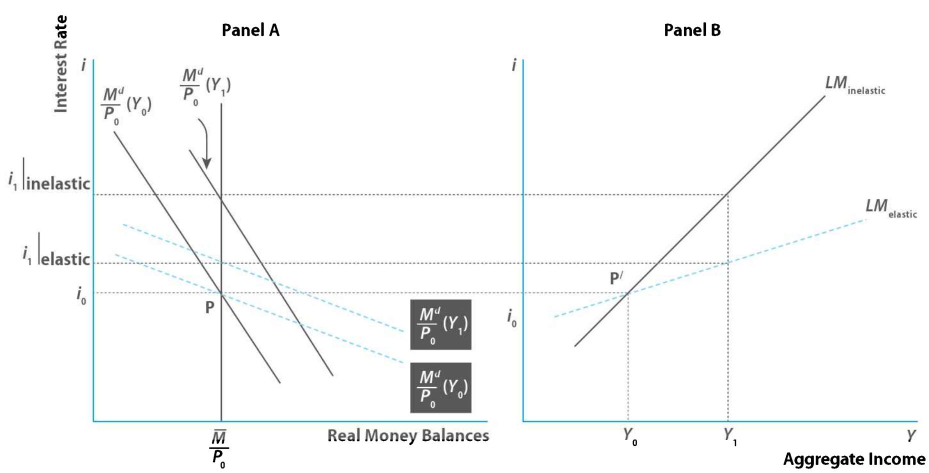

**› Figure 8.10**
The Effect of Increase in Money Supply. An increase in money supply shifts the money supply curve to the right—the dashed line in panel A—and results in a decline in the interest rate in the money market. The decrease in the interest rate associated with a higher money supply shifts the *LM* curve to the right (panel B).

A small rise in the nominal interest rate from $i_0$ to $i_1$ causes a relatively small reduction in the quantity of real money balances demanded. By contrast, the interest sensitivity of real money demand is high for the curve in panel B of Figure 8.11—an *interest elastic* money demand schedule. A small rise in the nominal interest rate, from $i_0$ to $i_1$, induces a relatively large decline in the quantity of real money balances demanded from $(M/P)_0$ to $(M/P)_1$.

Suppose we begin at the point where money demand associated with real income $Y_0$ equals money supply and this results in a nominal interest rate, $i_0$ (see Figure 8.12, panel A). Suppose the relevant money demand curve is the interest




**› Figure 8.11**
The Interest Elasticity of Money Demand. When the demand for money is not sensitive to the interest rate (interest inelastic) the curve is steep (panel A). When the demand for money is sensitive to interest rate changes (interest elastic) the curve is flat (panel B).




**› Figure 8.12**
The Interest Elasticity of the *LM* Schedule. When the demand for money is sensitive to changes in the interest rate (an Interest-elastic dashed curve) an increase in the demand for money for transactions that accompanies an increase in income does not result in a large change in the interest rate (panel A). The resulting *LM* curve (dashed line) is also relatively flat—panel B. If the demand for money is interest inelastic—solid curve—the *LM* curve is steep.

inelastic solid curve and there is an increase in income to $Y_1 > Y_0$. The money demand curve then shifts out resulting in interest rate $i_1|_{\text{inelastic}}$ associated with $Y_1$.

The *LM* CURVE corresponding to this money demand curve is also relatively interest inelastic, as seen in panel B. If the relevant money demand curve is the dashed interest elastic one in panel A, an increase in real income to $Y_1$ results in a smaller rise in the nominal interest rate to $i_1|_{\text{elastic}}$ and the corresponding *LM* schedule is relatively more elastic. Hence, a steep or interest inelastic *LM* curve is associated with a money demand curve that is interest inelastic. Similarly, a flat or interest elastic *LM* curve is associated with a money demand schedule that is interest elastic.

› The *LM* curve is inelastic (elastic) when the demand for money is interest inelastic (elastic).

## 8.5 *IS* and *LM*—Fiscal and Monetary Policy

Combining the *IS* and the *LM* schedules in a diagram enables us to find a single combination of real income and the nominal interest rate that is simultaneously equilibrium in the commodities market and in the market for real money balances.

**MACROFOCUS 8.1**

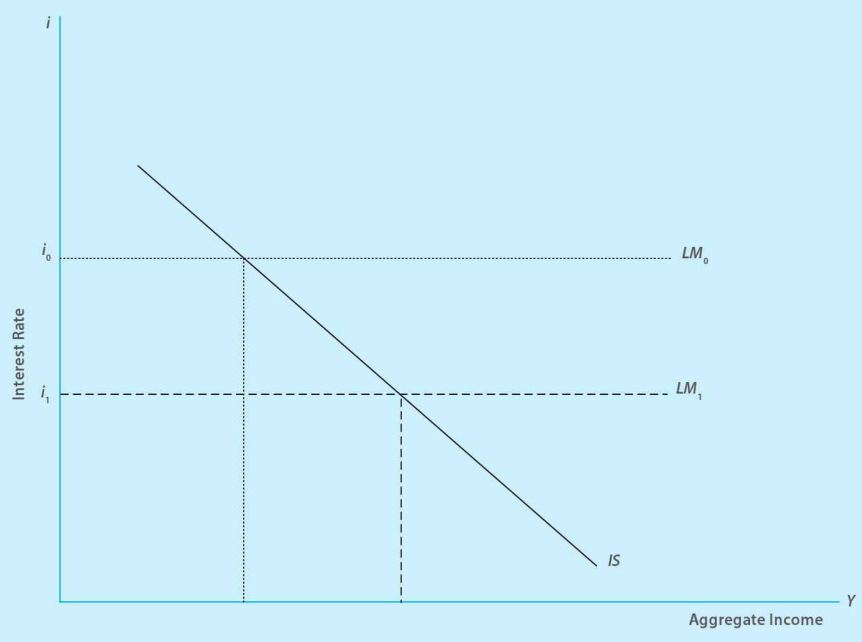
### Interest Rate Targets and Monetary Policy

Monetary policy involves the variation in money supply by monetary authorities for achieving a given level of income. However, in many economies, setting interest rates is much more common among central banks than fixing the money supply. In the United States, for instance, the Federal Reserve strives to achieve a target for the Federal Funds interest rate through open-market operations—the purchase and sale of government securities. In this case the *LM* curve is flat at the targeted level of the interest rate. The difference arises because the supply of money is fixed for an upward-sloping *LM* curve. As a result, when incomes, and the demand for money for transactions purposes along with it, rise (or fall), the interest rate rises (or falls) to keep the demand for money at the constant level of the money supply. A rise (or fall) in income is thus associated with a rise (or fall) in the interest rate, and the *LM* curve is upward sloping.

In the case of an interest rate target (Figure 1), however, the money supply fluctuates to maintain a given level of the interest rate and there is no relationship between *Y* and *i* —the *LM* curve is flat. In this case the monetary authorities change the price of holding money balances in a way that induces people to want to hold that amount of money. When the monetary authorities use an interest rate target, an expansionary monetary policy reduces the target level of the interest rate from $i_0$ to $i_1$ and the impact on the economy is the same as for the case of increasing the money supply in an upward-sloping *LM* case. Output rises (and the interest rate declines) as the economy moves down its *IS* curve. In a sense the distinction between using money supply targets and interest rate targets is of significance only when the demand for money schedule is not stable, and so, the corresponding position of the *LM* curve is also variable (see Chapter 15).

**› Figure 1**

Monetary Policy with an Interest Rate Target. When central banks directly set interest rates rather than control the money supply the *LM* curve is flat since the money supply is varied by the central bank to maintain the interest rate target it has set.



**› Figure 8.13**
*IS–LM* Equilibrium. **The *IS* and *LM* curves intersect at a point E and result in a nominal interest rate and real income outcome where the goods market (*IS*) and the money market (*LM*) simultaneously are in equilibrium.**

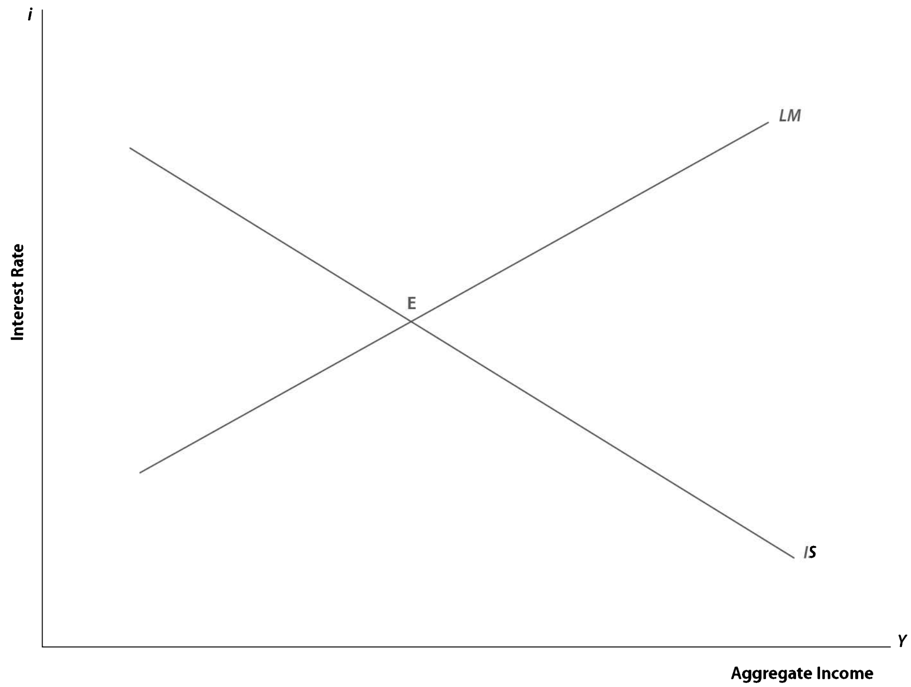



Figure 8.13 combines the *IS* and *LM* schedules in a single diagram. The two schedules intersect at point E, which is an *IS–LM* equilibrium, where real income and the nominal interest rate achieve goods market and money market equilibrium simultaneously. At any point other than E in the diagram, three conditions can occur:

(i) The goods market is not in equilibrium (any point other than E on the *LM* curve),
(ii) The money market is not in equilibrium (any point other than E on the *IS* curve),
(iii) Both the markets are not in equilibrium (points off the *IS* and *LM* curve).

In that case interest rates and expenditure must change to re-attain the real income–nominal interest rate combination that concurrently achieves equilibrium in the goods and money markets.

We now analyse the effect of *monetary policy* in an *IS–LM* set-up. Monetary policy results in a change in the nominal money stock and a corresponding change in the real money supply (when the price level is unchanged). The *IS–LM* MODEL allows us to understand the immediate impact of monetary and fiscal policy actions under the following assumptions:

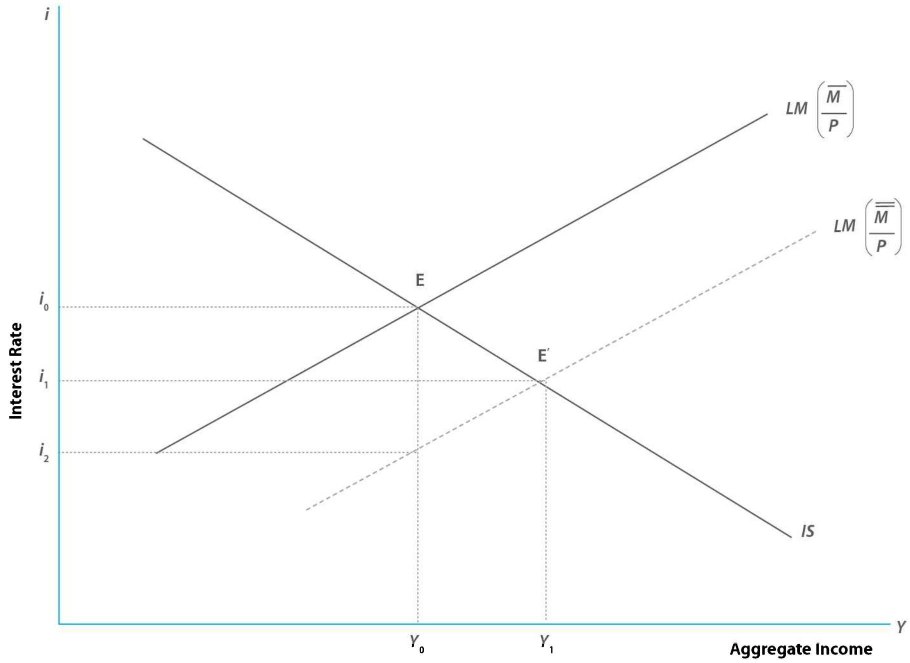
› The *IS–LM* model is appropriate for the short run where wages and prices do not change, and there is underutilization of the capacity.

1. wages and prices are fixed
2. resources such as labour and capital are not fully utilized.

In such a set-up, we can ignore the labour market and the supply of output, as firms are thought to not have supply constraints on the production of goods, but rather the constraint on output is the lack of demand. Any additional demand results in firm's increasing output, and a reduction in demand leads them to reduce output.

› The more interest inelastic the private expenditures, and the more interest elastic the demand for money, the more effective is the expansionary monetary policy in the *IS–LM* model.

Suppose now the central bank increases the nominal money stock. As we know from Figure 8.10, this shifts the *LM* curve to the right. This is depicted in Figure 8.14 as the dashed *LM* curve. The immediate impact of the INCREASE IN MONEY SUPPLY is that at the initial real income $Y_0$, the increased money balances are not required for transactions purposes, and so, individuals give up this liquidity in favour of interest earnings on bonds.




**› Figure 8.14**
The Increase in Nominal Money Stock. **An increase in money supply induces individuals to shift their portfolio towards bonds and the interest rate declines to $i_2$. This stimulates investment and consumption expenditure in the goods market, which raises the demand for money and results in the interest rate ultimately settling at $i_1$. An increase in the money supply, therefore, reduces the interest rate and increases aggregate income.**

The increase in the demand for bonds increases the price of bonds, or what amounts to the same thing, the interest rate declines to $i_2$. The decline in the interest rate, in turn, induces two effects. One effect is that individuals prefer to consume now rather than later, and consumption expenditure increases. This is because of the present savings' lower return owing to a decline in interest rates.

Second, the reduction in the interest rate makes more projects, which earlier were unprofitable at the higher interest rate, profitable and investment expenditure increases. This increased expenditure on consumption and investment is reflected in a movement along the *IS* curve from E towards E′ and causes a rise in real income. With increased expenditure and incomes, the demand for real money balances rises for facilitating transactions, and this causes the interest rate to rise to $i_1$with goods and money market equilibrium at E′.

To sum up, the rise in the nominal stock of money causes a reduction in the equilibrium nominal interest rate and an increase in equilibrium real income. The extent of the increase in real income depends on the interest elasticity of the *LM* schedule, as we discuss in Section 8.5.1.

### *8.5.1* Crowding Out

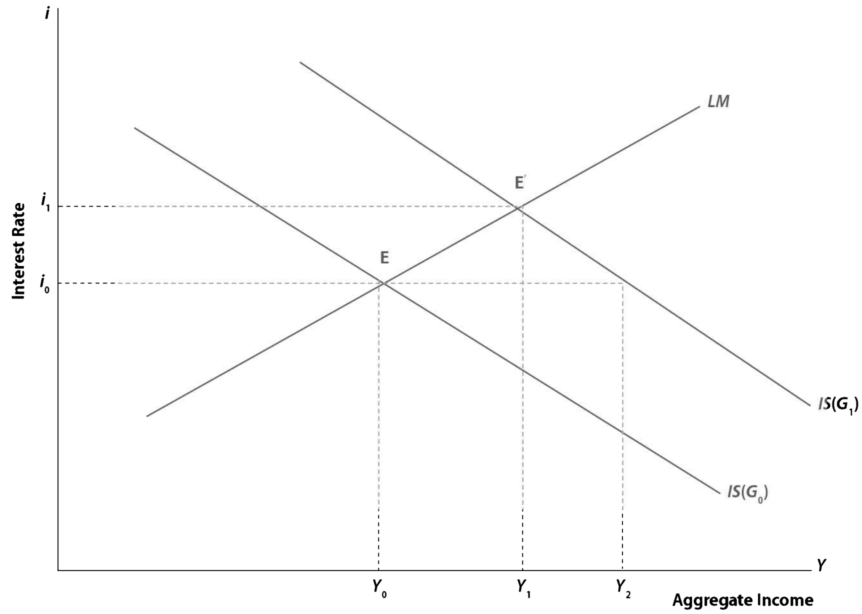
What if the government decides to use fiscal policy by, for instance, increasing government expenditure? As we know from Figure 8.3, the *IS* curve shifts outwards to $IS(G_1)$ as depicted in Figure 8.15.

The increased government expenditure raises real income initially to $Y_2$. The increased income causes a rise in the demand for money for transaction purposes. In the face of a fixed supply of money, the increased money demand (see Figure 8.9) causes the interest rate to rise. In symbolic form, the increase in the interest rate in what follows will be written as $\Delta i$. The rise in the interest rate induces a DECLINE IN INVESTMENT EXPENDITURE ($\Delta I|_{\Delta i}$) [see Eq. (8.12)] due to a decline in the present value of investment projects, and a decline in the interest-sensitive component of consumption expenditure ($\Delta C|_{\Delta i}$) [see Eq. (8.11)] as the attractiveness of future consumption increases.

› The change in a variable $V$ that occurs in response to a change in the interest rate is written as $\Delta V|_{\Delta i}$. In usual calculus notation this would be written as $\Delta V = (\partial V/\partial i)\,\Delta i$.

The decline in the interest-sensitive elements of private expenditures reduces income from $Y_2$ to $Y_1$ in Figure 8.15. In the process, despite the rise in interest rates, which choke off some investment and consumption, the final equilibrium

**› Figure 8.15**
The Increase in Government Expenditure. **An increase in government expenditure shifts the *IS* curve to the right. It increases the income and the demand for money for transactions purposes. The increased demand for money results in an increase in the interest rate in the money market. The rise in the interest rate crowds out—reduces—some interest-sensitive private expenditures on investment and consumption.**




level of income is higher after the increased government spending, and this rise in income induces an increase in the income-sensitive component of consumption expenditure ($\Delta C|_{\Delta Y}$) [see Eq. (8.11)]. The constrained increase in income—$Y_0Y_2$ in Figure 8.15—can then be thought to be the net outcome of:

1. An expansion of income arising from the increase in government spending —$\Delta G$
2. A decline of income arising from the induced decline in investment and that part of consumption which is sensitive to interest rates—$\Delta I|_{\Delta i}$, $\Delta C|_{\Delta i}$
3. An expansion in income arising from an induced increase in consumption expenditure as income rises —$\Delta C|_{\Delta Y}$

Hence, $$\Delta Y = \Delta G + \left(\Delta I|_{\Delta i} + \Delta C|_{\Delta i}\right) + \Delta C|_{\Delta Y} \qquad \textbf{(8.19)}$$

From Eq. (8.11), $$\Delta C|_{\Delta Y} = c_1\,\Delta Y$$

Substituting this expression into Eq. (8.19),

$$\Delta Y = \Delta G + \left(\Delta I|_{\Delta i} + \Delta C|_{\Delta i}\right) + c_1\Delta Y$$

Taking the terms in $\Delta Y$ on the left-hand side, we can write

$$\Delta Y = \frac{\Delta G}{1 - c_1} + \frac{\Delta I|_{\Delta i} + \Delta C|_{\Delta i}}{1 - c_1} \qquad \textbf{(8.20)}$$

The first term on the right-hand side of Eq. (8.20)—$\Delta G/(1 - c_1)$—is the unconstrained increase in income due to a rise in government expenditure—$Y_0Y_2$ in Figure 8.15. The second term on the right-hand side of Eq. (8.20) is the induced fall in investment and consumption expenditure when the interest rate increases—$Y_1Y_2$ in Figure 8.15.

The rise in interest rates discourages investment spending as investment is inversely related to the interest rate, and discourages consumption spending as well because future consumption (or savings) becomes more attractive with a rise in the interest rate. The decline in investment and consumption is given from Eqs. (8.11) and (8.12) by

$$\Delta I|_{\Delta i} = -\,b\Delta i$$
$$\Delta C|_{\Delta i} = -\,c_2\Delta i$$

The induced decline in investment and consumption is, therefore,

$$\Delta C\big|_{\Delta i} + \Delta I\big|_{\Delta i} = -\,(c_2 + b)\,\Delta i \tag{8.21}$$

What is the extent of the rise in interest rate? The initial increase in income at $Y_2$ induced an increase in the real demand for money for transactions purposes. With unchanged money supply this demand is met through the sale of bonds, which reduces the price of bonds and raises the interest rate. With money supply given as $M^s = \overline{M}$ and money demand as in Eq. (8.19), we have,

$$\frac{M^d}{P} = \frac{\overline{M}}{P} = h + nY - ei$$

or,

$$i = \frac{h + nY - (\overline{M}/P)}{e} \tag{8.22}$$

Hence, the rise in the rate of interest when income rises by $\Delta Y$ is

$$\Delta i = \frac{n}{e}\,\Delta Y \tag{8.23}$$

This rise in the interest rate is represented by the distance $i_0 i_1$ in Figure 8.15. Substituting Eq. (8.23) into Eq. (8.21) gives us the decline in consumption and investment that results from a rise in the interest rate as

$$\Delta C\big|_{\Delta i} + \Delta I\big|_{\Delta i} = -\,(c_2 + b)\,\frac{n}{e}\,\Delta Y \tag{8.24}$$

Substituting the expression Eq. (8.24) into Eq. (8.20), we obtain

$$\Delta Y = \frac{\Delta G}{1 - c_1} - \frac{(c_2 + b)\,(n/e)}{1 - c_1}\,\Delta Y$$

Multiplying throughout by $(1 - c_1)$

$$(1 - c_1)\Delta Y = \Delta G - (c_2 + b)\,\frac{n}{e}\,\Delta Y$$

or,

$$\Delta Y\left[(1 - c_1) + (c_2 + b)\,\frac{n}{e}\right] = \Delta G$$

$$\Delta Y = \frac{\Delta G}{(1 - c_1) + (c_2 + b)\,(n/e)} \tag{8.25}$$

This represents the distance $Y_0 Y_1$ in Figure 8.15.[15] Even though the rise in government expenditure caused real income to rise initially by a multiple amount $1/(1 - c_1)$, the induced rise in the interest rate caused investment and consumption expenditures to decline by a multiple amount $(c_2 + b)\,(n/e)\,\Delta Y/(1 - c_1)$.

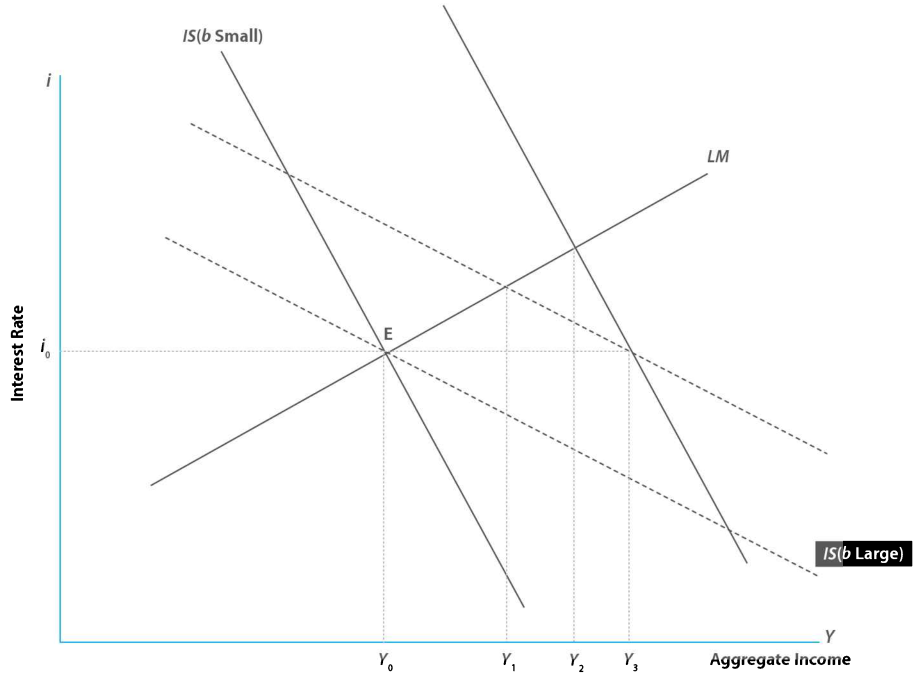
The induced reduction in consumption and investment, which offsets the effect of the increase in government spending, is referred to as CROWDING OUT. The tendency for the stimulative effects of fiscal policy to be offset by an induced decline in private expenditures is called crowding out. As can be seen from Eq. (8.25), CROWDING OUT IS LARGER (the final impact on income is smaller):

› Crowding out occurs when the stimulative effects of fiscal policy are offset by a reduction in the interest-sensitive consumption and investment due to a rise in the interest rate.

a. thelarger is $b$—the interest sensitivity of investment is large
b. the larger is $n$—the income sensitivity of money demand is large
c. the smaller is $e$—the interest sensitivity of money demand is relatively small
d. the larger is $c_2$—the interest sensitivity of consumption expenditure is large

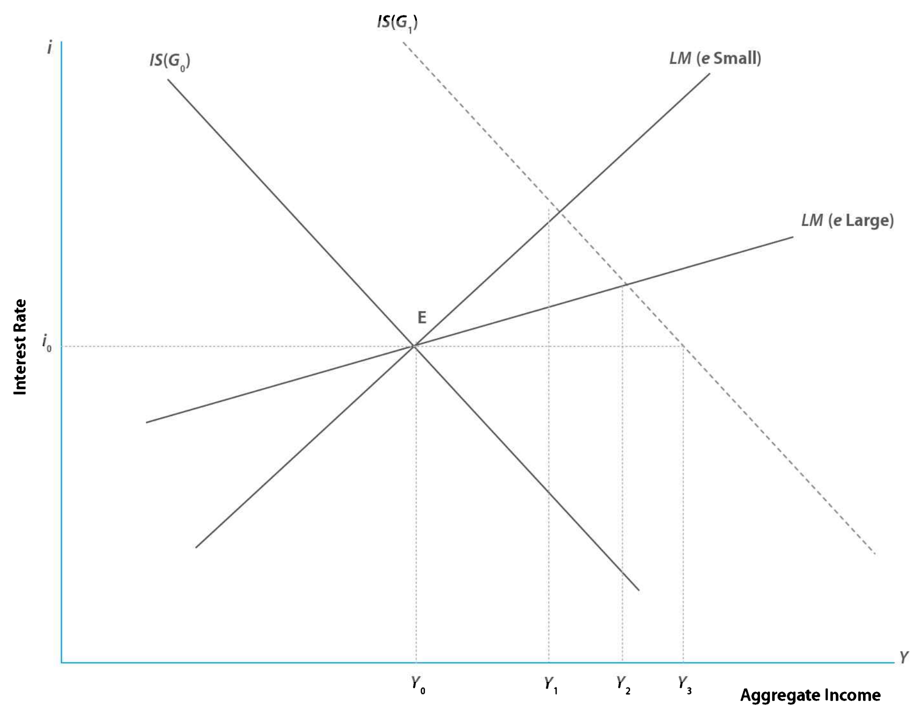
› Larger the crowding out (when the interest sensitivity of investment and consumption expenditure is large), the larger is the income sensitivity of money demand, and the smaller is the interest sensitivity of money demand.

As an illustration, we depict the extent of crowding out in the case when $b$ is large, and also when it is small in Figure 8.16, where the interest sensitivity of the *IS* curve has been previously derived in Figure 8.7. The increase in government expenditure at the initial interest rate of $i_0$ expands income from $Y_0$ to $Y_3$. However, the crowding out when $b$ is large is $Y_1 Y_3$, which is greater than the crowding out when $b$ is small at $Y_2 Y_3$.

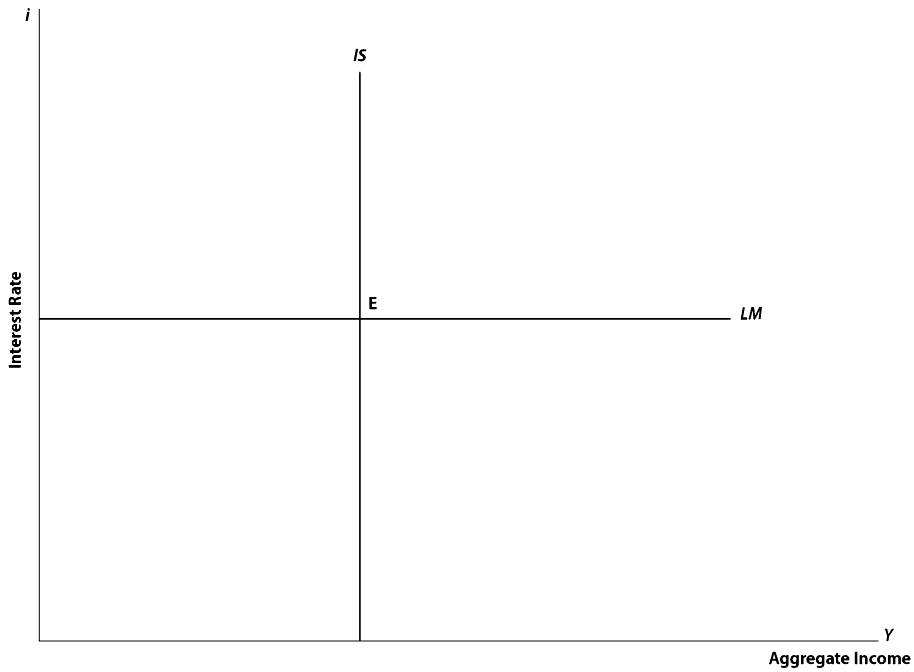



**› Figure 8.16**
Crowding out Effect of Interest Sensitivity of Investment. **When investment is not sensitive to interest rates (*b* is small), shown by the solid *IS* line, crowding out is given by distance $Y_2Y_3$. The decline in investment caused by an increase in government expenditure is smaller than when investment is sensitive to interest rate changes (*b* large)—the dashed *IS* line where crowding out is the distance $Y_1Y_3$.**

Similarly, Figure 8.17 depicts the effect of crowding out when the demand for money is interest elastic (that is $e$ is large) counterposed with the case when the demand for money is interest inelastic (that is $e$ is low), where the interest sensitivity of the *LM* curve has been previously derived in Figure 8.12. The increase in government expenditures at the initial interest rate expands income from $Y_0$ to $Y_3$. However, crowding out when money demand is interest inelastic (that is $e$ is small) is $Y_1Y_3$, which is larger than the crowding out of $Y_2Y_3$ when money demand is interest elastic—$e$ is large.




**› Figure 8.17**
Crowding out Effect of Interest Sensitivity of Money Demand. **When an increase in the interest rate does not have much effect on the demand for money (*e* is small) the *LM* curve is steep or interest inelastic. The interest rate must then rise by a relatively large amount to induce individuals to sell bonds and shift their portfolio towards money that is required for transacting expenditures at higher incomes. This crowds out—reduces—private expenditures to a larger extent than when the *LM* curve is flat (*e* is large) as there the interest rate need not increase that much to induce a reduction in bond holdings.**




**› Figure 8.18**
Fiscalist Case. When investment is not sensitive to interest rates (*b* is very small), the *IS* curve is vertical. When the demand for money is very sensitive to interest rates (*e* is very large) the *LM* curve is horizontal. In this case only fiscal policy can influence the level of aggregate income, a Fiscalist scenario.

### *8.5.2* Fiscalist and Monetarist Intervention

It is useful to consider two special cases that help in summarizing the discussion on fiscal and monetary policy. In the first instance, which we label the *Fiscalist case*, investment is insensitive to interest rates ($b$ is very small, approximately zero), and the *IS* curve is vertical. Also, the demand for money is very sensitive to the interest rate and is interest elastic ($e$ is very large). In the limit when investment is interest inelastic ($b$ is approximately zero), the *IS* curve is vertical.[16] When the money demand is perfectly interest elastic ($e$ is very large), the *LM* curve is horizontal (see Figure 8.18).

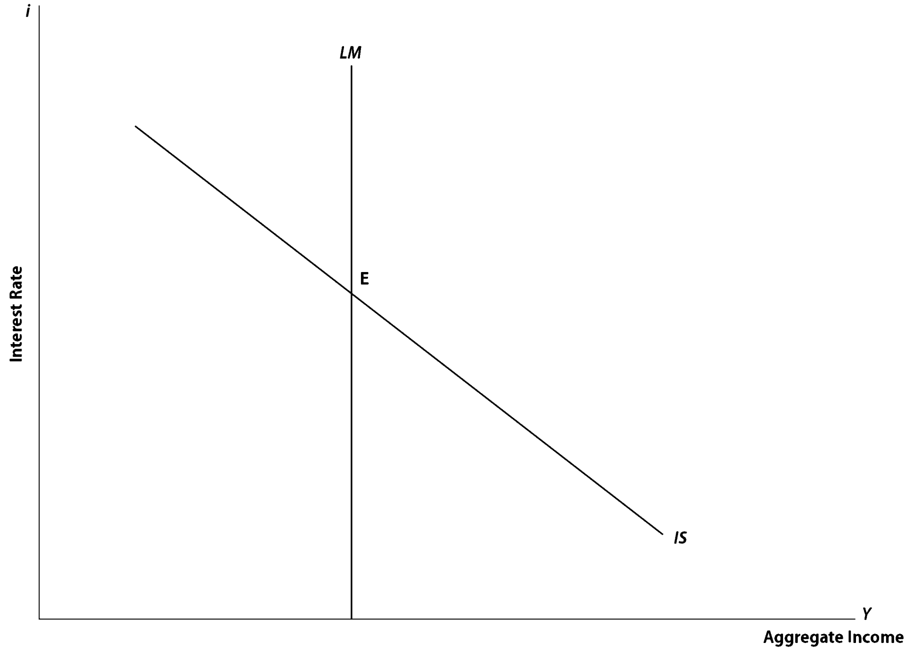
In this situation fiscal policy is powerful and monetary policy has no impact on the economy. An increase in government expenditure shifts the *IS* curve to the right and there is no crowding out as the interest rate never rises to reduce investment and interest-sensitive consumption spending. Money and bonds are perfect substitutes when the *LM* curve is horizontal, which means selling bonds to acquire money balances for facilitating increased transactions has no impact on the interest rate. When the economic configuration is as depicted in Figure 8.18, deploying monetary policy is meaningless.

The second limiting case that we consider can be labelled the *Monetarist case* (see Figure 8.19). In this situation, the demand for money is perfectly interest inelastic ($e$ is approximately zero) and the *LM* curve is vertical. The interest elasticity of the *IS* curve is immaterial as long as it is not perfectly interest inelastic.[17] In this situation, fiscal policy is powerless as any increase in government expenditure shifts the *IS* curve outwards but raises interest rates so dramatically that the decline in investment expenditure fully offsets the increased government expenditure. Government expenditure is considered to be fully crowded out. An increase in money supply, on the other hand, has beneficial impacts on real income, which increases as the *LM* curve shifts out and intersects the *IS* curve at a lower nominal rate of interest.

Crowding out is, therefore, greater, the steeper the *LM* curve and the flatter the *IS* curve—the Monetarist Case.

Conversely, when crowding out is smaller, the flatter is the *LM* curve, and the steeper the *IS* curve—the Fiscalist Case. These results are summarized in Table 8.2.

**› Figure 8.19**
Monetarist Case. When the demand for money is not at all sensitive to interest rates (*e is* approximately zero), the *LM* curve is vertical. In this case monetary policy is very effective in influencing the level of aggregate income.




MACROFOCUS 8.2

## Zero Interest Rates and Macro Policy in Japan

The Japanese economy experienced a period of deflation and slow growth after the stock and real estate markets collapsed in the late 1980s. During the 1980s, the Japanese economy was allowed to overheat with real estate speculation combined with a loose lending policy by banks. In 1987, capital gains from securities and real estate transactions were 40 per cent larger than GDP.* By 1989, the Bank of Japan started to dampen the speculation by raising interest rates and introducing lending regulations for real estate. This speeded up the burst of the speculative bubble and the Nikkei index fell from 38,915 in December 1989 to 7,000 in early 2003. It has been estimated that the decline in wealth was equal to two years' GDP. The decline in asset prices resulted in a massive deflation. During the 1990s the recession in Japan resulted in a decline in goods prices but this was insufficient to restore the demand for goods. The deflationary episode in Japan was broad-based as prices for all major items, including food, clothing, transportation, and durable goods, declined. The year-on-year consumer price inflation rate has been negative for about six years since the second quarter of 1998. In an effort to stimulate the economy, the Bank of Japan board made a decision on February 12, 1999 to lower the interest rate to be "as low as possible". The overnight call interest rate in the money market has been virtually zero since this decision. When the nominal interest rate is zero the economy is said to be in a liquidity trap. When prices start falling (a deflation), nominal interest rates begin to decline but the real cost of taking out loans rise and investment declines. With prices falling, real interest rates—nominal interest rates minus inflation—are high. Even though the nominal interest rate is zero per cent, loans are expensive in real terms and businesses find interest payments high. With falling prices, consumers also tend to delay purchases and this reduces demand further. A decline in consumption is a decline in the demand for goods produced by firms, and this reduces firm profitability and the incentive to invest further. This leads to a vicious spiral because banks, realizing that business profitability is in doubt, would be reluctant to lend. In any case, with low-lending rates and a zero lower bound on deposit rates, the profit margins on commercial lending are exceedingly low.

The interest rate offered by banks or for government bonds cannot be set below zero, as in that case, everyone would hold cash, which has a nominal interest rate of zero per cent and nobody has an incentive to hold financial instruments with a negative nominal return. Monetary policy is then incapable of stimulating the economy when the zero bound has been reached. Can fiscal expansion be the solution? Since the bubble burst in 1990, the Japanese government has spent in excess of 120 trillion yen on more than 12 different programmes aimed at stimulating economic activity.** Tax revenues declined with the burst of the bubble and fiscal deficits rose. Government expenditure also increased as part of a countercyclical fiscal package. Government debt accumulated to 108 per cent of GDP by 1999 and the deficit was 10 per cent of GDP.†

Yet the economy did not recover from the deflationary spiral. Why? The Japanese had been saving a little more than 30 per cent of GDP, which had financed investment when the collapse of the 1990s occurred. With the decline in domestic investment demand by firms, the excess savings were placed abroad (a current account surplus), and they financed the fiscal deficit (a reduction in public savings). However, when the government borrows money, individuals realize that the loans have to be paid back one day. The expectations of higher taxes in the future mean that the expected lifetime post-tax earnings are lower and this induces lower consumption. Also, Japan has a large segment of older people and their concerns about rising health burdens and anxiety regarding nursing care has caused them to increase their savings. From 1990, therefore, household savings rates have gradually risen†† and fiscal policy has been ineffective in raising aggregate income.

* A. Bigsten, "Can Japan Make a Comeback?" *The World Economy* 28, no. 4 (2005): 595–606.
** Tzong-Shian Yu, "Can East Asia rise again?", *Journal of Asian Economics* 13 (2003): 715–729.
† T. Ihori, T. Nakazato, and M. Kawade, "Japan's Fiscal Policies in the 1990s," *The World Economy* 26, no. 3 (2003): 325S–338S.
†† Nakagawa, "Why Has Japan's Household Savings Rate Remained High even during the 1990s?" Bank of Japan Monthly Bulletin (April 1999).

**› Table 8.2**
Effectiveness of the Fiscal and Monetary Policy

| | Fiscal Policy Effective | Monetary Policy Effective |
|---|---|---|
| Interest Sensitivity of Private Sector Expenditures | Private Expenditures Interest Inelastic ($b, c_2$ small) | Private Expenditures not Perfectly Interest Inelastic ($0 < b, c_2 < \propto$) |
| Interest Sensitivity of the Demand for Money | Demand for Money Interest Elastic ($e$ large) | Demand for Money Interest Inelastic ($e$ small) |

# 8.6 *IS–LM* in India

Prior to the liberalization of the Indian economy in the 1990s, fiscal policy was the dominant tool of macroeconomy policy and monetary policy played a subservient role.[18] Prior to the 1990s, increases in government expenditure—fiscal policy—were financed through two main mechanisms. The increased government borrowing was accommodated by hikes in the STATUTORY LIQUIDITY RATIO (SLR) imposed on commercial banks. This allowed the placement of government securities with commercial banks at below market rates of interest. By 1990, the SLR was 39 per cent. Second, the residual funding required was accommodated by the central government unilaterally accessing the central bank (the RBI) through the medium of *ad hoc* TREASURY BILLS. This mechanism was routinely used in the 1980s to meet the government's requirement of funds and this resulted in a monetization of the deficit which in turn expanded reserve money—the changes in net RBI credit to the central government accounted for 93 per cent of the variation in reserve money in the 1980s.[19] Such an automatic monetization of the deficit circumscribed the scope for discretionary monetary policy severely.

› The statutory liquidity ratio specifies the required holding of government securities as a fraction of the assets of the commercial bank.

› A Treasury bill or T-bill is a government bond with a maturity period not exceeding one year.

The growing fiscal deficit raised concerns about its SUSTAINABILITY and impact on the economy. Fiscal adjustment that proved difficult to sustain over the 1990s was finally sought to be achieved through the enactment of the Fiscal Responsibility and Budget Management Act in 2003 that institutionalized fiscal consolidation by a legislative mandate (see Chapter 12). Similarly, the autonomy of monetary policy was instituted through the phasing out of *ad hocs* and shifting government borrowing towards market rates of interest (see Chapter 14). As the suppression of the allocative role of interest rates in the money market was removed, the central bank could begin to conduct OPEN-MARKET OPERATIONS. This raised the ability of the central bank to influence liquidity and removed the constraint on DISCRETIONARY MONETARY POLICY.

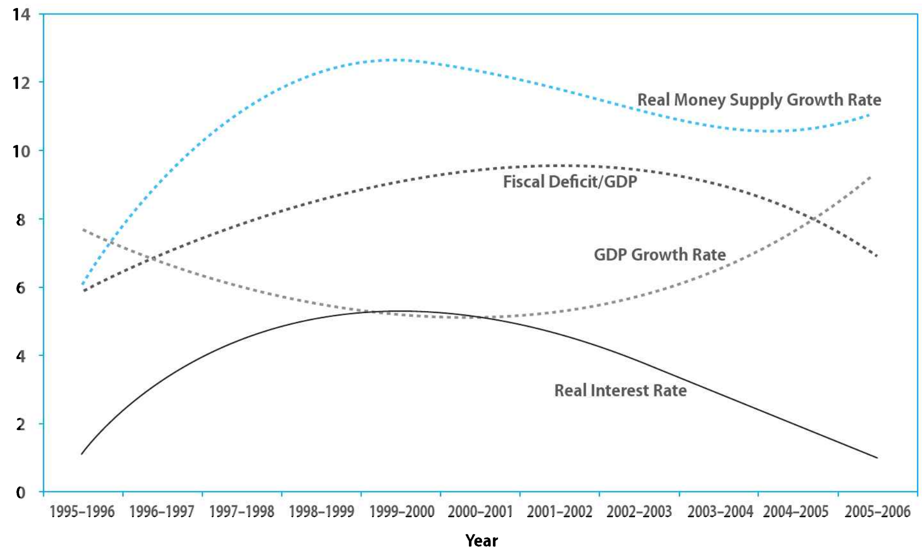
› The **sustainability** of fiscal deficits is analysed in Chapter 12.

› An open-market operation is a purchase (or sale) of government bonds from (to) the public by the central bank for the purpose of increasing (decreasing) the supply of bank reserves and the money supply.

› Discretionary monetary policy results from the deliberate and conscious choices by the central bank on money supply (or the interest rate) in order to influence output and prices in the economy.

Figure 8.20 depicts the trends[20] in the fiscal deficit, real money supply, and output (GDP) growth, and the real interest rate over the period 1995–1996 to 2005–2006. The *IS–LM* approach informs us that the interest rate rises when the *IS* curve shifts to the right (fiscal policy). Also, an increase in income (output) raises the demand for money for transactions purposes and raises the interest rate in the money market as well. The interest rate declines, however, if there is an increase in real money supply that shifts the *LM* curve to the right (monetary policy).

Figure 8.20 reveals that from 1995–1996 to 2000–2001 both fiscal and monetary policy were expansionary with a rise in the fiscal deficit and an increase in the real money supply growth rate. However, the impact of fiscal policy was larger than that of monetary policy in this period. The fiscal deficit (the extent of government borrowing) grew at a trend rate of 20.9 per cent during 1995–1996 to 2000–2001 compared to the money supply trend growth rate of 16.1 per cent. In contrast, during the period of 2001–2002 to 2005–2006, the fiscal deficit growth rate decelerated by a factor of 7 times to 3 per cent, whereas money supply growth decelerated by a factor of just 1.2 times to 13.7 per cent. The large government




› **Figure 8.20**
Trendlines—Macroeconomic Policy Stance. During 1995 to 2000, fiscal policy was more expansionary than monetary policy and was accompanied by a rise in the interest rate and a diminishing rate of increase in income (GDP). From 2000 onwards monetary policy was more expansionary resulting in a decline in interest rates and an increase in the rate of growth in income.

borrowing programmes in the period up to 2000–2001 shifted the *IS* curve more than the rightward shift of the *LM* curve and raised the real interest rate of the economy. The high interest rate dampened the output increase in the period with a trend deceleration in output growth (Figure 8.20). In the latter period from 2001–2002 to 2005–2006, government borrowing retreated whereas the expansionary character of monetary policy was not curtailed extensively. During this period, the *LM* curve shifted more to the right than the rightward shift of the *IS* curve. As a result an interest rate decline accompanied the rise in output. Moreover, this is the period when improvements in the technology of financial transactions (such as credit and debit cards) reduced the demand for money for transactions purposes. The demand for money grew at a smaller pace despite the swifter growth in output and the *LM* curve shifted rightwards. An easy monetary policy that accommodated the output increases and a gradual reigning in of fiscal policy has thus accompanied the recent boom in the INDIAN ECONOMY.

› During the first half of this decade the *LM* curve shifted out to the right to a larger extent than the rightward shift of the *IS* curve. The **Indian economy** witnessed a decline in interest rates and an increase in output.

Consistent with this explanation, Dua and Pandit[21] found that the money supply and government expenditures' impact on the interest rate in the Indian economy was as predicted by *IS–LM* theory. However, as they emphasize, the interest rate is not only affected by domestic factors but also by external factors. In fact, they show how deviations of the domestic interest rate from the interest rate abroad such as $i > i^*$ leads to a capital inflow, as predicted by the interest rate parity approach. In fact, capital inflows have become such a dominant component of money supply that the Economic Survey, 2002–2003, of the Government of India remarked "from 9.1 per cent as at end March 1991, the share of net foreign exchange assets in reserve money ... became 100.7 per cent on 24 January 2003, which is close to a currency board situation".[22] It is this extent of capital flows that has been the major factor behind the robust increase in money supply in the Indian economy. Managing capital flows has therefore become a predominant concern of monetary policy, an issue we examine in Chapter 11.

# 8.7 Ricardian Equivalence

One of the implications of the *IS–LM* analysis is that bond financing of government expenditures is more expansionary than tax financing of government expenditures. We have already seen that bond financing initially increases real income by $\Delta Y = \Delta G$ and the increased real income induces a further rise in consumption expenditure of $c_1 \Delta Y$ making the horizontal shift of the *IS* curve equal to $\Delta Y = \Delta G + c_1 \Delta Y$ or, $\Delta Y = \Delta G / (1 - c_1)$.

In contrast, if the increase in government expenditure is financed by raising tax revenues, $\Delta G = \Delta T$, the initial increase in real income due to the increase in government expenditure induces an increase in consumption expenditure given by $c_1 \Delta Y$. However, now there is an offsetting effect due to a decline in consumption expenditure as the increased taxes reduce disposable income available for spending.

From the consumption function Eq. (8.11) the reduction in consumption is given by $c_1 \Delta T$, which is equal to $c_1 \Delta G$ as $\Delta T = \Delta G$. Then, the rightward shift of the *IS* curve, when increased government expenditures are tax financed, is given by

$$\Delta Y = \Delta G + c_1 \Delta (Y - T)$$

$$= \Delta G + c_1 \Delta Y - c_1 \Delta G$$

or,
$$\Delta Y (1 - c_1) = \Delta G (1 - c_1)$$

or,
$$\Delta Y \big|_{\text{tax}} = \Delta G < \frac{\Delta G}{1 - c_1} = \Delta Y \big|_{\text{bonds}}$$

The rightward shift of the *IS* curve is larger under bond financing than tax financing and so bond financing of government expenditure is more expansionary than tax financing. This occurs because bond financing does not reduce the disposable income that is available for spending whereas tax financing does reduce disposable income and lowers aggregate expenditure.

The Ricardian equivalence (RE) proposition named after the classical economist David Ricardo (1772–1823) argues that the above reasoning is faulty and that essentially the impact on real income is the same whether government expenditure is tax financed or bond financed. In other words, the rightward shift of the *IS* curve is identical whether government expenditures are tax financed or bond financed. Ricardo's reasoning, which was resurrected in an influential article by Robert Barro,[23] is as follows:

› Ricardian equivalence (RE) is the proposition that in terms of the impact on the aggregate income it is irrelevant whether increased government expenditure is financed by taxes or bonds.

A government that issues bonds to finance an increase in expenditure must in the future pay interest on these bonds. To pay this interest taxes will eventually have to be increased. Far-sighted tax payers realize that the government has substituted a current tax that it could have imposed to finance its expenditures with a future tax. Thus rational, self-interested individuals will respond by saving the implicit tax cut of bond financing until the future, when the government increases taxes to pay back the debt. At that time, the savings will be used to pay the increased taxes. As a result government financing by borrowing has the same impact as financing by taxation since far-sighted individuals save an amount that is equivalent to the future taxation by the government.

For Ricardian equivalence to hold, two assumptions are required:

1. Individuals are forward looking and have perfect foresight.
2. The government budget is intertemporally balanced.

› Ricardian equivalence occurs when the government budget is intertemporally balanced and individuals are forward looking with perfect foresight about the future consequences of current actions.

Both assumptions imply that when the government finances a given level of expenditures by selling bonds, households, recognizing that the budget is intertemporally balanced, will have the foresight that higher taxes will be levied in the future to pay back the borrowings on the bonds sold today. Hence, the tax burden for individuals who have foresight is not just current taxes, but rather current taxes plus the present value of the taxes in the future that will be required to pay the interest on current bond sales.

For simplicity assume that the government does not resort to money financing and each government bond pays a coupon of INR 1 per annum.

Then, the market price of a bond is

$$P_B = \frac{1}{i}$$

Let $\Delta B$ be the number of bonds sold to finance the increased government expenditures. Then, the nominal value of these bonds is

$$\text{Value of bonds} = P_B \Delta B = \frac{\Delta B}{i}$$

The annual interest payments on the bonds sold will be

$$\text{Interest Payments on the bonds} = i\left(\frac{\Delta B}{i}\right) = \Delta B$$

The government will have to make interest payments of INR $\Delta B$ in each period, and it will make these payments by increasing taxes. Hence, the present value (PV) of implied future taxes is

$$\begin{aligned} \text{PV (Future Taxes)} &= \text{PV (Interest Payments)} \\ &= \frac{\Delta B}{1+i} + \frac{\Delta B}{(1+i)^2} + \cdots \\ &= \frac{\Delta B}{i} \end{aligned}$$

The real present value of implied future taxes will then be $\Delta B / iP$. Thus, the current period tax burden on individuals who have perfect foresight and realize the inter-temporal nature of the budget, individuals who follow the propositions of RE, is

$$T_{\text{RE}} = T + \frac{\Delta B}{iP} \tag{8.26}$$

The current period government budget equation states that current government expenditure is financed by current taxes and the real value of debt issued to finance current expenditures as given by

$$\frac{P_B \Delta B}{P} = \frac{\Delta B}{iP}.$$

The current period budget constraint of the government can then be written as

$$G = T + \frac{\Delta B}{iP} \tag{8.27}$$

But the right-hand side of this government budget constraint is none other than the current period tax burden experienced by individuals subject to RE. This makes the appropriate definition of disposable income, when RE propositions hold, as

$$Y^d_{\text{RE}} = Y - T_{\text{RE}} = Y - G \tag{8.28}$$

In the *IS–LM* approach described earlier, the current burden of the government is the current taxes levied, *T*, which affects the amount of disposable income in the hands of individuals. In the Ricardian approach, the *current burden of government* by contrast is the *goods and services it absorbs, G,* as given in Eq. (8.22). Accordingly, equilibrium in the goods market can be rewritten as

$$\begin{aligned} Y &= C + I + G \\ &= c_0 + c_1 (Y - G) - c_2(i - \pi^e) + a - b(i - \pi^e) + G \end{aligned}$$

or, $$Y(1 - c_1) = c_0 + G(1 - c_1) - c_2(i - \pi^e) + a - b(i - \pi^e) \tag{8.29}$$

This leads to two propositions:

*Proposition I:* It is irrelevant whether the increased government expenditure is financed by taxes or bonds.

From Eq. (8.29),

$$\frac{\Delta Y}{\Delta G} = \frac{1 - c_1}{1 - c_1} = 1$$

or, $$\Delta Y = \Delta G$$

Hence, the *IS* curve, under RE, shifts out by the same distance irrespective of whether government expenditures are tax financed or bond financed. This is because bond sales impose the same burden on individuals as the taxes that would have been imposed in their place.

*Proposition II:* A tax cut financed by bond sales has no effect on output.

This is due to the bond sales having the implication that higher future taxes will have to be levied to pay back the interest on the bonds. The present value of the higher future taxes is equal to the present tax cut. Moreover, disposable income as defined in Eq. (8.28) has not changed and the *IS* curve has consequently not shifted out. Recall that the disposable income measured in a standard way is used for consumption expenditure or saved:

$$Y^d = Y - T = C + S$$

or,

$$S = Y - C - T$$

Under RE, $C = C(Y - G, i - \pi^e)$ rather than as defined in Eq. (8.10) so that

$$S = Y - C(Y - G, i - \pi^e) - T$$

With $G$ unchanged when current taxes are substituted by debt,

$$\Delta S = -\Delta T$$

The change in taxes thus has an equal and opposite effect on savings. Tax cuts are saved completely unlike in conventional *IS–LM* theory where a tax cut ($\Delta T < 0$) increases disposable income by the amount ($Y - \Delta T$) and a fraction $c_1(Y - \Delta T)$ of this is spent by individuals. This shifts the *IS* curve outwards.

In the RE approach, governments should not bother too much about the deficits financed by debt because in macroeconomic terms what the government dis-saves, the private sector saves and aggregate savings in the economy do not get altered. People just save to pay for the anticipated higher taxes in the future when the government dis-saves.

Two reasons advocated for why this proposition may not hold are myopia and liquidity constraints. Keynesians, in particular, argue that those who advocate RE give people too much credit for looking far ahead into the future. Most people are short sighted and the time horizon of their decision making is about a year within which an attempt is made to make ends meet rather than concern themselves with tax burdens far into the future. Moreover, most people face constraints on the availability of credit to meet spending needs at various points of time in their lives.

Instead of saving the proceeds of a tax cut individuals, who are subject to liquidity constraints, may allocate the tax cut to higher consumption that they are compelled to curtail due to constraints on the availability of credit. For these and other reasons[24] many economists hold the view that RE holds true only to a limited extent, and that budget deficits and the mode of their financing do indeed matter.

In developing countries, where financial systems are underdeveloped capital markets lack depth and are subject to financial repression, and where there is considerable uncertainty regarding the incidence of taxes, the Ricardian propositions may not hold. Haque and Montiel[25] were unable to find evidence of debt neutrality for 15 out of a group of 16 developing countries. Veidyanathan,[26] using a larger sample of 60 countries and data covering three decades, was unable to detect any significant effects of public sector deficits on private savings and attributed this to individuals in these countries being subject to borrowing constraints. However, though individuals may not have perfect foresight, they are forward looking as any policymaker who is smart realizes fast on the job. Individuals anticipate policy, figure out if they are credible, and take offsetting action if necessary.

## SUMMARY

- » Walras' Law states that the sum of the excess demand for money, bonds, and current output must equal zero.
- » Walras' Law allows macroeconomics to explicitly analyse the markets for goods, output, and money markets and ignore the bond market.
- » The nominal interest rate is the sum of the real interest rate and the expected rate of inflation.
- » Consumption is a positive function of aggregate income and is negatively related to the real interest rate.
- » Investment is positively related to the productivity of capital and negatively to the real interest rate.
- » Equilibrium in the product market is given by the condition that savings—private plus government savings—equal investment.

$$[(Y - T) - C] + (T - G) = I$$

- » An increase in income increases disposable income and aggregate savings.
- » An increase in taxes reduces private savings {by $(l - c_1)\Delta T$} but increases government savings {by $\Delta T$}, with the net effect being a rise in aggregate saving {by $c_1\ \Delta T$}.
- » An increase in expected inflation reduces the real interest rate and makes savings for the purpose of future consumption less attractive.
- » An increase in government expenditure reduces public savings and aggregate savings.
- » The *IS* curve is downward sloping in the nominal interest rate-aggregate income space as a rise in income raises the supply of savings to the capital market and causes the interest rate to decline.
- » The *IS* curve shifts to the right with an increase in government expenditure since the increase in government bonds causes bond prices to decline and the interest rate to rise.
- » An increase in taxes increases aggregate savings and the nominal interest rate declines, thereby shifting the *IS* curve leftwards.
- » An increase in the productivity of capital increases the demand for investment and the interest rate thereby shifting the *IS* curve out to the right.
- » An increase in expected inflation leaves aggregate savings and investment unchanged but shifts the *IS* curves out to the right as the nominal interest rate rises.
- » The lower the responsiveness of aggregate investment to the interest rate, the steeper is the *IS* curve.
- » The *LM* curve is upward sloping in the nominal interest rate—aggregate income space as an increase in income raises the demand for real money balances which results in a rise in the interest rate.
- » The *LM* curve shifts to the right if there is an increase in the money supply as interest rates decline in response to individuals rebalancing their financial portfolio by purchasing bonds when this occurs.
- » The *LM* curve is inelastic (elastic) when the demand for money is interest inelastic (elastic).
- » The *IS–LM* model is appropriate for the short run where wages and prices do not change and there is underutilization of capacity.
- » An expansionary monetary policy in the *IS–LM* is more effective the more interest inelastic is the demand for money as long as private expenditures on consumption and investment are not interest inelastic.
- » An expansionary fiscal policy in the *IS–LM* model is more effective the more interest inelastic are private investment and consumption expenditures and the more interest elastic is the demand for money.
- » Crowding out occurs when the stimulative effects of fiscal policy are offset by a reduction in consumption and investment due to the rise in the interest rate.
- » Crowding out is larger the larger is the interest sensitivity of investment and consumption expenditure, the larger is the income sensitivity of money demand, and the smaller is the interest sensitivity of money demand.
- » Ricardian equivalence is the proposition that whether increased government expenditure is financed by taxes or bonds is irrelevant in terms of the impact on aggregate income.
- » For Ricardian equivalence to hold the government budget must be intertemporally balanced and individuals should be forward looking with perfect foresight about the future consequences of current actions.

## NOTES

1. We are for obvious macroeconomic reasons ignoring raw material costs which are cost of intermediate goods.
2. Alternatively, agents could spend less than their income and accumulate financial assets.
3. Recall that money demand has a role as a medium of exchange.
4. The real money stock is written as $M/P = m$, and the real stock of bonds is $b = P_B\, B/P$.
5. Modern macroeconomics has ignored the general equilibrium spillovers across market by deriving aggregate demand and supply separately. However, the wage and price setting behavior of firms will affect consumer's demand for goods through their effects on the labour market. Hence, aggregate demand cannot in general be derived separately from aggregate supply. However, if we presume that the marginal rate of substitution between consumption across periods and between consumption (income) and real money balances within a period is independent of leisure, then, aggregate supply is independent of aggregate demand.

Then, demand and supply of goods do not depend on conditions in the labour market. We explicitly assume these conditions hold. Otherwise, the income on the left-hand side of Eq. (8.5) depends on the labour market, and we would explicitly have to include excess demand for labour in what follows.

6. We are, therefore, assuming the substitution effect dominates the income effect; see the section on effects of interest rates in Chapter 3.
7. The lifetime budget constraint will then be

$$C_1 + [C_2/(1 + r)] = (Y_1 - T) + [(Y_2 - T)/(1 + r)]$$

8. For simplicity, taxes in this chapter are lump-sum taxes. This assumption does not affect the basic results.
9. If some government expenditure is investment expenditure, such as on infrastructure, then, that it can increase the productivity of private investment expenditure. We would then write *a* which represents the productivity of investment as a function of government investment and the value of *a* in Eq. (8.12) would be increasing with government investment expenditure. It is well known that infrastructure expands market opportunities and results in gains to the economy in terms of increasing the scale and efficiency of production and output.
10. In this chapter, we also do not model expectations endogenously. Expectations will be deemed to be exogenous and do not alter in response to specific policy changes. Expected inflation is, therefore, treated as exogenous.
11. The investment function is $I = a - b(i - \pi^e)$, which can be rewritten as $i = (a - I) / b + \pi^e$. Hence, a small *b* makes the investment schedule steep in $(i, I)$ space.
12. Recall from Chapter 2 that $M^s = C^P + D_R$, or, money supply is currency in circulation plus deposits.
13. The increase in the demand for money also follows directly from Eq. (8.18).
14. John R. Hicks, "Mr. Keynes and the Classics: A Suggested Interpretation," *Econometrica* (April1937): 147–159.
15. Alternatively, we could obtain Eq. (8.25) by setting the interest rate in the *IS* Eq. (8.16) equal to the interest rate in the *LM* Eq. (8.22) and solving that expression for *Y*. Then, Eq. (8.25) follows directly from this solution to the *IS–LM* model for *Y*.
16. Strictly speaking we require consumption too to be interest inelastic. However, if we interpret *b* broadly as the interest sensitivity of private sector expenditures, which is what is the case here, the outcome we report in the Fiscalist case goes through.
17. This requires $0 < b < \infty$.
18. These issues are further discussed in Chapters 12 and 14.
19. Chakravarthi Rangarajan "Monetary Management: The Changing Framework," printed in "The Indian Economy: Essays on Money and Finance," UBS Publishers' Distributors, 1998.
20. These trend lines in Figure 8.20 are the result of fitting a polynomial to each data series.
21. Pami Dua and B. L. Pandit, "Interest Rate Determination in India: Domestic and External Factors," *Journal of Policy Modeling* 24 (2002): 853–875.
22. Government of India, Economic Survey, 2002–2003: Paragraph 1.9. http://indiabudget.nic.in/es2002-03/esmain.htm
23. Robert J. Barro, "Are Government Bonds Net Wealth?" *Journal of Political Economy* 82 (1974): 1095–1117.
24. For instance, the tax changes considered in the theory are lump sum tax changes such as tax rebates. Changes in marginal tax rates, however, influence decisions on production and the supply of labour.
25. Nadeem U. Haque and Peter Montiel, "Consumption in Developing Countries: Test for Liquidity Constraints and Finite Horizons," *Review of Economics and Statistics* 71 (August 1989): 408–415.
26. Geetha Veidyanathan, "Consumption, Liquidity Constraints, and Economic Development," *Journal of Macroeconomics* 15 (Summer 1993): 591–610.

## TEST YOURSELF

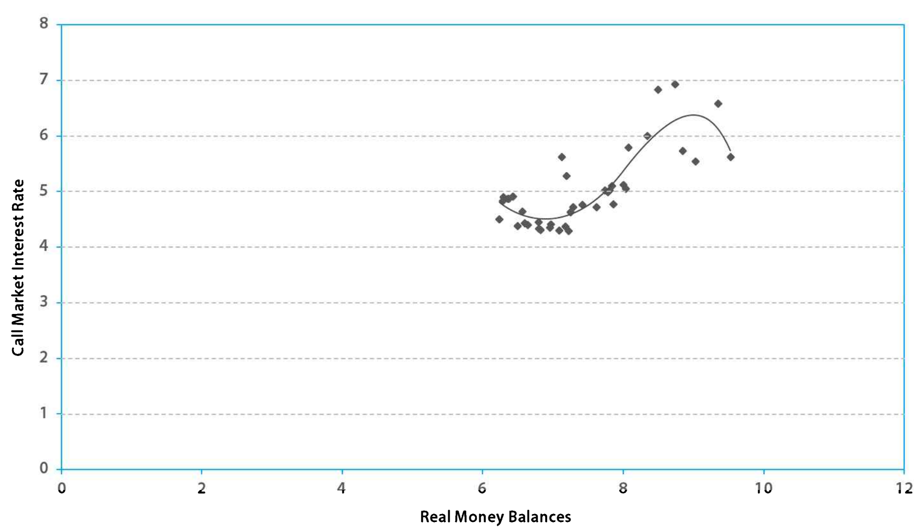
1. What is Walras' Law and how does macroeconomics employ it to eliminate the explicit treatment of the bond market?
2. Why is the *IS* curve a depiction of equilibrium in the product market? What factors cause the *IS* curve to shift and in what direction (outwards or inwards)?
3. Why is the *LM* curve a depiction of equilibrium in the money market? What causes the *LM* curve to shift to the right or left?
4. What makes the *IS* curve and the *LM* curve steeper, or more inelastic, and why?
5. What factors restrict the effectiveness of using fiscal policy as a means to increasing aggregate income? When is fiscal policy most effective?
6. Under what circumstances is fiscal policy more effective in influencing the level of income in the economy than monetary policy and vice versa?
7. "Tax financing of government expenditure is less expansionary than bond financing." Examine the conditions under which this proposition is true.
8. Demonstrate using the *IS–LM* model whether it may be necessary

to have negative interest rates in order that the economy achieves a given level of income that policymakers seek.

9. Suppose that in an economy with a balanced budget the government, in order to increase voter support, decides to temporarily reduce tax rates. After a year the tax rate will return to its original level. Meanwhile the government issues perpetual bonds to cover the cost of the tax cut. What would be the impact of the tax cut on domestic saving and consumption?
10. Suppose that the responsiveness of investment and consumption to the interest rate reduces in the *IS–LM* set-up. Would the change in the interest rate be required to be larger or smaller in order to have the same final impact on aggregate income?

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
2. Then under the classification of annual data, click on "Handbook of Statistics on Indian Economy"
3. Click on "Table 181: Components of the Money Stock" and create a spreadsheet of monthly data on narrow money (M1) for the period April 2003 to June 2006.
4. Then click on "Table 178: Wholesale Price Index—Monthly Average" and in your spreadsheet insert monthly data on the wholesale price index from April 2003 to June 2006.
5. The wholesale prices are in the form of indices. To create an indexed data series of real money balances, index the data you inserted in Step 3. For example, let the M1 data for April 2003, INR 4,933.33 billion, equal the index value of 1,100. In that case the index for May 2003 for INR 4,955.35 billion is 1,105.
6. Divide the index for M1 by the price index of the Wholesale Price Index in that month. For instance, in April 2003, the WPI index is 173.1. Thus, the index for real money balances in April 2003 is (1,100/173.1) = 6.35.
7. Next go back to the RBI Handbook of Statistics on Indian Economy. Click on "Table 188: Weighted Average Call Money Rates". Take the data for the corresponding period April 2003 to June 2006 on the average call money interest rate each month.
8. Create a graph depicting the real money balance index on the horizontal axis and the interest rate on the vertical axis. Does the real money balance vary inversely with the interest rate?
9. You would have noticed that there is no clear-cut relationship between the money balances and the interest rate. In fact, if you fit a trend to the series of dots in your graph obtained in Step 9 you would obtain a graph like that in Figure 8.21, where initially there is a negative relationship followed by a positive and then again negative relationship indicating a tilted *S* curve. Also, if you create a graph between the change in the interest rate in a year and the change in the real money balances in that year and fit a linear trend to that graph, you would observe a mildly negative relationship that seems to indicate a negative relationship between changes in money balances and changes in the interest rate.
10. Give reasons for why the relationship between real money balances and the interest rate may not be strictly negative.




› **Figure 8.21**
Money–Interest Rate Relationship (April 2003 to June 2006).

# Appendix 8.1 *IS-LM* in Japan

The *LM* curve is depicted as upward sloping because as the interest rate rises, the demand for money falls and equilibrium in the money market, which brings the demand for money in line with the available money supply is obtained by an increase in income. The equation for the *LM* curve is where real money supply ($M/P$) equals the demand for money:

$$\frac{M}{P} = h + nY - ei \qquad \textbf{(A8.1.1)}$$

or,

$$i = \frac{n}{e}Y + \frac{h - (M/P)}{e} \qquad \textbf{(A8.1.2)}$$

The model requires non-negative interest rates and income, or, $i \geq 0$, and $Y \geq 0$. For $i \geq 0$, we must have,

$$i = \frac{n}{e}Y + \frac{h - (M/P)}{e} \geq 0 \qquad \textbf{(A8.1.3)}$$

or,

$$Y \geq \frac{(M/P) - h}{n}$$

Similarly, rewrite Eq. (A8.1.1) in terms of the income $Y$ on the left-hand side

$$nY = \frac{M}{P} + ei - h$$

or,

$$Y = \frac{(M/P) - h}{n} + \frac{e}{n}i$$

In order that $Y \geq 0$ we must have

$$\frac{M}{P} - h + ei \geq 0$$

or,

$$i \geq \frac{h - (M/P)}{e} \qquad \textbf{(A8.1.4)}$$

Hence, if $M/P \leq h$, then, a minimum interest rate that satisfies the *LM* model for non-negative values of the variables is given by

$$i_{\min} = \frac{h - (M/P)}{e} \geq 0 \qquad \textbf{(A8.1.5)}$$

Suppose the economy is characterized by the following parameters:

$$\frac{M}{P} > h > 0 \qquad \textbf{(A8.1.6)}$$

Since $M/P > h$ and $i_{\min}$ cannot be less than zero (negative nominal interest rates are ruled out), then it must be the case that $i_{\min} = 0$.

Substitute $i = i_{\min} = 0$ in the equation for the *LM* curve as given by Eq. (A8.1.2)

$$0 = \frac{n}{e}Y + \frac{h - (M/P)}{e}$$

or,

$$Y = \frac{(M/P) - h}{n}$$

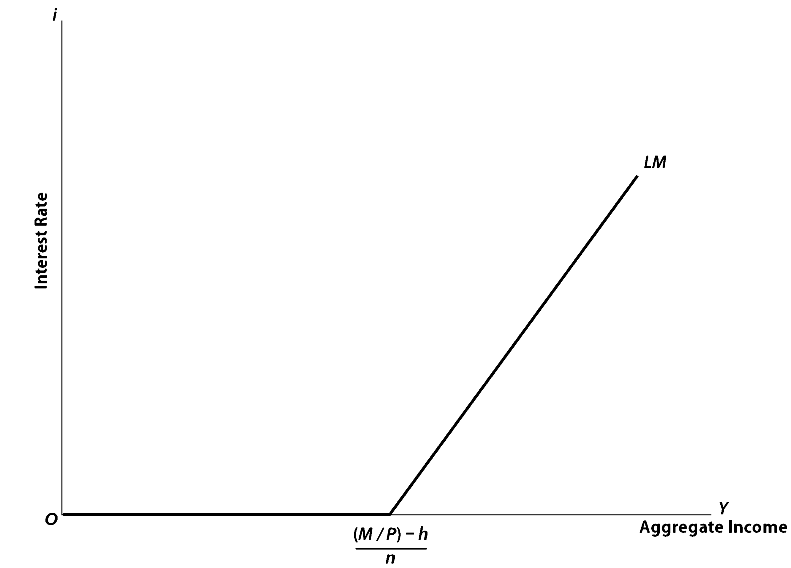
We can then conclude that a nominal interest rate of zero is accompanied by a range of values for $Y$ given by

$$i = i_{\min} = 0 \textbf{ implies that } 0 \leq Y \leq \frac{(M/P) - h}{n} \qquad \textbf{(A8.1.7)}$$

This implies an *LM* curve, as depicted in Figure A8.1.1.

When the *LM* curve is horizontal at a zero-nominal interest rate the economy is said to be in a liquidity trap. Recall that the nominal return on money is zero whereas the return on a bond is inversely related to its price. With a known coupon, say INR 1, on the bond, the return on the bond is given by $i = 1P_B$ where, $P_B$ is the price of the bond. Hence, the interest rate is the return on bonds. When the interest rate is zero, then, the nominal return on bonds ($i = 0$) and on money (0) are identical. With both bonds and money paying the same nominal interest rate—zero—people are indifferent about whether to hold their financial wealth as money or bonds: the liquidity trap.

› **Figure A8.1.1**
The *LM* Curve with a Liquidity Trap. **When the nominal interest rate is zero, individuals hold enough money for transactions purposes, and are willing to hold any increase in money supply as both bonds and money yield the same nominal return of zero. When this occurs, the *LM* curve is horizontal along the income axis and the economy is in a liquidity trap.**

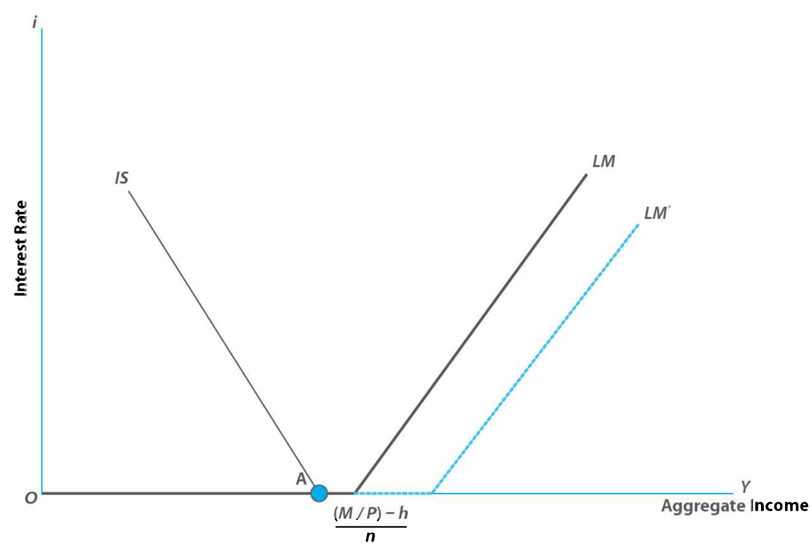


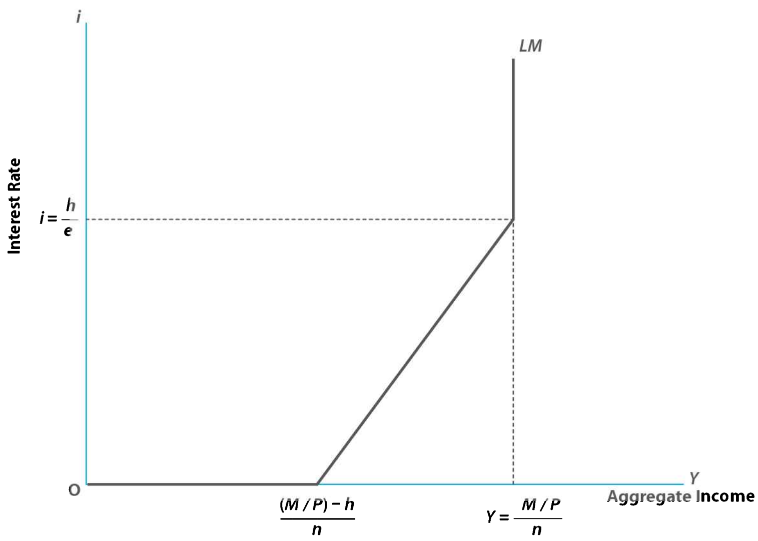

Now, suppose the *IS* and *LM* curves intersect at point A in Figure A8.1.2. If the central bank increases the money supply the *LM* curve shifts out rightwards to *LM′*. The *IS* and *LM* curves still intersect at the same point A, and the expansionary monetary policy has had no effect on output. So a central bank is powerless in a liquidity trap to affect output. This arises because the additional money is readily held by individuals at an unchanged interest rate of zero. There is no sense in reallocating the portfolio to purchase bonds with the increased money balances as the return on bonds is also zero.

Even though the central bank may attempt to cut interest rates (the implication of an expansionary monetary policy) prices declined at a faster rate in Japan (see MacroFocus 8.2)—deflation. A rise in deflation shifts the *IS* curve to the left—in Eq. (8.16) in Section 8.3.2 the term $\pi^e$ declines as deflation rises—and output declines. A combination of a liquidity trap and deflation led the Japanese economy into a recessionary phase in the 1990s.

For completeness it should be indicated that there is also a vertical portion to the *LM* curve. As income increases, there is a demand for cash balances for the purposes of transactions in goods and services. The money supply is fixed and so the increased demand for money can only be met if the velocity of money

› **Figure A8.1.2**
Monetary Policy in a Liquidity Trap. **In a liquidity trap, the expansionary monetary policy that shifts the *LM* curve to the right has no impact on income since individuals willingly hold the extra money and there is no impact on the interest rate. Interest-sensitive private expenditures such as investment expenditures are thus unaffected and aggregate income does not change.**






**› Figure A8.1.3**
General Form of the *LM* Curve. As income increases there is an increase in the demand for money for transactions purposes. This is satisfied by selling bonds. However, beyond an interest rate $i = h/e$ the demand for money is so high that virtually all money balances are being held for transactions purposes alone and money is being circulated in transactions at a very high rate. The *LM* curve is vertical at this point.

increases. Velocity ($V$) is a measure of the speed with which money circulates in transactions involving the purchase and sale of goods and services. Technically, $V = (PY/M)$ where $PY$ is the nominal GDP and $M$ is the nominal money supply. Velocity then is the rate at which the stock of money changes hands to support the transactions involving $PY$, the number of transactions for a given quantity of money. There is logically an upper bound to velocity. This occurs when there is such a great demand for money balances for the purposes of transactions that people transfer all their financial wealth from bonds into money. Bonds are purchased when there is an expectation that bond prices will rise, which occurs as interest rates increase resulting in an expectation that they would eventually decline. (Recall that bond prices and interest rates are inversely related.) As the interest rate increases, the asset or portfolio demand for money declines as individuals seek to purchase bonds and eventually at very high interest rates the asset demand for money goes to zero. The requirement for money as a means of payment, when this occurs, dominates the demand for money as a store of wealth. In that case, $M^d/P = nY$. Then, the asset demand no longer responds to increases in the interest rate as all the demand for money is for transactions. The demand for money that is solely for transactions purposes being $M^d/P = nY$, money market equilibrium is given by

$$\frac{M}{P} = nY$$

or,

$$Y = \frac{M/P}{n}$$

Substituting this into the equation for the *LM* curve

$$\begin{aligned} i &= \frac{n}{e}Y + \frac{h - (M/P)}{e} \\ &= \frac{n}{e}\frac{M/P}{n} + \frac{h}{e} - \frac{(M/P)}{e} \\ &= \frac{h}{e} \end{aligned}$$

Thus, when all money is for transactions purposes only—at $Y = [(M/P)/n]$—the interest rate is $i = h/e$. Beyond this point there is no more money available for transactions purposes and the *LM* curve becomes vertical. The general shape of the *LM* curve then is that depicted in Figure A8.1.3. In the chapter we have omitted the horizontal and vertical portions of the *LM* curve for simplicity.

9

# Aggregate Demand and Aggregate Supply

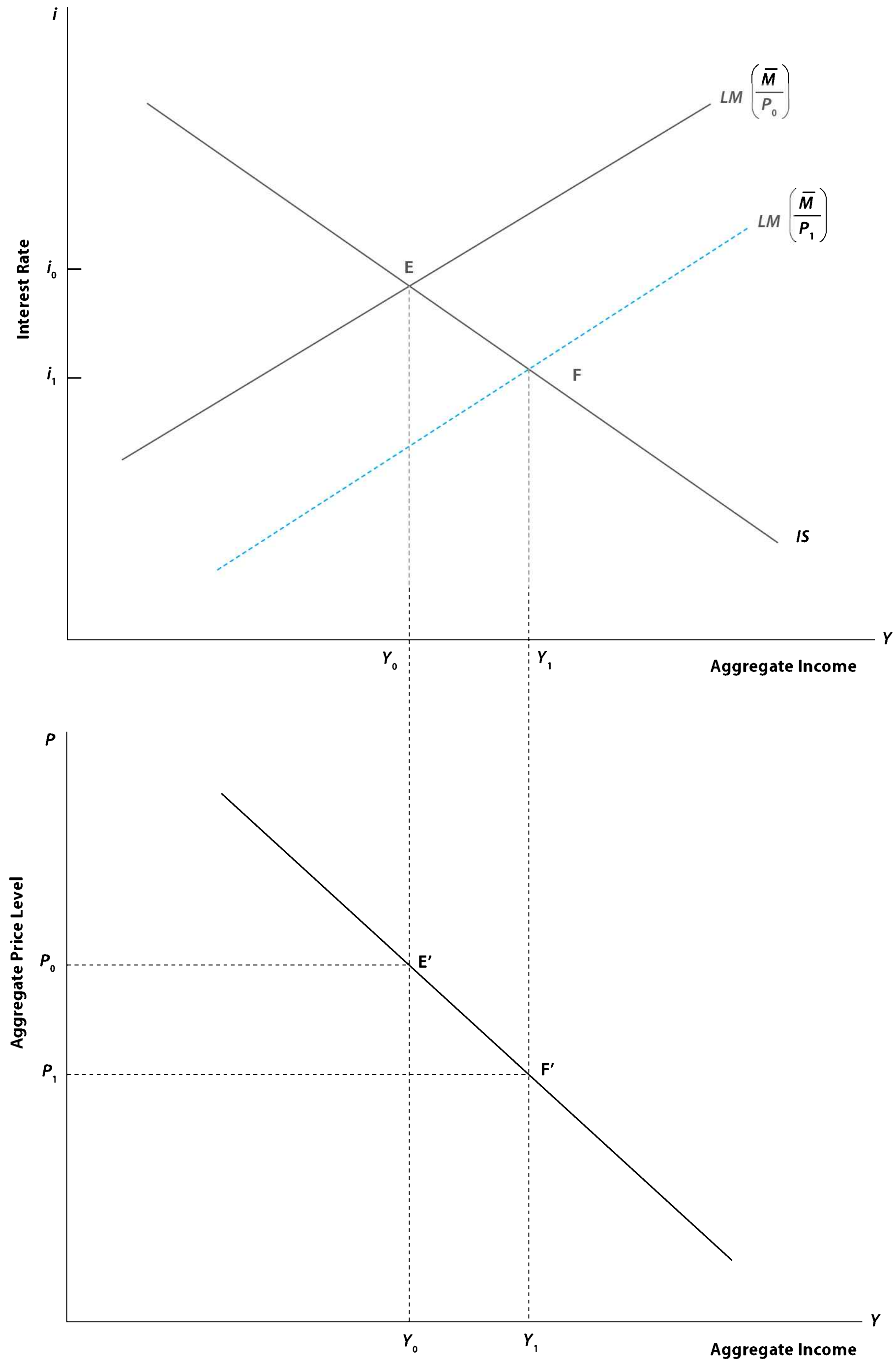
CRITICAL QUESTIONS

- *What is aggregate demand?*
- *How are the equilibrium output and price level determined when the aggregate supply is Keynesian and when they occur as a result of mistaken expectations?*
- *How does the presence of policy lags affect activist (discretionary) policymaking?*
- *Who are real business cycle theorists, and why do they privilege real supply side shocks over demand shocks in understanding business cycles?*

# 9.1 The Aggregate Demand Curve

We use the *IS–LM* model to develop the aggregate demand (*AD*) for the economy as a whole. The *IS–LM* model relates the nominal interest rate to output, whereas *AD* relates the price level to output. The *IS–LM* model is useful for examining the effect of various shocks on the interest rate and on variables such as savings and investment that depend on the interest

**› Figure 9.1**
The Aggregate Demand (*AD*) Curve. A reduction in the price level increases the money supply and shifts the *LM* curve to the right (the upper panel). The reduction in the price is, therefore, associated with an increased aggregate income (the lower panel).



rate. However, when discussing issues related to the price level or inflation and relating that to unemployment, using an *AD* and supply framework is more convenient.

› The aggregate demand (*AD*) curve is the relationship between the real quantity demanded of final goods and services and the general price level.

The AGGREGATE DEMAND CURVE depicts the relation between the aggregate quantity of goods demanded,[1] $C + I + G$, and the price level, $P$. Although the *AD* curve slopes downward like the curve for a single product, the *AD* curve relates the aggregate quantity of output demanded to the general price level, whereas the demand for a single good relates the demand for that good to the price of that good relative to the prices of other goods. Thus, when the prices of all goods increase by 5 per cent, the price level, $P$, also increases by 5 per cent, while the relative prices of goods remain unchanged. Despite unchanged relative prices, the increase in the price level reduces the aggregate quantity of goods demanded.

The reason that an increase in the price level reduces the aggregate quantity of output demanded is illustrated in Figure 9.1. For a given price level, the aggregate output that households, firms, and government demand is where the *IS* and *LM* curves intersect.

Suppose that the nominal money supply is $\bar{M}$, and the initial price level is $P_0$. Then the real money supply is $(\bar{M}/P_0)$, and the initial *LM* curve is $LM\ (\bar{M}/P_0)$. The *IS* and $LM\ (\bar{M}/P_0)$ curves intersect at point E where the amount of output demanded is $Y_0$. If the price level decreases to $P_1$, the real money supply shifts the *LM* curve to the right (see Figure 8.10 of Chapter 8). The *IS* and $LM\ (\bar{M}/P_1)$ curves now intersect at point F where the aggregate quantity of output demanded is $Y_1$. Thus, the decrease in the price level from $P_0$ to $P_1$ increases the aggregate quantity of output demanded from $Y_0$ to $Y_1$. This negative relation between the price level and aggregate output demanded is shown as the downward-sloping *AD* CURVE in the lower panel of Figure 9.1. (The equation of the *AD* curve is derived in Appendix I.) Points E′ and F′ in the lower panel of Figure 9.1 correspond to points E and F in the upper panel of the same figure, respectively.

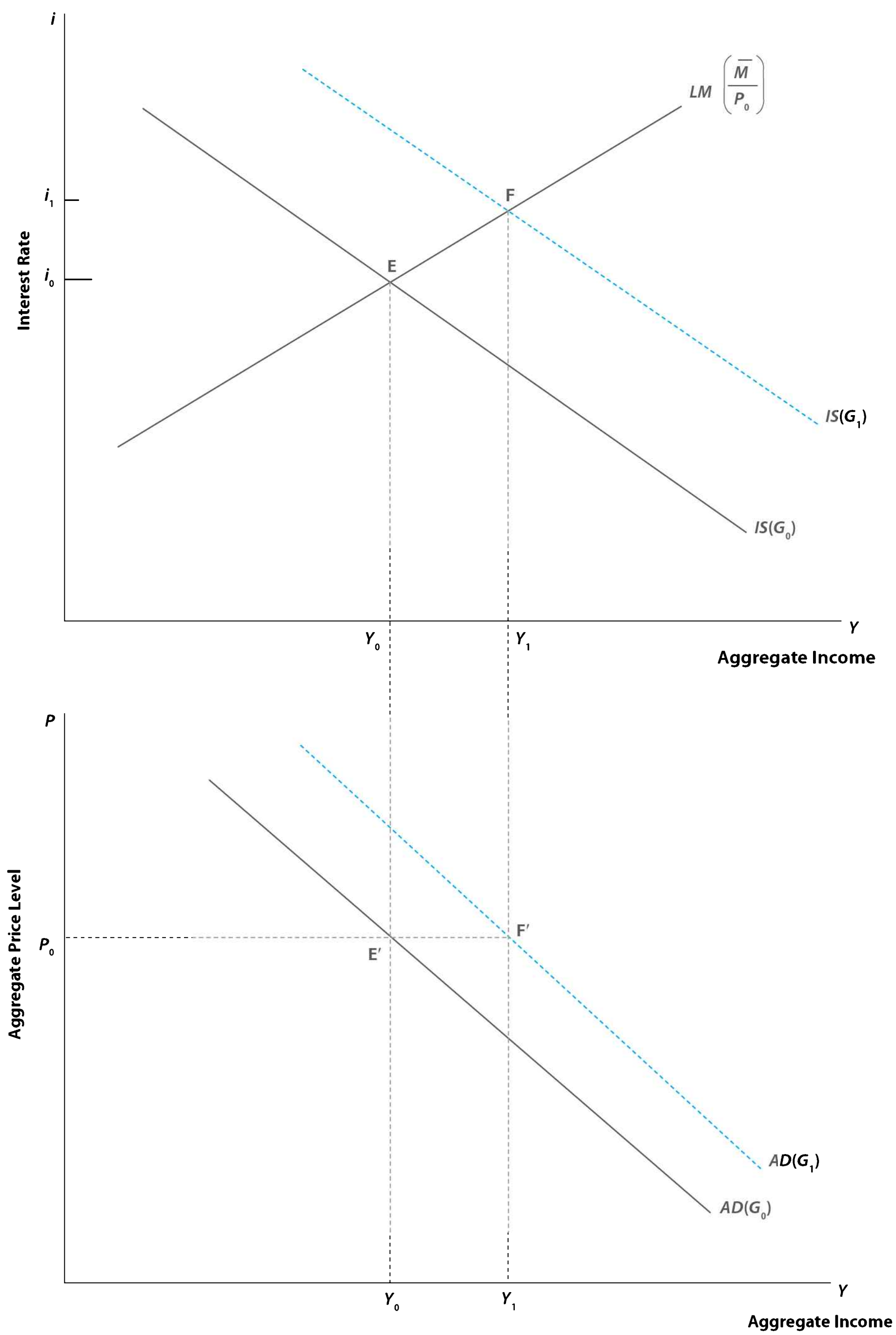
› The *AD* curve relates a given price level to the demand for output and is derived from the simultaneous equilibrium in the goods (*IS*) and money (*LM*) markets.

Why is the *AD* curve downward sloping? A reduction in the price level when income is $Y_0$ increases the supply of real money balances. As the demand for real money balances is still associated with income $Y_0$, the discrepancy between money supply and demand is adjusted by agents purchasing bonds with the excess money balances, which causes the nominal rate of interest to decline at any given level of real income. At the same time, a reduction in the nominal rate of interest induces a rise in investment and consumption expenditure along the *IS* curve. Consumption and investment demand are, therefore, higher at points F and F′, and this increased demand is associated with a lower price level.

What causes the *AD* curve to shift? As *AD* is determined by the intersection of *IS* and *LM*, this implies that holding the price level constant, any factor that causes the interaction of the *IS* curve and *LM* curve to shift to the right will raise *AD* and shift the *AD* curve to the right. In a similar vein, for a constant price level, any factor that causes the intersection of the *IS* and *LM* curves to shift to the left SHIFTS THE *AD* CURVE as well to the left. A factor that shifts the *IS* curve to the right is an increase in government expenditure (see Figure 8.3 in Chapter 8).

› Changes in any variables other than the price level, which shift the *IS* and *LM* curves, such as the fiscal and monetary policy, consumption spending, or investment expenditure, shift the *AD* curve.

This effect of the increased government expenditures on the *AD* curve is illustrated in Figure 9.2. The increase in the government expenditure from $G_0$ to $G_1$ with the price level held constant at $P_0$ results in the intersection of the *IS* and *LM* curves at point F where the aggregate output demanded has increased to $Y_1$. As the increase in government purchase

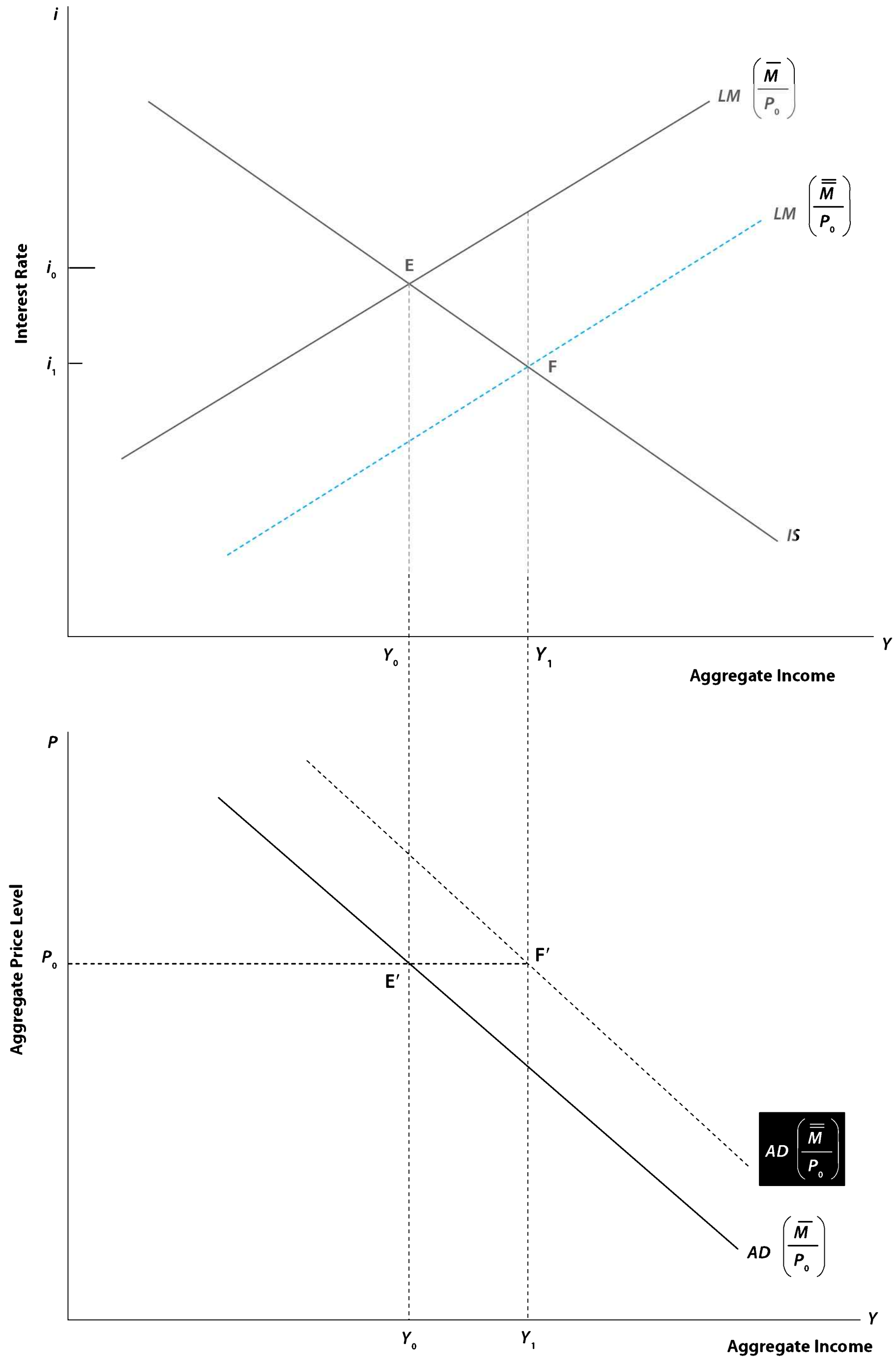



› **Figure 9.2**
An increase in Government Expenditure. **An increase in the government expenditure at a given price level shifts the *IS* curve out to the right and increases the aggregate income and output (upper panel). At the given price level, the increased government expenditure shifts the aggregate demand (*AD*) curve to the right (lower panel).**

raises the output demanded at any price level, the entire *AD* curve shifts to the right to $AD\,(G_1)$.

If the central bank increases the nominal money supply to $\overline{\overline{M}}$, this is associated with a rightward shift of the *LM* curve (see Figure 8.10 of Chapter 8). The new *IS*–*LM* equilibrium, with the price level constant at $P_0$, is at the point F in Figure 9.3. The *AD* curve at $P_0$ and at any price level shifts to the right in response to the increased nominal money supply. With Figures 8.3–8.6 and 8.10 in Chapter 8 as references, we summarize the factors that shift the *AD* curve in Table 9.1.

**› Figure 9.3**
An increase in Nominal Money Supply. An increase in money supply by the central bank at a given price level shifts the *LM* curve to the right and increases the aggregate income and output (upper panel). At the given price level, the increase in money supply shifts the *AD* curve to the right (lower panel).




**› Table 9.1**
Factors that Shift the *AD* Curve

| Factor | Direction of Shift |
|---|---|
| 1. An Increase in Government Expenditures | *IS* and *AD* Shift to the Right |
| 2. An Increase in Taxes | *IS* and *AD* Shift to the Left |
| 3. An Increase in the Productivity of Capital | *IS* and *AD* Shift to the Right |
| 4. An Increase in Expected Inflation | *IS* and *AD* Shift to the Right |
| 5. An Increase in Expected Future Output (Rise in Permanent Income) | *IS* and *AD* Shift to the Right |
| 6. An Increase in Nominal Money Supply | *LM* and *AD* Shift to the Right |
| 7. A Reduction in Real Demand for Money due to Increased Use of Credit Cards[2] | *LM* and *AD* Shift to the Right |

# 9.2 *AD–AS* Equilibrium in the Keynesian Case

› The aggregate supply (*AS*) curve is the relationship between the real quantity of final goods and the services produced in an economy and the general price level.

As we saw in the chapter on labour markets, the AGGREGATE SUPPLY CURVE derives from the workers and firms pursuing their rational self-interest. Workers seek to maximize their preference for consumption and leisure, whereas firms seek to maximize profits. In the Keynesian set-up, money wages are inflexible or sticky at $\overline{w}$. There are many possible reasons for this.

One possibility is that there exist minimum-wages laws that result in an artificial floor to the wages that firms can pay. Another possibility is that organized groups of workers seek to keep their nominal wages at levels they think to be appropriate relative to other occupations: workers seek to maintain their position in the job ladder hierarchy. Yet another possibility is that workers and firms adopt implicit contracts. This is a tacit agreement that firms will not reduce workers' wages during an economic downturn and neither will workers demand a raise in wages when business conditions improve. This is an insurance scheme with workers paying an insurance premium in the form of lower than feasible wages in good times in return for insurance coverage in the form of higher than feasible wages in bad times. Keynesians who argue that nominal wages are inflexible offer one of these rationales in support of their position. When wages are sticky, we derived the KEYNESIAN AGGREGATE SUPPLY CURVE in Chapter 7, Figure 7.10. This *AS* curve is upward sloping as long as there is an excess supply of labour $N^D < N^S$. Once $N^D = N^S$, there is full employment, and the *AS* curve becomes vertical.

› The **Keynesian *AS* curve** in the short run is drawn for a fixed money wage and is upward sloping up to the point where full employment occurs, where it is vertical.

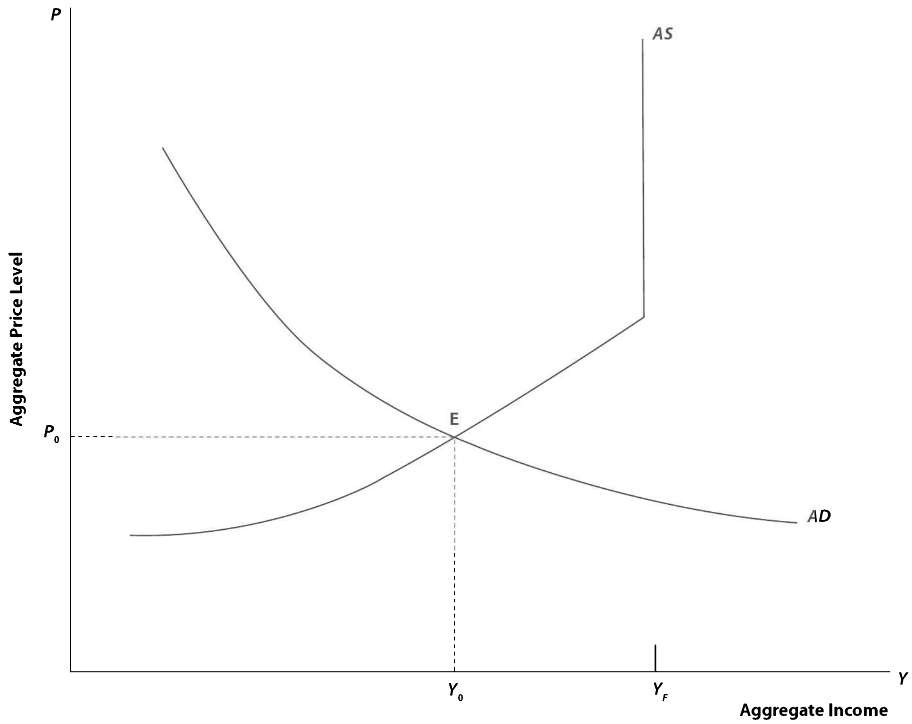
The *AD* and the Keynesian *AS* schedules are depicted in a single diagram now illustrating the market for real output (see Figure 9.4). The equilibrium price level and real output are at the intersection of the *AD* and *AS* schedules at point E. As point E is on the *AD* schedule, the economy is at a point where *IS* equals *LM*. Thus, at price level $P_0$ in Figure 9.4, aggregate expenditure is equal to output (*IS* equilibrium), and the quantity of real money balances demanded equals the real value of nominal money balances supplied (*LM* equilibrium). As point E is on the aggregated supply schedule, workers and firms in pursuit of their rational self-interest produce the level of output $Y_0$ at the price level $P_0$. Of course, point E results from the given stationary levels of all those variables that result in the given position of the *AD* curve as summarized in Table 9.1. A change in any of the factors listed in Table 9.1, such as government expenditure or in the nominal money stock, will shift the *AD* curve (Figure 9.4) and result in a different equilibrium in the market for real output. Note also that *AD* does not intersect the *AS* curve on its vertical portion. Point E is, thus, an equilibrium where unemployment prevails.

The complete Keynesian system may accordingly be written[3] as follows:

$$\left.\begin{array}{r l}
\textbf{\textit{IS} Curve:} & Y = C(\underset{(+)}{Y-T}, \underset{(-)}{i-\pi^e}) + I(\underset{(-)}{i-\pi^e}) + G \\
\textbf{\textit{LM} Curve:} & \dfrac{M^s}{P} = L(\underset{(+)}{Y}, \underset{(-)}{i}) \\
\textbf{Labour Supply:} & \dfrac{W}{P} = MRS_{c,l}
\end{array}\right\} \textit{AD} \text{ Curve}$$

$$\left.\begin{array}{r l}
\textbf{Labour Demand:} & \dfrac{W}{P} = MP_N \\
\textbf{Output Supply:} & Y = F(\overline{K}, N),\ N = \min[N^D, N^S] \\
\textbf{Nominal Wage Ridigity:} & w = \overline{w}
\end{array}\right\} \textit{AS} \text{ Curve}$$

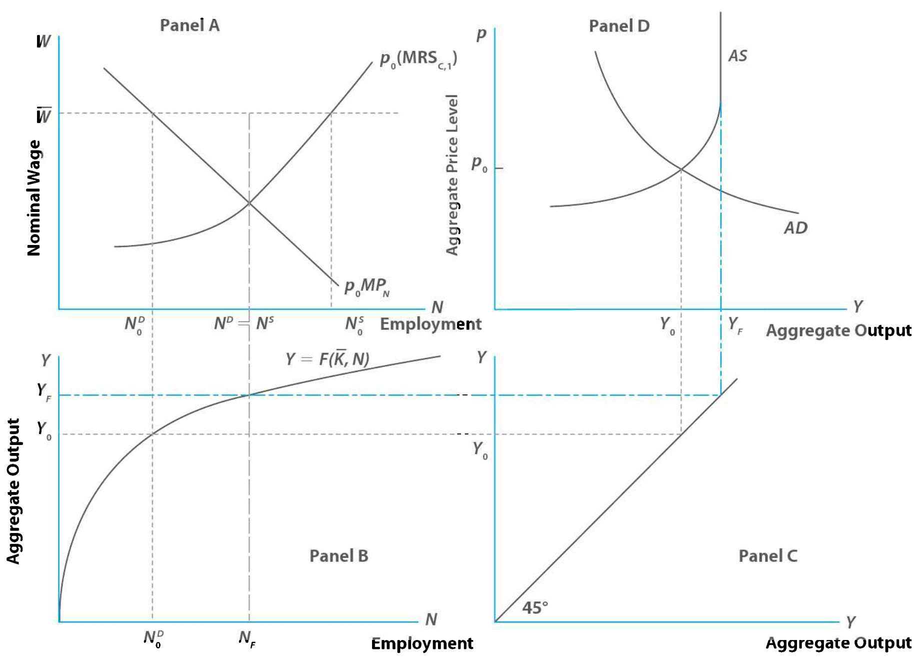
**› Figure 9.4**
The Market for Real Output (Keynesian). **The *AD* curve representing equilibrium in the goods and money markets intersects the Keynesian Aggregate Supply (*AS*) curve at point E, which determines the equilibrium in the goods, money, and labour markets, and the aggregate price level and output/income.**




These six equations together determine the six macroeconomic variables of interest—$w$, $N^D$, $N^S$, $i$, $P$, and $Y$. This represents a complete macroeconomic description of the closed economy. In diagrammatic terms, the solution to the equations above is depicted in Figure 9.5.[4] As depicted, the equilibrium outcome results in unemployment as $N_0^D < N_0^S$. $(N_0^S - N_0^D)$ individuals are willing to work at the going wages but there are no job offers forthcoming, a situation referred to as INVOLUNTARY UNEMPLOYMENT. What can be done in such a situation?

› Involuntary unemployment refers to a situation where the workers are willing to work at the going wage and yet firms do not employ them.

**› Figure 9.5**
Aggregate Demand and Supply in the Keynesian System. **The nominal wage results in the demand for labour that is less than the supply of labour (panel A). Employment is determined by the demand side of the market and the level of employment determines the aggregate output produced (panel B). The price level at which the firm determined its demand for labour is associated with the output produced as a point on the *AS* curve in panel D. The *AD* curve intersects the *AS* curve on its upward-sloping part in panel D, resulting in involuntary unemployment.**



**MACROFOCUS 9.1**

## Fiscal Policy and Working Capital

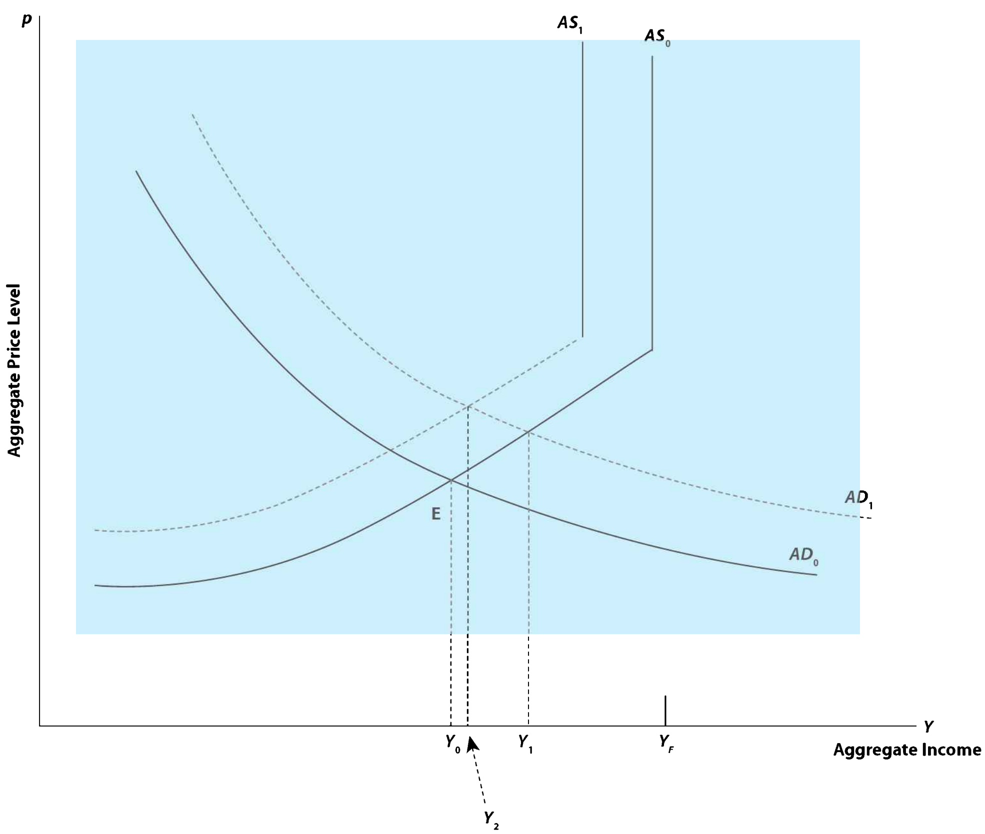
In developing countries, working capital finance is crucial for doing business. It is required to make payments in advance for the services of inputs once a production process is initiated. Suppliers of inputs to firms are often financially constrained and cannot wait for payments to be made after the production cycle has been completed.*

If inputs uniformly precede outputs by a period of $T$ units of time, labour services and capital stock that are introduced into the production process at time $t$ will account for a finished output at time $t + T$. Workers then have to be paid in advance of the sale of output and working capital expenses involve the borrowing costs at time $t$ when production is initiated. These working capital funds are tied up till the firm receives a cash flow when the finished output is sold. In such a situation, the firm hires labour up to the point where the cost or the wages equal the present value of the benefit—the discounted revenue received from the output sold by the labour hired, or, $pMP_N/(1 + i)^T$. (Recall from Chapter 7 (Section 7.1) that when revenues and expenditures are synchronized, the demand for labour is described by the present cost equals present benefit condition, $w = pMP_N$.) If the interest rate rises, the present value of the benefit of hiring labour declines and firms reduce their demand for labour. If the price of output increases, the present value of the benefit of hiring labour rises and firms increase their demand for labour. With employment determined by the demand for labour (the short side of the market), the $AS$ of output will accordingly increase with the price, but will decrease with a rise in the interest rate.

The $AS$ curve is upward sloping but shifts to the left with a rise in the interest rate as this increases the cost of working capital required to hire inputs like labour. Similarly, a decrease in interest rates shifts the $AS$ curve out to the right as the cost of financing working capital inputs decreases. Figure 1 illustrates the impact of fiscal policy when the $AS$ responds to the interest cost of working capital finance.

An expansion in government spending that shifts the $AD$ curve to $AD_1$ initially increases output to $Y_1$ along the aggregate supply curve $AS_0$. Fiscal policy, however, involves some crowding out, and the interest rate rises. This makes working capital costly, and the $AS$ curve shifts leftwards to $AS_1$. Output then decreases to $Y_2$. If the interest sensitivity of $AS$ is strong, $Y_2$ can even be to the left of $Y_0$, and increased government expenditure can contract the economy. Higher interest rates due to fiscal policy in addition to crowding out private investment on the demand side, also make working capital more costly and reduce $AS$.** In economies with poor infrastructure and underdeveloped financial systems, the production and distribution of goods and services takes time requiring firms to make large advances for working capital. Expansionary fiscal policy has a cost-push effect that causes firms in these countries to reduce output.

**› Figure 1**
Fiscal Policy and the Working Capital.




* L. Taylor, *Structuralist Macroeconomics—Applicable Models for the Third World* (New York: Basic Books, 1983).
** D. R. Shaller, "Working Capital Finance Considerations in National Income Theory," *American Economic Review* 73, no. 1 (1983): 156–165.

Keynesians are of the view that the low *AD* is due to low confidence of investors regarding the future prospects of the economy.[5] Investment demand is thought by the Keynesians to be unresponsive to the rate of interest, and the *IS* curve is considered to be steep,[6] implying that crowding out of private expenditures would be low (see Figure 8.16 in Chapter 8). The Keynesians accordingly advocate a FISCAL POLICY to reduce the existing involuntary unemployment. Keynes had taken this position after considering the facts of the Great Depression when the stock markets declined by 85 per cent between September 1929 and June 1932. Between 1929 and 1933, the GNP in the US economy fell by nearly 30 per cent and the unemployment rate rose from 3 to 25 per cent. The net investment was negative from 1931 to 1935.

› Keynesians advocate the use of **fiscal policy** to reduce the involuntary unemployment as they believe that crowding out of private investment expenditures is low.

The view that the Great Depression was due to a leftward shift of the *AD* curve is accepted by Keynesians and Monetarists alike. However, whereas Keynesians emphasize a leftward shift in the *IS* curve due to a collapse of investor confidence, the Monetarists emphasize a leftward shift of the *LM* curve. With the decline in the interest rates, depositors' confidence in the economy also declined, and they rushed to withdraw cash from their bank accounts. With everyone rushing to withdraw money before other depositors, banks were unable to meet depositor demand for cash and many banks failed. The increase in the desired currency-deposit ratio resulted in a reduction in the money multiplier,[7] which reduced the money stock. The money stock declined by 26.5 per cent between 1929 and 1933. Friedman and Schwartz fault the Federal Reserve Board in the United States for failing to increase the money supply and offsetting the decline in *AD*. Monetarists argue that the *LM* curve is steep and that an increase in the money supply is the effective policy in shifting the *AD* schedule to the right.

## 9.3 *AD* and the Mistaken Expectations Aggregate Supply

Besides the disagreement regarding demand management—whether to use fiscal or monetary policy—Friedman and Lucas also disagree about the appropriate construction of the *AS* curve. As we saw in the chapter on the labour market, Friedman argued that with workers imperfectly informed about the price level, they supply labour on the basis of an expected price and, hence, an expected real wage. When EXPECTATIONAL ERRORS occur, the *AS* curve for a given expected price $AS\,(p^e \neq p)$ is upward sloping. When expectations are realized, $p = p^e$, the natural rate of employment and output is achieved with a vertical aggregate supply $AS\,(p = p^e)$ (see Figure 7.11 in Chapter 7).

› The mistaken expectations *AS* curve is upward sloping when the expected price diverges from the actual price. When expectations are realized, the natural rate of output is achieved with a vertical *AS* curve.

In the 1960s, nominal wages in most OECD countries showed great flexibility in the upward direction, and this made it difficult to maintain an analysis of unemployment based on sticky wages. Moreover, Friedman argued that if there is unemployment, those who are unemployed should be going around knocking on the doors of firms and expressing a willingness to work at a lower wage, which would still be better than the zero wages associated with unemployment. In this way, money wages would move to equilibrate the labour market. The unemployment that exists in the Friedman set-up is due to mistaken expectations, and as individuals revise their expectations in the light of errors in past expectations—an adaptive process—they reach the natural rate of output ($Y^*$).

When expectations are wrong, however, the market can equilibrate at a different level of employment from the natural rate temporarily. In this Friedman mistaken expectations framework suppose that *AD* intersects the

*AS* curve when price expectations are $p_0^e$ at point M to the left of the natural rate of output $Y^*$ (see Figure 9.6, panel D). At point M, workers have expected a price higher than the realized price, $p_0^e > p_1$, and they realize a nominal wage $w_1$ (depicted in panel A) that is lower than what would have been realized ($w_0$) if expectations had been correct. With mistaken expectations, employment and output are lower than the natural rate of employment and output, $Y_1 < Y^*$.

As workers adapt to this outcome, they will revise their expectation of prices downwards, and the supply of labour will shift down. This results in pushing the money wages down as well. Initially, the downward revision of price expectations from $p_0^e$ to, say $p_i^e$, where $p_i^e < p_0^e$, will result in the operative labour supply function $p_i^e\,(MRS_{C,l})$, which is not depicted in panel A so as not to clutter up the diagram. The corresponding mistaken expectation aggregate supply function $AS\,(p_i^e \neq p_i)$ is shown in Figure 9.6, panel D. However, workers are still expecting a price level above the one that actually obtains at the intersection of $AS\,(p_i^e \neq p_i)$ and the *AD* curve in panel D. There is also a continued presence in the labour market of individuals who are volunteering to work at lower wages. The revision of prices will, therefore, continue until price expectations are again consonant with experienced prices, $p_n^e = p_n$, and the operative mistaken expectations *AS* curve is $AS\,(p_n^e \neq p_n)$. The adaptive expectations mechanism restores employment to its natural rate, and there is NO NEED FOR GOVERNMENT INTERVENTION in the form of *AD* management. If, in this Monetarist set-up, the government manipulated *AD* and shifted the *AD* curve to the right, it would push the price level up, exactly at a time when workers are revising their expectation of prices in the reverse direction, that is, downwards. An expansionary demand policy would cause workers to revise their price expectations upwards and shift the mistaken expectation *AS* curve up, resulting in an offset to the direction set by *AD* policy. Demand management is counterproductive in this framework. Friedman's macroeconomic system can be summarized as follows:

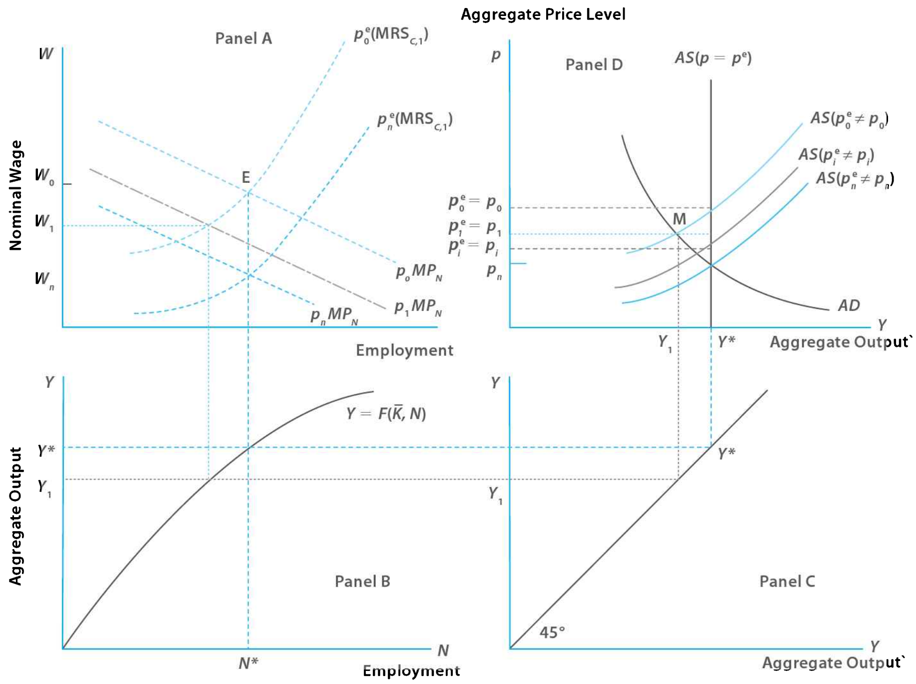
› In the Mistaken Expectations set-up, *AD* management could end up destabilizing the economy. Given enough time, the economy tends to self-correct and return to the natural rate of employment and output without changes in the monetary and fiscal policy.

$$\left.\begin{array}{r}\textit{IS}\text{ Curve}: Y = C(\underset{(+)}{Y - T}, \underset{(-)}{i - \pi^e}) + I(\underset{(-)}{i - \pi^e}) + G \\ \textit{LM}\text{ Curve}: \dfrac{M^s}{P} = L(\underset{(+)}{Y}, \underset{(-)}{i})\end{array}\right\}\ \textit{AD}\text{ Curve}$$

$$\left.\begin{array}{r}\text{Labour Supply}: \dfrac{W}{P^e} = MRS_{c,l} \\ \text{Labour Demand}: \dfrac{W}{P} = MP_N \\ \text{Output Supply}: Y = F(\bar{K}, N),\ N^D = N^S \\ \text{Price Expectations}: P_{t+1}^e = P_t^e + \lambda(P_t - P_t^e)\end{array}\right\}\ \textit{AS}\text{ Curve}$$

The essence of the Keynesian–Monetarist debate, then, is not so much about whether fiscal policy is more appropriate than monetary policy. Rather, it is about whether governments should intervene by manipulating *AD* judiciously to achieve full employment (see Figure 9.5), or should market forces be relied on to bring about eventually the necessary adjustment of the real and nominal wages (Figure 9.6). Should states intervene or should markets be left to adjust and solve the problem of unemployment?

Macroeconomists prefer state intervention if their prior belief is that money wages are sticky and respond only sluggishly, or with large inertia to the excess supply in the labour market. In such a situation, it is expedient to manipulate demand. Otherwise, if markets self-correct with reasonable accuracy and speed, it is expedient to allow markets to solve economic problems such as employment.




**› Figure 9.6**
Mistaken Expectations and *AD–AS*. The *AD* equals the mistaken expectation aggregate supply at point M in panel D when workers expect prices to be $p_0^e$. Tracing out through panels C and B to panel A, the realized nominal wages at $w_1$ are lower than what was expected ($w_0$), and, in panel D, the expected price is higher than that realized, $p_0^e > p_1$. This makes workers revise their expectations regarding prices downward until the expectations are realized at $p_n^e = p_n$.

# 9.4 Policy Lags

Friedman went even further to claim that there are limits to *AD* policy because of the presence of lags. If *AD* results in an economic situation, such as that depicted in Figure 9.5, where full employment does not prevail, it must first be verified whether the current position of the *AD* curve is due to some *temporary* factor, or whether the position of the curve is *permanent*.

For instance, if the *AD* curve takes its position because of a temporary shock that results in a one-period reduction in consumption spending, it is wise not to initiate any activist demand management policy because, even though income in this period is affected, it will revert back the next period and not have any permanent effect.

› **Policy lags** can make attempts to fine-tune the economy and reduce the impact of business cycles counterproductive.

As policy operates with a lag, it will start having effects on an economy subject to temporary shocks when the *AD* schedule approaches the vertical portion of the *AS* curve, and it will tend to generate inflation. Friedman[8] argued that the LAGS IN MONETARY POLICY are so long that they aggravate cyclical fluctuations. After studying the behaviour of the US economy and money supply since 1880, he concluded that a long and variable lag exists between the high and low points in the rate of change in the money supply and the peaks and troughs in overall business activity. The money supply peaked an average of 16 months before the overall economy, and troughed an average of 12 months before business activity in general did the same. Friedman[9] observed before the Joint Economic Committee of the 86th Congress that there have been a large number of occasions "on which major fluctuations have been a consequence of changing and, at times, erratic governmental policies with respect to money."

On this basis, he urged to keep monetary changes from being the costly, destabilizing force they have been through much of history. Friedman identified the inside and outside lags, which are important to consider for activist policymakers. INSIDE LAGS in macroeconomic policy refer to the delay between the time a policy change is needed and the time the policy change is implemented. The inside lag comprises three types of lags:

› Inside lags refer to the delay between the time a policy change is needed and the time the policy change is implemented.

1. Data lag
2. Recognition lag
3. Implementation lag

The data lag involves the time taken before the data that indicates the state of the economy becomes available. For instance, GDP data is available with a time lag. Quick estimates are revised and eventually become available as actual figures, and this process can take a couple of years.

After the data is available, a recognition lag comes into effect. This is the time period policymakers take to recognize that a disturbance has occurred, which requires action on their part in terms of stimulating or restraining aggregate spending. Kareken and Solow[10] showed that, on average, the recognition lag is approximately 5 months. This lag was found to be a bit longer when restrictive policy was necessary and a bit shorter when expansionary policy was required.

The implementation lag is the delay between the recognition of the need for action and the policy decision. Here, monetary policy has a much shorter decision lag. Even though there are quarterly monetary policy reviews, central bank officials can and do meet on a daily basis to discuss and decide on policy. Fiscal policy that requires changes in government spending and tax programmes has to be passed in the parliament. This occurs mainly during its budget session. Thus, many months can pass before fiscal policy changes are put into effect. The inside lag for monetary policy is, therefore, substantially shorter than the inside lag for fiscal policy.

After implementation, a policy is subject to an OUTSIDE LAG—the time required for a change in policy to affect the economy, or, an effectuation lag.

› The outside lag is the time required for a policy change to affect the economy.

Fiscal policy, despite having a larger inside lag than monetary policy, has a shorter outside lag. Changes in government spending have an immediate impact on the real output by adding to or subtracting from the *AD*, and individuals respond to tax changes by varying consumption expenditures immediately. Monetary policy, by affecting the money supply and interest rate, affects investment expenditure, which responds slowly to such a change. This is because the interest rate is just one of the many factors that influence the decision to order new capital equipment, and it takes time for the capital goods sector to begin to produce new items on its order books.

Even though monetary policy may face shorter implementation delays, its initial impact could be small and the significant impact of the policy change on output and income builds up only gradually over a few quarters. There are six main channels through which monetary policy, which changes interest rates, affects economic activity:

1. *Intertemporal substitution*: Interest rates represent the price of expenditure in the present relative to the future. If the substitution effect is small, private expenditures do not respond strongly to changes in interest rates.
2. *Pass-through effects of interest rates*: While the central bank is able to affect rapidly the short-term money market interest rates, there are delays in the pass through of changes in the overnight cash rate to other interest rates such as the deposit and lending rates of financial intermediaries. As the interest rates of financial intermediaries are important determinants of cash flow, asset prices, and the incentive to postpone expenditure, a slow pass through contributes to a transmission lag from the cash rate to economic activity.

3. *Cash-flow effects on liquidity constrained borrowers*: An important source of lags arises from the gradual response of investment—business investment as well as household investment in durables and housing—to changes in monetary policy. Here, adjustment costs associated with changing the level of the relevant capital stock are responsible for the lag. However, changes in interest rates also affect the incentive to postpone investment when returns are uncertain. The largely irreversible nature of many investments means that there is an option value to waiting to invest in a world of uncertainty.[11] When a firm or individual makes an irreversible investment, this option is exercised eliminating the possibility of waiting for the arrival of new information that might have affected the timing or the desirability of the investment. A change in interest rates affects this option value, and will, therefore, affect the timing of the investment.
4. *The effect of induced changes in the exchange rate on the tradeable sector*: One would expect the full implications of a change in monetary policy to be incorporated into asset prices as soon as the change becomes apparent. This does not appear to be the case for the exchange rate. It has been found, for instance, that contractionary monetary policy that raised the interest rate leads to a prolonged gradual appreciation of the exchange rate with the maximal appreciation occurring after 2 to 3.5 years.[12] A consequence of this is that exchange rate effects on the tradable sector of the economy are gradual and prolonged.
5. *Credit supply effects*: As lenders do not have perfect knowledge of borrowers,[13] they build risk premia into their interest rates. When the central bank raises the cash rate, these risk premia may also rise altering the supply of credit and, thereby, influencing investment. Financial intermediaries may also, instead of raising interest rates to reflect the higher risk, impose some form of rationing by raising the loan-to-valuation ratio or simply requiring more collateral. They might even just not lend to riskier customers. If interest rates rise, some of the most creditworthy borrowers may not go ahead with their projects, while borrowers who have a greater readiness to default on their borrowing will continue to want to borrow—banks then respond to higher cash rates by rationing credit so as to ensure that their average default rates remain low.
6. *Wealth/Asset effects*: Asset prices often have a life of their own, separate from the influence of monetary policy. They are, thus, a channel of monetary policy only to the extent that they have a relatively predictable relationship with interest rates. In theory, there is a link between cash rates and bond prices, and this, in turn, affects equity prices. Equity prices, in turn, feed into the cost of capital and asset values affect collateral and the readiness of banks to lend. These relationships often get submerged when there are more powerful forces driving asset prices, such as financial and industrial deregulation, tax and tariff reforms, which significantly dampen inflationary expectations, and rampant animal spirits.

In deregulated financial markets, the central bank can influence liquidity in the payments clearing system and the cash rates at the short end of the yield curve is the operating instrument to influence activity and prices. Good policymaking requires an appreciation of the dynamic relationship between the monetary policy instrument—the overnight cash rate—and the final objectives of policy—inflation and output.

The transmission channels referred to above imply that the influence of changes in interest rates on economic activity requires policymakers to act with considerable skill in interpreting the transmission process. Otherwise, their attempts at correcting fluctuations in economic activity might lead to further problems that set off a chain of corrective actions. Since the stabilization policy has potential destabilizing effects on the economy due to the presence of lags, some macroeconomists like Friedman advocated that policy should not respond to fluctuations. With no role for countercyclical policy, Friedman argued that money supply should rather grow at a constant rate.[14] Advocates of stabilization policy, on the other hand, argue that lags are unimportant and that the government should maintain a high and stable level of *AD*.

## 9.5 Real Business Cycles

The repeated sequence of economic expansion giving way to a decline in economic activity followed by recovery is known as the business cycle. When an economy goes through a downturn, many sectors experience declining sales and production, and workers are laid off or forced to work only part time. Should policymakers respond to such cyclical fluctuations?

The Keynesians argue that because wages and prices adjust slowly, disturbances in spending or production can drive the economy away from a better utilization of productive resources for long periods of time. The Keynesian view is that due to this sluggishness in market response, the government should intervene to smooth business cycle fluctuations. The Monetarists, on the other hand, do not see much need for government intervention to counteract economic fluctuations, and have greater faith in the equilibrating power of market forces.

MACROFOCUS 9.2

### Nominal Anchors and Intertemporal Policy Constraints

The power of producing good or bad economic effects is often attributed to policymakers. Yet, these policymakers are distinct: fiscal policy is the responsibility of the Ministry of Finance, and monetary policy is under the jurisdiction of the central bank (the Reserve Bank of India). However, they are connected through the arithmetic of the government budget constraint. In each period, the government must finance its expenditure plans. To do this, it can use taxation, the sale of new government bonds, or printing money. When it sells bonds, the government accumulates public debt that is the stock of government bonds sold to the private sector. It must pay interest on this debt so that the government budget constraint is (see Chapter 6) as follows:

Government expenditure + Interest on public debt = Tax revenues net of transfers + New bonds + New money

where the money referred to is base money. What is important for the economy, therefore, is the net impact of monetary and fiscal policy—not the impact of either in isolation, as has been our focus. Suppose the Finance Ministry, for instance, increases its spending by issuing bonds to the central bank and spending the newly issued money. If this rightward shift of the *AD* curve kicks up inflation, the central bank may need to offset this and pull in the opposite direction by engaging in restrictive monetary policy. However, if the growth of currency is controlled with a view to stabilizing the price level, other elements of the budget constraint must be adjusted. Either government expenditures must be scaled down, or tax collections or the net proceeds from new issues of interest-bearing debt must be stepped up to compensate for the loss of a source of revenue.

As a result, printing new money to finance government expenditure is ruled out by the constitution of the central bank in many countries. This denial of money financing of the government expenditures allows monetary policy to provide the nominal anchor for the economy. If agents in the economy are confident that the central bank will not change its announced policy about the target inflation rate, their expectations of future inflation are unaffected by any temporary changes in the inflation rate. This is described as a case of anchored inflationary expectations, where people believe that the central bank will act to ensure that any temporary inflation will be reversed, and so do not change their expectation of the future inflation even though the current inflation has risen. When expectations do not change, workers, for instance, will not press for wage increases, and firms will not be pushed to raise prices. As a result, future inflation will be controlled, and the short-run *AS* curves will not drift in response to inflation, allowing output to return to the natural rate more quickly.

The fiscal authorities must somehow manage their affairs without seigniorage so that the fiscal deficit (government expenditure + interest on debt – net tax revenues) equals the net proceeds from new debt. This would require that the government has a fiscal policy satisfying the condition that the present value of the government's net of interest surplus equals the current stock of interest-bearing government debt (see Chapter 12).

The government must balance any current deficits with the prospect of future surpluses sufficient to service the debt created by those deficits. These long-run effects of fiscal and monetary policy are distinct from their short-run impact on *AD* where both can be used to expand or contract *AD* in the short run.

There are three types of impulses or shocks that result in economic fluctuations. One is policy shocks that follow the decisions made by policymakers and that affect the demand side of the economy. These policy impulses include changes in the money supply, the exchange rate, and fiscal policy.

A second type of shock is private demand shocks, such as modifications in consumption and investment spending, which are influenced by expectations about the future course of the economy.

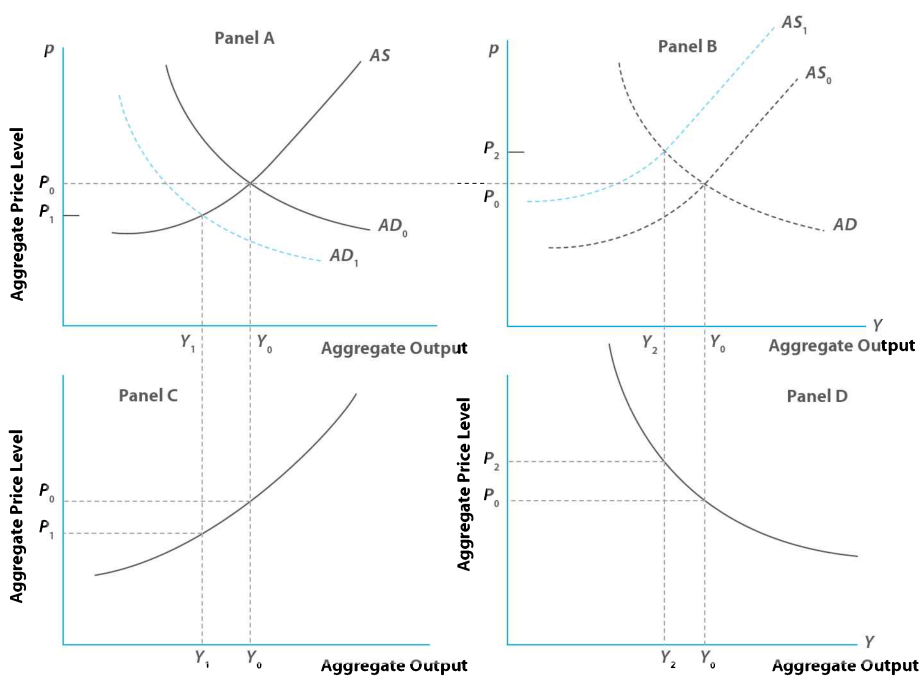
Finally, there are supply shocks that affect the production side of the economy, which include natural disasters, climatic change, and innovations in technological knowledge.

› Keynesians and Monetarists emphasize that fluctuations in economic activity are determined by demand shocks such as a collapse of investor confidence, a change in consumer spending, or changes in the government expenditure and money supply.

Both KEYNESIANS and MONETARISTS tend to emphasize the fluctuations in economic activity as being determined by demand shocks such as a collapse of investor confidence, a change in consumption spending as consumers become more optimistic or pessimistic about the future, a change in government expenditure, or changes in money supply. However, the two oil-price increases during the 1970s made macroeconomists more aware of the importance of supply-side factors in explaining macroeconomic fluctuation. Supply shocks result in a relationship between the price level and output (or inflation and growth) that differs from the relationship between these variables when demand shocks occur.

In panel A of Figure 9.7, an adverse demand shock is depicted as a leftward shift of the *AD* curve from $AD_0$ to $AD_1$, resulting in a decline in price and output. In panel B of Figure 9.7, an adverse supply shock, say, due to a drought or oil-price hike, reduces the production possible from given levels of labour and capital (the production function shifts inwards) and shifts the *AS* curve up from $AS_0$ to $AS_1$.

If we trace out the effects of pure demand shocks in panel C of Figure 9.7, the data would depict a positive relation over time between price and output. If most shocks in a period are supply shocks, however, the observation between price level and output would be a negative one as traced out in panel D of Figure 9.7. The Great Depression resulted in data observations as in panel C




**› Figure 9.7**
Demand and Supply Shocks.
An adverse demand shock shifts the *AD* curve to the left (panel A) and results in a decline in the price level and output. Tracing out shifts in the *AD* curve due to shocks gives us the relationship between the price level and the output, as depicted in panel C. An adverse supply shock reduces the output at each price level and shifts the *AS* curve to the left (panel B). The decline in the output is accompanied by a rise in the price level. Tracing out shifts in the *AS* curve due to shocks gives us the relationship between the price level and output, as depicted in panel D.

of Figure 9.7. During the decades of the 1970s and 1980s, however, the data show price levels (or inflation) and output (or growth) moving together as in panel D of Figure 9.7. These decades were noted for their supply shocks predominantly.

In the 1970s and 1980s, a group of economists, who came to be called REAL BUSINESS CYCLE THEORISTS, suggested that real shocks may be far more important than monetary shocks in explaining the path of aggregate output overtime.[15] Real business cycle theorists argue that nominal variables, such as the money supply and the price level, have no role in explaining fluctuations in real variables such as output and employment.

› Real business cycle theorists emphasize that real shocks to supply such as oil price hikes, changes in the government regulations, and productivity changes stemming from technological changes are far more important in explaining economic cycles.

Edward Prescott[16] of the University of Minnesota does state that the claim is not that monetary institutions never matter, for he admits that periods of financial breakdown, such as the Great Depression, are not accounted for by real business cycle models. His argument rather is that post-war macroeconomic time series can be represented without explicit attention to monetary institutions.

The supply-side shocks that the real business cycle theorists refer to include the following:

1. Unfavourable developments in the physical environment, which adversely affect output. This type of shock would include natural disasters such as earthquakes, droughts, and floods.
2. Significant changes in the price of energy such as the oil-price hikes of 1973 and 1979, and the subsequent reduction in 1986.
3. War, political upheaval, or labour disputes, which disrupt the existing performance and structure of the economy.
4. Government regulations such as import quotas, which damage incentives and divert entrepreneurial talent towards rent-seeking activities.
5. Productivity shocks generated by changes in the quality of the capital and labour inputs, new management practices, the development of new products, and the introduction of new techniques of production.

Apart from stressing productivity shocks stemming from technological progress and so real over nominal shocks, REAL BUSINESS CYCLE THEORISTS also view aggregate economic variables as the outcomes of the decisions made by many individual agents acting to maximize their utility subject to production possibilities and resource constraints. Real business cycle models, thus, have an explicit and firm foundation in microeconomics. Typically, the economy is populated by many identical long-lived agents. In what follows we simplify and adopt our standard two-period framework. Each agent has a utility function over the consumption and leisure they expect to enjoy over their lifetime.

› Real business cycle theorists assume that individuals maximize intertemporal utility and firms maximize profits and there are no market imperfections. As a result, market equilibrium is efficient and there is no need for government intervention.

$$U = U(C_1, H - N_1) + \frac{U(C_2, H - N_2)}{1+\rho}$$
$$= \log C_1 + \beta \log(H - N_1) + \frac{\log C_2}{1+\rho} + \frac{\beta \log(H - N_2)}{1+\rho} \qquad \textbf{(9.1)}$$

where, $H$ is the fixed-time endowment of the individual, and $N_i$ is the labour supplied in time period $i$. Thus, $H - N_i = l_i$ is the leisure enjoyed by an agent in time period $i$. The utility function is separable in consumption and leisure, and $\rho$ is the intertemporal discount rate. As utility is separable in consumption and leisure (labour), the optimal labour supply is independent of consumption and can be derived without explicit consideration of the optimal consumption decision. Each agent has access to a constant return to scale production technology for the single commodity in the economy.

$$Y_t = F(K_t, N_t) = A_t K_t^{1-\alpha} N_t^{\alpha} \tag{9.2}$$

where, the term $A_t$ summarizes the effect of technological progress. The production technology is, thus, subject to temporary productivity shits or technological changes, which are the underlying source of variation in the economic environment to which agents respond. Technological progress takes the form of a first-order autoregressive process.

$$A_t = \alpha + \theta A_{t-1} + \epsilon_t \tag{9.3}$$

Technology in period $t$, $A_t$, equals a fraction $\theta$ of its previous period's value, $A_{t-1}$, plus a random term, $\epsilon_t$. In empirical work on real business cycles, it has been found that for technology shocks to result in movements in output that are persistent and indicative of actual output fluctuations, the value of $\theta$ must be large, usually above 0.95. $\epsilon_t$ is a random variable that is independently and identically distributed. Finally, the model involves an identity and a behavioural equation. The identity is that the capital stock evolves through time according to Eq. (9.4).

$$K_{t+1} = (1-\delta)K_t + I_t \tag{9.4}$$

where, $0 < \delta < 1$ is the rate of depreciation. The identity that the sum of consumption and investment equals the output allows us to substitute $I_t$ in Eq. (9.4) with $Y_t - C_t$.

$$K_{t+1} = (1-\delta)K_t + Y_t - C_t \tag{9.5}$$

With aggregate consumption as the difference between aggregate output and savings, $C_t = (1-s)Y_t$, where $s$ is the marginal propensity to save. We can write Eq. (9.5) as follows:

$$K_{t+1} = (1-\delta)K_t + sY_t$$

or,

$$K_{t+1} = (1-\delta)K_t + sA_t K_t^{1-\alpha} N_t^{\alpha} \tag{9.6}$$

The agent's problem then boils down to choosing how to allocate hours between work and leisure, and how to allocate the supply of the goods between investment and consumption by maximizing the utility function subject to the constraints (9.3) and (9.6) and subject to the lifetime budget constraint (9.7) written as the following:

$$C_1 + \frac{C_2}{1+r} = \omega_1 N_1 + \frac{\omega_2 N_2}{1+r} = Y_{\text{lifetime}} \tag{9.7}$$

where $\omega$ is the real wage $(w/p)$. Under most specifications of preferences and production functions, it is not possible to solve analytically this maximization problem. Real business cycle theorists accordingly compute approximate solutions to the choice problem in the neighbourhood of the steady state.

We will proceed, however, by teasing out the economic intuition underlying the resulting optimal decisions. Because the utility function has as one of its arguments leisure and not directly labour,[17] we recast the budget constraint in terms of leisure. The intertemporal budget constraint, then, is as follows:

$$\omega_1(H - l_1) + \frac{\omega_2(H - l_2)}{1+r} = Y_{\text{lifetime}}$$

Multiplying throughout by $(1+r)$,

$$(1+r)\omega_1(H - l_1) + \omega_2(H - l_2) = (1+r)Y_{\text{lifetime}}$$

or,

$$(1+r)\omega_1 H + \omega_2 H - (1+r)Y_{\text{lifetime}} = (1+r)\omega_1 l_1 + \omega_2 l_2$$

As usual, there are two polar situations:

1. The agent does not enjoy leisure in period 1, $l_1 = 0$.

   Then, $$l_2 = \frac{(1+r)\omega_1 H + \omega_2 H - (1+r)Y_{\text{lifetime}}}{\omega_2} = \frac{\Psi}{\omega_2}$$

   where, $\Psi$ is the numerator of the expression for $l_2$.[18]
2. The agent does not enjoy leisure in period 2, $l_2 = 0$, in which case,

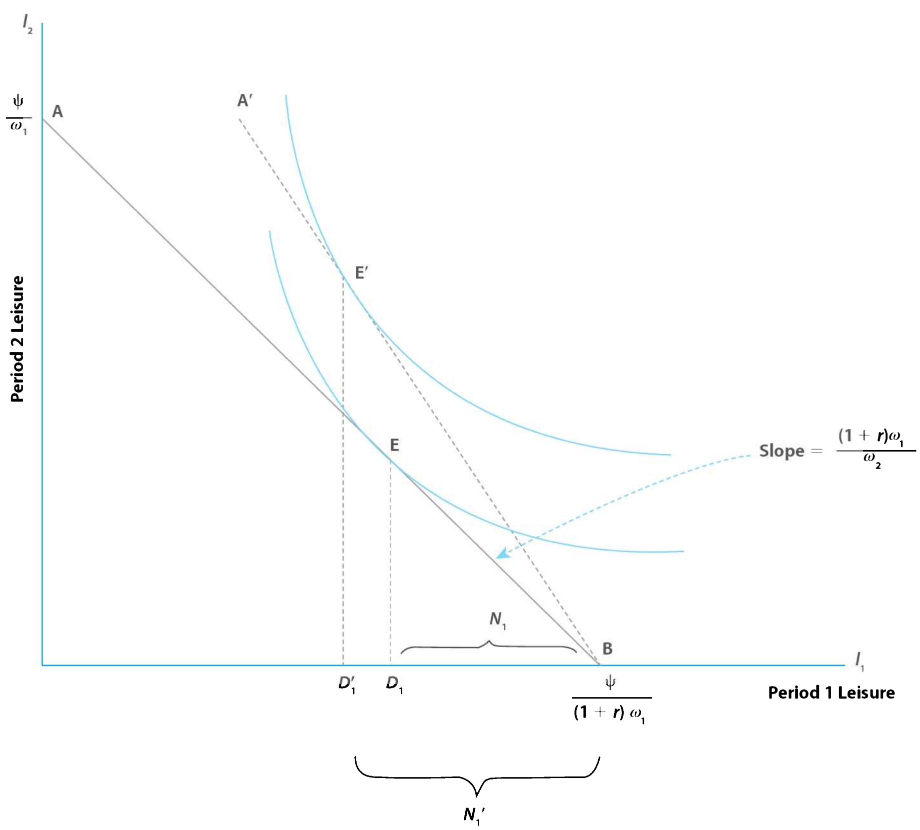
$$l_1 = \frac{(1+r)\omega_1 H + \omega_2 H - (1+r)Y_{\text{lifetime}}}{(1+r)\omega_1} = \frac{\Psi}{(1+r)\omega_1}$$

These corner situations, $l_1 = 0$ and $l_2 = 0$, are represented by points A and B in Figure 9.8. Joining points A and B gives the budget constraint for leisure. The slope of this constraint is the ratio of the vertical to horizontal distance, or,

$$\text{Slope of budget line AB} = \frac{\text{Distance OA}}{\text{Distance OB}}$$

$$= \frac{\Psi/\omega_2}{\Psi/(1+r)\omega_1} = \frac{(1+r)\omega_1}{\omega_2}$$

Which point on the lifetime constraint AB will the agent choose? Optimality demands that the agent select that point on AB which is tangential to an indifference curve reflecting the preferences of the agent for leisure (and, therefore, labour) in the two time periods as described in the utility function. This occurs at point E in Figure 9.8. The agent then enjoys $OD_1$ hours of leisure and spends $N_1 = D_1B$ hours in work in period 1. Suppose the first-period wage, $\omega_1$, increases. Then, the slope of the budget constraint AB becomes steeper—the dashed line A′B. As is usual, there are substitution and income effects. Current




**› Figure 9.8**
The Intertemporal Labour Supply Decision. **In real business cycle theory, individuals trade work in one period for work in another period, or, leisure in one period for leisure in another, as depicted by the budget constraint AB. An increase in the wage in period 1 makes period 1 leisure more expensive and the budget constraint A′B operates with the increase in the wage in this period. The individual then takes less leisure or works more, $N_1' > N_1$, in response to the increased wage in period 1. Less leisure today is substituted for more leisure in the future.**

productivity and income is high, which encourages intertemporal substitution of current for future work.[19] On the other hand, lifetime income is higher that acts to reduce current and future work effort. Real business cycle theorists demonstrate that for plausible parametrizations of the model, the substitution effect dominates so that current labour supplied rises. $N_1' = D_1'B$ hours are now worked in response to the higher current wage. Similarly, any other factor that increases the slope of the budget constraint, such as a rise in interest rates, will result in an increase in current labour supply. The labour supply function can then be written as follows:

$$N^s = N^s\left(\underset{(+)}{\frac{\omega_1}{\omega_2}}, \underset{(+)}{r}\right) \tag{9.8}$$

The demand for labour arises from the standard profit–maximization condition,[20] and labour is demanded up to the point where the real wage equals the marginal product of labour:

$$\omega = MP_N$$

We know that labour is subject to diminishing marginal productivity, giving us a downward-sloping demand curve[21] for labour:

$$N^D = N^D(MP_N) \tag{9.9}$$

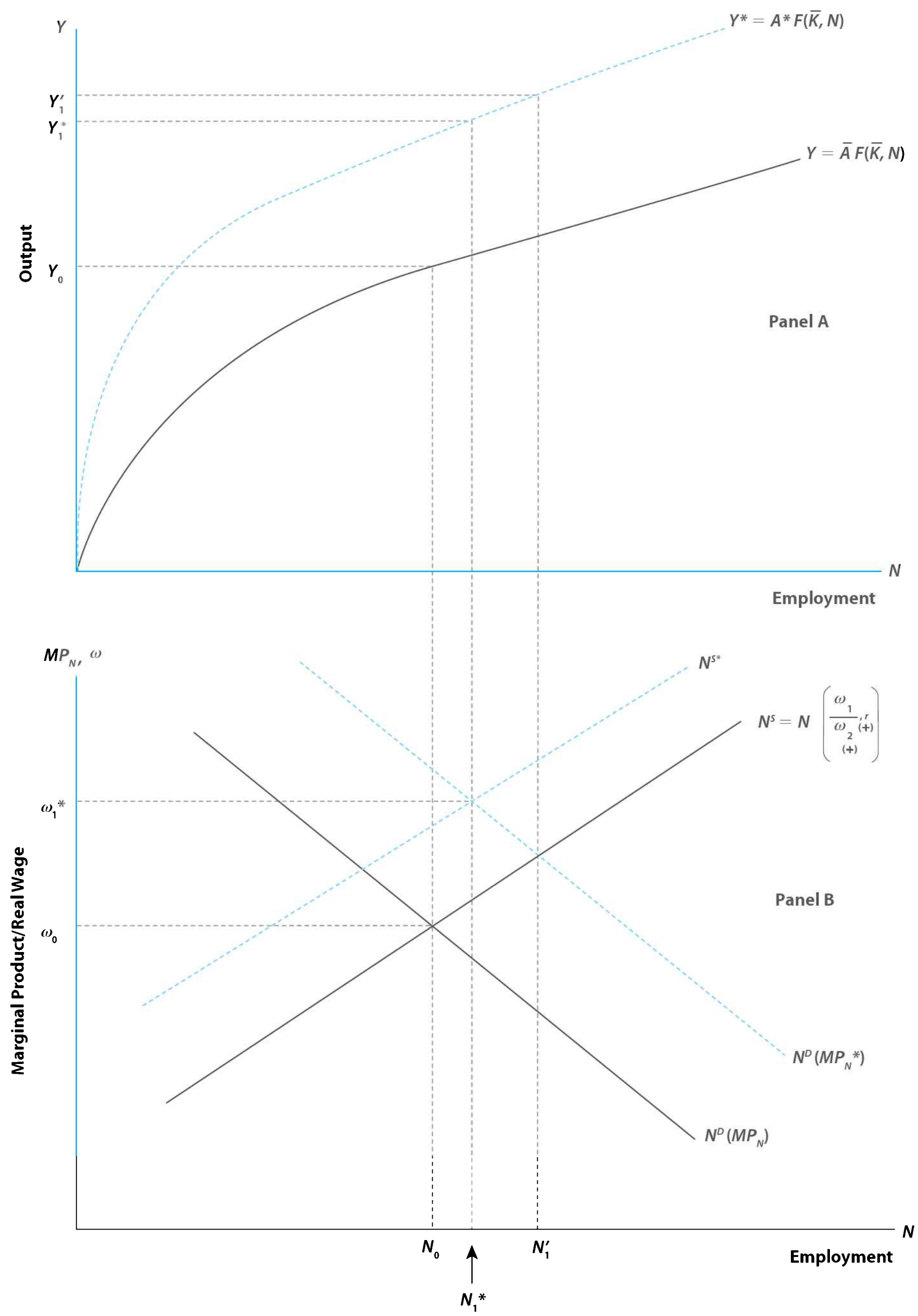
The demand and supply curve for labour are drawn in panel B of Figure 9.9 for an instant of time where the capital stock and technology are constant at $K_t = \bar{K}$, $A_t = \bar{A}$, making output $Y = \bar{A}\bar{K}^{1-\alpha}N^\alpha = \bar{A}F(\bar{K}, N)$ a function of labour input. This production function is drawn in panel A of Figure 9.9, where the gradually diminishing slope of the production function reflects the diminishing marginal productivity of labour.

Now suppose there is a beneficial technological shock that shifts the production function outwards to $Y^* = A^*F(\bar{K}, N)$. Improvements in technology enable more product or output to be produced with the same inputs. As the productivity of labour goes up with this beneficial technological shock, the demand curve for labour shifts out to $N^D(MP_N^*)$. With the increased demand for labour raising the current wage, individuals supply more labour[22] in response to the productivity shock at $N_1'$. However, now that output and income have increased from $Y_1$ to $Y_1'$, this increase in output raises savings, and the rise in savings, in turn, lowers the interest rate. This shifts the labour supply curve up and to the left: a reduction in the interest rate increases the present value of future income and makes the return from current labour supply unattractive. The output and income, thus, finally increase to $Y_1^*$ and the real wage increases to $\omega_1^*$.

› Real business cycle theorists assert that the supply of labour is strongly and positively related to transitory movements in the real wage rate and that the real wage rate is procyclical.

In real business cycle theory, the REAL WAGE is, therefore, procyclical with output. When an economic variable moves in the same direction as the aggregate economic activity—up in expansions and down in contractions—it is procyclical. A variable that moves in the opposite direction to the aggregate economic activity is countercyclical. When demand shocks are the source of fluctuations, as in the Keynesian approach, real wages are countercyclical. When the economy goes into a boom in real business cycle theory, leisure falls (labour supply rises), while consumption rises because output and income rise.

Explaining why consumption rises and leisure declines during a boom is problematic for real business cycle theory; as both consumption and leisure are normal goods, we would expect these variables to move together.[23] The only reason why agents in a boom find it rational to decrease leisure and, at the same time, increase the quantity of goods they demand is because the price of leisure




**› Figure 9.9**
Employment and Wages in Real Business Cycle Models. A favourable technological shock shifts the production function outwards, with more output produced per unit of labour (panel A). The increased productivity of labour shifts the demand for labour curve out to the right in panel B. This results in a rise in the real wage. However, due to the increase in income and savings, the interest rate declines and reduces the return to the current labour supply. The supply of labour curve then shifts to the left in response to this reduced return to the labour in panel B. The reduction in the labour supply increases the real wage further. An improvement in the productivity is, therefore, associated with a rise in real wages.

relative to goods—the real wage—increases during a boom. In response to the high return to working, agents increase the labour supply and decrease leisure.

A procyclical real wage is, therefore, possible only in two ways: (1) If, in a recession, the production technology is relatively adverse, which makes the marginal product of labour and the real wage low; and (2) If, in a boom, the production technology is relatively favourable, resulting in a high real wage and marginal product of labour.

Real business cycle theorists, therefore, have to assume substantial fluctuations in the rate of technological change, which is not what the data on total factor productivity reveal.[24] Moreover, in actual fluctuations, the real wage appears to be at most only moderately procyclical.

Yet another weakness of the real business cycle approach is that it requires agents to willingly reallocate leisure (and labour) over time. Agents appreciably reduce the quantity of labour they supply in response to a reduction in the real wage or in response to a decrease in the real interest rate. Real business cycle theorists require labour to be exceedingly intertemporally substitutable. Individuals facing an increase in the real wage today relative to future real wages should work a lot today and a little in the future.

Yet, studies of individual labour supply over time have not been able to find that individuals take action in response to expected real wage changes by considerably reallocating leisure over time.[25,26] A recent study[27] finds some evidence, though, in favour of intertemporal substitution of labour. Another significant behavioural response envisaged by real business cycle theorists is that labour supply responds significantly to the real interest rate. As Mankiw[28] argues, this just does not seem believable:

> The real interest rate is simply not a significant consideration when individuals decide to leave their jobs or to accept new employment. While economists can easily convince laymen and students that the quantity of apples depends on the price of apples, it is much harder to convince them that labour supply depends on the real interest rate.

Perhaps the most disturbing feature of real business cycle theory is that by portraying work and leisure to be highly substitutable over time, it treats fluctuations in employment as reflecting *voluntary* changes in the number of hours people want to work. Any unemployment that we see around us is, thus, voluntary and to the extent that it is the outcome of optimizing behaviour, there is no room for any corrective policy or state intervention. The invisible hand of the market results in outcomes such as employment and unemployment, and the outcome is PARETO-EFFICIENT, that is, any intervention designed to improve someone's lot can only be at the expense of someone else's lot. Prescott's[29] statement that "economic fluctuations are optimal responses to uncertainty in the rate of technological change," builds on the premise that markets are complete and that information asymmetries are nonexistent. In this world, however, individuals are never rationed out of jobs involuntarily, which is not plausible.

› Pareto efficiency or optimality is a condition in which no change in outcome that will make some members of society (such as the unemployed) better off without making some other members of society (such as the employed) worse off is possible.

It is useful to document that real shocks affect the economy differently from nominal stocks but then to allow real shocks to result in optimal outcomes for the economy with no role for stabilization policy through state intervention is to develop a world of art.

## SUMMARY

- The *AD* curve is the relationship between the real quantity demanded of newly produced final goods and services and the general price level. It is a downward-sloping curve.
- The *AD* curve is derived from the equilibrium conditions in the goods (*IS* curve) and the money market (*LM* curve).
- Changes in any variable other than the price level that shift the *IS* and *LM* curves, such as fiscal and monetary policy, consumption spending or private investment, shift the *AD* curve.
- The *AS* curve is the relationship between the real quantity supplied of newly produced final goods and services in an economy and the general price level.
- The Keynesian *AS* curve in the short run is drawn for a fixed money wage and is upward sloping up to the point where full employment ($N^D = N^S$) occurs. At this point, it is vertical.
- Involuntary unemployment refers to a situation where workers are willing to work at the going wage rate, and, yet, firms do not employ them.
- Keynesians advocate the use of fiscal policy to reduce involuntary unemployment as they believe that crowding out of private investment expenditures is low.
- The mistaken expectations short-run *AS* curve is upward sloping when the expected price diverges from the actual price

($p^e \neq p$). When expectations are realized ($p^e = p$), the natural rate of output is achieved with a vertical *AS* curve.

- In the mistaken expectations set-up, *AD* management could end up destabilizing the economy because with a revision of expectations, employment reverts to its natural rate. According to Friedman, given enough time, the economy tends to self-correct and return to the natural rate of employment and output without changes in monetary or fiscal policy.
- Even when it is recognized that policy changes are needed, policy lags can make attempts to fine-tune the economy counterproductive. The inside lag is the time required to recognize a need for a policy change and then to implement that policy change. It is shorter for monetary policy due to shorter implementation lags associated with it. The outside lag is the time required for the change in policy to affect the economy and is shorter for fiscal policy.
- Keynesians and Monetarists emphasize that fluctuations in economic activity are determined by demand shocks such as a collapse of investor confidence, a change in consumer spending arising from changes in confidence regarding the state of the economy, or sudden unanticipated changes in government expenditure or money supply.
- Real business cycle theorists emphasize that real shocks to supply such as oil price hikes, changes in government regulations, and productivity changes stemming from technological progress are far more important in explaining economic cycles.
- Real business cycle theorists assume that individuals maximize intertemporal utility while firms maximize profits, and there are no market imperfections. As a result, market equilibrium is efficient, and there is no need for government intervention.
- Real business cycle theorists assert that the supply of labour is strongly and positively related to transitory movements in the real wage rate—the intertemporal substitution hypothesis—and that the real wage rate is procyclical.

## NOTES

1. We are restricting ourselves to a closed economy for the time being.
2. This is a reduction in parameter *h* in Eq. (8.18) of Chapter 8.
3. *L* in the *LM* curve represents the real demand for money or the demand for liquidity.
4. Figure 9.5 depicts the equilibrium values of $w$, $N^D$, $N^S$, $P$, and $Y$. Panel D of this diagram depicts equilibrium output as $Y_0$, and this information can be used to derive the equilibrium nominal interest rate from the intersection of the *IS* and *LM* curves in Figure 9.1.
5. This can be interpreted as a low value of parameter *a* in the investment function.
6. Parameter *b* in the investment function also assumes a low value.
7. See Chapter 2 on the money multiplier.
8. Milton Friedman, "A Monetary and Fiscal Framework for Economic Stability," *Essays in Positive Economics* (Chicago: Chicago University Press, 1953).
9. Milton Friedman, "Employment, Growth, and Price Levels," 86th Congress, First Session, 25–28 May 1958, p. 609.
10. John Kareken and Robert Solow, "Lags in Monetary Policy," in *Stabilisation Policies*, Commission on Money and Credit (Prentice Hall, 1963).
11. See the section on irreversibility and investment in the chapter on investment.
12. Martin Eichenbaum and Charles L. Evans, "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates," *Quarterly Journal of Economics* 110, no. 4 (1995): 975–1009.
13. These are referred to as information asymmetries. For more on this, see Joseph E. Stiglitz and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 71, no. 3 (1981): 393–410.
14. See Milton Friedman, "The Role of Monetary Policy," *American Economic Review* (March 1968).
15. The seminal paper is that of C. R. Nelson, and C. I. Plosser. See Charles R. Nelson and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Times Series: Some Evidence and Implications," *Journal of Monetary Economics* 10 (September 1982): 139–162.
16. Edward C. Prescott, "Theory Ahead of Business Cycle Measurement," Federal Reserve Bank of Minneapolis, *Quarterly Review* (Fall 1986): 9–22.
17. This is because it allows the drawing of standard convex to the origin indifference curves. Labour in the utility function would provide negative utility.
18. When $l_1 = 0$, we still have $0 \le l_2 \le H$, that is, it does not necessarily imply that, in period 2, the agent must spend all the hours in leisure just because he spent all the time working in period 1 ($l_1 = 0$). If $l_2 = H$, then, the expression for
$$Y_{\text{lifetime}} = \omega_1 N_1 + \frac{\omega_2 N_2}{1+r} = \omega_1(H - l_1) + \frac{\omega_2(H - l_2)}{1+r} = \omega_1 H$$
Then,
$$l_2 = \frac{\Psi}{\omega_2} = \frac{(1+r)\omega_1 H + \omega_2 H - (1+r)\omega_1 H}{\omega_2} = H$$
Hence,
$$l_2 = \frac{\Psi}{\omega_2} \text{ is the general expression for } 0 \le l_2 \le H$$
19. Simultaneously, there will also be an intertemporal substitution of the current consumption for leisure that we are not foregrounding right now.

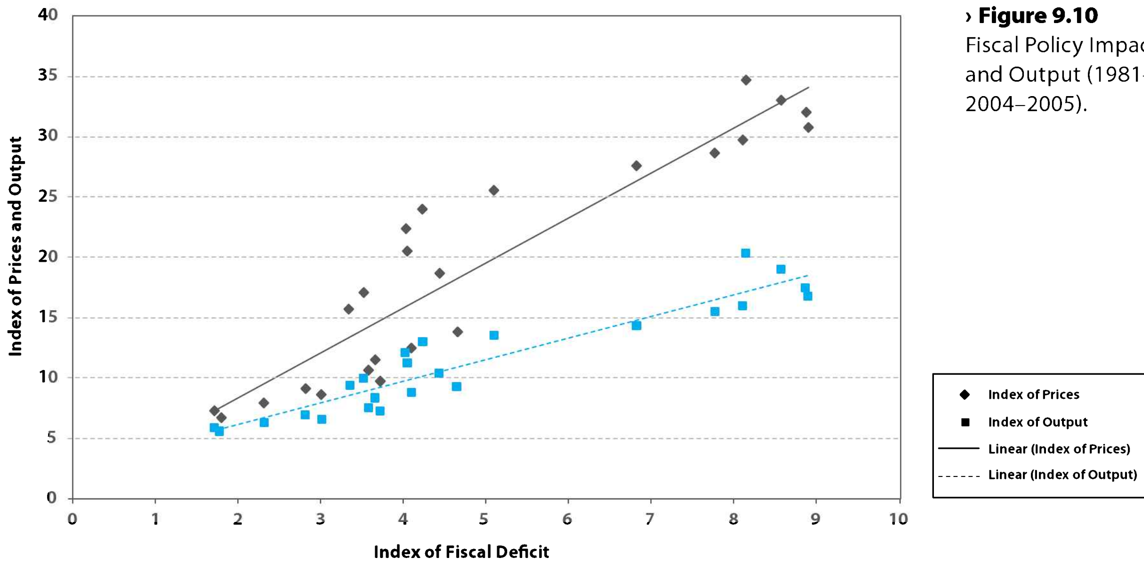
20. Real Profits are given by $\Pi = q - wN - (r + \delta)K$, where, $q = AK^{1-\alpha}N^{\alpha}$.
21. With $Y = AK^{1-\alpha}N^{\alpha}$, we have $MP_N = \alpha(Y/N)$. The average product of labour ($Y/N$) diminishes as employment increases resulting in a downward-sloping $MP_N$ curve.
22. If productivity shocks are persistent and long lived, there would be less incentive for agents to increase the current labour supply because the future output would also rise and increase the income effect while reducing the substitution effect.
23. Nicholas G. Mankiw, "Real Business Cycles: A New Keynesian Perspective," *Journal of Economic Perspectives* (Summer 1989): 79–90.
24. Ibid.
25. Joseph G. Altonji, "Intertemporal Substitution in Labour Supply: Evidence from Micro Data," *Journal of Political Economy,* Part 2, 94 (June 1986): S176–S215.
26. Lawrence Ball, "Interpersonal Substitution and Constraints on Labour Supply: Evidence from Panel Data," *Economic Inquiry* 28 (October 1990): 706–724.
27. Casey B. Mulligan, "Substitution over Time: Another Look at Life Cycle Labour Supply," *NBER Macroeconomics Annual* 13, 1998.
28. Ibid.
29. Edward C. Prescott, "Theory Ahead of Business Cycle Measurement," Federal Reserve Bank of Minneapolis, *Quarterly Review* (Fall 1986): 21.

## TEST YOURSELF

1. What is aggregate demand and how is it determined?
2. How can fiscal and monetary authorities affect the position of the *AD* curve of an economy?
3. Suppose the *AS* curve of the economy is Keynesian. How does monetary and fiscal policy affect the level of investment in the economy? As investment adds to capacity, comment on the possible long-run impacts.
4. Draw a mistaken expectations aggregate demand–aggregate supply diagram with the economy away from the natural rate of output. Describe how the economy gets to long-run equilibrium over time.
5. How does the presence of lags inherent in the policy process affect the use of monetary versus fiscal policy?
6. It is often argued that active policymaking is unnecessary as the economy has a tendency to self-correct. Elaborate on the pros and cons of this view.
7. What are the main channels through which monetary policy may affect economic activity?
8. What are shocks? What differential impacts do demand and supply shocks have on the economy?
9. Explain how productivity shocks affect real wages in the real business cycle model.
10. The real business cycle approach to macroeconomics ignores the demand side of the economy and portrays business cycles as a pure supply-side phenomenon. What is the advantage of this approach to macroeconomics?

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India Web site (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
2. Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy".
3. Click on "Table 121: Combined Deficits of Central and State Governments," and create a spreadsheet of annual data on the fiscal deficit of the nation from 1980–1981 to 2004–2005.
4. Then click on "Table 1: Macro-economic aggregates (at current prices)," and in your spreadsheet, insert annual data on GDP at market prices (base 1999–2000) from 1980–1981 to 2004–2005.
5. Finally click on "Table 2: Macro–economic aggregates (at constant prices)," and in your spreadsheet, insert annual data on GDP at market prices (base 1999–2000) from 1980–1981 to 2004–2005.
6. Find the inflation rate as implied by the GDP price deflator. The GDP price deflator is the GDP at current market prices divided by the GDP at constant market prices and the result so obtained multiplied by 100. For instance, the GDP at current prices in 1980–1981 is INR 1,453.70 billion, and at constant prices, it is INR 6,953.61 billion. The GDP deflator for 1980–81 is (1,453.70/6,953.61 = 0.2091). The GDP deflator then is 20.91. Similarly, the GDP deflator for 1981–1982 is 23.17. The implied inflation rate between 1980–1981 and 1981–1982, then, is [(23.17 − 20.91)/20.91] × 100 = 10.81 per cent.
7. Find the output growth rate over the period. The output growth in 1981–1982, for instance, is the growth in GDP at constant prices over the previous year. This is given by [(7,370.78 − 6,953.61)/6,953.61] × 100 = 6.00 per cent.
8. Now create an index of the overall fiscal deficit. First, find the growth rate of the fiscal deficit per period. The




› **Figure 9.10**
Fiscal Policy Impact on Price and Output (1981–1982 to 2004–2005).

fiscal deficit in 1980–1981 was INR 107.8 billion, and in 1981–1982, it was INR 106.08 billion. The growth rate, then, is $[(106.08 - 107.8)/107.8] \times 100 = -1.6$ per cent. For each period, subtract the growth rate for the nominal fiscal deficit by the inflation rate you calculated in Step 6 above. The inflation rate in 1981–1982 was 10.85 per cent. Hence, the real growth in the fiscal deficit in 1981–1982 is $-1.6 - 10.85 = -12.45$ per cent. Let the real value of the fiscal deficit in 1980–1981 have an index value equal to the value 2. Then, the index of the real value of the fiscal deficit in 1981–1982 is equal to $2 \times [1 + (-12.45/100)] = 1.75$. Similarly, calculate the index of the fiscal deficit per period.

9. Similarly, create an index of output. Let the output at constant prices in 1980–1981 of INR 6,953.61 billion have an index value of 5.33. Then, the index value of 1981–1982 output at constant prices of INR 7,370.78 billion will have an index value of $[(5.33 \times 7{,}370.78)/6{,}953.61] = 5.65$.
10. Finally, create an index of the price level of the economy. The inflation rate, as measured by the GDP deflator in 1980–1981 was 10.81 per cent. If the price level in 1979–1980 is indexed at 5.5, then, the price index in 1980–1981 is $5.5 \times [1 + (10.81/100)] = 6.09$.
11. Now, create a graph with the index of the fiscal deficit on the horizontal axis and the index of the price level and output on the vertical axis. What does your graph depict about the impact of the fiscal deficit on the aggregate price level and aggregate output in the economy?
12. The graph you created in Step 12 should result in a series of points as depicted in Figure 9.10. Note that the linear trend line of the impact of the fiscal deficit on prices is higher than that on output. In fact, as the fiscal deficit increases, it has a larger impact on prices and does not impact as much on output. If you closely analyse the slopes of the linear trends, you will find that a unit increase in the index of the fiscal deficit increases the index of prices by 3.7 (the slope is 3.7), while it increases the index of output by 1.8. Why does the increased government expenditure that is financed by borrowing (the fiscal deficit) result in a larger increase in prices than on output in the Indian economy?

# Appendix 9.1 The Aggregate Demand Curve

The *AD* curve is the outcome of equilibrium in the goods and financial markets. Hence, the *AD* curve is derived by considering the determinants of the variables that influence the equilibrium in these markets. In the goods market for a closed economy, we have the familiar income equals expenditure identity:

$$Y = C + I + G \tag{A9.1.1}$$

Consumption expenditure depends on disposable income and the interest rate, as given by Eq. (8.11) in Chapter 8.

$$C = c_0 + c_1(Y - T) - c_2(i - \pi^e) \tag{A9.1.2}$$

We have taken direct taxes to be a lump sum, but now we allow them to vary with income:

$$T = t_0 + t_1 Y \tag{A9.1.3}$$

where, $t_0$ is the autonomous component of taxes that does not vary with income; and $t_1$ is the marginal tax rate. Of course, $0 < t_1 < 1$.
Investment is inversely related to the rate of interest, as given by Eq. (8.12) in Chapter 8:

$$I = a - b(i - \pi^e) \tag{A9.1.4}$$

Government expenditure is exogenous, $G = \bar{G}$. Substituting Eqs. (A9.1.2), (A9.1.3), and (A9.1.4) into Eq. (A9.1.1), we obtain,

$$\begin{aligned} Y &= c_0 + c_1(Y - t_0 - t_1 Y) - c_2 i + c_2 \pi^e + a - bi + b\pi^e + G \\ &= c_0 - c_1 t_0 + a + G + (b + c_2)\pi^e + c_1(1 - t_1)Y - (b + c_2)i \end{aligned}$$

Solving for the interest rate,

$$i = \frac{c_0 - c_1 t_0 + a + G - [1 - c_1(1 - t_1)]Y}{(b + c_2)} + \pi^e \tag{A9.1.5}$$

Similarly, equilibrium in the money market is given by the condition that money demand equals money supply. The supply of money is controlled by the central bank and is given by Eq. (8.17) of Chapter 8:

$$M^S = \bar{M} \tag{A9.1.6}$$

The demand for money is inversely related to the interest rate and directly related to aggregate income [see Eq. (8.18)] of Chapter 8:

$$\frac{M^d}{P} = h + nY - ei \tag{A9.1.7}$$

Equating Eqs. (A9.1.6) and (A9.1.7) gives us

$$\frac{\bar{M}}{P} = h + nY - ei$$

or,

$$i = \frac{h + nY - (\bar{M}/P)}{e} \tag{A9.1.8}$$

Equating the interest rates in Eqs. (A9.1.5) and (A9.1.8),

$$\frac{c_0 - c_1 t_0 + a + G - [1 - c_1(1 - t_1)]Y}{(b + c_2)} + \pi^e = \frac{h + nY - (\bar{M}/P)}{e}$$

Taking terms in *Y* on the left-hand side,

$$Y\left[\frac{n}{e} + \frac{1 - c_1(1 - t_1)}{(b + c_2)}\right] = \frac{c_0 - c_1 t_0 + a + G}{b + c_2} + \pi^e - \frac{h - (\bar{M}/P)}{e}$$

or $$Y\left\{\frac{(b+c_2)n+[1-c_1(1-t_1)]e}{(b+c_2)e}\right\}=\frac{c_0-c_1t_0+a+G}{b+c_2}+\pi^{e}-\frac{h-(\overline{M}/P)}{e}$$

Multiplying throughout by the inverse of the coefficient of *Y*,

$$Y=\frac{c_0-c_1t_0+a+G}{b+c_2}\frac{(b+c_2)e}{(b+c_2)n+[1-c_1(1-t_1)]e}+\pi^{e}\frac{(b+c_2)e}{(b+c_2)n+[1-c_1(1-t_1)]e}$$
$$+\frac{(\overline{M}/P)-h}{e}\frac{(b+c_2)e}{(b+c_2)n+[1-c_1(1-t_1)]e}$$

We may write this as

$$Y=\left[c_0-c_1t_0+a+G\right]\frac{e}{(b+c_2)n+[1-c_1(1-t_1)]e}+\pi^{e}\frac{(b+c_2)e}{(b+c_2)n+[1-c_1(1-t_1)]e}$$
$$+\left[(\overline{M}/P)-h\right]\frac{(b+c_2)}{e}\frac{e}{(b+c_2)n+[1-c_1(1-t_1)]e}$$

Alternatively,

$$Y=[c_0-c_1t_0+a+G]\frac{1}{(b+c_2)(n/e)+[1-c_1(1-t_1)]}$$
$$+\pi^{e}(b+c_2)\frac{1}{(b+c_2)(n/e)+[1-c_1(1-t_1)]}$$
$$+\left[(\overline{M}/P)-h\right]\frac{(b+c_2)}{e}\frac{1}{(b+c_2)(n/e)+[1-c_1(1-t_1)]}$$

To solve this equation for *P*, take the term in real money supply to the left-hand side.

$$\left[(\overline{M}/P)-h\right]\frac{(b+c_2)}{e}\frac{1}{(b+c_2)(n/e)+[1-c_1(1-t_1)]}$$
$$=Y-[c_0-c_1t_0+a+G]\frac{1}{(b+c_2)(n/e)+[1-c_1(1-t_1)]}$$
$$-\pi^{e}(b+c_2)\frac{1}{(b+c_2)(n/e)+[1-c_1(1-t_1)]}$$

Then, the expression for $\overline{M}/P$ is given by,

$$\frac{\overline{M}}{P}=Y\frac{e}{(b+c_2)}\left\{(b+c_2)\frac{n}{e}+[1-c_1(1-t_1)]\right\}-(c_0-c_1t_0+a+G)\frac{e}{(b+c_2)}-e\pi^{e}+h$$

Then, the expression for *P* is

$$P=\frac{\overline{M}}{Y[e/(b+c_2)]\{(b+c_2)(n/e)+[1-c_1(1-t_1)]\}-(c_0-c_1t_0+a+G)[e/(b+c_2)]-e\pi^{e}+h}$$

**(A9.1.9)**

This is the equation of the *AD* curve, where *P* and *Y* are inversely related, and, thus, give a downwardly sloping curve. By observing the signs of the parameters in this equation, we can write the effect of a change in a parameter on the position of the *AD* curve as follows:

$$P=f\left(\underset{(-)}{Y};\underset{(+)}{M},\underset{(+)}{G},\underset{(-)}{t_0},\underset{(-)}{t_1},\underset{(+)}{\pi^{e}},\underset{(+)}{a}\right) \quad \textbf{(A9.1.10)}$$

The *AD* curve—the relation between *P* and *Y*—shifts to the right when

a. the money supply increases: *M* increases;
b. government expenditure increases: *G* increases;
c. there is an increase in inflation expectations: $\pi^e$ increases; and
d. there is an increase in business confidence or productivity: *a* increases.

The *AD* curve shifts to the left when either the marginal income tax rate $t_1$ or the base rate of tax $t_0$ increases.

# 10

# More on Unemployment

CRITICAL QUESTIONS

- *What accounts for the fact that wages are higher than the market-clearing level?*
- *What are efficiency wages?*
- *How do efficiency wages improve a firm's productivity?*
- *How is the aggregate supply curve related to efficiency wages in the short run and in the long run?*
- *Why are wage contracts staggered over time?*
- *Who are the insiders and outsiders in the labour market?*
- *How does the presence of insiders generate unemployment?*

## 10.1 Shirking and Efficiency Wages

In any economy, there are always unemployed people who are available for work and are, in fact, seeking work. Up to now we have looked at two explanations for this aggregate unemployment or excess supply of labour. One has been that the money wage has been set above the equilibrium rate, and as actual employment is determined by the short side of the labour market (the demand side), this results in unemployment.

The other explanation advanced argues that workers form mistaken expectations about prices, and when expected prices are higher than actual prices, wages are set too high, which again results in unemployment. What is the basis of this non-market-clearing wage that results in the outcome of unemployment? There are two approaches that explain this.

Keynes argued that employees were more concerned with relative wages rather than real ones and that various institutional factors kept the wage rates high. Friedman argued that workers are imperfectly informed about prices and with adaptive expectations, they revise the nominal wage as the information about prevailing prices becomes available.

But with nominal wages set too high either because workers are concerned about their relative position in the labour market or because workers expect prices to be high, what prevents unemployed workers from knocking on the doors of firms and offering to work for lower wages? And what prevents the firms from hiring such workers, thus bidding down the wage and balancing supply and demand? After all, a worker who is otherwise identical to the firm's current workers and who offers to work for less than what the firm currently pays its workers should be acceptable to a firm that seeks any opportunity to increase profits. The efficiency wage theory explains why firms may not find it profitable to hire more workers at reduced wages.

It is the hallmark of the efficiency wage approach that even if workers offer to work for lower wages, firms do not want to reduce wages. In fact, firms pay higher wages because it is beneficial to them. Firms are willing to PAY MORE than the market-clearing wage in order to recruit good workers, to discourage quitting, or to stimulate employees to work effectively.[1] Not just that, in developing countries, higher wages increase workers' food consumption, thereby causing them to be better nourished and more productive. At the subsistence, competitive wage workers might be undernourished, which can affect their health and cause them to resort to absenteeism in addition to reducing their potential productivity on the job. Higher wages can allow workers to afford a more nutritious diet, and be better nourished, healthier, and more productive, which is beneficial to the firm. Thus, the increase in labour costs, in terms of higher wages, can be compensated for by a greater increase in the productivity of the workforce.[2] The efficiency wage is that wage where the marginal cost of increasing the wage exactly equals the marginal gain in the productivity of the firm's workers. The source of efficiency wages that we focus on here, however, is not nutrition or a reduced turnover of workers, but the fact that firms have a limited ability to monitor workers, which causes them to provide workers with an incentive to exert effort. This approach is due to a famous paper by Shapiro and Stiglitz.[3]

› Firms pay higher than the market-clearing wages to elicit effort, reduce turnover, recruit better workers, and improve worker morale.

A problem arises when firms cannot easily MONITOR WORKERS' EFFORTS at their jobs as monitoring activities are expensive. Workers can shirk and spend time on activities unrelated to the position for which they have been hired, such as reading the newspaper, updating their gossip from the tabloids, or sending

› When firms have difficulty **monitoring employees**, they have an incentive to offer a wage-employment package that deters workers from shirking on the job.

SMSes to their friends. Shirking workers result in a loss of productivity to the firm, which would like to offer a wage-employment package that encourages workers not to shirk at all.

Let there be $\bar{\bar{N}}$ workers in the labour market, and let the labour supply curve be inelastic. Workers are either employed or unemployed. If $w/p$ is the real wage and $e$ is the workers' effort, the instantaneous benefit to a worker is

$$B = \begin{cases} \frac{w}{p} - e & \text{if employed} \\ 0 & \text{if unemployed} \end{cases}$$

There are only two possible effort levels. Thus, either the worker exerts effort, $e$, or the worker shirks, $e = 0$ . Hence,

$$B_{\text{employed}} = \begin{cases} \frac{w}{p} - e & \text{if effort is expended by worker} \\ \frac{w}{p} & \text{if worker shirks} \end{cases}$$

At any moment, a worker could then be (1) employed and exerting effort; (2) employed and shirking; or (3) unemployed.

Let $\mu$ be the probability of unemployment. Unemployment occurs because firms detect workers who are shirking and detection occurs with probability $\mu$. Workers who are caught shirking are fired.[4] What is the net benefit gained by the employee by cheating on the job?

$$\text{Net benefit of shirking} = B\text{ (shirking)} - B\text{ (exerting effort)}$$

$$= \frac{w}{p} - \left(\frac{w}{p} - e\right)$$

$$= e$$

The downside to cheating on the job is that the worker will be detected with probability $\mu$ and that will result in their being fired and the loss of their income. The expected loss from shirking is, therefore, the probability that the cheating is detected ($\mu$), times the loss of income from being fired *(w/p)*.

$$\text{Expected loss of shirking} = \mu \frac{w}{p}$$

The expected loss of wages from cheating is, therefore, equal to the gain from cheating when

$$\mu \frac{w}{p} = e$$

or,

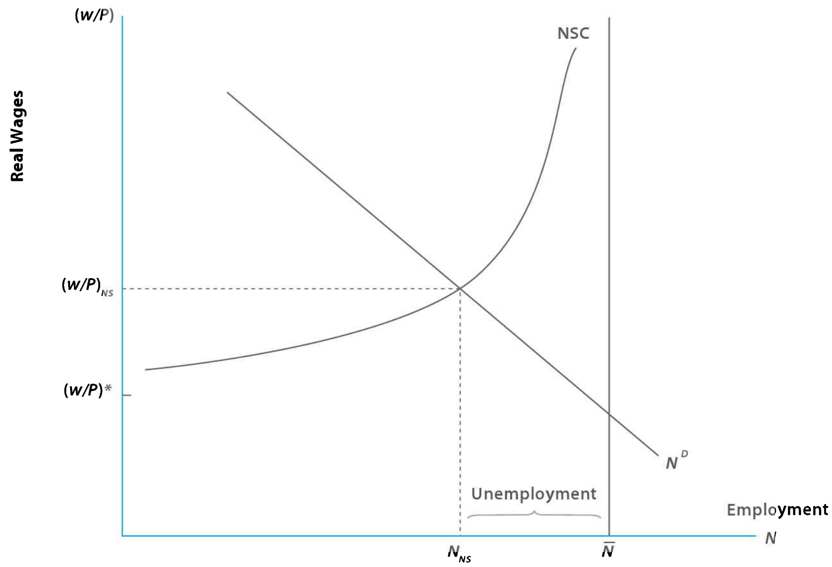
$$\frac{w}{p} = \frac{e}{\mu} \qquad \text{[no shirking condition (NSC)]}$$

This equation is the no shirking condition (NSC), which gives the wage the firm must pay to ensure that workers do not shirk. If the unemployment rate is high—$\mu$ is large—shirking is costly for a worker since once a shirking worker gets caught and fired, the worker faces a long unemployment spell. Firms can pay a relatively low wage in such a situation and will still be able to attract workers who will not shirk.

On the other hand, if the unemployment rate is low—$\mu$ is small—shirking workers who are caught face a short unemployment spell that is not too costly for them, and firms must offer a relatively high wage in order to make shirking costly for such workers. This generates an upward-sloping no shirking supply curve as shown in Figure 10.1. As the probability of unemployment declines, that is as we move closer to the full employment level $\bar{\bar{N}}$, the firm must pay a higher no shirking wage to deter workers from shirking.

Firm behaviour is otherwise unchanged: It seeks to maximize profits. With $e$ as the effort per employee and $N$ the number of employees who are exerting effort, its workforce measured in efficiency units is $eN$. Its output is given by[5]

$$Y = F(eN)$$




**› Figure 10.1**
Unemployment in a Shirking Model. There are $\overline{N}$ workers offering to work in the labour market. Firms seek to deter shirking workers. When the rate of unemployment is high, the firm can pay close to competitive wages $(w/p)^*$ and deter shirking. As the unemployment rate falls, higher wages have to be paid to deter shirking. This results in the upward-sloping NSC. The demand for labour is given by the downward-sloping marginal productivity of the employment curve. Equilibrium unemployment is determined at the point where the NSC curve intersects the demand-for-labour curve.

With $S$ as the number of employees who are shirking, the firm's profits are

$$\Pi = F(eN) - \frac{w}{p}(N + S)$$

Profit maximization results in a downward-sloping demand-for-labour curve $N^D$ depicted in Figure 10.1, along which the wage is equal to the marginal product of employment, or

$$\frac{w}{p} = eMP_{eN} = eF'(eN)$$

The equilibrium wage is where the no shirking supply curve and the labour demand curves intersect. At this equilibrium, $(w/P)_{NS}$ is the EFFICIENCY WAGE and $N_{NS}$ workers will be employed, leaving $(\overline{N} - N_{NS})$ WORKERS UNEMPLOYED.[6] If there were perfect monitoring, equilibrium would occur at the intersection of $N^D$ and $\overline{N}$. It is because of imperfect monitoring that equilibrium occurs at the intersection of the labour demand curve and the NSC. Three interesting implications arise:

› Efficiency wages are the wages that maximize firm profits when workers may work or shirk.

› Efficiency wages are higher than the competitive market wage and generate **involuntary unemployment**.

1. There is no market pressure to bid the wage down to the competitive wage $(w/P)^*$ by unemployed workers. If firms paid less than $(w/P)_{NS}$, $(w/P) < (w/P)_{NS}$, the number of workers willing to work and not shirk will be less than the number of workers demanded $N_{NS} < N^D$, and this would force the wage to rise. Similarly, if the wages were higher than the efficiency wage, $(w/P) > (w/P)_{NS}$, the number of workers willing to work and not shirk would be greater than the demand for workers at that wage, $N_{NS} > N^D$, and the wage would fall to $(w/P) = (w/P)_{NS}$. Workers cannot bid down wages to the competitive level $(w/P)^*$.
2. Workers strictly prefer to be employed at the prevailing wage $(w/P)_{NS}$ and to exert effort rather than to shirk.
3. There is involuntary unemployment with $(\overline{N} - N_{NS})$ unemployed workers wanting to work at the current wage but being unable to find jobs. Firms do not have an incentive to employ these workers because full employment encourages workers to shirk. Full employment would imply that a fired worker can walk across to the next firm and get another job. Unemployment is the stick that is essential in order to keep the workers who earn higher than the market-clearing wages in line. Unemployment keeps the worker from shirking on the job.

Why do firms pay a higher wage? The New Keynesians such as Shapiro and Stiglitz offer an explanation of unemployment, which emanates from the demand side of the labour market. They argue that firms choose to pay a wage higher than the market-clearing wage because of their limited ability to detect shirking on the job and that the higher wage is intended to provide an incentive to workers to exert more effort and raise the output of the firm. Even if workers, in Friedman fashion, offer to work for lower wages, the efficiency wages approach argues that firms do not want to reduce wages. In the Friedman version of why wages may be higher than market-clearing levels, it is the workers who set them that way to neutralize the expected rise in cost of living. The nominal wage is chosen by workers in the Friedman model, whereas it is chosen by firms in the New Keynesian model. In the classical model, by contrast, the firm chooses employment to maximize profit, taking the real wage as given.

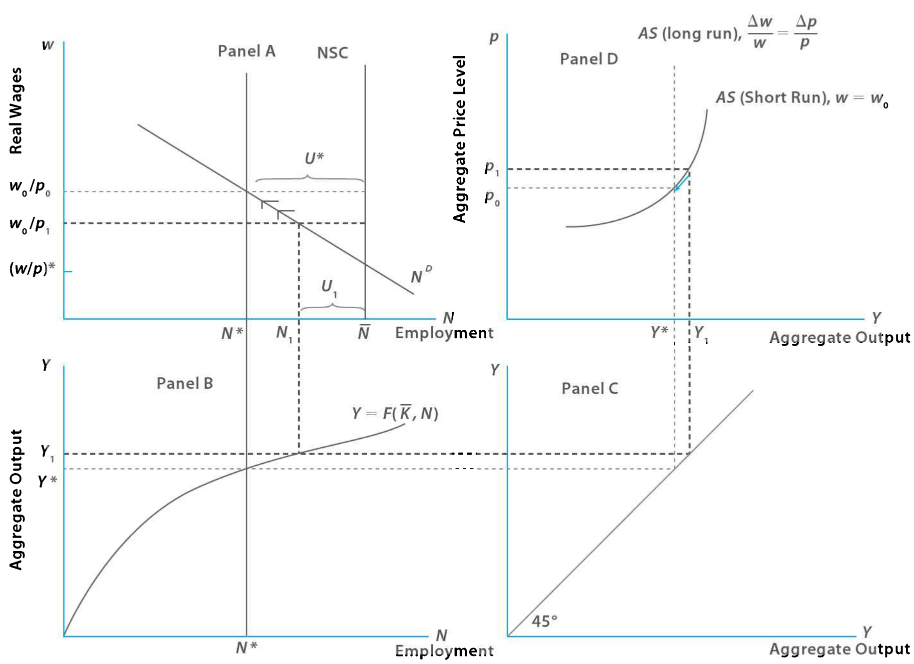
## 10.2 The Long-Run and Staggered Wage Setting

Efficiency wages account for the existence of unemployment in equilibrium just as the Friedman mistaken expectations model does. However, it does not result in an upward-sloping supply curve of output,[7] and to get this type of outcome, the Keynesians assume that nominal wages are chosen less frequently than employment is chosen by the firm. The nominal wage is said to be STICKY. The firm sets its wage rate less frequently than it hires or fires workers.

› The New Keynesians argue that nominal wages are **sticky** in the short run.

Let the initial nominal wage and price be such that the real wage is $w_0/P_0$ and let unemployment be at the natural rate where $U^*$ individuals are unemployed (see Figure 10.2, panel A).

Now suppose the price level increases to $p_1$. As nominal wage adjustment takes time, in the short run the wage will remain at its nominal level $w_0$. In real terms, the real wage has declined to $w_0/p_1$, and the firms respond by recruiting more labour lower down on their demand-for-labour curve. Unemployment declines




**› Figure 10.2**
The New Keynesian Supply Curve. Unemployment is at its natural rate $U^*$ at the prevailing real wage $w_0/p_0$ in panel A. With sticky nominal wages, firms increase their demand for labour from $N^*$ to $N_1$ when the price level rises to $p_1$ in panel A. The increase in employment increases the output in panel B. The higher price level $p_1$ is associated with an increased output $Y_1$ in panel D. As prices rise, the output increases resulting in the upward-sloping *AS* (short-run) curve in panel D. Over the long run, as the nominal wages adjust, employment and output move towards the natural rate as given by the vertical *AS* (long-run) curve.

to $U_1$ and so the output rises as traced out via the production function part of the figure (see panel B in Figure 10.2). Conversely, if there is a decline in the price, the real wage rises, firms employ less, and there is a rise in unemployment along with a decline in output. In the short run then, we have an upward-sloping supply curve, which results due to nominal wage stickiness, $AS\ (w = w_0)$.

When the real wage equals the efficiency wage $w_0/p_0$, unemployment is at its natural rate $U^*$ and the quantity of output produced is the natural rate of output $Y^*$. New Keynesians attempt to differentiate themselves by referring to point $(w_0/p_0, N^*)$ in Figure 10.2, panel A, as the QUASI-EQUILIBRIUM RATE OF UNEMPLOYMENT. It is an equilibrium because the point is a solution to the profit maximization and NSC equations, and the quasi term comes about because there is aggregate excess supply of labour. Such a quasi-equilibrium point is the only point at which the demand for labour is consistent with wage-setting behaviour that deters shirking. With deference to the New Keynesians, we refer to the point as the Keynesian natural rate of unemployment. This natural rate of output is an outcome of the fundamentals of the economy—preferences, endowments, and technology—and is unaffected by the money supply or any other factor that affects aggregate demand.

› The New Keynesian natural rate of unemployment is an equilibrium labour market outcome where wages deter shirking behaviour, firms maximize profits, and an excess supply of labour exists.

Suppose the price level is $p_1$ and unemployment is below the natural rate. The New Keynesians assume that adjustment begins with some fraction of firms adjusting their nominal wages over any given period of time.

### *10.2.1* Staggered Wage Contracts

As emphasized by Taylor,[8] contract decisions are STAGGERED in that all wage contract decisions in the economy are not made at the same point in time. First, wages are set for a predetermined length of time (contract period) because the cost of frequent adjustments to wages and salaries is high. Firms set wages by keeping in mind labour productivity and forecasts of revenues in the near future. It would be prohibitive to frequently review worker's performance and seek information about wages paid elsewhere and renegotiate wage contracts. Second, the wage setting is also staggered, that is decisions on wages across the economy are not synchronized or coordinated, and workers in different industries set and adjust wages at different points in time. There is an informational benefit to this process as the staggered wage setting provides workers and firms information about wages and prices in other parts of the economy. In a decentralized market economy, if all wages were set at the same time, there would be a lack of information about wages elsewhere and workers and firms would not be in a position to know what their wages are relative to other industries and occupations. When wage decisions are influenced by relative wage considerations, a non-sychronized wage setting is advantageous, as then the wage setting can be based on prevailing wages in comparable reference industries.

› The New Keynesians argue that the nominal wage setting is staggered so that wages are sticky in the short run.

Adjustments to wage contracts are generally staggered throughout the year. Suppose that wage contracts last a year and decision dates for renegotiating contracts are evenly staggered—half the contracts, say, are set in January this year, and half in January next year. There are then two cohorts of workers, each negotiating a wage contract for two periods in a staggered manner.

At time $t$, the first cohort sets its nominal wage for two periods, $t$ and $t + 1$, while at time period $t + 1$ the second cohort sets its wage for periods $t + 1$ and $t + 2$. In such a framework, the first period of the first cohort's contract will overlap with the second cohort's contract set at time $t - 1$. Furthermore, the last period of the first cohort's contract will also overlap with the contract of the second cohort which will be set at time $t + 1$. In such a two-period staggered wage setting process, firms face average wage costs[9] given by the simple average of the current contract wage ( $X_t$ ) signed in period $t$ and previous contract wages signed in period $t-1$ ( $X_{t-1}$ ),

or,
$$w_t = \frac{1}{2}(X_{t-1} + X_t)$$

where $X_i$ is the contract wage settlement signed in period $i$. The contract wage set in a period need not be a function of just currently observed variables. If agents are forward looking, they will estimate the average wage in the next period as well as the tightness of the labour market as given by future unemployment. If workers expect wage settlements in the next period $t + 1$ to be moderate, they will also attempt to conclude moderate wage settlements in period $t$. Also, if there is an expectation that the economy will enter into a downturn next period, which raises future unemployment, there will be a tendency to restrain wage increases in this period.[10]

### *10.2.2* Wage Adjustment

If wage contracts were *synchronized* rather than *staggered,* it would allow aggregate wage responses to be induced faster when prices change, and with all firms responding quickly to price changes, it would be akin to a classical supply curve. As firms adjust their wages, they will also adjust their employment levels. As time passes, more firms adjust their nominal wages and the economy-wide average wage slowly SHIFTS BACK TO EQUILIBRIUM. The sequences of events are illustrated by the arrows in Figure 10.2. A similar sequence of adjustment occurs when the prices decline and the real wage temporarily rises.

› In the long run, as wages adjust, the labour market moves towards the natural rate of employment and output.

Firms renegotiate nominal wages when the unemployment rate moves away from the natural rate because at a real wage less than that ($w/p < w_0/p_0$), there are fewer workers who are willing to work and not shirk than the number demanded by firms in the industry and the wage would rise. Similarly, if the wage was higher than the efficiency wage, there would be more workers willing to work and not shirk than are being demanded and the wage would fall. The efficiency wage after all is set above the market-clearing wage so as to make it costly for workers to shirk. Unemployment, to reiterate, is a disciplining device that keeps employed workers in line.

**MACROFOCUS 10.1**

**Henry Ford and Efficiency Wages**

The Ford Motor Company that was founded in 1903 had a highly skilled workforce. Ford began by assembling cars rather than manufacturing them, and the parts were produced by outside machine shops with low tolerances. This required a lot of shaping and fitting to get them together properly, which, in turn, required skilled metal-working craftsmen. Ford then decided in 1908 to make the Model T, where one automobile would be just like another automobile. He decided to do this by having parts made with high tolerances so that skilled fitting would no longer be required. The workshop was also re-designed to reduce unnecessary movements of workers and parts. The scale of production increased tremendously from 450 employees in 1908 to 14,000 in 1913, and the output increased 25-fold over this period. In the process, tasks got routinized and the work became menial. The requirement for workers to be in lockstep so that the assembly line moved smoothly increased the pressure on them. By 1913, workers' dissatisfaction and turnover increased drastically so that Ford hired 50,448 men in the course of the year to maintain the average labour force at 13,623. In January 1914, Ford announced a reduction in the working day from 9 to 8 hours and a rise in daily payments from $2.34 to $5.00 a day. He also made it clear that inefficient workers would be discharged and that if the company fortunes dipped, then it would not be bound to continue such wage payments.

Soon after the announcement, there were long queues for jobs outside the plant and the crowd had to be managed with water cannons. An excess supply of labour seeking jobs became the norm at Ford. There was a pronounced decline in absenteeism, and the turnover rate of 370 per cent declined to 16 per cent. The productivity per worker increased by over 40 per cent and profits by about 20 per cent. Fifteen per cent more cars were produced per day in the span of a year.

Thus, the wage hike was sufficient to ensure that it exceeded workers' opportunity cost of working and allowed Ford to ration jobs. It provided substantial productivity benefits as well. An increase in the wage reduced turnover among workers and extracted more productive effort from them. To sum up, ford derived the benefits of paying efficiency wages.*

* D. M. G. Raff and L. H. Summers, "Did Henry Ford Pay Efficiency Wages?" *Journal of Labour Economics* 5, no. 4, Part 2 (1987): 557–585.

## 10.3 The Insider–Outsider Approach

In the efficiency wage approach, it is not in the interest of the firm to reduce wages. In the insider–outsider approach, it is not in the interest of incumbent workers to do so. Insider–outsider theories argue that various types of labour turnover costs create rents and market power for incumbent workers in existing firms, who are called the insiders. This allows insiders to push their wages above the market-clearing wage without them losing their jobs.[11] INSIDERS are experienced incumbent employees whose positions are protected by labour turnover cost. Outsiders are taken to be the unemployed. Rather than being used as the descriptive terms employed and unemployed, the terms insider and outsider are used as analytical categories. These highlight the asymmetric position of incumbent workers who have market power emanating from labour turnover costs unlike the unemployed who lack market power.

› Insiders are incumbent employees who have market power in the determination of wages. This market power arises from the insiders' ability to exploit the turnover costs of hiring and firing in the wage bargain compared to outsiders.

The traditional labour turnover costs are the costs associated with hiring and firing labour. Hiring costs include the costs of searching, screening, negotiating with, and training newly hired workers. Firing costs include costly firing procedures and severance pay that are often part of a job security legislation. Labour turnover costs also arise when insiders refuse to cooperate with outsiders who try to get jobs by underbidding the wages of insiders. If an unemployed worker offers to work for a lower wage, insiders can THREATEN NOT TO COOPERATE with the underbidder, thus making their productivity so low that the firm does not find it worthwhile to hire them. Apart from making it unprofitable to the firm to hire someone who offers to work at a lower wage, incumbents can also make life disagreeable for the underbidder by harassing them for breaking into the firm by underbidding. Outsiders then face less favourable working conditions than insiders do, and these discriminatory conditions result in a disinclination on the part of the outsiders to engage in underbidding because they know that if they did so, the insiders would HARASS them and raise their cost of accepting employment with the firm.

› Insiders can also **refuse to cooperate** with outsiders who attempt to underbid prevailing wages and thereby reduce labour productivity. Firms then find outsiders unattractive to hire.

› **Harassment** also raises the opportunity cost of employment and deters outsiders from offering their labour services.

Let $N_I$ and $N_E$ represent the labour input of the insider and the entrant, respectively, and their real wage be given by $\omega_I$ and $\omega_E$. The production function of the firm[12] is given by $F(N_I + N_E)$. Standard demand-for-labour curves are downward-sloping curves representing the marginal product of labour. For the case of an insider's demand curve (IDC), the demand for labour will be the sum of the marginal product of labour, $MP_{N_I} = F'$, plus the marginal cost of firing ($F$) that labour, $MC_F$. Similarly, the demand curve for entrants (EDC) is the marginal product of entrants, $MP_{N_E} = F'$, minus the marginal cost of hiring ($H$) workers, $MC_H$. Thus,

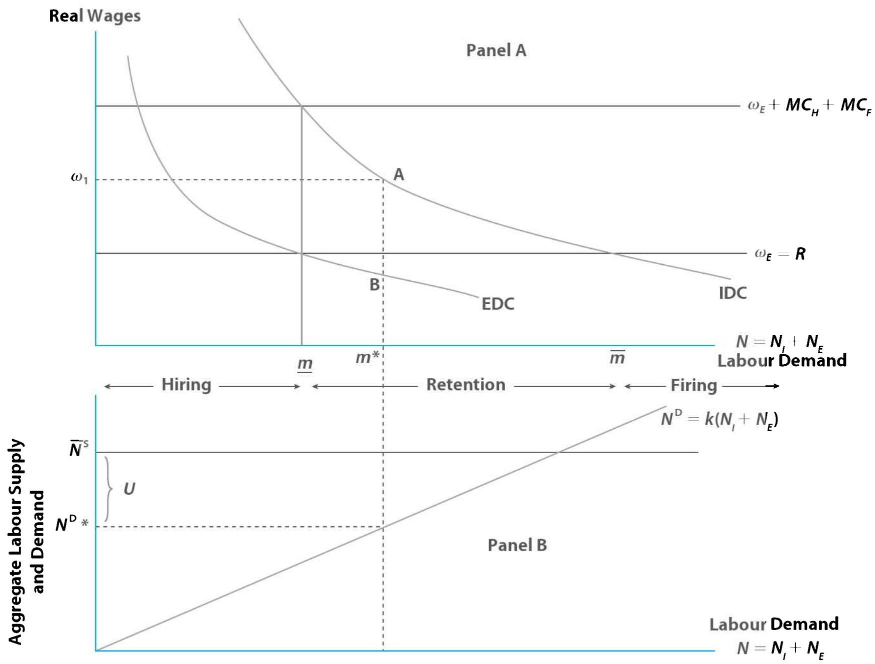
$$\text{IDC} = F' + MC_F$$

$$\text{EDC} = F' - MC_H$$

If there were no labour turnover costs, $MC_F = MC_H = 0$, and the traditional demand curve for labour would be given by $F'$, which lies in between the IDC and the EDC curves in Figure 10.3.

Let the wages of entrants be equal to their reservation wage,[13] $\omega_E = R$, as shown by the horizontal line in panel A of the figure.[14] What is the maximum wage, above the reservation wage of entrants, that the insiders can demand without losing their jobs?

The maximum wage that insiders may obtain, without being replaced by outsiders, is the entrant wage plus the sum of the marginal costs of hiring and firing labour, $\omega_E + MC_H + MC_F$. Any insider wage above this makes




**› Figure 10.3**
Employment with Insiders and Outsiders. **The EDC is the marginal productivity of labour minus the marginal cost of hiring in panel A. The IDC is the marginal productivity of labour plus the marginal cost of firing. With reservation wage rate $\omega_E$ for an entrant, incumbents are profitable to the firm and will not be replaced by entrants up to a wage given by $\omega_E$ plus the marginal costs of hiring and firing in panel A. In this case, insiders will retain their jobs (between $\underline{m}$ and $\overline{m}$). When the productivity of insiders falls below the reservation wage rate (after $\overline{m}$), it is profitable to the firm to fire them. Hiring will occur to the left of $\underline{m}$. The employment per firm at $m^*$ multiplied by the number of firms gives the aggregate demand for labour in panel B. Subtracting the aggregate demand from the aggregate supply $N^S$ gives the extent of unemployment.**

it profitable for the firm to replace insiders by outsiders. We can refer to $\omega_E + MC_H + MC_F$ as the relative profitability constraint or the maximum wage that insiders may obtain without being replaced by outsiders.

Let the number of insiders at the time of analysis be $m^*$. Instead of bargaining over the rents with firms, assume that insiders acquire the entire rent generated by the marginal turnover costs ($MC_H + MC_F$). With $m^*$ insiders, the wages that they are able to demand will be given by $w_1$, giving point A on the IDC. At this wage, all insiders get to keep their jobs and no entrants find jobs. This is because their marginal productivity at point B, after deducting for labour turnover costs, is below the entrant wage $\omega_E$.

Hence, if there are any unemployed workers between $\underline{m}$ and $\overline{m}$, they will remain unemployed.

Another case could be that if instead the number of insiders currently in the firm is lesser than $\underline{m}$, the firm will hire more workers, a hiring situation.

Finally, if the number of insiders currently in the firm is greater than $\overline{m}$, some insiders will be fired even if they accept an entrant's wage, because their productivity, as given by the IDC, will be below $w_E$, a firing situation. For any initial employment level between $\underline{m}$ and $\overline{m}$, the number of employees will remain unchanged, a retention situation.

Let us look at an aggregate level now. The simplest way to represent the situation is to assume that there are $k$ identical firms so that the aggregate demand for labour is given by $N^D = k(N_I + N_E)$. At the labour demand $m^*$ of the individual firm, the aggregate labour demand will be $N^D*$, which can be seen in panel B of Figure 10.3. With aggregate labour supply in the economy given by $\overline{N}^S$, unemployment is given by $U = \overline{N}^S - N^D*$.

› Established firms that set wages on an insider–outsider basis are susceptible to competition from the **new firms**, which have no insiders and can employ workers at lower reservation wages.

The insider–outsider theory provides an alternative explanation for why underbidding of wages by the unemployed may not help them gain entry into the labour market. It is not in the interests of incumbents since they exploit the rents generated in connection with the turnover costs to keep potential entrants out. However, the insider–outsider theory fails to account for NEW FIRMS, which having no insiders could hire unemployed workers at their reservation wages, and undersell established firms. The potential entry of a new firm could then cause insiders to lose their jobs and lead them to temper their wage demands and shift the IDC curve to the left. However, such entry may not be able to ALLEVIATE UNEMPLOYMENT substantially if there are obstacles to the entry of firms, such as limited entrepreneurial skills, risk aversion, and credit rationing, which restricts the availability of equity capital and borrowed funds. Apart from the new firms' entry, established firms may go in for plants at other locations and pit the insiders in one plant against those in another location. It can do this by making an offer that the plant that settles for a lower wage would remain open while the other would shut down if their insiders insist on keeping their wages the same. Of course, firms lose out on economies of scale if they split production across various units of small size, and centralized unions operating above the level of individual plants could prevent any wage concessions sought by the firm in this manner.

› The rate of unemployment is dependent on the barriers to entry faced by new firms in the economy.

**MACROFOCUS 10.2**

## Macroeconomic Policy with Insiders

What macroeconomic policy is appropriate when employment is determined by insider–outsider effects? Suppose there is a supply-side shock—a sudden increase in the prices of intermediate imported inputs such as oil or a natural disaster such as a drought—that adversely affects the marginal productivity of labour. This shifts the marginal productivity insider demand curve to the left, the dashed IDC curve. If the shock occurs after insider wages have been set, employment declines at the constant insider wage to point C. Insiders, whose initial number is $m^*$, would normally negotiate wages on the basis of market power and obtain wage rate $\omega_1$.

Suppose the shock occurs prior to wage setting by insiders. Then, all insiders could club together out of anxiety thinking they may lose their jobs and they would accept a wage reduction to $\omega_2$—point B where employment is still $m^*$.

The other extreme possibility is that insiders insist on maintaining the wage prior to the shock, $\omega_1$. Some workers lose their jobs. If there is a seniority system that governs firing decisions, then it would be a common knowledge as to who the laid-off workers would be. If a majority of workers continue to be employed at an unchanged wage, this outcome is likely. Laid-off workers are not in a position to save their job by bidding for lower wages as insiders would threaten not to cooperate with them, which reduces their productivity and attractiveness to the firm. The creation of a hostile work environment for those who underbid also raises the reservation wages of laid-off workers who underbid.

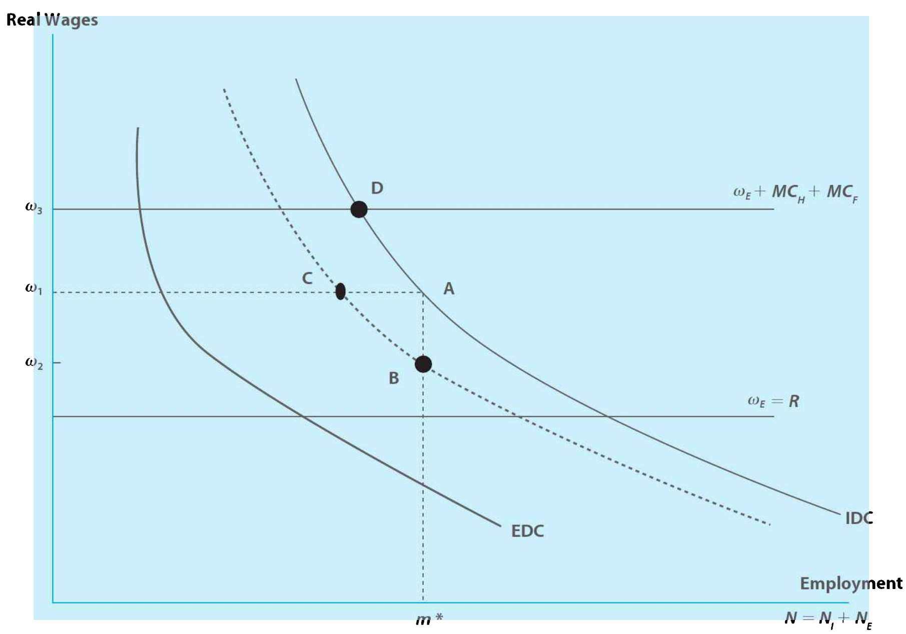
If later a favourable supply-side shock occurs, the labour market scenario might return to that depicted by the curve IDC. In that case, if insiders exploit the improved labour market situation and push for higher wages, they can attain wage rate $\omega_3$ without being displaced by other workers. This is point D in Figure 1 where the employment increase is modest. A favourable shock may not increase employment, while an unfavourable shock reduces employment. Wages, however, show an upward trend over time, even as employment reduces.

In such a situation, reducing the barriers to entry of firms may be an important correlate of demand-management policy.

Lower barriers to entry by firms in the United States compared to firms in Western Europe could be a reason for the recovery in US employment from the recession in the late 1970s and early 1980s, compared to European employment. (An increase in the number of firms $k$ in Figure 10.3 reduces unemployment.) Also, government expenditures on infrastructure by raising the marginal productivity of labour could have a larger impact on the labour market than spending by governments, which do not have such supply-side effects such as the transfer payments that have increasingly dominated European government budgets.*

**› Figure 1**

Employment and Insider-Outsider Effects.




*A. Lindbeck and D. J. Snower, "Long-Term Unemployment and Macroeconomic Policy," *American Economic Review* 78, no. 2 (1988): 38–43.

## SUMMARY

- Firms pay higher than the market-clearing wages to elicit effort, reduce turnover, recruit better workers, and improve worker morale.
- When firms have difficulty monitoring employees, they have an incentive to offer a wage-employment package that deters workers from shirking on the job.
- Efficiency wages are the wages that maximize firm profits when it is difficult to monitor whether workers work or shirk.
- Efficiency wages are higher than the competitive market wage and generate involuntary unemployment.
- The New Keynesian natural rate of unemployment is an equilibrium labour market outcome where wages deter shirking behaviour, firms maximize profits, and there is an excess supply of labour.
- The New Keynesians argue that nominal wage setting is staggered so that wages are sticky in the short run.
- In the long run, as wages adjust, the labour market moves towards the natural rate of employment and output.
- Insiders are incumbent employees who have market power in the determination of wages. This market power arises from insiders' ability to exploit turnover costs of hiring and firing in the wage bargain compared to outsiders.
- Insiders can also reduce labour productivity by refusing to cooperate with outsiders who attempt to underbid prevailing wages. Firms then find outsiders unattractive to hire. Harassment by insiders also raises the opportunity cost of employment and deters outsiders from offering their labour services.
- Established firms setting wages on an insider–outsider basis are susceptible to competition from new firms that have no insiders and can employ workers at lower reservation wages.
- The rate of unemployment depends on the barriers to entry faced by new firms in the economy when insider effects prevail in the labour market.

## NOTES

1. Lawrence F. Katz, "Efficiency Wage Theories: A Partial Evaluation," *NBER Macroeconomics Annual* 1 (1986): 235–276.
2. Harvey Leibenstein, "The Theory of Underemployment in Backward Economies," *Journal of Political Economy* 65 (April 1957): 91–103.
3. Carl Shapiro and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74 (1984): 433–444.
4. When firms want to hire workers, they choose at random from the pool of unemployed workers. Since workers are considered to be identical, the chance of finding a job does not depend on how workers become unemployed, or on how long they are unemployed.
5. The marginal product is positive and subject to the usual condition of diminishing marginal productivity. Thus, the production function has the following properties: $F' > 0$, $F'' < 0$.
6. We have presented the results as a result of decision by one firm. For ease of exposition we have concentrated on employment per firm. The number of workers unemployed is the number of firms times the number of unemployed per firm.
7. Since efficiency wages are real efficiency wages, an increase in the price would be consistent with a rise in the nominal wage such that all real variables are the same. Any price change in the efficiency wage model can result in no change in the employment and output, a vertical supply curve.
8. John B. Taylor, "Staggered Wage Setting in Macro Model," *American Economic Review* 69, no. 2 (1979): 108–113.
9. Taylor argues that the average real contract wage is increasing in the level of real economic activity, and he uses it to argue further that inflation dynamics in an economy is influenced by the interaction of staggered contracts apart from the well-known expectations mechanism.
10. Wage settlements in period $t$ can then be written as $X_t = (1/2)\{[(1/2)(X_{t-1} + X_t)] + [(1/2)(X_t + X_{t+1})]\} - (\theta/2)[(U_t - U^*) + (U_{t+1} - U^*)]$, where $\theta$ is the sensitivity of wages to unemployment, and $U^*$ is the natural rate of unemployment. Solving for $X_t$, $X_t = [(1/2)(X_{t-1} + X_{t+1})] - \theta[(U_t - U^*) - (U_{t+1} - U^*)]$. Hence, wages may be determined on a forward-looking basis.
11. Assar Lindbeck and Dennis J. Snower, *The Insider-Outsider Theory of Employment and Unemployment* (Cambridge, MA: MIT Press, 1988); "Interactions Between the Efficiency Wage and Insider-Outsider Theories," *Economics Letters* 37 (1991): 193–196.
12. Insiders can be presumed to be more productive than outsiders without any damage to the basic results.
13. The reservation wage is the minimum increase in income that makes a person indifferent between working and not working. A person will not work at all if the market wage is less than the reservation wage, and the person will enter the labour market if the market wage exceeds the reservation wage.
14. We do not deal with the cooperation-harassment version of the theory. However, lack of cooperation with outsiders who try to underbid wages keeps down their productivity and shifts the EDC curve to the left. Similarly, harassing underbidders is akin to pushing up the reservation wage of outsiders.

## TEST YOURSELF

1. Briefly outline the various approaches to explaining the phenomenon of involuntary unemployment.
2. Why do firms pay efficiency wages, and what prevents the market from clearing in an efficiency wage labour market?
3. Why are nominal wages sticky in the short run? How would the aggregate supply curve differ in the situation when the wage adjustment is slow as compared to the case where the wage adjustment is fast?
4. What is staggered wage setting and why is it prevalent?
5. Why is the long-run aggregate supply curve inelastic even though the short-run aggregate supply curve is elastic?
6. Differentiate between insiders, entrants, and outsiders in the labour market.
7. What sort of market power do insiders wield, and how does it impact on employment in the firm?
8. Distinguish between the hiring, retention, and firing phases of employment within a firm that is subject to insider power.
9. How does a supply shock affect the relationship between wages and employment in an insider–outsider model?
10. How successful would aggregate demand management be in dealing with unemployment when firms pay efficiency wages or wage setting is influenced by insider power?

## ONLINE APPLICATION

This chapter has argued that the labour market cannot be exclusively identified with productive relations due to the complexity of what is being exchanged in this market. For instance, the cooperation offered by senior workers to new employees plays a crucial role in inducting new workers into a firm and affects its productivity. Insider power has many manifestations. One such example is the implicit understanding in many public-sector institutions in India that the child or the spouse of an employee should be hired in their place in case of death, injury, or retirement. If the public sector is subject to insider–outsider effects, then there could be limits to its ability to generate employment. To understand this more, try the following:

1. (a) Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
   (c) Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy".
2. Click on "Table 14: Employment in Public and Organized Private Sectors" and create a spreadsheet of data on public sector and organized private sector employment for the years 1983–1984, 1987–1988, 1993–1994, 1999–2000, and 2004–2005. (In case the data for 2004–2005 is not entered on the Web site, take it to be 23.45 million for the public sector and 10.56 million for the organized private sectors.)
3. From the National Sample Survey Employment–Unemployment Surveys and the National Commission for Enterprises in the Unorganised Sector (NCEUS) Report on Conditions of Work and Promotion of Livelihoods in the Unorganised Sector (http://nceus.gov.in/), we have the following data on total employment:

| Year | Employment (millions) |
|---|---|
| 1983 | 303 |
| 1987–1988 | 324 |
| 1993–1994 | 374 |
| 1999–2000 | 397 |
| 2004–2005 | 457 |

4. Subtract the data on organized sector employment, as obtained in Step 2, from total employment to obtain the estimate of unorganized employment in India. The unorganized sector is an enterprise-based concept and is defined as "all unincorporated private enterprises owned by individuals or households engaged in the production and sale of goods and services operated on a proprietary or partnership basis and with less than ten total workers." (The NCEUS argues that an enterprise-based definition is insufficient as it does not reflect the characteristics of jobs. For instance, some workers in the organized sector may not enjoy employment security [casual workers], and may not be eligible for paid sick/annual leave or other social security benefits. The commission argues that though organized sector employment increased during 1999–2000 and 2004–2005, the entire net increase of 8.5 million employment was due to workers without social security benefits. Thus, the organized sector may have informal employment, an issue we sidestep for now.)
5. Estimate (a) the share of the public sector and unorganized sector employment in total employment and (b) the growth rate of employment between 1983–1984 and 2004–2005 for the public sector and the unorganized sector. What do these statistics reveal about the structure of employment in India? Could the relatively faster growth of public sector employment of the 1980s, which decelerated subsequently compared to unorganized sector employment, be due to insider–outsider effects?
6. Recall that in Chapter 1 we had found that the contribution of the public sector to GDP was about 25 per cent. Can you explain the very low employment contribution compared to that in output that occurred in the public sector?

# 11

# Open Economy I: The Mundell–Fleming Model

CRITICAL QUESTIONS

» *How does an open economy change the nature of the* IS–LM *analysis?*

» *How do controls on capital affect the balance of payments?*

» *What is the impact of the monetary policy under fixed and flexible exchange rates?*

» *What is the impact of the fiscal policy under fixed and flexible exchange rates?*

» *How can an economy deal with capital inflows?*

# 11.1 Accounting for an Open Economy

We now turn to the determinants of output and income when we have an open economy. Let us first recall some macroeconomic accounting that we had done initially. For an open economy, the aggregate output can be written as

$$Y = C + I + G + X - M \tag{11.1}$$

The aggregate spending by domestic residents is called absorption ($A$) and is defined as

$$A = C + I + G$$

Some of the output of the economy is absorbed by foreigners who spend on our goods (exports $= X$), and imports ($M$) are deducted because that part of consumption, investment, and government expenditures that we import are not part of domestic production. Exports less imports is also referred to as net exports ($NX$), or mathematically,

$$NX = X - M$$

Hence, the aggregate output in an open economy can be written as

$$Y = A + NX \tag{11.2}$$

With these identities in place, we can proceed to some theory of the open economy. The goods market now includes expenditure by residents (foreigners) on foreign (domestic) goods. If consumption is deemed to be a positive function of income, and investment and consumption a negative function of the interest rate, we can write $A(r, Y)$ as that part of the domestic absorption that depends on the aggregate output (income) and the interest rate. The other component of absorption is the government expenditure, which we take to be an exogenous policy variable. As the real exchange rate $(EP^*)/P$ depreciates, exports become cheaper and imports more expensive so that net exports increase with a rise in the real exchange rate.[1] Any increase in income also results in an increase in the demand for foreign goods (imports) so that net exports decline with a rise in income. The open-economy goods market equilibrium condition or *IS* curve is then given by

$$Y = A(r, Y; G) + NX\left(Y, \frac{EP^*}{P}\right) \tag{11.3}$$[2]

where $G$ is the exogenous level of government expenditure. In what follows, we treat the terms of trade, $P^*/P$, as fixed. We focus on the SHORT RUN during which price levels are treated as given. We scale the domestic and foreign price levels so that $P = P^* = 1$. Since the price level is fixed, the reference interest rate in this chapter is the real interest rate, $r$. Also, the exchange rate that affects net exports is now the nominal exchange rate, $E$.

› The Mundell–Fleming model focuses on the short run where prices are fixed. With fixed prices, the reference interest rate is the real interest rate and the reference exchange rate is the nominal exchange rate.

Equation (11.3) equating aggregate income to expenditure is an expression for the aggregate savings equals the investment condition for an open economy. To see this, note that we can write $Y = A + NX = C + I + G + NX$ as

$$Y - C - G - NX = I$$

If we subtract and add taxes less transfers—$T$—to the left-hand side,

$$(Y - T) - C + (T - G) - NX = I$$

Recall that $(Y - T)$ is private disposable income, and $(Y - T) - C$ is private savings ($S_{pvt}$). Also, $(T - G)$ is government savings ($S_{govt}$). Thus,

$$S_{pvt} + S_{govt} - NX = I$$

› The investment equals savings (*IS*) condition for an open economy is given by the combination of interest rate and income, where aggregate savings (private, government, and foreign) equals investment, or $S - NX = I$.

or, with NATIONAL SAVINGS written as $S = S_{pvt} + S_{govt}$

$$S - NX = I \tag{11.4}$$

The slope of an open-economy *IS* curve then is the ratio of two terms: In the numerator is the responsiveness of savings minus net exports to the level of income, and in the denominator is the responsiveness of domestic investment and consumption to the interest rate.[3] An increase in income increases savings, which increases the supply of funds in the financial market, and this reduces the interest rate. A reduction in the interest rate makes investment projects more profitable, and the investment rises to match the increased savings. However, the rise in income increases import demand and reduces net exports as well. This decrease in net exports increases the value of $S - NX$ relative to what it would be, $S$, if we were dealing with a closed economy. Hence, the reduction in the interest rate for a given increase in income must be larger when we consider an open economy so that the rise in $S - NX$ is matched by a corresponding rise in $I$. This implies that the *IS* CURVE has a steeper slope for an open economy than a closed economy.[4]

› The open-economy ***IS* curve** is steeper than the closed economy *IS* curve.

The *IS* curve of an open economy is also drawn with reference to a given exchange rate. As the exchange rate depreciates and induces increased exports and a consequent rise in net exports and income, the *IS* curve shifts to the right. The open-economy *IS* curve, therefore, shifts to the right with an increase in government expenditure as well as with a depreciation of the exchange rate.

The money market or *LM* curve for an open economy is given by the *equality* between money demand and money supply. Again, let us recall some of the accounting associated with the money supply. The money supply is the outcome of the interplay between the central bank and the private sector. The balance sheet of the central bank, as discussed in Chapter 2, is replicated in Table 11.1.

› **Table 11.1**
A Central Bank Balance Sheet

| Assets | Liabilities |
|---|---|
| Credits to Domestic Banks | Currency in Circulation, $C$ |
| Central Bank Credit to Government = Holdings of Government Debt (Including Loans to Government) *Less* Government Deposits, $G_{b/CB}$ | Reserves and Central Bank Balances of Domestic Banks |
| Advances (Credits) to Commercial Private Sector, $A_{CB}$ | Net Non-Monetary Liabilities of Central Bank,* $NML_{CB}$ |
| Foreign Exchange Assets, $FE_{CB}$ | |

*** Non-monetary liabilities include contingent liabilities, guarantees, and other claims, *less* fixed and other assets.**

Reserve money or base money or high-powered money as it is sometimes known is written as

Reserve money ($M_0$) = currency in circulation ($C$) + reserves and central bank balances of domestic banks

Similarly, we can aggregate the central banks' credit to the government, commercial banks, and the private sector as the domestic credit advanced by the central bank or

$$DC = G_{b/CB} + A_{CB} + \text{credits to domestic banks}$$

Hence, ignoring the non-monetary liabilities of the central bank, its balance sheet can be written as

**› Table 11.2**
Alternative Exposition of a Central Bank Balance Sheet

| | |
|---|---|
| Foreign Exchange Assets, $FE_{CB}$ | Reserve Money, $M_0$ |
| Domestic Credit Advanced, $DC$ | |

Hence,

$$M_0 = FE_{CB} + DC$$

is an alternative expression for RESERVE OR BASE MONEY. In a fractional banking system, the MONEY SUPPLY, as we saw in Chapter 2, is a multiple of the stock of high-powered or reserve money. In fact, when $m$ is the money multiplier, then

› The reserve or base money in an open economy is the sum of the domestic credit advanced by the central bank and its foreign exchange assets.

› The money supply is the money multiplier times the reserve money.

$$M^S = mM_0 = m[FE_{CB} + DC] \qquad \textbf{(11.5)}$$

Thus, any increase in the country's net foreign exchange with the central bank ($\Delta FE_{CB}$), if it is not sterilized through a countervailing reduction in domestic credit to the government and the private sector, $\Delta DC \neq -\Delta FE_{CB}$, would result in an increase in reserve money, $\Delta M_0 > 0$. This through the money multiplier would result in an increase in the money supply. It is because the central bank has the tools to offset the foreign exchange inflow via a pursuit of sterilization that we focus on the effect of changes in foreign exchange on the balance sheet of the central bank. The money demand function is unchanged in an open economy. Thus, money market equilibrium is given by

$$M^S = m[FE_{CB} + DC]$$

$$\frac{M^D}{P} = L(r, Y) \qquad \textbf{(11.6)}[5]$$

$$M^D = M^S = M \qquad \textbf{(11.7)}$$

The $LM$ curve is upward sloping as before, with the difference that the changes in the stock of net foreign assets of the central bank change the money supply.

# 11.2 Capital Mobility and the Balance of Payments

We need to distinguish how financially open an economy is before we look at open-economy $IS$–$LM$ issues. To do that, we consider the balance of payments (BP) of an open economy. We follow the Mundell–Fleming model,[6] where the current balance is determined independently of the capital account so that the achievement of overall balance requires adjustment in the domestic economy. The current account, which we take to be the net export function as already described, depends positively on the exchange rate and negatively on income, that is, $NX = NX\left(\underset{(-)}{Y}, \underset{(+)}{E}\right)$. The capital account depends on the extent of CAPITAL MOBILITY. We expect capital flows to be attracted to destinations that offer higher rates of return.

› Capital controls are legal restrictions on the ability of citizens to hold and exchange assets denominated in the currencies of other nations. This affects capital mobility, the degree to which finance can flow across a country's borders.

Thus, if $r^*$ is the exogenously given foreign interest rate that includes any expected depreciation of the domestic currency, then the NET CAPITAL INFLOW ($NKI$) will be an increasing function of the extent to which the domestic interest rate is greater than the one ruling abroad, inclusive of any depreciation expected in the value of the domestic currency. We have,

› Capital flows to destinations that offer higher returns. The net capital flows to an economy is, therefore, an increasing function of the extent to which the domestic interest rate is greater than the interest rate abroad (inclusive of the expected depreciation of the domestic currency).

$$NKI = K(r - r^*) = K(r) \qquad K'(r) > 0 \qquad \textbf{(11.8)}$$

We assume that exchange rate expectations are static.[7] If $NKI$ is positive, it means that domestic residents are selling more financial assets, such as bonds, to foreigners than buying from them. The BP is the sum of the current account and the capital account. If we ignore international transfers comprising net factor incomes (NFI) abroad, the current account coincides with net exports, $CA = NX$, and we can write

$$BP = NX(Y, E) + NKI(r - r^*) \tag{11.9}$$

The BP equilibrium obtains when the flow of capital is just sufficient to finance the current account deficit or absorb the surplus. Under a pure floating exchange rate regime, the sum of the surplus on capital and current account must be zero, that is, a surplus on one account must be balanced by a deficit on the other. A pure float thus requires the following to hold true at all times:

$$NX(Y, E) + NKI(r - r^*) = 0 \tag{11.10}$$

We now distinguish between three regimes of financial openness or capital mobility:

› The extent of capital mobility affects the slope of the $BP$ curve. When capital is completely immobile, the $BP$ curve coincides with the equilibrium on the current account.

1. A country that has capital controls and completely restricts the MOBILITY OF CAPITAL is in a regime of *capital immobility*. A change in the domestic interest rate vis-à-vis the foreign interest rate has no impact on foreign capital flows. We may write $(\Delta NKI/\Delta r) \rightarrow 0$. In the limit, $NKI(r - r^*) = 0$ and the $BP$ coincides with equilibrium on the current account. Thus,

$$BP = NX(Y, E)$$

› Perfect capital mobility results in a domestic interest rate equal to the foreign interest rate, $r = r^*$, and a horizontal $BP$ curve.

2. A country that does not place any restrictions whatsoever on the flow of capital is in a regime of PERFECT CAPITAL MOBILITY. In such a case, capital flows to the location where it earns the highest return. With domestic and foreign financial assets being perfect substitutes, portfolio adjustment is instantaneous and rates of return are equated across the world. With arbitrage in the capital markets and capital flows ensuring that $r = r^*$, perfect capital mobility implies large flows of capital for even a small deviation of $r$ from $r^*$. Thus,

$$\frac{\Delta NKI}{\Delta r} \rightarrow \propto$$

› Imperfect capital mobility results in finite flows in and out of a country due to interest rate differentials and an upward-sloping $BP$ curve.

3. In a regime of IMPERFECT CAPITAL MOBILITY, international interest rate differentials result in finite flows into or out of a country. In this case, small deviations from interest rate parity, $r \neq r^*$, do not result in a potential flood into or out of a country. Imperfect capital mobility can be due to capital controls, a limited supply of arbitrage funds, or where risk aversion results in a risk premium that increases with the flow of funds into the home country. As differences between $r$ and $r^*$ can exist, we have

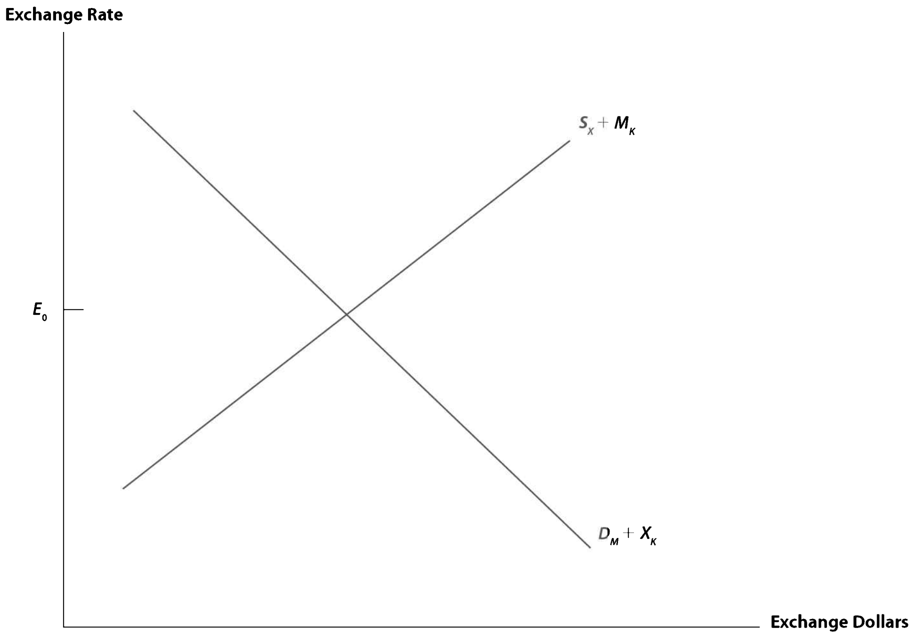
$$0 < \frac{\Delta NKI}{\Delta r} < \propto$$

The BP locus for imperfect capital mobility is derived in Figure 11.2 when exchange rates are flexible.[8] Here, variations in the value of the domestic currency ensure equilibrium in the BP. The exchange rate is determined by the BP equilibrium, as we saw in Chapter 2, where the demand for and supply of foreign exchange are equated. This occurs when the exchange rate adjusts until

$$NX(Y, E) + NKI(r - r^*) = 0$$

or

$$NX(Y, E) = -NKI(r - r^*)$$




› **Figure 11.1**
Exchange Rate Determination. The supply of dollars is by exporters $S_X$ and due to capital inflows or the import of capital $M_K$. The demand for dollars is by importers $D_M$ and due to capital outflows or the export of capital from the economy $X_K$. The exchange rate is determined where the supply of dollars equals the demand for dollars.

Under flexible exchange rates, the central bank need not intervene in the foreign exchange market. With no variation in the stock of net foreign exchange assets, changes in the money supply are solely caused by changes in domestic credit as in Eq. (11.5). We reproduce the BP diagram of Chapter 2 as Figure 11.1. To recapitulate the notation, let

$D_M$ : demand for foreign exchange by importers
$S_X$ : supply of foreign exchange by exporters
$X_K$ : demand for foreign exchange due to a capital outflow
$M_K$ : supply of foreign exchange due to a capital inflow

Recall that $NX = S_X - D_M$ and $NKI = M_K - X_K$. In the figure, the equilibrium exchange rate is $E_{o,}$ where the supply and demand for foreign exchange are equal.

The BP Eq. (11.9) has three unknowns—$r, Y, E$. The equation is hence graphed in three quadrants in Figure 11.2. Take panel B of the figure first, which plots combination of $r$ and $Y$ that are consistent with the BP equilibrium for a given value of the exchange rate $E$. The $BP(E)$ line is upward sloping. This is because as income increases, there is a growth in import demand, which reduces net exports, $NX$. As $NX\ (Y, E) = -NKI\,(r - r^*)$, preservation of equilibrium requires that the capital account must improve, which is achieved if domestic interest rates rise and there is an increase in capital inflows.

Thus, higher income is associated with higher interest rates for the BP equilibrium. The amount by which interest rates rise in order to offset a rise in income (the slope of the $BP$ line) depends on the interest elasticity of net capital flows.[9] The more elastic are capital inflows with respect to interest rates, the flatter is the $BP(E)$ line. If there were perfect capital mobility, a marginal interest rate increase is sufficient to stimulate an infinite flow of capital, and the $BP(E)$ line is completely flat. If the EXCHANGE RATE DEPRECIATES, net exports increase as domestic goods become cheaper for foreigners. This must be offset by a decline in capital inflows, as $NX = -NKI$, which requires a lower interest rate. A rise in $E$ (a depreciation of the currency) thus shifts the $BP(E)$ line down to the right. The exception is in the case of perfect capital mobility, where the $BP(E)$ line is flat and unaffected by changes in the exchange rate.

› A **depreciation of the currency** increases the net exports and shifts the *BP* curve down and to the right. An appreciation of the currency reduces the net exports and shifts the *BP* curve up and to the left.

› **Points to the left of the *BP* (*E*)** curve represent a BP surplus.

› **Points to the right of the *BP* (*E*)** curve represent a BP deficit.

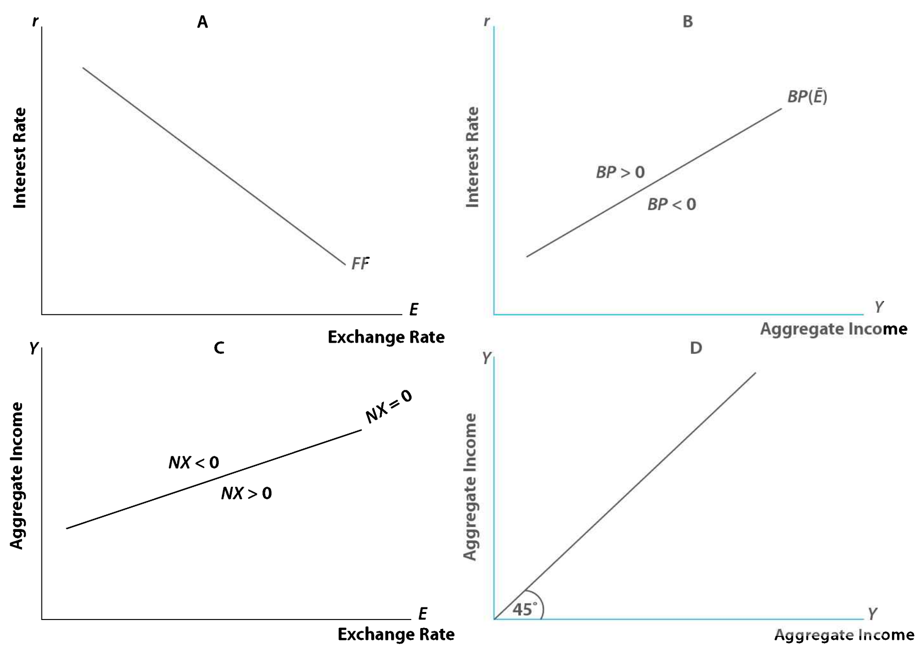
POINTS TO THE LEFT OF THE *BP*(*E*) curve are consistent with a BP surplus ($BP > 0$). For instance, for a given ($\bar{r}, \bar{Y}$) on the *BP* curve, if for the same level of income ($\bar{Y}$), the rate of interest rises to $\bar{\bar{r}} > \bar{r}$, then there is a capital inflow without any change in net exports resulting in a BP surplus. Similarly, if for a given interest rate ($\bar{r}$), income decreases to $\bar{\bar{Y}}$, there is an increase in net exports as imports decline. Without any change in the interest rate and net capital inflows, this results in a BP surplus. Symmetrically, POINTS TO THE RIGHT OF THE *BP*(*E*) curve are consistent with a BP deficit ($BP < 0$).

Panels A and C keep track of other aspects of the BP. In panel C, the line *NX* plots net exports. In the external sector of the economy, income along with the exchange rate determines the current account balance. To see why the *NX* line is upward sloping, ask what happens if the economy is at some point on the *NX* line and the income increases.

An increase in income gets partly spent on imports, which reduces net exports. For net exports to be in balance, we require the imports to be offset by an increase in exports, which occurs when the currency depreciates. Alternatively, we can, with reference to Figure 11.1, note that a rise in imports shifts the $D_M + X_K$ curve to the right and results in a depreciation of the currency.

It is important to keep in mind that there is no requirement for the economy to settle on the *NX* line in Figure 11.2 where the current account is in balance. The only requirement is that any current account deficit (surplus) is offset by a capital account surplus (deficit) of the same size so that the BP is in equilibrium. For any given fixed level of income, if the exchange rate depreciates, exports increase and the given level of income is associated with a current account surplus where $NX > 0$. Hence, points to the right of the $NX = 0$ curve are associated with surpluses, $NX > 0$, and points to the left with deficits, or $NX < 0$.

Panel A in Figure 11.2 plots the BP locus in ($r, E$) space. Its slope is derived by noting that at any point on the line, an increase in the interest rate results in a capital inflow that shifts the $S_X + M_K$ line in Figure 11.1 to the right and results in a decline in *E* or a currency appreciation. The *FF* curve in Figure 11.2




› **Figure 11.2**
The Balance of Payments under Imperfect Mobility of Capitals. In panel B, a rise in the interest rate induces capital inflows, which require an increase in income and offsetting imports to maintain the BP equilibrium. The *BP* curve is, therefore, upward sloping in the interest rate–income space. In panel C, a depreciation of the exchange rate increases exports, which requires an increase in the aggregate income and offsetting imports to maintain the balance of net exports. In panel A, an increase in the interest rate induces capital inflows and an increase in the supply of dollars that causes the exchange rate to appreciate, giving rise to the downward-sloping *FF* curve.

is, therefore, downward sloping. The *FF* curve is drawn for a given exogenous income *Y*. This is because the BP equilibrium requires higher interest rates and a more favourable capital account to be accompanied by a lower price of foreign currency and a less competitive foreign trade sector, as reflected in a deterioration in the current account. The *FF* curve thus slopes downwards for a given income level. As income increases, the rise in import demand shifts the $D_M + X_K$ curve to the right in Figure 11.1, which results in a depreciation of the currency. A given level of the interest rate is, therefore, associated with a higher price of foreign currency when income increases, and the *FF* curve in Figure 11.2 shifts to the right when income increases. As income increases, equilibrium requires depreciation to compensate for the additional import demand and *FF* shifts to the right. If capital were perfectly mobile, the *FF* line would be flat with no shift of the line when income increases.

We now consider the case of an economy where capital is completely immobile. In this case, $BP = NX(Y, E) = 0$. The economy must always maintain a trade balance where exports equal imports because the foreign currency required for financing international trade cannot be borrowed in international capital markets. This is because capital is not allowed to enter or exit the country. In this situation, the exchange rate clears the foreign exchange market for any given value of output and income that is determined in the domestic economy.

**MACROFOCUS 11.1**

## Capital Account Convertibility

If capital flows freely across national borders, interest rate differentials would be minimal, reducing the cost of capital, and making investment more efficient. So why are there capital controls? All advanced economies have open capital accounts. Most developing countries do not. Due to government policy, developing countries have pervasive distortions in markets. If the removal of these distortions is not practical, then it could be beneficial to have an additional distortion in the international capital market. Important policies in developing countries that impede the removal of capital controls include

1. Protection and subsidies—If certain industries are protected from global competition, it increases the price domestic producers receive relative to world prices. A liberal capital account would then attract investment into these protected industries where the value of investment at world prices is negative. A similar argument applies for sectors that are subsidized.
2. Moral hazard—In economies where there is substantial financial repression (see Chapter 14), the banking system is substantially owned by the government, which implicitly ensures the stability and viability of the banking system. Bankers do not pursue efficient screening and risk management since they are secure in the knowledge that the threat of bankruptcy is weak due to the protection of government. This creates an incentive for overlending and overinvestment (expected return less than expected cost), as banks do not factor in the cost if an investment does not generate sufficient returns as the government will step in if the downside risk occurs.* An open capital account in such a case leads to a greater-than-efficient investment.
3. Fixed exchange rates—A country that uses fixed exchange rates will have its inflation moderated by the inflation abroad. An economy that employs monetary policy to regulate economic cycles cannot have an open capital account. This is the impossible trinity of fixed exchange rates, independent monetary policy, and capital mobility (Chapter 5). Opening the capital account and continuing the practice of discretionary monetary policy requires letting go of pegging the exchange rate. This requires that inflation expectations, which were influenced by the inflation rate abroad during a pegged regime now have to be anchored in the credibility of the central bank and its monetary policy (Chapter 15). This, in turn, requires strong monetary institutions, including independence from the fiscal authority for the central bank. Government policies that are subject to short-term effects due to political pressures impose costs on the economy that forward-looking capital markets factor in and respond to via inflows and outflows. An independent monetary authority along with a well-supervised and regulated financial system ensures that macroeconomic policy is not jeopardized by the scrutiny of forward-looking capital markets. Capital account liberalization, the move to a floating exchange rate, and reforms for strong monetary institutions must be pursued at the same time.

For a review of capital account convertibility in the Indian context go to the Reserve Bank of India [RBI] Web site: (http://www.rbi.org.in/home.aspx).

Click on "Publications".

Click on "Reports".

Click on "Report of the Committee on Fuller Capital Account Convertibility," submitted in July 2006, under the chairmanship of S. S. Tarapore.

* E. D'Souza, "Prudential Regulation in Indian Banking," in S. Arumugam (ed.) *Indian Capital Markets—Modern Perspectives and Empirical Evidence* (Mumbai: Allied Publishers, December 2001), pp. 486–511.

**› Figure 11.3**
Balance of Payments under Capital Immobility. With perfect capital immobility, the interest rate plays no role in attracting capital inflows and the *BP* curve in panel B is vertical at the level of aggregate income, as determined in the domestic market. The economy must maintain a trade balance as the financing of deficits through capital flows is ruled out and the economy is always on the $NX = 0$ line in panel C. An increase in the aggregate income increases imports that must be matched by an increase in exports for trade balance. This requires the exchange rate to depreciate and the *FF* curve shifts to the right in panel A.

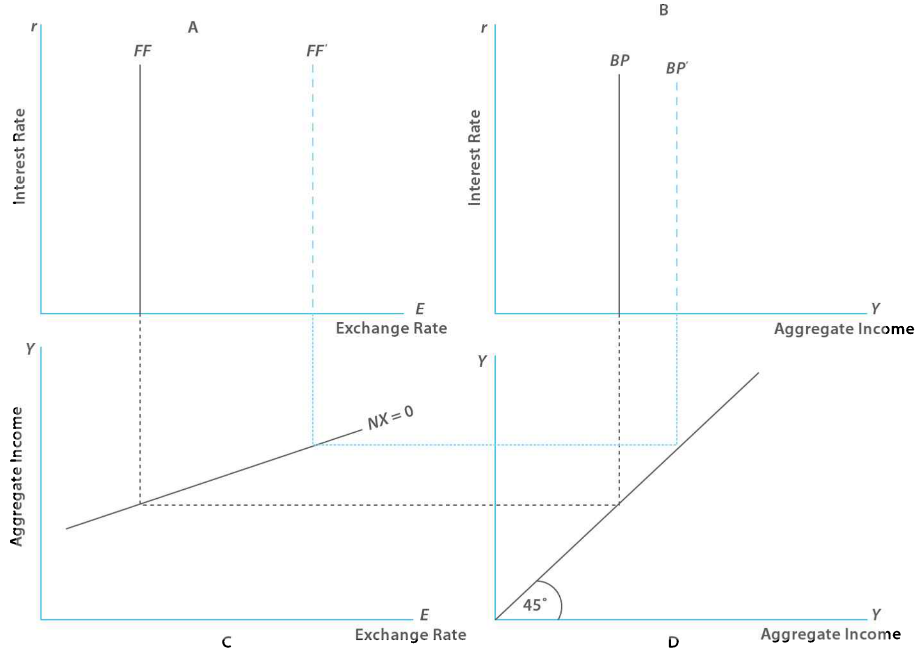
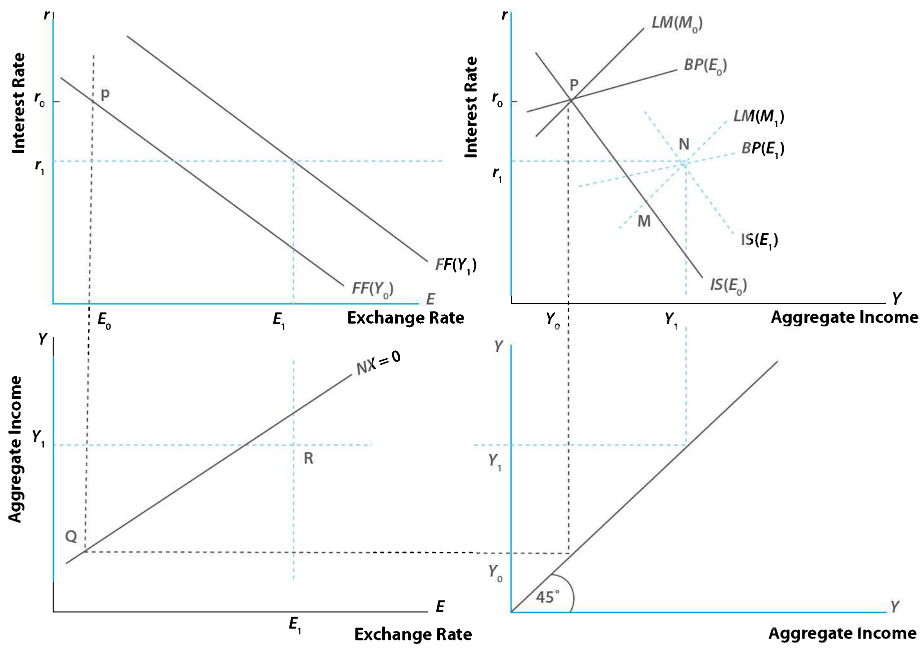



Figure 11.3 portrays the case of capital immobility. The *BP* curve has no role in the determination of domestic income and output, which is why it is vertical at the existing level of output $Y$.[10] If output increases, the *BP* curve shifts to the right to *BP′*, and tracing out via panel D, we can see that in panel C, it results in a more depreciated currency so that more exports are generated to finance the increase in imports induced by the rise in output. This rise in the exchange rate causes the *FF* curve in panel A to shift to the right to *FF′*.

## 11.3 Monetary Policy under Flexible Exchange Rates

Essentially the Mundell–Fleming model is the open-economy extension of the Keynesian model. It deals with a demand-constrained economy where the output adjusts passively. It thus requires the burden of adjustment to aggregate demand fluctuations to fall on the level of economic activity rather than on the price level. It thus suffices to work in an *IS–LM* framework rather than in an *AD–AS* framework. It implies that the aggregate supply curve is taken to be flat, and with the price level fixed, we can simplify matters by setting $P = 1$.

The starting point of the analysis of an open economy is one of equilibrium where the *IS–LM–BP* curves intersect so that the goods, money, and foreign exchange markets are in equilibrium. We then isolate a policy variable—monetary or fiscal—and ask ourselves the question as to what will occur when this policy variable is changed. A change in monetary or fiscal policy is a shift in the *LM* or *IS* curve and has an immediate impact on the interest rate in the domestic financial market. The analysis, therefore, proceeds by identifying the direction of change in the interest rate in the money market as a result of the change in the policy variable. We know that a change in the interest rate affects the differential returns between domestic and foreign financial assets, $r - r^*$, and this influences capital flows. The capital flows in turn have an




› **Figure 11.4**
The Monetary Policy with Flexible Exchange Rates. The goods (*IS*), money (*LM*), and foreign exchange (*BP*) markets are initially in equilibrium at point P. An increase in the money supply shifts the *LM* curve to the right with a temporary equilibrium at point M. The reduction in the interest rate reduces capital inflows and the increased income increases imports and results in a deficit in the net exports. The BP is, therefore, in deficit, which causes a depreciation of the currency to $E_1$. The *BP* curve shifts downwards and the *IS* curve shifts to the right as competitiveness improves with the currency depreciation. Final equilibrium then is at point N. The increase in income and imports depreciates the exchange rate and causes the *FF* curve to shift to the right as well.

impact on the foreign exchange market. If exchange rates are fixed, then the capital flow affects the stock of foreign exchange assets and the money supply. If the exchange rate is allowed to vary, then the exchange rate will change in response to the capital flow. Similarly, the policy change affects the level of economic activity. A change in income results in a changed demand for imports and net exports. The change in demand for imports affects the demand for foreign exchange and also has an impact on the foreign exchange market. The third part of the analysis accordingly identifies the impact of the change in the interest rate in the money market on the foreign exchange market—foreign exchange reserves or the exchange rate.

Finally, the change in the foreign exchange market impacts on the goods or money market. If the exchange rate changes then this affects the competitiveness of exports and this shifts the *IS* curve. If foreign exchange assets rise then reserve money on the balance sheet of the central bank increases and this changes the money supply and shifts the *LM* curve. The final part of the analysis then identifies the impact of the CHANGE IN POLICY on the goods or money market and the corresponding final equilibrium.

› For those who would rather not keep track of many variables, it suffices to concentrate on the impact of a change in policy on the *IS*, *LM*, and *BP* curves alone. After this understanding has been garnered, the implications of the **policy change** on the *NX* and *FF* curves may be imbibed on a subsequent reading.

We begin the analysis of monetary policy with an initial equilibrium in Figure 11.4 at points P, p, and Q. At interest rate $r_0$ and exchange rate $E_0$, the BP is in equilibrium as points p and P lie on the *FF* and *BP* lines, respectively. Also, at the associated exchange rate $E_0$, the goods market clears along the curve $IS(E_0)$. Further, tracking the initial level of income from $Y_0$ down via the 45-degree line over to the *NX* line, we see that there is a zero current account deficit at the exchange rate–income combination of $E_0$, $Y_0$, with no net tendency for capital to leave or enter the country.[11]

We now consider the case of expansionary monetary policy. Since the price level is fixed in the *IS*–*LM* analysis, an increase in the nominal money stock is equivalent to a rise in the real money stock. The *LM* curve shifts down to $LM(M_1)$ as a result of the increase in money supply and the interest rate declines. At point M, interest rates have declined (resulting in a drop in net capital inflows), and there is a higher level of economic activity (resulting in a higher current account deficit). Hence at point M, the economy is facing a BP deficit. The increased import demand and the reduction in capital inflows cause the $D_M + X_K$ curve to shift to the right and the $S_X + M_K$ curve to the left, as in Figure 11.1, resulting in a depreciation of the currency. The *FF* curve in Figure 11.4 also shifts to the right as income increases, as equilibrium requires a currency depreciation to compensate for the additional import demand at any given interest rate as discussed previously. A BP deficit puts a pressure on the exchange rate to depreciate. The *BP* curve then begins to shift down to the right.

As the exchange rate depreciates, the competitiveness of the domestic economy improves, and with exports increasing, the *IS* curve shifts to the right to say $IS(E_1)$. The surge in demand pushes interest rates part of the way back to their original level at $r_0$. Finally, the economy settles at point N where the interest rate is $r_1$ and the BP is back in balance along the line $BP(E_1)$. Two effects influence the return of BP equilibrium. The partial rise of the interest rate back towards its original level reduces the deficit in the capital account to an extent and brings it to a level such that a surplus in the current account can cover it and restore the BP equilibrium. The current account surplus, which is point R in the diagram, in turn is the outcome of the depreciation in the currency being sufficiently large that it boosts exports more than the rise in imports influenced by the increase in income. The INCREASE IN MONEY SUPPLY in a floating exchange rate regime with imperfect capital mobility thus results in:

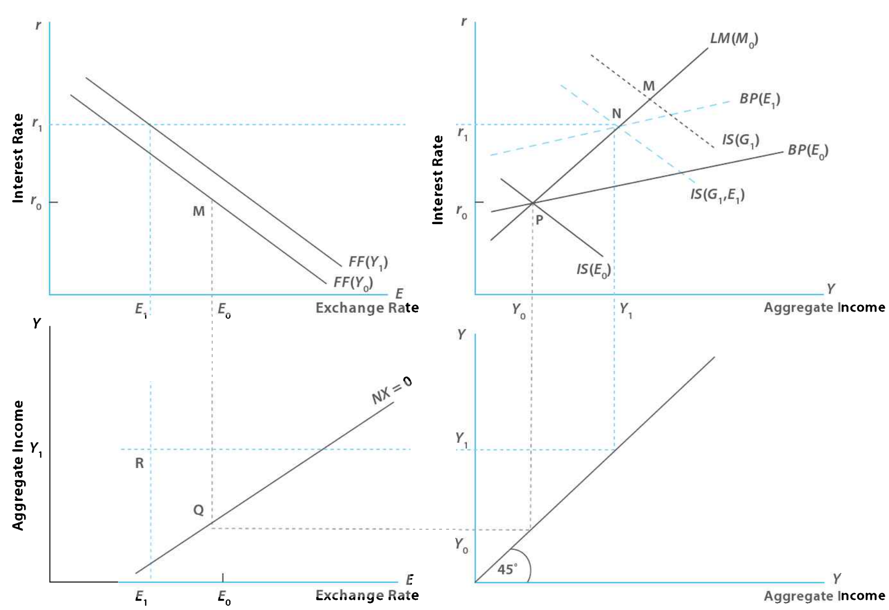
› An expansionary monetary policy under floating exchange rates causes a depreciation of the exchange rate, an increase in income, a decline in the interest rate, and an improved current account balance.

1. a depreciation of the exchange rate
2. an increase in income
3. a decline in the interest rate
4. an improvement in the current account balance

# 11.4 Fiscal Policy under Flexible Exchange Rates

A fiscal expansion results in an increase in government expenditure from $G_0$ to $G_1$ and shifts the *IS* curve to the right to $IS_1(G_1)$, see Figure 11.5. The impact is to raise both the interest rates and income initially to point M. As the economy

**› Figure 11.5**
Fiscal Policy with Flexible Exchange Rates. **The goods (*IS*), money (*LM*), and foreign exchange (*BP*) markets are initially in equilibrium at point P. A fiscal expansion shifts the *IS* curve to the right, and there is temporary equilibrium at point M. The increase in the interest rate results in a capital inflow, and the increase in income increases imports and causes a deficit in net exports. The capital inflow is greater than the deficit in the trade balance, and the BP is in surplus, causing the exchange rate to appreciate to $E_1$. The *BP* curve then shifts up as the currency appreciates and the *IS* curve shifts to the left as the competitiveness of the economy is adversely affected. The final equilibrium is at point N.**



is above the $BP(E_0)$ curve at point M, there is a BP surplus. This is due to the rise in the interest rate causing a capital inflow and a capital account surplus that is greater than the current account deficit due to the rise in imports associated with a rise in income. The rightward shift of the $S_X + M_K$ curve in Figure 11.1 must be greater than the rightward shift of the $D_M + X_K$ curve resulting in an appreciation of the exchange rate.

The *FF* curve also shifts to the right as income has increased. The appreciation of the exchange rate ($E_1 < E_0$) makes domestic goods more expensive to foreigners and reduces exports, thus shifting the *IS* curve down to the left to $IS(G_1, E_1)$ and reversing some of the rise in the interest rate. With the appreciation of the currency, the *BP* curve also shifts upwards to $BP(E_1)$ and the final outcome is at point N. On the trade front, there is a current account deficit at point R due to the decline in exports from an appreciation of the currency and a rise in imports due to a rise in income. In contrast to a closed economy, where the rise in interest rates crowds out some investment expenditure, in an open economy, there is an additional CROWDING OUT from the appreciation of the currency that crowds out spending by foreigners on domestic goods and reduces net exports. The ultimate expansion of demand is, therefore, smaller than that in a closed economy.

› A fiscal expansion in an open economy results in two types of crowding out: (a) crowding out of private investment and (b) crowding out of exports.

THE INCREASE in government expenditure in a floating exchange rate regime with imperfect capital mobility thus results in

1. an appreciation of the exchange rate
2. an increase in income
3. a rise in the interest rate[12]
4. a deterioration in the current account balance

› An expansionary fiscal policy in a floating exchange rate regime causes an appreciation of the currency, a deterioration of the current account, a rise in income, and a rise in the interest rate.

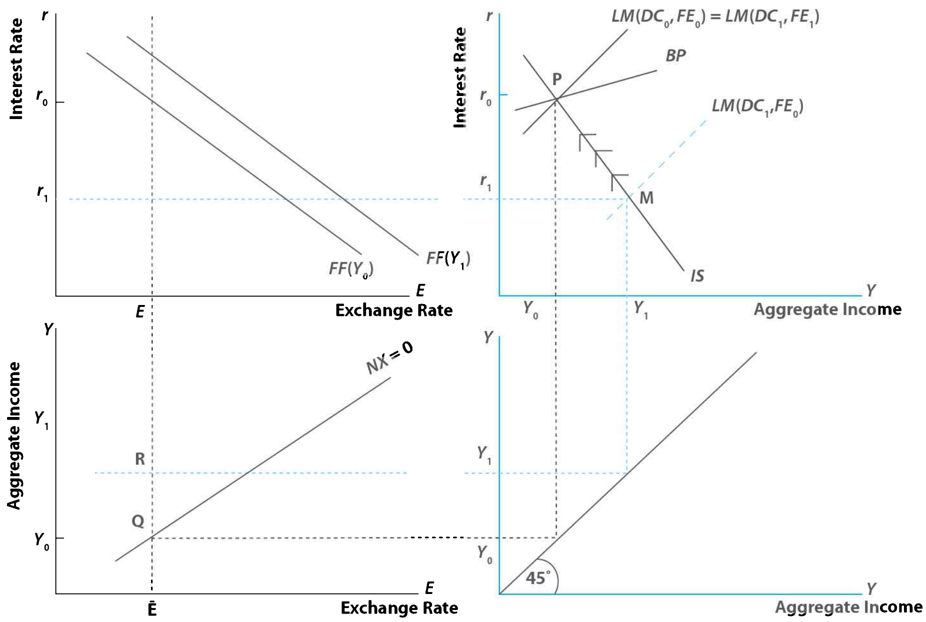
It can be concluded that for a floating exchange rate regime when capital is imperfectly mobile, MONETARY POLICY is more expansionary than fiscal policy because fiscal policy is associated with a deterioration of the current account balance. We would, therefore, expect that countries with a floating exchange rate regime would concentrate on monetary policy and deemphasize fiscal policy in economic policymaking for managing economic cycles.

› In a flexible exchange rate regime, monetary policy is more effective than fiscal policy in influencing aggregate income.

## 11.5 Fixed Exchange Rates and the Monetary Policy

We analyse the case of monetary policy with fixed exchange rates in Figure 11.6 where the initial equilibrium, as before, is at P. The exchange rate, however, is now fixed at $\bar{E}$. We now consider the case of a monetary expansion, which takes the form of an expansion of domestic credit[13] from $DC_0$ to $DC_1$. This shifts the *LM* curve to the right to $LM(DC_1, FE_0)$. The decline in the interest rate worsens the capital account, and the increase in income causes a deterioration in the current account, which goes into deficit at point R in the diagram. There results an overall BP deficit, which requires to be financed, and since the exchange rate is fixed, this financing comes from a running down of foreign exchange reserves with the central banks.

As the foreign exchange component of the monetary base reduces, the money supply reduces and the *LM* curve begins to shift back to its original position as indicated by the arrows in the diagram. Equilibrium is restored when the economy is back at point P. The only difference is that the composition of the money stock has changed. In the new equilibrium, there is a greater




**› Figure 11.6**
Monetary Policy with Fixed Exchange Rates. The goods (*IS*), money (*LM*), and foreign exchange (*BP*) markets are initially in equilibrium at point P. A monetary expansion shifts the *LM* curve to the right, and there is a temporary equilibrium at the point M. The interest rate has declined, which results in a decline in capital inflows, and the increase in income induces additional imports and a deficit in net exports. The *BP* is accordingly in deficit at point M. This deficit is financed by running down foreign exchange reserves at the central bank. The reduction in foreign exchange reserves results in a decline in the money supply, and the *LM* curve shifts leftwards. Finally, the economy achieves equilibrium back at point P where the *IS* and *BP* curves intersect.

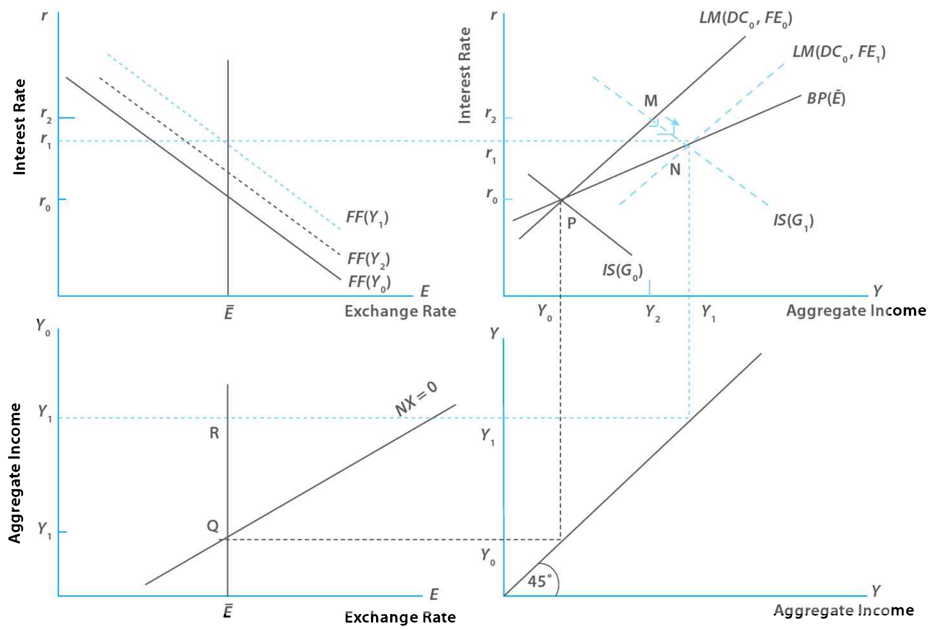
quantity of domestic financing of assets ($DC_1 > DC_0$), and a lower quantity of foreign exchange ($FX_1 < FX_0$). Of course, the central bank may attempt to prevent the money supply from declining at point M by further expanding domestic credit. This is called *sterilization*—the process of neutralizing the effect of a BP deficit (surplus) by creating (retiring) enough domestic credit to offset the fall in foreign exchange reserves. It achieves this by purchasing domestic government bonds to add to its stock of domestic assets and prevent the decline in its asset base as a result of the erosion in foreign exchange reserves. However, this cannot be continued indefinitely, as with reserves declining currency markets will build expectations as to whether the central bank has adequate reserves to defend the exchange rate or will resort to devaluation of the currency, and such expectations have destabilizing effects on the economy. Hence, a MONETARY EXPANSION in a fixed exchange rate regime results in

› The monetary expansion in a fixed exchange rate regime has no impact on macro variables such as the interest rate or income in the long run.

1. a decline in the interest rate, an increase in income, and a deterioration in current and capital accounts of the BP in the short run
2. a decline in foreign exchange reserves with no change in the interest rate, income, or the BP in the long run

## 11.6 Fixed Exchange Rates and the Fiscal Policy

An expansionary fiscal policy shifts the *IS* curve upwards, as in Figure 11.7, and the interest rate and income rise to point M where the new *IS* curve $IS(G1)$ intersects the unchanged *LM* curve. As point M is above the $BP(\bar{\bar{E}})$ line, it represents a BP surplus. Hence at M, the higher interest causes the capital account to improve by a greater amount than the deterioration in the current account caused by a rise in imports that is spurred by the increase in income. The interest rate is too high at $r_2$ for BP equilibrium to take place. However, as foreigners accumulate domestic financial assets to reap the benefits of the high rate of return on them, there is an increase in capital inflow that causes foreign exchange reserves to rise. As this component of the money supply




**› Figure 11.7**
Fiscal Policy with Fixed Exchange Rates. The goods (*IS*), money (*LM*), and foreign exchange (*BP*) markets are initially in equilibrium at point P. An expansionary fiscal policy shifts the *IS* curve to the right, and there is a temporary equilibrium at point M. The increase in the interest rate induces capital inflows, and the increase in income results in a deficit in net exports. However, the increased capital inflows more than offset the increased trade deficit and the *BP* is in surplus. This results in an accumulation of foreign exchange reserves at the central bank, and the *LM* curve shifts to the right as indicated by the arrows. Eventually, the final equilibrium is at point N, where the *IS*, *LM*, and *BP* curves intersect.

increases, the *LM* curve shifts to the right in the direction of the arrows in the diagram. Eventually, the *LM* curve settles at point N where the deficit on the current account at point R is financed by a capital inflow that is induced by a relatively high domestic interest rate. Hence, a FISCAL EXPANSION in a fixed exchange rate regime results in

1. a rise in the interest rate and income and a surplus in the BP in the short run
2. a further increase in income and a decline in the interest rate from its short-run level, with no change in the overall BP and accompanied by a current account deficit and an expansion in foreign exchange reserves in the long run

Thus, in a fixed exchange rate regime, with the capital imperfectly mobile, FISCAL POLICY IS MORE EXPANSIONARY than monetary policy and is associated with a current account deficit.

› An expansionary fiscal policy in a fixed exchange rate regime increases the foreign exchange reserves and results in the long-run income level becoming even higher than the short-run impact of the policy.

› In a fixed exchange rate regime, the fiscal policy is more effective than the monetary policy in influencing the aggregate income.

# 11.7 Capital Flows and Macro Policy

We would expect countries with a fixed exchange rate to keep fiscal policy in the forefront of economic policymaking for managing economic cycles. However, *IS*, *LM*, and *BP* passing through a common point only ensures BP equilibrium. The aggregate demand at this point, however, may be associated with unemployment. Economic policy is required to deal with *demand deficiency* in addition to *balance of payments disequilibria*. The intersection at a common point of the *IS*, *LM*, and *BP* curves results in a full equilibrium in the goods market, the money market, and the BP.

If the three curves do not intersect at a common point, which is highly likely in a fixed exchange rate regime, then the disequilibrium that results is likely to be concentrated most in that market, which bears the least influence on the rate of interest and the level of income. Thus, if the BP is not likely to be as much income and interest rate determining as are *IS* and *LM*, then the intersection of the *IS* and *LM* curves is the point towards which *r* and *Y* will strongly gravitate towards. Given a demand-constrained economy, let income be at a point that is less than the full-employment income,[14] $Y < Y_f$. Then, the economy could be subject to a dual problem of unemployment and a surplus in the BP, as depicted in Figure 11.8, which can be argued to characterize the current Indian economic conjuncture. In Figure 11.8, the *IS* and *LM* curves intersect above and to the left of the *BP* curve, resulting in an income and interest rate combination that results in a BP surplus. Simultaneously, *IS* and *LM* intersect (at point A) to the left of the full-employment income line $Y_f$, resulting in unemployment.

**MACROFOCUS 11.2**

## The Pressure on China to Float

Protectionism in the United States has been increasingly targeting China in an attempt to get it to revalue its currency. In fact, the US Senate is considering a bill that would require the Treasury to identify undervalued currencies and allow US firms to ask for protective anti-dumping duties if such countries do not revalue. This follows the infamous Schumer–Graham bill, which was withdrawn in 2006, that had proposed a 27.5 per cent tariff on all Chinese goods to offset the alleged undervaluation of the yuan.

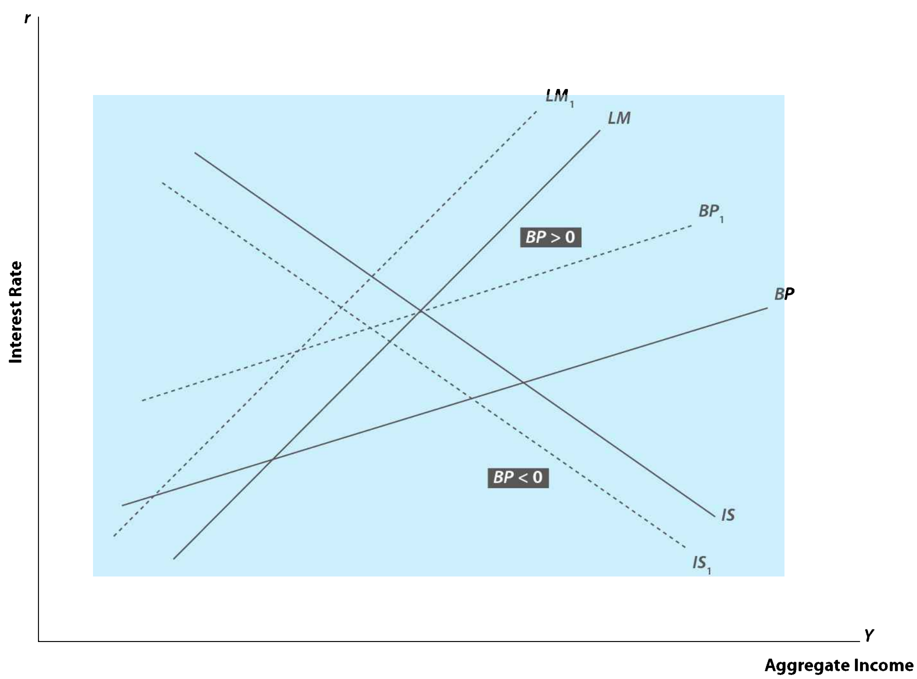
The United States argues that China has been artificially depreciating its currency and causing the US trade deficit to increase. China's trade surplus with the United States last year accounted for approximately 30 per cent of the US trade deficit. For a decade or so, China had pegged the yuan to the dollar, and in July 2005, it abandoned this policy. However, since then, the yuan has appreciated by less than 10 per cent against the dollar because the dollar itself has weakened since then. The yuan is still undervalued—China's foreign exchange reserves, which are in excess of $1.2 trillion, are the highest in the world, and this indicates a massive purchase of dollars by the People's Bank of China to keep the currency from appreciating. Many people believe that a revaluation of at least 25–40 per cent is required. The current Chinese economy predicament is depicted in the accompanying Figure 1.

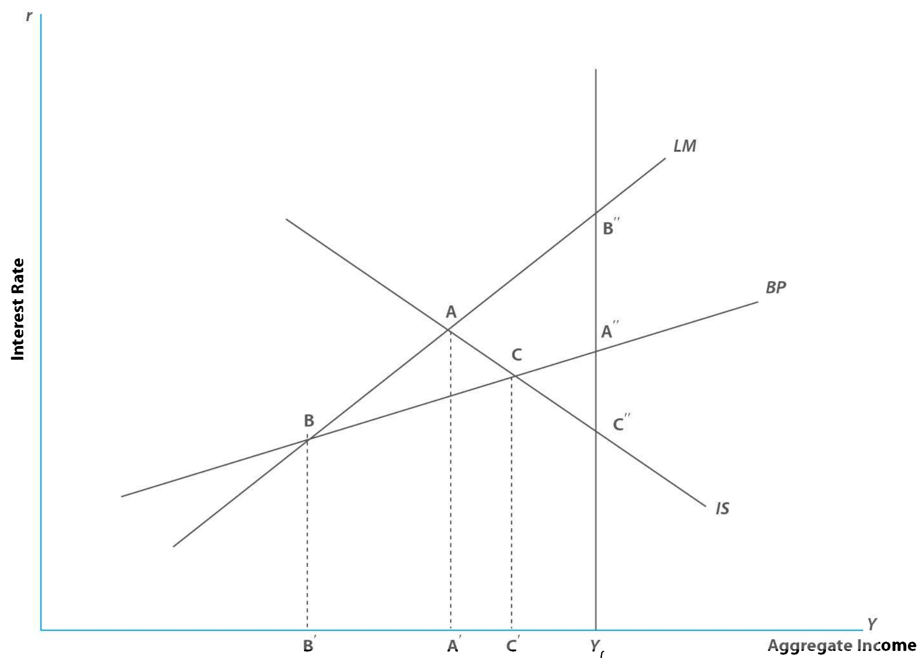
It is hoped that a large revaluation will shift the *BP* curve to $BP_1$, reduce the BP surplus, and control the overheating Chinese economy (the fastest growing economy in the world in the past decade). The revaluation, however, has second-round effects. Currency appreciation reduces net exports and shifts the *IS* curve leftwards to $IS_1$. The reduction in foreign exchange reserves reduces money supply and shifts the *LM* curve leftwards to $LM_1$. The economy overcorrects and the BP surplus continues, which requires further revaluation.

If the government engages in small revaluations at irregular intervals, this could introduce speculation regarding the currency and increase volatility. Controlling outcomes in an economy is made difficult by second-round effects and by expectations that agents form.

In 2005, when China unexpectedly declared a 2.1 per cent revaluation and it changed from the dollar peg to a basket peg with a managed trading band, it disturbed market expectations and induced speculation. Markets overreact to unanticipated changes in policy. Transparency of a policy reform process is important in avoiding destabilizing expectations.

**› Figure 1**
The Pressure on China to Float..






› **Figure 11.8**
Demand Deficiency and a Payments Surplus. If the economy is at the point A, where the goods and money markets are in equilibrium, there is nevertheless a *BP* surplus as point A is above the *BP* line. The output at A is also below the full-employment income level $Y_f$. The economy is characterized by demand deficiency and *BP* disequilibrium. Demand deficiency can be overcome by an expansionary fiscal policy that shifts the *IS* curve rightwards to pass through point B″. But this further improves the *BP* situation. Expansionary monetary policy can overcome the demand deficiency by shifting the *LM* curve rightwards to pass through point C″, but this worsens the *BP* situation.

External equilibrium can be restored via a contractionary fiscal policy that shifts the *IS* curve to pass through point B, which decreases income from point A′ to B′. Alternatively, an expansionary monetary policy restores external equilibrium while increasing income from A′ to C′.

Suppose, however, that full employment, and not external equilibrium, is the primary policy goal. Then, expansionary monetary policy restores full employment at point C″ but causes a BP deficit. Expansionary fiscal policy restores full employment at point B″ while simultaneously making the initial BP surplus even larger. In short, monetary policy that was effective in meeting the external equilibrium objective is not as effective for meeting the full-employment goal. MONETARY POLICY thus has a comparative advantage over FISCAL POLICY for the purpose of meeting the objective of external equilibrium, but fiscal policy has a comparative advantage of meeting the full-employment goal. Of course, the twin problems of unemployment and a payments surplus can be resolved by simultaneously applying an expansionary monetary and fiscal policy. This will result in the *IS* and *LM* curves intersecting at the same point where the *BP* intersects the full-employment line at A″. An economy must, of course, be willing to accept lower capital inflows for this to be an acceptable policy situation. Otherwise, the economy must trade off external and full-employment equilibrium and accordingly choose the mix of monetary and fiscal policy.

› In an economy experiencing demand deficiency and a BP surplus, monetary policy is better able to achieve the objective of BP equilibrium whereas fiscal policy is better able to achieve the objective of attaining full employment.

From 2002 onwards, emerging markets began to be favoured with huge capital inflows. The lessons of the 1997 East Asia crisis and similar currency crises in Russia in 1998, Turkey in 2001, and Argentina in 2002 made countries cautious about these inflows. Rather than using it to finance huge current account inflows, as they had previously done, they used the inflows to build up reserves. Many economies since 2002 have found themselves in a position as depicted by point A in Figure 11.8. This outcome was the result of internal and external factors.

An external factor was the fall in US interest rates due to the Federal Reserve attempting to reduce the impact of a recession in the US economy. This increased the interest rate differential ($r - r^*$) and encouraged capital flows to emerging markets, which caused an appreciation of the exchange rate and pushed the *BP* curve down and to the right.

An internal factor responsible for the emergence of the economic configuration, as shown by point A, was the increase in aggregate spending following liberalization in these countries. As controls were removed and financial markets liberalized, consumption increased in response to the availability of a variety of goods and the easy availability of finance. Investment also increased as the climate of investment improved. The increase in aggregate spending resulted in an outward shift of the *IS* curve to a point like A from a point like B. As we have already mentioned, many economies used these capital inflows to build up reserves. As reserves have built up, however, there is an increasing pressure on policymakers to contain the capital inflows and remove the disequilibrium in the economy, as depicted by point A. There are four strategies advocated to deal with the inflow problem:

1. Sterilization—Recall that the monetary base comprises foreign exchange reserves plus domestic credit advances. Sterilization involves an equally sized decline in domestic credit advances via an open market sale of government securities so as to maintain an unchanged monetary base. A purchase of foreign exchange by the central bank involves an increase in its assets, which requires an increase in its liabilities or a decline in some other asset. When the central bank sells government securities, it is reducing its stock of domestic assets (stock of government bonds) and the total assets of the central bank remain the same. As the central bank reduces its stock of domestic government bonds, it would require replenishing its stock of domestic assets in order to continue with sterilized intervention.

   In India, as the central bank lost its asset base from sterilization, the government added to the stock of the bank's domestic assets by introducing the Market Stabilization Scheme (MSS) in 2004. Under the MSS, the government issues securities to mop up rupee liquidity and then parks the proceeds in a separate cash account maintained and operated by the RBI. These funds are then used to redeem or repurchase securities issued under the MSS. This results in a decline in the RBI credit to government, which nullifies the impact of the increased stock of foreign exchange assets with the RBI due to capital inflows.[15] To induce the commercial banks and the public to hold these bonds, the central bank must pay a high interest rate on these bonds—often an interest rate higher than its returns on the holding of international reserves.[16] The central bank payment of interest on these sterilization bonds amounted to INR 296.9 million in 2004–2005, for instance.

   Sterilization allows the economy to remain at point A in Figure 11.8 without a corresponding increase in the money supply. The inflationary pressure that increased money supply causes is thus curtailed. This advantage, however, is counterbalanced by the cost of open market operations.
2. Appreciation of the currency—Allowing the currency to appreciate in response to a capital inflow is another policy response to inflows. This makes imports more costly and reduces net imports and shifts the *IS* curve to the left to point B and restores the BP disequilibrium. This reduction in demand makes governments hesitate in implementing such a policy. However, when inflation is caused due to high import prices, such as when oil prices are high, this is an attractive option to governments as it reduces the price of imports and curtails inflationary pressure.
3. Allow the inflow of money—As reserves increase, the central bank could just allow the increase in assets to result in an increase in its liabilities and the money supply. The *LM* curve shifts outwards towards point C in Figure 11.8. The BP, as we saw, reverts to equilibrium in such a situation. However, an

increase in the money supply can cause inflationary pressure to build up over time, especially if the economy is close to the $Y_f$ curve. Given the lags in monetary policy, the increase in money supply tends to result in inflation over a longer period of time. Thus, the achievement of BP equilibrium comes at the cost of inflation.

4. Capital controls—The imposition of capital controls results in a steeper *BP* curve. As the slope of the *BP* curve increases, say, a counterclockwise rotation around point C, it can result in a curve that goes through point A where there is no BP surplus. However, the cost here is that interest rates are higher and the economy is financing investment at high cost.

Clearly, policymaking is a matter of trade-offs that involve hard choices about what to do when faced with a situation such as that represented in Figure 11.8.

## SUMMARY

» The aggregate output equals expenditure *IS* condition for an open economy is given by the sum of domestic absorption and net exports, $Y = A + NX$, or equivalently, $S - NX = I$, that is aggregate savings [private, government, and foreign ($NX$)] equals investment.

» Given any combination of interest rate and income where aggregate savings equals investment, $S - NX = I$, if there is a rise in the interest rate, the accompanying decline in income is smaller for an open economy as net exports increase. The open-economy *IS* curve is, therefore, steeper than the closed-economy *IS* curve.

» Reserve or base money in an open economy is the sum of domestic credit ($DC$) advanced by the central bank (to the government, the commercial private sector, and commercial banks) and foreign exchange assets ($FE_{CB}$). Money supply is a multiple of reserve money given by $M^S = m[FE_{CB} + DC]$.

» Capital flows to destinations that offer higher returns. The net capital inflows ($NKI$) to an economy is, therefore, an increasing function of the extent to which the domestic interest rate is greater than the interest rate abroad, $NKI = K(r - r^*)$, where $r^*$ is inclusive of expected depreciation of the domestic currency.

» The balance of payments (BP) schedule is the set of interest rate income combinations that are consistent with a BP equilibrium in which the current account balance and capital account balance sum to zero.

» Capital controls are legal restrictions on the ability of citizens to hold and exchange assets denominated in the currencies of other nations. These affect capital mobility—the degree to which finance can flow across a country's borders.

» The extent of capital mobility affects the slope of the *BP* curve. When the capital is completely immobile, the *BP* curve coincides with the equilibrium on the current account.

» Perfect capital mobility results in a domestic interest rate equal to the foreign interest rate, $r = r^*$ (the uncovered interest parity condition), and a *BP* curve that is horizontal.

» Imperfect capital mobility results in finite flows in and out of a country due to interest rate differentials and an upward-sloping *BP* curve.

» Points to the right of the *BP* curve represent a BP deficit. Any point to the southeast of a point on the *BP* curve depicts a lower interest rate-higher income combination. The lower interest rate reduces net capital inflows, and the higher income reduces net exports. $BP = NX + NKI$ must then be in deficit at this point off and to the right of the *BP* curve.

» Points to the left of the *BP* curve represent a BP surplus, $BP > 0$.

» A depreciation of the currency increases net exports and the aggregate income and shifts the *BP* curve down and to the right.

» An appreciation of the currency reduces net exports and the aggregate income and shifts the *BP* curve up and to the left.

» A monetary expansion under flexible exchange rates results in
» 1. a depreciation of the currency
» 2. an increase in income
» 3. a decline in the interest rate
» 4. an improvement in the current account balance

» A fiscal expansion under flexible exchange rates results in two types of crowding out: (1) crowding out of private investment as the interest rate increases and (2) crowding out of exports as the currency appreciates.

» A fiscal expansion under flexible exchange rates results in
» 1. an appreciation of the currency
» 2. an increase in income
» 3. a rise in the interest rate
» 4. a deterioration in the current account balance

» In a flexible exchange rate regime, monetary policy is more effective than fiscal policy in influencing aggregate income.

» An expansionary monetary policy under fixed exchange rates and no sterilization of the foreign exchange assets accumulated results in
» 1. the short run in
» (a) a decline in the interest rate

» (b) an increase in income
» (c) a deterioration in the current and capital accounts of the BP
» 2. the long run in
» (a) no change in the interest rate, income, or BP
» (b) a decline in foreign exchange reserves
» An expansionary fiscal policy when the exchange rate is fixed results in
» 1. the short run in
» (a) a rise in the interest rate
» (b) a rise in income and
» (c) a surplus in the BP
» 2. the long run in
» (a) a further decline in the interest rate
» (b) a further rise in income
» (c) no change in the overall BP
» (d) a current account deficit and
» (e) an expansion of foreign exchange reserves
» In a fixed exchange rate regime, fiscal policy is more effective than monetary policy in influencing aggregate demand.
» In an economy experiencing demand deficiency and a BP surplus, monetary policy is better able to achieve the objective of BP equilibrium whereas fiscal policy is better able to achieve the objective of attaining full employment.
» The recent capital inflows to emerging markets are the result of internal factors, such as liberalization of the economy, and external factors, such as the low rates of return in advanced economies, such as the United States which went through a recession.
» Capital inflows can be managed by (1) sterilization, (2) allowing the currency to appreciate, (3) allowing money supply to increase, and (4) imposing capital controls. Sterilization maintains the inflow episode at the cost of high interest rates and/or the introduction of sterilization bonds that have an interest cost. Appreciation of the currency can solve the BP disequilibrium at the cost of reducing aggregate demand as net exports decline. Increasing the money supply runs the cost of increasing inflation. Imposing capital controls has the opportunity cost that it makes the economy a high cost economy. Each response to capital inflows involves trade-offs that makes policymaking no clear-cut exercise.

## NOTES

1. The presumption here is that the Marshall–Lerner condition holds.
2. $A_r < 0$ as a rise in the interest rate reduces consumption and investment. $0 < A_Y < 1$ as a rise in income induces a rise in consumption expenditure. $NX_Y < 0$ as a rise in income increases imports and reduces net exports. $NX_E > 0$ as a depreciation of the currency increases export competitiveness and increases net exports.
3. $\partial r/\partial Y|_{IS} = [\partial(S - NX)/\partial Y]/[\partial A/\partial r] = [\partial(S - NX)/\partial Y]/[\partial(C + I)/\partial r]$ for an open economy.
4. $\partial r/\partial Y|_{\text{open economy}} = [(1 - A_Y) - NX_Y]/A_r < (1 - A_Y)/A_r = \partial r/\partial Y|_{\text{closed economy}}$
5. $L_r < 0$, $L_Y > 0$, that is a rise in the interest rate reduces the demand for money as bonds become attractive, and a rise in income increases the demand for money to facilitate transactions.
6. See Robert A. Mundell, "The Appropriate Use of Monetary and Fiscal Policy Under Fixed Exchange Rates," IMF Staff Papers, no. 9, 1962, pp. 70–77; Also see J. Marcus Fleming, "Domestic Financial Policies Under Fixed and Floating Exchange Rates," IMF Staff Papers, no. 9, 1962, pp. 369–377. As Mundell himself once stated, "I am not quite sure when the term 'Mundell-Fleming model' first appeared in the literature … (my work) preceded his (Fleming's) in publication but not necessarily in conception." See Robert A. Mundell, "Notes on the Development of the International Macroeconomic Model," in A. Arnon and W. Young (eds.) *The Open Economy Macromodel: Past, Present and Future* (Boston, MA: Kluwer Academic Publishers, 2002), pp. 1–16.
7. The Dornbusch sticky price model looks at the expectation mechanism with respect to exchange rates. See Rudiger Dornbusch, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy* 84 (1976): 1161–1176.
8. We deal with fixed exchange rates in a subsequent section.
9. Slope of the *BP*(*E*) line is given by $\partial r/\partial Y|_E = -(\partial NX/\partial Y)/(\partial NKI/\partial r)$.
10. Slope of the *BP* line is given by $\partial r/\partial Y = -(\partial NX/\partial Y)/(\partial NKI/\partial r) \to \infty$ as $\partial NKI/\partial r = 0$.
11. We address the question as to why the *BP* curve is less steep than the *LM* curve in the various diagrams in the appendix.
12. With perfect capital mobility, the *BP* curve is flat and the whole of the increase in government expenditure is offset by the fall in exports making crowding out complete. This is because with the interest rate determined in the international capital market, $r = r^*$, the burden of adjustment is borne entirely by the exchange rate that has to appreciate to such an extent that the current account deficit that results is as great as the increase in fiscal spending. Therefore, external borrowing funds the whole of the increase in government expenditure.
13. Recall that $M^S = m[FE_{CB} + DC]$.
14. We can also think of $Y_f$ as the point of potential output where expectations are realized as in the Friedman–Lucas approach.
15. See RBI Annual Report, Box III.3, 2004–2005, pp. 110.
16. When the central bank pays a higher interest rate on sterilization bonds than it receives on the holdings of international reserves, it is said to incur a quasifiscal deficit.

## TEST YOURSELF

1. What is the *BP* schedule for an open economy? What determines its slope? What causes shifts in the *BP* curve?
2. Suppose an economy maintains a fixed exchange rate and its monetary policy involves no sterilization. How does a contractionary monetary policy affect the balance of payments and the aggregate income?
3. Suppose an economy with a fixed exchange rate allows capital flows without any controls. What is the impact of a contractionary fiscal policy on the economy?
4. Suppose an economy with a flexible exchange rate has perfect capital mobility. Explain how a contractionary fiscal policy affects the value of its currency and the composition of the balance of payments.
5. An economy with a flexible exchange rate and perfect capital mobility is expanding too fast and overheating. Should the monetary or the fiscal authority take the lead in attempting to contain income growth?
6. As an economy increasingly integrates with the world economy during this era of globalization, should it maintain fixed or floating exchange rates? Why?
7. Currently the US economy is subject to a consumption boom that has increased it trade balance substantially. Suppose capital mobility is perfect and the dollar–renminbi exchange rate floats. The US government is trying to prevail upon Beijing to reduce government spending. Is this a good policy from the US point of view?
8. An economy's exchange rate is fixed and has imperfect capital mobility. Analyse the effect on the economy in the short and long runs of a reduction in taxes.
9. What results in the *LM* curve becoming endogenous in the case of fixed exchange rates and the *IS* curve becoming endogenous in the case of flexible exchange rates?
10. Suppose there is an increase in the price level in an open economy with flexible exchange rates. What happens to the nominal exchange rate and the real exchange rate? Can an appreciation of the nominal exchange rate be associated with a depreciation of the real exchange rate?

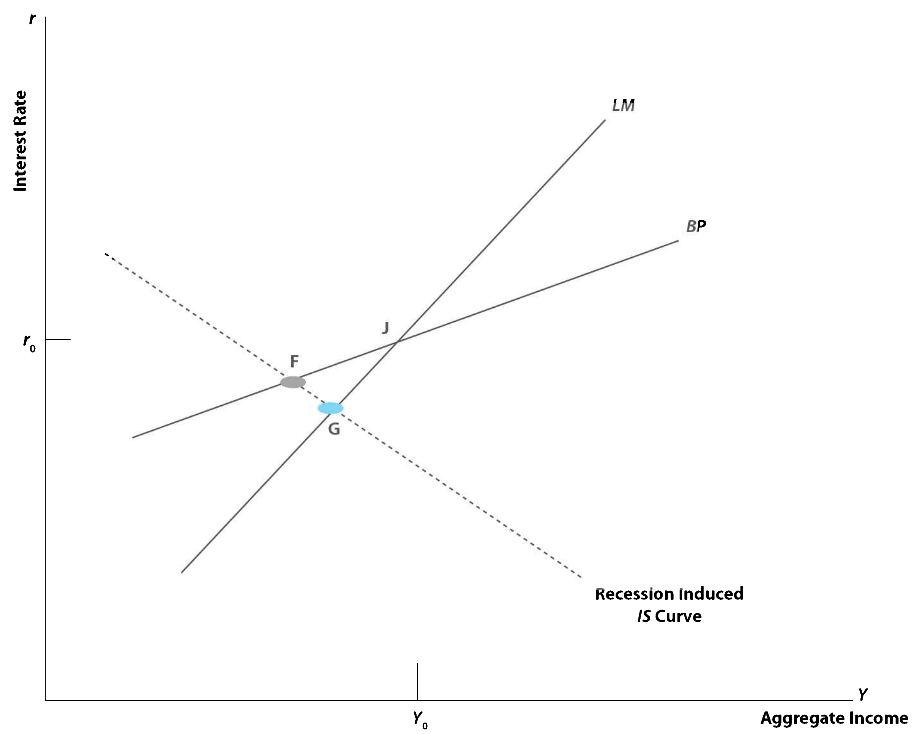
## ONLINE APPLICATION

1 (a) Go to the home page of the Reserve Bank of India Internet (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
   (c) Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy."
2. Click on "Table 118: Combined Deficits of central and state Governments" and create a spreadsheet of annual data on the fiscal deficit of the nation from 1980–1981 to 2004–2005.
3. Then click on "Table 132: India's Foreign Trade—Rupees" and in your spreadsheet, insert annual data on the trade deficit from 1980–1981 to 2004–2005.
4. Finally, click on "Table 162: Foreign Exchange Reserves" and in your spreadsheet, insert annual data on foreign currency reserves in INR from 1980–1981 to 2004–2005.
5. Now index the various data series. Let the base value of the fiscal deficit of INR 959.207 billion be 100. Then, the index of the fiscal deficit in 1980–1981 of INR 107.8 billion is $[(107.8 \times 100)/95.9207] = 112.38$.
6. Similarly, let the base value of the trade deficit of INR 30.1515 billion be 100. Then, the index of the trade deficit in 1980–1981 of INR 58.3844 billion is $[(58.3844 \times 100)/30.1515] = 193.63$.
7. The base value of the foreign exchange reserves can be set at INR 19.7487 billion equals to an index value of 100. Accordingly, the index of foreign exchange reserves of INR 55.45 billion in 1980–1981 is $[(55.45 \times 100)/19.7487] = 280.77$.
8. Now, create a graph with the index of the fiscal deficit on the horizontal axis and the index of the trade deficit and foreign exchange reserves on the vertical axis. You will notice that fiscal policy has a positive impact on the trade deficit—a linear trend will depict that a unit increase in the index of the fiscal deficit increases the index of the trade deficit by 0.88 units. The impact of the fiscal deficit on foreign exchange reserves is much larger—a unit increase in the index of the fiscal deficit increases the index of foreign exchange reserves by 8.71 units. In the text, we concluded that fiscal policy results in a deterioration of the trade balance and an expansion of foreign exchange reserves if the exchange rate is fixed. Relate your graph to the conclusions of the text.
9. The time period up to 1993–1994 was characterized by an administered exchange rate regime. The exchange rate flexibility in India gradually increased after that. The text argued that under flexible exchange rates, fiscal policy should result in an appreciation of the currency. Insert data on the real (effective trade-weighted) exchange rate from 1993–1994 onwards (base year 1993–1994 = 100) by clicking on "Table 153: Indices of Real Effective Exchange Rate (REER) and Nominal Effective Exchange Rate (NEER) of the Indian Rupee (36 country bilateral weights)." If you create a graph from this, you will see that the real exchange rate has apparently no trend and seems to be unrelated to fiscal policy. Explain this apparent lack of relationship.

# Appendix 11.1 Slopes of the *BP* and *LM* Curves

The *BP* curve has a positive slope as has the *LM* curve, and the *BP* curve has been drawn flatter than the *LM* curve throughout the chapter. The absolute slope of the *BP* curve is the ratio of the responsiveness of net exports to the level of income and the responsiveness of net capital inflows to the rate of interest (see Note 9). The absolute slope of the *LM* curve is the ratio of the responsiveness of the transactions demand for money to the level of income and the responsiveness of the asset demand for money to the rate of interest.[1] There is no way we can *a priori* state a generalization that one of these ratios exceeds the other for all economies. We need a justification for the slopes as drawn. One such justification is to consider an intersection of *LM* and *BP* at some point J, resulting in ($r_0$, $Y_0$) in the accompanying Figure A11.1.1.

We then consider what would happen if a recession that shifts the *IS* curve to the left from point J to pass through points F and G were to occur without this being anticipated by the monetary and fiscal authorities. There would be a decline in income and the interest rate. We would expect portfolio readjustment to be rapid as financial markets react fast to economy-wide conditions, and individuals would adjust their holdings of money vis-à-vis other financial assets (bonds) so that the lower ($r$, $Y$) combination pushes them lower down on the *LM* curve. What would happen to the external equilibrium in the BP? We could expect that international capital markets react faster to information regarding a recession and the market for international trade takes time to renegotiate contracts and delivery schedules. Thus, the recession's interest rate effect of decreasing net capital inflows outweighs its income effect of increasing




› **Figure A11.1.1**
Slopes of *LM* and *BP* Curves.
A recession that was not forecast by monetary and fiscal authorities shifts the *IS* curve to the left. Speedy portfolio readjustment in the money market ensures an interest rate prevailing at the intersection of the *IS* and *LM* curves. With goods trade involving a longer adjustment period than capital flows, the *BP* should be in deficit, which implies that the *BP* curve is flatter than the *LM* curve.

[1] $dr/dy\,\big|_{LM} = -(\partial L/\partial Y)/(\partial L/\partial r) = -L_Y/L_r$

the net export of goods and services.[2] This being the case, it means the new interest rate is too low (point G is lower than point F) and the new income too high (at point G, income is higher than at point F) to prevent a payments deficit. The *BP* curve thus lies above the *LM* curve for income values to the left of the intersection of *BP* and *LM*. Similarly, in an unanticipated boom, one would expect a BP surplus to develop, implying that the *BP* curve lies below *LM* for values of *Y* to the right of the intersection of *BP* and *LM*.

[2] $\partial NKI / \partial r > |\partial NX / \partial Y|$ and $|\Delta r| < |\Delta Y|$ for a reduction in $r$ and $Y$.

# 12

# Stabilization and Government Deficits

CRITICAL QUESTIONS

- » *What are the roles of automatic stabilizers and discretionary fiscal policy?*
- » *How does the long-run goal of fiscal policy constrain discretionary fiscal policy?*
- » *What is the intertemporal government budget constraint?*
- » *When is the government deficit sustainable?*
- » *What are the measures adopted to cut fiscal deficits?*

# 12.1 Limits to Discretionary Policy

› The statutory liquidity ratio specifies the required holding of government securities as a fraction of the assets of a commercial bank.

› The government requirement for funds to finance the expenditure above the level mandated by the holding of government securities (statutory liquidity ratio) was met by the government unilaterally accessing the RBI by issuing T-bills or bonds with a maturity not exceeding one year. The RBI monetized these T-bills and increased reserve money.

› Automatic stabilizers are expenditures that automatically increase, or taxes that automatically decrease, when the aggregate income and output decline.

› Automatic stabilizers result in a steeper aggregate demand curve and reduce the variability of the output arising from shocks to the aggregate demand.

Up until 1990 the concept of the fiscal deficit was absent from official government documents, such as the budget and the Economic Survey.[1] It was only after the economic crisis of the early 1990s that fiscal consolidation made the fiscal deficit a focus of policy discussions. Before 1991, as we will see in the Chapter 14 on financial reforms, government borrowing to finance the deficit was accommodated by hikes in the STATUTORY LIQUIDITY RATIO imposed on banks, which allowed the placement of government securities at below market rates of interest. This move to market rates of interest through the auctioning of government securities requires the fiscal deficit to be sustainable so that interest payments and deficits do not run out of control.

In India, the government formalized an agreement with the Reserve Bank of India in 1994 to phase out *ad hoc* TREASURY BILLS over a three-year period. Prior to this the government used to finance its deficits routinely through the issue of *ad hoc* Treasury Bills that provided it unrestricted access to RBI credit and enabled it to monetize its deficits. With the delinking of budget deficits from their monetization any fiscal shock had to *per se* be absorbed through variations in borrowings. Fiscal reforms, therefore, became associated with making fiscal deficits sustainable.

As we saw in Chapter 8, fiscal policy refers to changes in government expenditures and taxes with a view to achieving policy goals such as full employment output. There are two ways in practice by which fiscal policy contributes towards macroeconomic policy goals—automatic stabilizers and discretionary fiscal actions. AUTOMATIC STABILIZERS are government expenditures that automatically increase or taxes that automatically decrease when the pace of economic activity diminishes. Direct taxes such as income taxes are a good example of automatic stabilizers. As incomes decrease individuals have to pay less taxes to the government and this limits the decrease in consumption spending that accompanies a decline in income. Let us take an example.

With a marginal tax rate of 30 per cent when income reduces by INR 1,000, your tax bill reduces by INR 300. Of the remaining INR 700 (the amount by which your disposable income has been reduced), with a marginal propensity to consume of 0.65, you would have consumed INR 455 and saved INR 245. So your consumption expenditure declines by INR 455. Without the income tax in place, you would have decreased your consumption spending by a much larger $0.65 \times$ INR 1,000, or INR 650.

As taxes rise and fall with income they reduce the variation in consumption spending and result in more stable expenditures in the economy. In advanced economies, transfer payments such as unemployment benefits also act as automatic stabilizers. A decline in economic activity accompanied by a rise in unemployment results in a rise in government transfers via unemployment benefit programmes. This income support to those affected by a downturn in the economy limits the decline in consumption spending that would otherwise have occurred and stabilizes expenditures.

Automatic stabilizers affect the slope of the aggregate demand curve—the inverse relationship between the price level and aggregate income/expenditure. A rise in the price level reduces real money supply and raises the interest rate. This, in turn, reduces consumption and investment expenditure and aggregate income. Tax payments decline along with income and this limits the decline in consumption spending. Hence, without income taxes the decline in consumption expenditure would have been greater. Thus the AGGREGATE DEMAND CURVE with automatic stabilizers has a STEEPER slope than one without such stabilizers in existence.[2]

Automatic stabilizers also cushion the impact of shocks—unanticipated sudden changes—on the economy. A reduction in business confidence, for instance, shifts the aggregate demand curve to the left. The magnitude of this shift is, however, reduced by the presence of automatic stabilizers. Such smaller shifts of the aggregate demand curve reduce the variability of economic fluctuations.

› Discretionary fiscal policy results from deliberate and conscious choices by the government on public expenditure and taxes that are made to achieve macroeconomic goals.

Discretionary fiscal policy is the deliberate and conscious use of government expenditure and tax decisions to influence the course of the economy. This aspect of fiscal policy is subject to long inside lags of recognition and implementation, which reduces its effectiveness in helping to stabilize the economy. To stabilize the economy fiscal policy must, therefore, work through automatic stabilizers. This means that as the economy expands, tax revenues increase and the government's budget deficit declines. By contrast, if the economy goes into a downturn the budget deficit increases. In addition to such cyclical impacts, the budget deficit may increase or fall as fiscal policy is used to provide an additional discretionary stimulus to steer the economy towards an equilibrium level of output.

So far we proceeded to analyse fiscal policy in the *IS–LM* model as an exogenous change in government spending or taxation that causes a shift in the *IS* curve. At that point we did not consider the financing of these changes in fiscal policy explicitly. A change in government expenditure for a given taxation and borrowing from the central bank (seigniorage) must be financed through borrowing and an increase in public debt. This imposes an obligation on the government to pay interest on the debt and eventually to repay the principal amount. Once we recognize this intertemporal aspect of fiscal policy, it is natural to ask ourselves as to whether the repayment of debt places limits to government borrowing that affects its ability to engage in discretionary fiscal policy. Only if the public debt is sustainable, in the sense of the government being able to repay the debt, does a government have the fiscal space to increase borrowings to finance increases in public expenditure. The discretionary element of fiscal policy is, therefore, limited by the sustainability of the government's debt.

Fiscal policy, which was associated with its stabilization short-run role of using government expenditure and tax policy to stabilize output over the business cycle, has a longer run objective as well. This is the objective of fiscal discipline or debt sustainability, which we address in this chapter.

Up until now we have explored the role of fiscal policy as a short-run output stabilization tool. Two concerns that we faced in using this tool were that the impact of fiscal policy can be rather slow and that the impact is also rather uncertain. The lags in fiscal policy action arise because such actions must be approved by the Parliament in India which could result in the loss of precious time before action is initiated. Discretionary fiscal policy is usually carried out with annual budgets, which is a major hurdle as it results in long lags. For discretionary fiscal actions to be useful as a countercyclical device, the decision and implementation lags need to be sharply reduced. This is a difficult task, however, as fiscal policy decisions on expenditure or taxation have redistributive impacts and such actions in a democratic set-up should be subject to parliamentary oversight.

Monetary policy, on the contrary, has a faster speed of reaction and can be implemented in a short time as it is not subject to as much *ex ante* political control with the monetary authorities shielded much more from political influence. The impact of fiscal policy is also uncertain in the short run because a lot depends on how economic agents perceive fiscal policy actions. For instance, if agents look at tax measures initiated by the government as being temporary and they are forward looking then they will adjust saving behaviour to counteract the impact of this measure. Also, given credit rationing and myopia by individuals, discretionary fiscal action's effectiveness is an open-ended issue.

Fiscal policy, which had been used as a short-run tool in the pursuit of stabilizing the business cycle in a liberalized economic environment, is increasingly being asked to deliver over time in the longer run on fiscal discipline. Fiscal discipline means that the government faces an intertemporal budget constraint and any fiscal action taken by the government does not violate this constraint. This does not require the doing away of deficits altogether. Some deficits may be caused by government expenditure that is an investment expenditure, which is growth promoting and so may well pay for itself through the taxes garnered from the increase in output. Such deficits are not a hazard and it is difficult to label them as problematic for long-term fiscal discipline.

Similarly a few years of deficits are not antithetical to fiscal discipline provided they are followed by a few years of surpluses. However, when future action to reverse a sequence of deficits is not predictable, the government may be seen as not respecting its intertemporal budget constraint and thereby reneging on fiscal discipline. Moreover, as governments are in office for limited time spans and they cannot commit to deficit reduction on behalf of their successors, there is a built-in bias for deficits. The reason for this is that the burden of the deficit is not faced by the incumbent government currently. Hence, it is important that while governments have the short-run flexibility to stabilize the economy, they also respect the long-run objective of fiscal discipline as given by their intertemporal budget constraints.

## 12.2 The Government Budget Constraint and Debt Dynamics

The intertemporal budget constraint of the government is

$$G_t - (T_T + T_N + T_D)_t + rB_{t-1} = (M_t - M_{t-1}) + (B_t - B_{t-1})$$

where the subscript $t$ indicates the time period, and

$G$: public expenditure—current plus capital expenditure
$T_T$: tax revenue (net of non-debt-related transfer payments, such as subsidies)
$T_N$: non-tax revenues, such as user charges on public utilities
$T_D$: revenues from disinvestment
$B_t$: end of period stock of domestic public debt which bears interest rate $r$
$M_t$: stock of credit allocated by the central bank

Let $T = T_T + T_N + T_D$ be the total government revenue. In what follows we restrict ourselves to debt financing of deficits so that we may write the budget constraint as

$$\Delta B = (B_t - B_{t-1}) = G_t - T_t + rB_{t-1}$$

The left-hand side of the above equation is the FISCAL DEFICIT. The PRIMARY DEFICIT is the non-interest component of the fiscal deficit:

› The fiscal deficit is the government expenditure that is not related to the repayment of the debt, $G$, plus repayment of the debt, $rB$, less government revenues, $T$.

› The difference between the fiscal deficit and interest payments is the primary deficit.

$$\text{Primary Deficit} = G_t - T_t = D_t$$

Another way of expressing the government-budget constraint is to write the budget constraint of the government in the following way:

$$\Delta B = B_t - B_{t-1} = rB_{t-1} + (G - T)$$

or,

$$B_t = (1 + r)B_{t-1} + D_t$$

where $D_t$ is the primary deficit. Dividing throughout by GDP, $Y_t$, we obtain

$$\frac{B_t}{Y_t} = (1 + r)\frac{B_{t-1}}{Y_{t-1}}\frac{Y_{t-1}}{Y_t} + \frac{D_t}{Y_t}$$

Let $b_t = B_t / Y_t$, the debt/GDP ratio, $d_t = D_t / Y_t$, the primary deficit/GDP ratio and the one-period growth rate of GDP be $g = (Y_t - Y_{t-1}) / Y_{t-1} = (Y_t / Y_{t-1}) - 1$ or, $1 + g = Y_t / Y_{t-1}$. Then, we can rewrite the above as

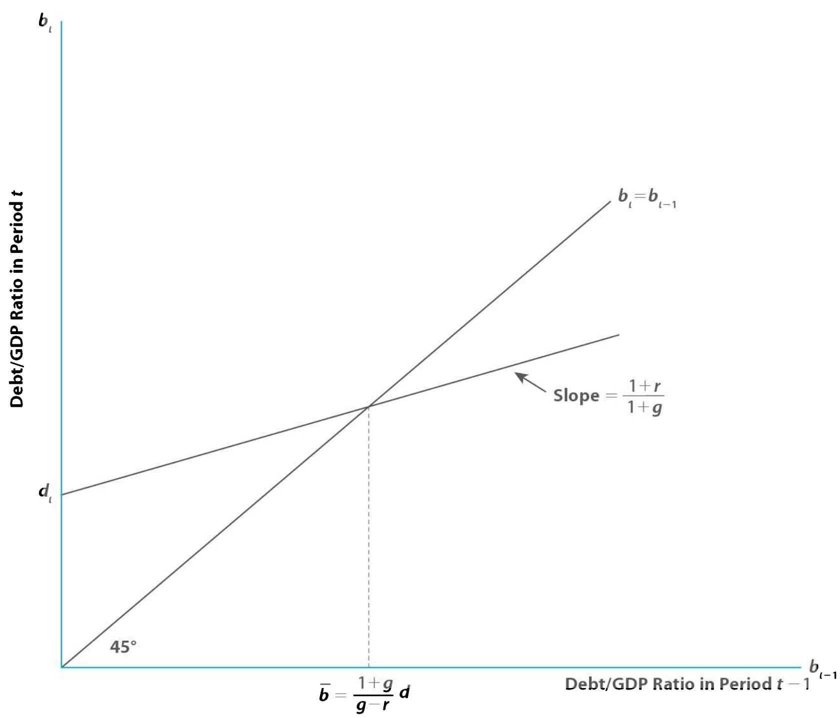
$$b_t = \frac{1 + r}{1 + g} b_{t-1} + d_t$$

› The debt/GDP ratio increases as the government issues new debt (1) to cover a primary deficit; (2) to pay interest on the existing debt. An increase in the growth rate of the GDP reduces the debt/GDP ratio.

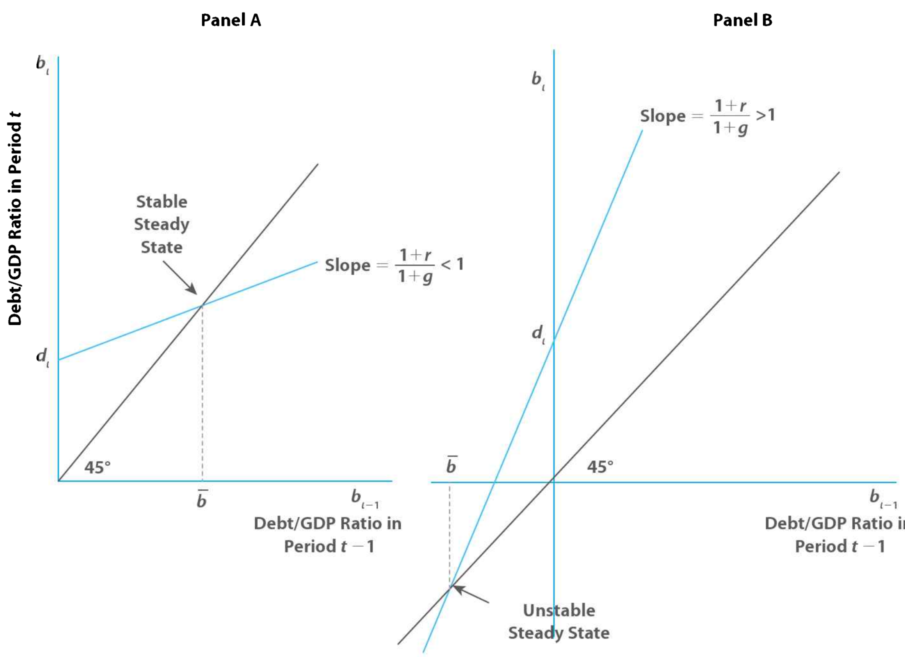
The DEBT/GDP RATIO grows for two reasons. First, the government issues debt to cover a primary deficit—the term $d_t$. Second, the government must pay interest on existing debt. This is reflected in the expression $1 + r/1 + g$.

To understand this second expression, consider what happens when the primary deficit is zero. In that case two factors affect the expression for the debt/GDP ratio $b_t = (1 + r/1 + g)\ b_{t-1}$. The government must increase debt by a factor of $(1+ r)$ to pay the interest on existing debt. This causes a rise in the debt/GDP ratio. But, GDP is itself increasing at the rate of $(1 + g)$, which increases the tax revenues of the government as taxes increase with GDP. This factor causes a decline in the debt/GDP ratio. The net effect of the two factors is given by the ratio $1+ r/1+ g$. If $g > r$, the debt/GDP ratio will not increase and in that sense the debt/GDP ratio is sustainable.

To analyse the dynamics of the debt and the government budget constraint, the above difference equation can be further analysed. If we graph $b_t$ as a function of $b_{t-1}$, it is a straight line with an intercept equal to $d_t$ and a slope equal to $1 + r/1 + g$. This graph appears in Figure 12.1.




› **Figure12 .1**
The Evolution of Debt/GDP. The debt/GDP ratio in period $t$, $b_t$, is the sum of two components. The first is the primary deficit/GDP ratio, $d_t$. The second is the ratio of one plus the interest rate to one plus the growth rate of GDP times the debt/GDP ratio of the previous period, $t - 1$. The steady-state debt/GDP ratio that does not change with time is given by $\bar{b}$.




**› Figure 12.2**
Stability of Debt/GDP. **When the interest rate is less than the growth rate of the GDP, as in panel A, the steady-state level of the debt is stable in the sense that if the economy at any moment of time has a debt/GDP value other than $\bar{b}$, it will gravitate towards $\bar{b}$. On the other hand, if the interest rate is greater than the growth rate of GDP as in panel B, the steady-state level of the debt is unstable and if the economy at any moment of time has a debt/GDP value other than $\bar{b}$, it will achieve a value of debt/GDP that recedes away from $\bar{b}$ over time.**

Now, consider the steady-state solution of the difference equation. A steady-state solution is a value for the debt/GDP ratio that satisfies the government budget equation and that is independent of time.[3] A STEADY STATE, $b_t = b_{t-1} = \bar{b}$, is a solution that gives that value of $b$ which if attained will not change without a shock to the system. The $b_t = b_{t-1} = \bar{b}$ line is the 45-degree line from the origin and the steady-state value is given by

› The steady-state debt/GDP ratio is that value of the debt/GDP ratio that satisfies the government budget constraint and is independent of time.

$$\bar{b} = \frac{1+r}{1+g}\bar{b} + d$$

or,

$$\bar{b}\left(\frac{g-r}{1+g}\right) = d$$

or,

$$\bar{b} = \frac{1+g}{g-r}d$$

### 12.2.1 Primary Deficits and Stability

We now examine a property of steady states called stability. In the two panels in Figure 12.2, the intercepts, $d$ is the same number in both the graphs. However, the slopes of the graphs in the two panels are different.

In panel A the slope is less than unity—$[(1+r)/(1+g)] < 1$, or $r < g$, and this depicts the situation of a STABLE STEADY STATE. The debt/GDP ratio at any time $t$, $b_t$, moves closer to the steady-state value $\bar{b}$ over time regardless of which value it starts from.

In panel B the slope is greater than unity, $[(1+r)/(1+g)] > 1$, or $r > g$, and this depicts an UNSTABLE STEADY STATE. The debt/GDP ratio, $b_t$, moves further away from the steady-state value $\bar{b}$ over time for any starting value of $b_t$ other than ($\bar{b}$ ) the steady-state value itself. [Note that we are for simplicity assuming $[(1+r)/(1+g)] > 0$ which is assured if $r > -1$, $g > -1$.]

› A steady state is stable if the debt/GDP ratio moves closer towards the steady-state value of the debt/GDP ratio over time, regardless of where it starts from. This occurs when $r < g$.

› A steady state is unstable if the debt/GDP ratio moves further away from the steady-state value of the debt/GDP ratio when we look at its trajectory starting from any point other than the steady-state point itself. This occurs when $r > g$.

From panel A of Figure 12.2 we can see that if the growth of GDP exceeds the interest rate, the economy would eventually settle into a steady state with a debt/GDP ratio that is constant at $\bar{b}$. A constant debt/GDP ratio is not unduly worrying to policymakers and politicians but a perpetual constant value of the debt/GDP ratio implies that the debt will never be repaid and $b_t$ will not tend to zero. Instead, the debt/GDP ratio will tend towards $\bar{b}$, $b_t \to \bar{b}$, and the government would be insolvent in the sense that it never repays the principal sum of debt borrowed.

In the case of panel B in Figure 12.2, where $r > g$, the GDP is growing at a slower rate than the interest rate and the debt/GDP ratio increases over time. Eventually debt will be larger than GDP and the debt is so large that the entire GDP is insufficient to pay the interest on the debt. The government will eventually become bankrupt. In practical terms, as the government's tax revenues are substantially less than GDP bankruptcy occurs well before the interest on the debt is larger than the GDP. The *steady-state debt*, however, is *negative*. A negative debt implies the government must lend to the private sector instead of borrowing from it. A policy that describes this involves increasing taxes above government expenditures and using the revenues to first pay off the existing government debt and then to purchase financial assets from the private sector.

To sustain the positive primary deficit $d$ the government must accumulate enough assets and this is obtained by running large primary surpluses for many years with an average value of $\bar{b}$. This is not politically an easy objective to attain. Governments often strive for an easier alternative instead, which is to continue the existing level of debt and to raise enough

**MACROFOCUS 12.1**

### Fiscal Surpluses

"The budget has been to our era what civil rights, communism, the depression, industrialization, and slavery were at other times.*"

Deficit reduction has become a major priority for governments and has captured many policymakers' attention. Those championing fiscal restraint were at one time at the margins, and intellectual forces championing higher spending and lower taxes dominated the policy discourse. Over the previous decade, however, there have emerged economies that have moved from reducing deficits, in order to bring the budget into balance, and towards sustaining surpluses. Surpluses are harder to find support for because they typically follow a painful period of deficit reduction during which expenditures have been cut and taxes raised and citizens begin to question whether further surpluses are needed at all. Instead it is natural for demands to be present for new spending and tax cuts follow a period of restraint. Yet some economies have managed to sustain surpluses over many years.

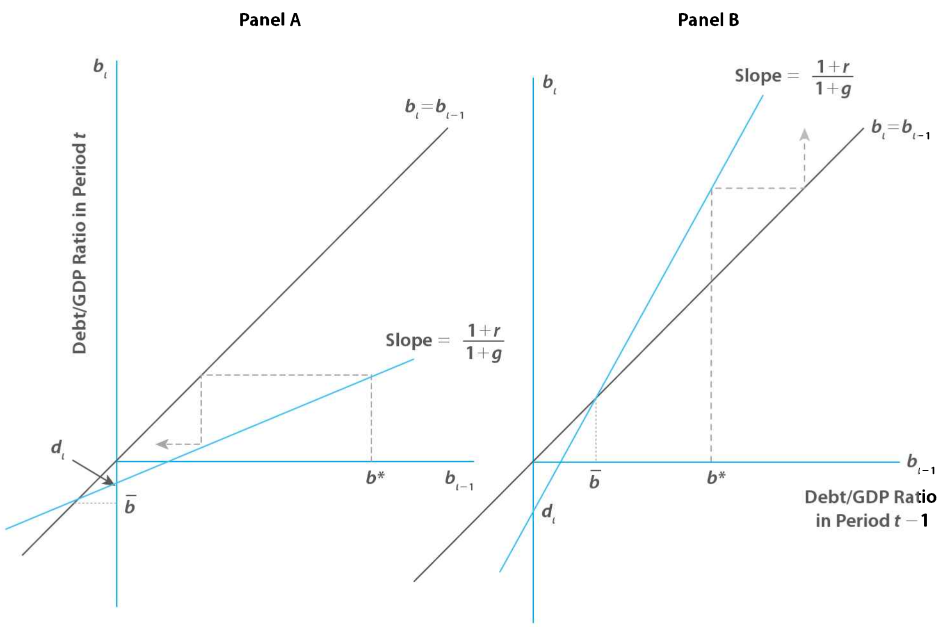
Norway is one country that has set itself a macro goal of continued budget surpluses in order to build up the savings required to address long-term economic concerns arising from an aging population and declining petroleum revenues. Projections by the Norwegian government show a doubling of retirement benefits from 7 per cent of the GDP in 2001 to about 15 per cent of GDP by 2030. Oil revenues are also projected to decline from about 8 per cent of GDP to less than one per cent during the same period. In the mid 1990s Norwegian decision-makers reached a consensus to save surpluses to pay for future budget needs and since 1994 the country has been maintaining a surplus. The government created the Government Petroleum Fund in which surpluses were deposited to help pay for future pension costs and the funds assets were invested in foreign stocks and bonds to help reduce inflation and depreciation pressure on the exchange rate for the Krone.

In New Zealand the public debt reached a peak of nearly 65 per cent of GDP in 1992. This led to a loss of investor confidence in the early 1990s and a credit downgrade. Keenly aware of the need to sustain foreign and domestic investor confidence in the nations' economic and fiscal policies, the Fiscal Responsibility Act (FRA) was passed in 1994 with a goal of reducing the country's net debt to between 20 and 30 per cent of the GDP. Under the FRA, New Zealand policymakers must consider the impact of fiscal policy on variables such as debt burden and national wealth. The entire budget surplus was committed to debt reduction with a promise to cut taxes once the debt target was achieved. When the debt target was attained, a tax cut was enacted and the government then reset its debt target to 20 per cent of the GDP and further lowered it to 15 per cent in 1998. This occurred during a period of coalition government. The minority partner in the coalition was an advocate of social spending and the larger National Party was an advocate of continued debt reduction and the parties agreed to implement a spending package which allowed for continued surpluses. Even when the Asian economic crisis occurred in 1997, which would have resulted in a budget deficit, the government enacted a package of spending cuts while scaling back previously promised spending increases.

Both these countries justified surpluses by expressing a clear set of fiscal targets and making changes to the budget process.** In Norway the budget was usually the outcome of the aggregate of spending decisions on individual items of expenditure. This was changed to a top-down approach where the Parliament agrees first on an overall fixed-expenditure limit and ceilings on 23 spending and two income areas at the beginning of the budget process. All spending and revenue proposals must then fit within these ceilings. In New Zealand policymakers began to focus on its balance sheet, assets and liabilities. This made them aware of long-term unfunded liabilities. The insurance premiums on the accident insurance programme, for instance, were increased to fully fund the programme. These examples prove that going beyond zero balances and towards fiscal surpluses is achievable.

* J. White and A. Wildavsky, *The Deficit and the Public Interest* (Berkeley, CA: University of California Press, 1990).

** P. L. Posner and B. S. Gordon, "Can Democratic Governments Save? Experiences of Countries with Budget Surpluses," *Public Budgeting and Finance* (Summer 2001): 1–28.




› **Figure 12.3**
Primary Surpluses. For the case of primary surpluses, $d_t < 0$. When the interest rate is less than the growth rate of GDP as in panel A, the steady-state level of the debt is stable and if the economy at any moment of time has a debt/GDP value other than $\bar{b}$, it will gravitate towards $\bar{b}$. This is where the debt is repaid and the government has accumulated a stock of financial assets that earn it a return. If the interest rate is greater than the growth rate of GDP as in panel B, the steady-state level of the debt is unstable. And if the economy at any moment of time has a debt/GDP value other than $\bar{b}$, that diverges away from $\bar{b}$ over time.

revenue to service the interest on it. This is a policy of pursuing budget balance. Under this balanced budget policy, the government would attempt to set its fiscal deficit as close to zero and in the limit to equal zero. If $(G - T) + rB = 0$, then because the interest payments on debt are positive, $rB > 0$, we must have $(G - T) < 0$—the government runs a primary surplus.

Primary surpluses generate the revenues that pay off the interest on the debt that was raised to finance deficits. They make the present value of the net worth of government more non-negative but do not ensure government solvency. A government which over long periods of time never generates a primary surplus in its budget can in this sense be considered to be prima facie insolvent.

### *12.2.2* Primary Surpluses

In Figure 12.3 the government runs a primary surplus. The situation is different, however, for the stable steady-state case (panel A) as compared to the unstable steady-state case (panel B).

Panel A describes the situation where if the government continues to run a small primary surplus given by $d$ and its initial debt/GDP ratio was $b^*$, it would eventually repay its debt (as the zigzag arrow depicts). It would also accumulate a stock of positive assets/GDP given by $\bar{b}$.

In panel B, in contrast, even though the government runs a small primary surplus, the initial debt/GDP ratio of $b^*$ diverges away from the steady-state debt/GDP ratio of $\bar{b}$. As $\bar{b}$ is positive, the government, if it manages to nudge the debt/GDP ratio to its steady-state value, will still perpetually be indebted and never pay back the debt of $\bar{b}$ raised—it is insolvent. (If $b^* < \bar{b}$, however, the government will get rid of debt and soon enter the third quadrant where it is accumulating assets at an increasing rate.)

## 12.3 Debt and Deficits in India

For most of the first half of the 1990s the Indian economy was in a panel A of Figure 12.2 type situation. During 1997–1998, 1999–2000, and 2000–2001, however, (see Table 12.1) the interest rates shifted up and were higher than the growth of GDP as depicted in panel B of Figure 12.2. This is a situation where the

**› Table 12.1**
Debt and Primary Deficits in India
*Source*: Reserve Bank of India, Handbook of Statistics on Indian Economy.

| Year | Debt Liabilities of Government as Ratio of GDP | Actual Primary Deficit as % of GDP | Growth Rates of GDP (Factor Cost) | Real Interest Rates on Government Securities |
|---|---|---|---|---|
| 1990–1991 | 0.66 | 5.59 | 5.57 | 0.93 |
| 1991–1992 | 0.65 | 2.52 | 1.30 | −2.02 |
| 1992–1993 | 0.65 | 2.37 | 5.12 | 3.84 |
| 1993–1994 | 0.66 | 3.58 | 5.90 | 3.18 |
| 1994–1995 | 0.64 | 2.11 | 7.25 | 2.57 |
| 1995–1996 | 0.63 | 1.73 | 7.34 | 4.77 |
| 1996–1997 | 0.61 | 1.38 | 7.84 | 6.27 |
| 1997–1998 | 0.64 | 2.34 | 4.79 | 5.49 |
| 1998–1999 | 0.65 | 4.00 | 6.51 | 4.02 |
| 1999–2000 | 0.70 | 4.22 | 6.06 | 7.85 |
| 2000–2001 | 0.75 | 3.94 | 4.37 | 7.47 |
| 2001–2002 | 0.80 | 4.04 | 5.79 | 5.99 |
| 2002–2003 | 0.85 | 3.37 | 3.98 | 3.18 |
| 2003–2004 | 0.88 | 2.25 | 8.51 | 2.82 |
| 2004–2005 | 0.89 | 1.50 | 6.91 | 1.12 |
| 2005–2006 | 0.86 | 1.79 | 8.43 | 3.32 |

N.B.: The interest rates on government securities is the weighted average of interest rates on Central and state government securities less the inflation rate as measured by the GDP (factor cost) deflator.

current debt/GDP is slowly exploding as the debt/GDP ratio increases over time. By May 2000, the RBI's Monetary and Credit Policy Statement had noted that "such high levels of fiscal deficit are not sustainable over the medium term".

For estimating an equation such as $b_t = (1 + r/1 + g)b_{t-1} + d_t$, it is advisable to look at a data period longer than just a decade.

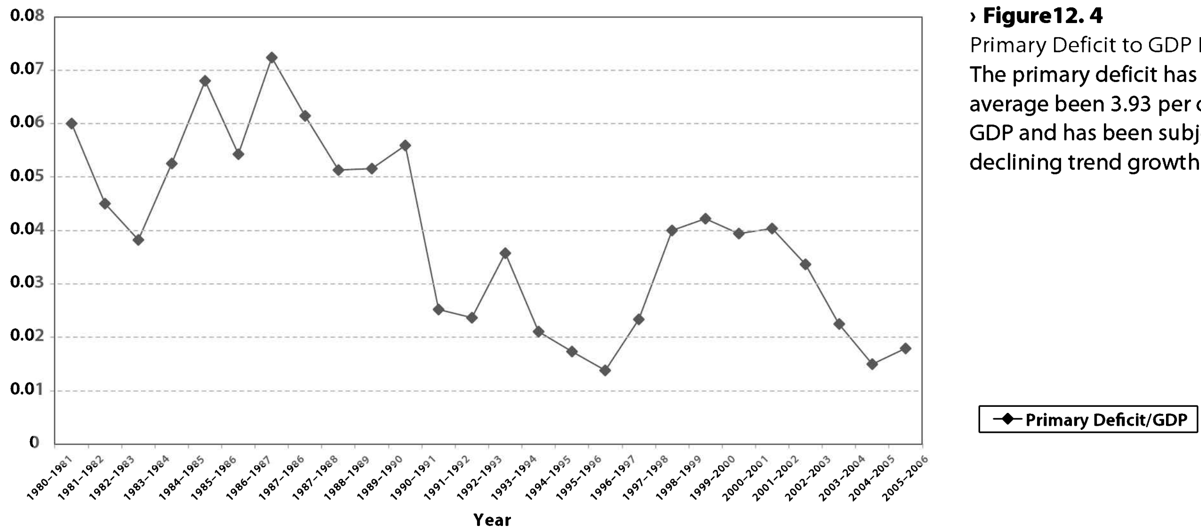
The summary statistics for the debt/GDP ratio, primary deficit/GDP ratio, the growth rate of GDP, and the real interest rate on government debt for the central and state governments combined for the period 1980–1981 to 2005–2006 are given in Table 12.2.

The accompanying diagrams—Figures 12.4 to 12.6—chart out the time series of these variables. As can be seen, the interest rate on government securities has been below the GDP growth rate for most of the period and only in the recent past has the real interest rate exceeded the growth rate of GDP (an unstable situation). As Table 12.2 reveals, the average real interest rates on government securities have risen in the nineties compared to the eighties and this has contributed to the rising debt as $b_t = (1 + r/1 + g)\, b_{t-1} + d_t$ increases with $r$.

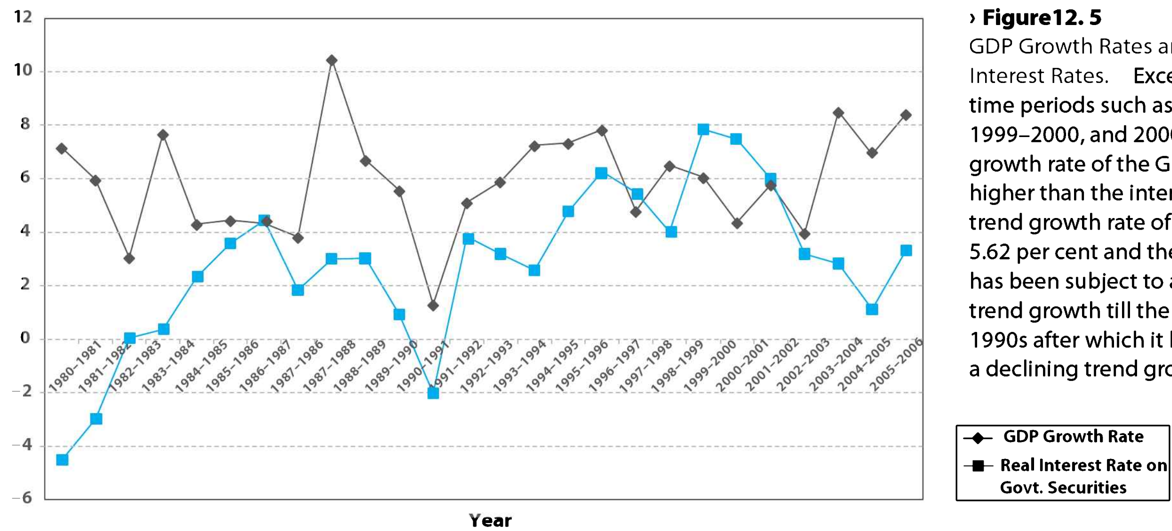
**› Table 12.2**
Summary Indicators of Debt and Deficits
*Source*: Reserve Bank of India, Handbook of Statistics on Indian Economy.

| | Primary Deficit/GDP 1980–1981 to 2005–2006 | Debt/GDP 1980–1981 to 2005–2006 | GDP Growth Rate 1980–1981 to 2005–2006 | Real Interest Rate on Government Securities 1980–1981 to 2005–2006 | Real Interest Rate on Government Securities 1990–1991 to 2005–2006 |
|---|---|---|---|---|---|
| Average | 0.0393 | 0.66 | 5.91 | 2.77 | 4.12 |
| Standard Deviation | 0.02 | 0.12 | 1.97 | 2.96 | 2.70 |
| Trend Growth Rate (%) | −4.20 | 2.29 | 5.62 | | |

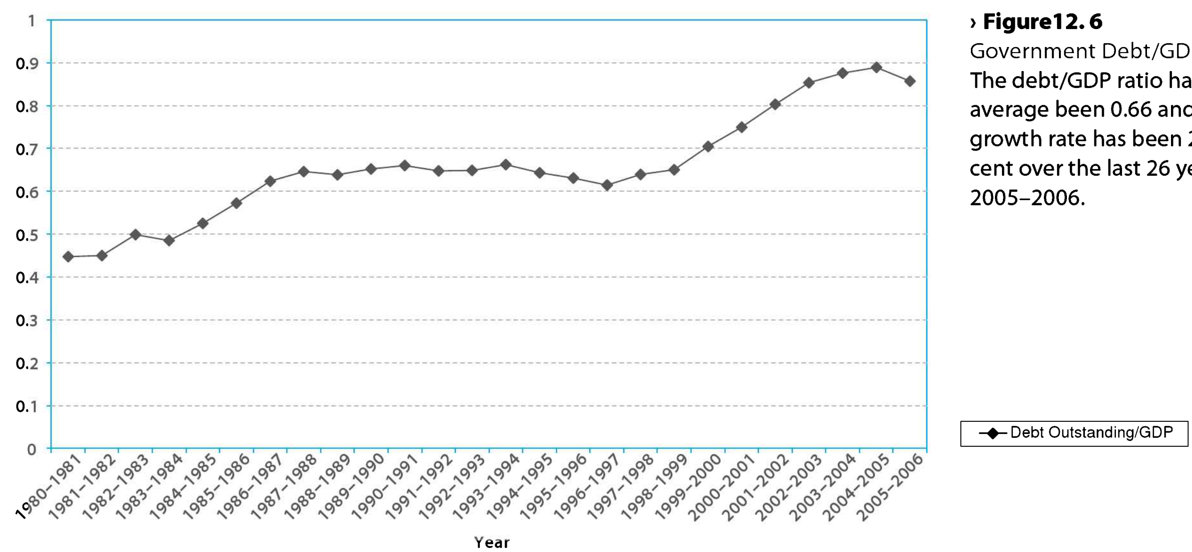
N.B.: Real interest rate on government securities series has a statistically insignificant linear trend. The trend growth rate is quadratic (inverse U shape) with a maximum around 1994–1995 for the entire sample and a maximum in 1999–2000 for the period 1992–1993 to 2005–2006. On a trend basis then interest rates rose over the period up to the latter half of the 1990s and then began to decline.




**› Figure12. 4**
Primary Deficit to GDP Ratio.
The primary deficit has on an average been 3.93 per cent of the GDP and has been subject to a declining trend growth rate.




**› Figure12. 5**
GDP Growth Rates and Real Interest Rates. Except for some time periods such as 1986–1987, 1999–2000, and 2000–2001, the growth rate of the GDP has been higher than the interest rate. The trend growth rate of the GDP is 5.62 per cent and the interest rate has been subject to an increasing trend growth till the end of the 1990s after which it has been on a declining trend growth path.




**› Figure12. 6**
Government Debt/GDP Ratio.
The debt/GDP ratio has on an average been 0.66 and its trend growth rate has been 2.29 per cent over the last 26 years till 2005–2006.

Using the data for the Indian economy for the period of 1980–1981 to 2002–2003, we estimate[4] the equation for debt as[5]

$$b_t = 0.0420 + 1.05 b_{t-1}$$

That is, the average primary deficit for the period 1980–1981 to 2002–2003 was $d = 0.0420$ and $[(1+r)/(1+g)] = 1.05$. This implies that as $[(1+r)/(1+g)] = 1.05$, which is greater than 1, the debt/GDP ratio up to 2002–2003 was slowly rising and the government was insolvent and would never pay back its debt. (This is a panel B of Figure 12.2 type situation.) In reaction to this kind of situation the government passed the Fiscal Responsibility and Budget Management Act (FRBMA) that stipulated numerical targets for the reduction of the deficit. Since then the primary deficit has reduced (see Figure 12.4) and if we re-estimate the equation for debt from 1980–1981 to 2005–2006 we obtain

$$b_t = 0.0393 + 0.97 b_{t-1}$$

The average primary deficit/GDP for the period 1980–1981 to 2005–2006 is 0.0393 (Table 12.2) and the coefficient of $b_{t-1}$ is now $[(1+r)/(1+g)] = 0.97 < 1$, which means that the government has been able to stabilize the deficit and put it on the path towards sustainability (panel A of Figure 12.2 type situation).

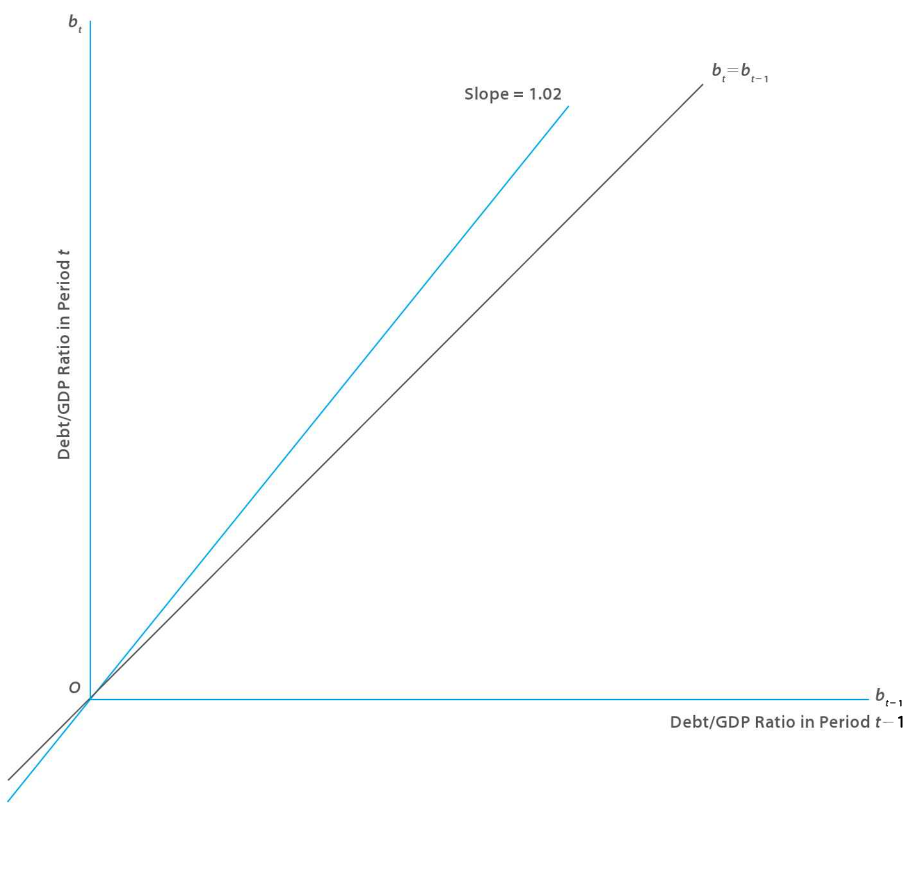
However, the steady-state level of debt is still large ($\bar{b} = 1.19$) and fiscal correction is still required to reduce this large steady-state level of debt.

Even if the primary deficit/GDP ratio is set equal to zero, the estimated equation for debt is

$$b_t = 1.02 b_{t-1}$$

where $[(1+r)/(1+g)] = 1.02 > 1$, therefore, the debt is slowly exploding. This evolution of the debt is graphed in Figure 12.7.

This is a situation for which corrective action is needed. That means running primary surpluses. If $r$ rises, the government will sink slowly into bankruptcy. From




› **Figure 12.7**
Zero Primary Deficit and Evolution of Debt. **A strategy of reducing the primary deficit/GDP ratio to zero will still not result in stabilizing the debt/GDP ratio, which will increase over time as the steady state associated with this strategy is unstable.**

the point of view of the present, however, the government as of today is insolvent since it will never repay its debt. It can and does roll over debt to repay interest on the debt and the growth in GDP enables it to pursue this strategy and put off repayments of the principal on the debt for a future government to pursue.

## 12.4 Sustainable Primary Deficits

Another way to analyse the long-term debt and deficit scenario is to ask at any point in time that, if the historically given debt at that time is to be continued at that constant level forever, and if there is no change in interest rates and the growth of GDP—the current configuration of the economy continues to prevail—what is the primary deficit that can sustain this time path of the economy? The answer is derived from the debt equation:

$$b_t = d_t + \frac{1+r}{1+g} b_{t-1}$$

or,

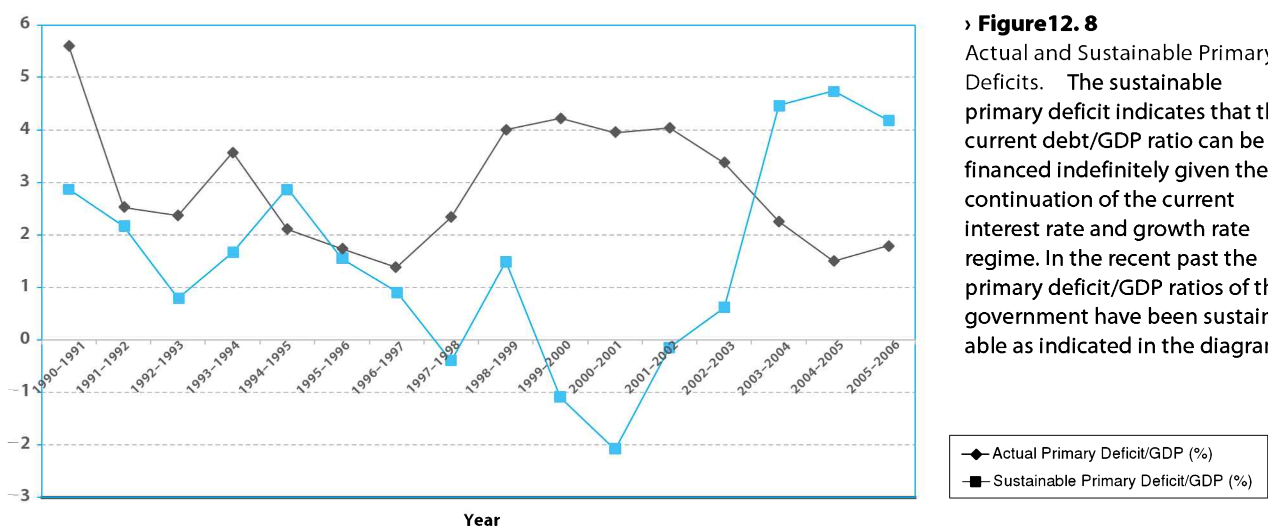
$$1 - \left[\frac{1+r}{1+g}\right] b_{t-1} = d_t \qquad \text{when} \qquad b_t = b_{t-1}$$

or,

$$d_t = \frac{g-r}{1+g} b_{t-1}$$

This gives the primary deficit/GDP ratio that is required in order to sustain the existing level of debt in the economy. Figure 12.8 denotes that the sustainable primary deficit/GDP ratio has been below the actual primary deficit/GDP ratio for most of the nineties. In 1994–1995 the primary deficit was sustainable due to a favourable circumstance that the growth of GDP at 7.3 per cent was among the highest in the decade and the interest rate at 2.6 per cent was among the lowest in the decade (See Table 12.1). This enabled the government to reduce the debt required to be raised to make interest payments and pay the principal on the debt. Thus, the debt/GDP ratio reduced in the next year, 1995–1996.

In fact, the primary deficit/GDP ratio should be below 1.5 per cent of GDP (the average sustainable primary deficit since the beginning of the 1990s) if current levels of debt are to be continued indefinitely. A fiscal correction cannot be delayed indefinitely. As the graph shows there was a sharp fiscal




**› Figure 12.8**
Actual and Sustainable Primary Deficits. The sustainable primary deficit indicates that the current debt/GDP ratio can be financed indefinitely given the continuation of the current interest rate and growth rate regime. In the recent past the primary deficit/GDP ratios of the government have been sustainable as indicated in the diagram.

correction in 1990–1991 after which there was a decline till 1996–1997 in the primary deficit/GDP ratio except for the solitary year 1993–1994. After 1996–1997, however, till 2002–2003, the fiscal situation worsened again and the increased primary deficits during this period required substantial fiscal correction policy initiatives afterward.

## 12.5 Fiscal Rules for Adjustment

Before expounding on what are the constituents of a fiscal correction it should be mentioned that apart from debt the deficit is financed by money creation. Thus,

$$(M_t - M_{t-1}) + (B_t - B_{t-1}) = rB_{t-1} + G - T$$

where $M_t$ is domestic base money.[6] Hence,

$$(B_t - B_{t-1}) = rB_{t-1} + (G - T) - (M_t - M_{t-1})$$

and debt creation is smaller to the extent that some of the deficit is financed by money creation. In a growing economy some extra real money balances are demanded to finance the growing level of transactions. Beyond that the demand for nominal money expands only to the extent that inflation erodes the purchasing power of existing real balances. To restore real balances the public adds at least partially to nominal money holdings. In this way inflationary finance automatically creates a demand for the money issue that finances the deficit. The inflation tax in India is estimated to be 1.5 per cent of GDP.[7] This means that the sustainable primary deficit/GDP ratio for 2000–2001, when we account for the inflation tax, is (−2.09 + 1.5) or −0.59 per cent of the GDP and not −2.09 per cent of the GDP, as Figure 12.8 depicts.

The analysis of the dynamics of the debt makes clear that a few years of debt decline is not enough to attain sustainability. What may be needed is a procedure that guarantees that the debt/GDP ratio is stabilized year after year. There are two important points to note about the INTERTEMPORAL BUDGET CONSTRAINT of the government.

› The **intertemporal budget constraint** is an accounting identity that does not recommend how adjustments to budget balance should be made. The **intertemporal budget constraint** is silent about the timing of any required adjustment to the budget.

First, it is not a behavioural equation and so it is silent on the question of which of the government's revenue and expenditure variables should be adjusted in the face of disequilibrium. Below we indicate what sort of changes in government expenditure and tax revenues can be initiated to bring the deficit under control but these policy recommendations are independent of the long-term discipline required for fiscal sustainability.

Second, the intertemporal budget constraint is also silent on the question of the timing of any required adjustment and simply specifies that an adjustment must occur, sooner or later. In practice, this keeps open the possibility of delay in undertaking what could be a costly adjustment programme for the government, and deficits may just balloon up. To keep out this possibility governments have been enacting rules that set strict limits on the size of deficits. Chile and Brazil, for instance, have set strict limits on the budget deficit.

› A fiscal rule is a restraint on the fiscal policy defined in terms of an indicator of the overall fiscal performance. In India the rule is defined in terms of the fiscal deficit (3 per cent of GDP by 2008–2009) and a public-debt rule disqualifying borrowing from the central bank.

In Chile the rule, in fact, calls for a structural surplus of 1 per cent of GDP. The Brazilian rule establishes that the Congress sets limits to spending on personnel over a three-year horizon while the government commits to a pre-announced primary budget balance.

In India, the Parliament passed the FISCAL RESPONSIBILITY AND BUDGET MANAGEMENT ACT (FRBMA) in August 2003, which requires the fiscal deficit of the central government to not exceed 3 per cent of GDP by 2007–2008 and that the deficit on the revenue account be eliminated by the same date. The specified annual reductions in the two measures are 0.5 percentage point of GDP or more for the revenue deficit and 0.3 percentage point of GDP or more for the fiscal deficit. The FRBMA was amended in July 2004 to shift the terminal date for achieving the numerical targets pertaining to the fiscal indicators by a year to 2008–2009. The Act also disqualifies the RBI from subscribing to government paper after 31 March 2006.[8]

In February 2004 the government constituted a task force to devise a strategy for implementing the FRBMA and the recommendations of the task force were made public in July 2004. The task force made recommendations to enhance direct taxes by 2 percentage points of GDP and to shift the revenue base of indirect taxes to include a greater share of services. In addition to the central government, several states in India have passed fiscal responsibility acts. Karnataka was the first to enact such legislation in 2002 and it was followed by Kerala, Tamil Nadu, Punjab, Uttar Pradesh, and Maharashtra.

In all these cases the common trend has been to include an imposition of quantitative and time-bound targets for fiscal consolidation. Most states have stipulated a reduction in the fiscal deficit as a per cent of the state domestic product (SDP) to 3 per cent by 31 March 2010.

**MACROFOCUS 12.2**

### Are Fiscal Rules a Good Way to Bind Policymakers?

A fiscal rule is a restraint on fiscal policy and is typically defined in terms of an indicator of overall fiscal performance.* Among the various types of fiscal rules that have been adopted are the following:

1. *Expenditure-limiting rules* that cap the level or growth in total government expenditure or specific categories of government spending. For example, Brazil has a rule that limits spending on the government wage bill and Argentina has a limit on primary expenditure. The United States also has a rule that limits discretionary expenditures.
2. *Current balance rules* stipulate that current government spending should not exceed current government revenue. This amounts to a rule that governments may not borrow to finance current expenditure. Such rules apply in Brazil and the UK.
3. *Overall-balance rules* prescribe limits to the extent of the budget deficit. Such rules are in force in India, Chile, Peru, and the European Union. The Chilean rule specifies in fact a structural surplus of 1 per cent of the GDP on an annual basis.
4. *Public debt rules* limit the extent of the public debt and in some cases prohibit borrowing from sources like the central bank. Such rules are in force in India, Indonesia, the UK, the European Union, and the CFA franc zone.

In practice, governments find it easy to circumvent, suspend, or even abandon fiscal rules. A good example is the Balanced Budget and Emergency Deficit Control Act—the Gramm-Rudman-Hollings legislation—introduced in the United States in 1985. This act set targets to reduce the projected budget deficit to zero by 1991, which was later extended to 1993. The act applied to a projected deficit and not an actual deficit and the authorities used this to generate optimistic forecasts and projections that met the targets even though the actual deficits consistently exceeded the limits.

Even in the European Union the Maastricht Treaty, which came into force in 1993 stipulated a condition that entry into the EU requires overall fiscal balances not exceeding 3 per cent of GDP. During 1993 to 1999 this fiscal balance rule was followed but thereafter many countries breached the rule. By the end of 2003, for instance, France and Germany which were in a recession, had breached the deficit limit for the third year running but did not implement the remedial measures suggested by the European Commission. Even though the rules stipulated that both countries should be fined heavily, in November 2003, the majority of the EU finance ministers voted not to impose fines.

Rules may not bind governments to sound fiscal policies. The sanctions for non-compliance such as reprimands by courts or financial penalties on officials (elected or appointed) are too weak to deter governments. Also policymakers often have limited control over fiscal outcomes. In the short run a GDP shock that impacts on tax revenues is not amenable to be controlled by changing tax rates due to the inside lags associated with policy.

Cowen, Glazer, and Zajc**, in fact, suggest that when the durability of rules is in doubt, discretionary policy regimes are better. This is because when there is a preconceived mistrust by economic actors that government may not follow the rule, then, the existence of a rule prevents a government from demonstrating its underlying preference for sound policies. A government with discretionary powers can actually signal its commitment to prudent policy and establish a reputation for sound policy that makes its policy credible. Hence, rules based policy cannot be credible unless buttressed by a pre-existing reputation of responsible policymaking.

* G. Kopits and S. Symansky, "Fiscal Policy Rules," IMF Occasional Paper No. 162 (Washington, D.C.: IMF, 1998).
** T. Cowen, A. Glazer, and K. Zajc, "Credibility may Require Discretion, not Rules," *Journal of Public Economics* 76 (2002): 295–306.

In all of the cases, whether at the centre or at the level of the states, the limits to the deficit are self-imposed and there are no sanctions when the deficit limits are reached. As the rules are generally non-binding there are chances that they will be set aside when economic and political expediency are required. This has already occurred in India in 2005–2006 when the target for the revenue deficit was put into a "pause" mode by the Finance Minister on account of a higher demand on resources arising out of the award of the Twelfth Finance Commission.

A problem with rules for deficits then is that they are not flexible enough in the face of unexpected events. This debilitates the discipline that the FRBMA is designed to deliver. Rules like the FRBMA work to promote the long-run quest for discipline in spending but at the cost of ignoring short-run exigencies. The limit of 3 per cent of GDP on the fiscal deficit is also somewhat arbitrary and seems to have been selected from the same limit put for European Monetary Union membership in the EU's Stability and Growth Pact.[9]

This means that it is considered alright if public investments, presumed to average 3 per cent of GDP, are financed through borrowing. Thus deficits are tolerated as long as they do not exceed the size of public investment spending.

There is no guarantee, however, that public investments generate a rate of return that matches the cost of borrowing. Also there is no economic rationale for what in budgetary accounting is referred to as public investment. Education is classified as a public consumption expenditure while the building of new government offices is considered an investment. However, public spending on education is likely to be far more socially productive than building a new government building. Similarly, defence and police expenditures are treated as government consumption expenditures while arguably they may be considered as an investment in security. Rules meanwhile do not leave room for flexibility in the face of unexpected events such as was the case in 2005–2006 and such instances can occur again if there is a contingency, such as a war or external threat and prolonged economic downturn. Also rules do not matter much if they are not backed by an enforcement mechanism. This leaves the task of deficit containment in the face of unsustainable levels of debt to the presence of competent and dedicated policymakers and politicians who have good judgment and know when to exercise discretion and when to press for fiscal discipline. Of course, that is a tall order.

## 12.6 Fiscal Restructuring

Two questions are really raised in budget balancing as an element of adjustment policy. The first has already been answered above: How large a deficit is consistent with stability? The second question is: How to cut deficits down to the required size?

› Fiscal adjustment to meet a rule requires expenditure restructuring and the introduction of a productive tax structure.

› Taxes are collected by applying a rate structure to a tax base. Income, consumption, and wealth are the major tax bases.

FISCAL ADJUSTMENT involved expenditure restructuring and the introduction of a productive tax structure. Prior to reforms, the Indian income-tax laws, for instance, had resulted in various distortions such as the exclusion of perquisites, the exemption of income up to a certain level from taxation and from particular sources such as agriculture. In addition, they treated certain items of income such as capital gains to a lower tax rate. This eroded the BASE OF THE INCOME TAX and it became necessary to have higher rates of tax to raise a given volume of revenue. High rates of tax adversely affect economic incentives as the after tax returns are reduced and they even induce evasion when enforcement by the authorities is weak. Tax reforms accordingly aimed at securing a broad tax base, without exemptions, and only a few rates of tax. The 1991 Tax Reforms Committee has laid the agenda for this. Some of the unfinished issues here are an *extension of the tax net* to the agricultural and services sectors, reductions in the corporate and personal income tax *exemptions*, such as for the export and small-scale sectors, strengthening *tax administration* to reduce corruption and evasion to increase compliance and reduce the undermining of the collection of taxes, and moving to a full-fledged VAT with some surcharges for luxuries.[10] Extension of the tax net has been significant but improvements in fiscal administration still have a long way to go.

The third important aspect of a productive tax structure is absence of significant *subsidies* of any form and establishment of *efficient public utility rates*. The economic approach advocated is for public sector prices to be set on the basis of microeconomic efficiency considerations. Income distribution consequences of prices set in this manner should be resolved through the general tax structure. During times of financial instability public sector pricing is a contentious issue. When inflation is high, policymakers go slow on public sector price increase to reduce inflation. This results in a deficit, which creates financial problems that are then resolved by raising public sector rates. For instance, the government in India subsidizes kerosene and domestic LPG as these are deemed to be consumed by economically weaker sections of society. Oil marketing companies cannot adjust retail selling prices of these products in line with international prices. The restraint on this pass through of international prices to domestic consumers adversely affected the margins of oil marketing companies as international oil prices began to rise from late 2003. Rather than allow these companies to raise prices, the government issued oil bonds to them to help tide over their financial problems. To repay these bonds in the future the government will have to raise domestic prices and if international prices do not soften, the inflationary impact on the economy could be substantially larger than if the prices were raised today.

This see-sawing is bad for the economy and instead public-utility rates should be indexed on a regular basis even when more indexation implies greater vulnerability to inflation. Inflation should be tackled by budget balance and not by temporary freezes on public sector price increases. The aim of the government should be to eliminate all subsidies as far as possible. The resulting revenue gains should be deployed partly to restore fiscal balance and partly to reduce inflation, which as inflation is a regressive tax does raise the real incomes of the poor. Part of the revenue gains should be used for targeted food and employment programmes for poor groups.

Currently the central bank's "quasi-fiscal" deficit should also be made part of the fiscal accounting so as to make the accounting system transparent. Quasi-fiscal deficits arise from loans by the central bank at subsidized rates, losses on foreign exchange operations in the form of guarantees, and subsidized credit that is not strictly an investment subsidy but rather a production subsidy that finances wages and profits.

On the expenditure side a number of reforms that are important are:

1. *Cuts in public sector employment*—Civil service reforms are needed to reduce the public sector wage bill and improve efficiency. These include the 30 per cent staff cuts recommended by the Fifth Pay Commission, increases in the retirement age to address long-term pressures on pension budgets, and improvements in mechanisms to adjust pay scales. These are difficult issues as employment in the public sector is bloated due to patronage as well as poor accountability.

   A second important issue on the expenditure side is *privatization* and the closing of public sector firms. Privatization is advocated for three reasons. First, the public sector often does not have the managerial capacity to administer in a cost-effective and quick-response manner. Second, there is a scarcity of resources with the public sector and it does not have the capital resources to provide all public services well. The provision of public and merit goods could be privately provided and the degree of public funding of this provision could be worked out in the interests of promoting equity. Third, privatizations are often delayed because of the perceived low prices received from privatization. However, if failure to privatize and obtain resources results in fiscal precariousness and inflationary pressures then even lower prices may be obtained from a delayed privatization programme.

2. *Maintenance and investment expenditures* on social and economic infrastructure—The fiscal problem is not just about firing public sector employees and privatization. There is an urgent need to restructure public sector spending from consumption to investment and productive services. Most spending should be diverted to infrastructure, health and social services for poorer groups. Available resources should be diverted to sectors in which private initiative is less willing or able to function. Infrastructure spending on telephone companies, steel mills, and airlines which need not be in the public sector is a misallocation when it is at the expense of spending on health and education.

   The extent and manner of fiscal adjustment is but one element of an adjustment programme. Other elements include an appropriate MONETARY POLICY and an appropriate level of the EXCHANGE RATE. In some countries with hyperinflation there is a fourth pillar of adjustment programmes, incomes policy, which becomes important. We discuss these other pillars of stabilization policy in a summary fashion.

› Fiscal adjustment must be accompanied by a monetary policy that keeps the real interest rates low. Real exchange rates should also not be overvalued as competitiveness gets reduced, which affects the aggregate demand. In high-inflation economies an incomes policy that indexes wages is also important for controlling inflation expectations.

3. *Monetary policy*—A tight monetary policy is not a substitute for balancing the budget which is the essential aspect of stabilization. Monetary policy does not, however, play an independent role in adjustment—it is dictated by the budget and exchange rate policy. A tight monetary policy that results in high real rates of interest is more a signal of serious misalignments in the budget, the real exchange rate, or both. Real interest rates should be low (US interest rates plus say 3 per cent) and such a situation is sustainable if fiscal and real exchange rate policies are sound. Following stabilizations, real interest rates usually are high and debtors often go into distress due to these high real interest rates. Many argue that some monetary expansion is required to reliquify the economy. Reliquification is, however, best done by monetizing reserve inflows rather than by deficit finance or domestic credit creation.
4. *Exchange rate policy*—This plays a strategic role in adjustment. "When fiscal austerity reduces demand, full employment growth requires an offsetting mechanism for crowding in. A competitive real exchange rate does provide such a mechanism."[11] Competitiveness requires reduced inflation and adjustment programmes that are undertaken during high inflation would do well to fix the exchange rate. The downside of this is that if inflation does not end soon the government will have to abandon the fixed rate and signal its acceptance of inflation as inevitable which has expectational effects. Governments, therefore, prefer a floating exchange rate but as they are concerned then with targeting low inflation there is a tendency for this to be accompanied by a tight monetary policy which raises interest rates sharply, attracts capital inflows, and leads to a real appreciation of the currency which becomes overvalued. The high interest rate increases debt servicing and worsens the budget and overvaluation creates an expectation of devaluation. If there is political resistance to the devaluation, the policy makers are forced to defend the overvalued exchange rate through tight money which puts pressure on the fiscal budget. If the government cannot absorb the fiscal pressure it succumbs by allowing the exchange rate to adjust which

is just as well because it is better to remove pending problems such as large fiscal deficits rather than to create new ones like an overvalued exchange rate that generates speculative expectations of reversal of tight monetary policy and an eventual collapse of the exchange rate. Real exchange rates should not be overvalued from the start of the adjustment programme as that sacrifices competitiveness and delays the onset of growth.[12]

5. *Incomes policy*—In periods of high inflation, without an incomes policy adjustment has lower chances of success. As inflation accelerates, contracts shorten as people seek to protect their real incomes and the shortening of contracts in turn causes inflation to accelerate. If built in inflation expectations are prevalent in contracts through indexation clauses then it is important to break the inflation inertia that prevails due to the indexation. Under wage indexation an increase in inflation results in an increase in wage costs that are often passed on into higher prices. Current inflation accordingly is tied to the past inflation in the economy. Exchange rate depreciation, a positive value of $E_{t+1} - E_t$, also has pass-through impacts on domestic prices as imports become costlier. Finally, inflation is also responsive to the tightness in the labour market or the unemployment rate—the Phillips curve. Current inflation, $\pi_t$, can then be considered to be a weighted average of past inflation, exchange rate depreciation, and the unemployment rate.

$$\pi_t = \tau\pi_{t-1} + (1 - \tau)(E_{t+1} - E_t) - \phi u_t$$

With current inflation linked to past inflation and depreciation of the currency, and incomes indexed to inflation, current inflation cannot break away from past inflation unless the government intervenes and controls the exchange rate depreciation and disallows renegotiation of wages and incomes. It is only in an economy where individuals are entirely forward looking and wages and prices are fully flexible will income freezes not be required as individuals will expect lower prices (presuming they believe the stabilization programme is credible) in the future and not renegotiate nominal incomes. However, this may be an extreme assumption to make for an economy and as long as people are not all that forward looking incomes policy has a role to play in shifting wages and prices to a new regime. However, incomes should not just be frozen but indexation on a periodic annual basis should be part of incomes policy. Without explicit indexation the government is under pressure to grant wage increases and the income setting process becomes politicized with the chances of large wage increases being sanctioned by the government. Indexation helps establishing the expectation of a low inflation.

There are two other points that need to be investigated in the context of large fiscal deficits. One is that large fiscal deficits are seen to generate pressure on the current account as the increased income is spent abroad (in the form of higher imports) and the higher inflation rate domestically relative to inflation abroad also adversely impacts on the current account. So to what extent can an economy finance an expansionary fiscal regime from abroad rather than in the domestic capital market? This requires us to look at the determinants of the current account and to indicate when current account deficits are "excessive" and should not be relied on further to finance domestic growth. The second issue concerns the financial sector (a supply side structural policy issue). High fiscal deficits in developing countries till recently were accompanied by controlling the interest rates of depository financial institutions. Apart from interest rate limitations, these institutions have also been required to hold reserves or government debt at controlled rates. This affected the efficiency of financial intermediation whereby savings were not allocated efficiently among the alternative investment projects. Financial sector reforms are thus tied to stabilization efforts that are usually discussed only around the following elements: (a) the extent and manner of fiscal stabilization, (b) the appropriate monetary policy, (c) the appropriate exchange rate policy, and (d) the use of an incomes policy. A more holistic view of the impact of stabilization policy needs to discuss the current account and financial liberalization as well. We turn to these issues in the following chapters.

## SUMMARY


- Automatic stabilizers are expenditures that automatically increase or taxes that automatically decrease when aggregate income and output decline.
- Automatic stabilizers result in a steeper aggregate demand curve and reduce the variability of output arising from economic shocks to aggregate demand.
- Discretionary fiscal policy results from deliberate and conscious choices by the government on public expenditure and taxes so as to achieve its macroeconomic goals.
- The fiscal deficit is government expenditure that is unrelated to the repayment of debt, $G$, plus interest payments on the debt, $rB$, less government revenues, $T$. Fiscal deficit $= G - T + rB$.
- The difference between the fiscal deficit and interest payments on the debt is the primary deficit. Primary deficit $= G - T$.
- The fiscal deficit is financed by increasing the stock of debt-related government securities, $B_t - B_{t-1} = \Delta B$.
- The debt/GDP ratio increases as (1) the government issues new debt to cover a primary deficit as well as (2) to pay interest on the existing debt. An increase in the growth rate of the GDP reduces the debt/GDP ratio.
- The steady-state debt/GDP ratio is that value of the debt/GDP ratio that satisfies the government budget constraint and which is independent of time, that is, $b_t = b_{t-1} = \bar{b}$, where $\bar{b} = [(1 + g)/(g - r)]d$.
- A steady-state is stable if the debt/GDP ratio moves closer towards the steady-state value of the debt/GDP ratio over time, regardless of where it starts from. This occurs when $r < g$.
- A steady-state is unstable if the debt/GDP ratio moves further away from the steady-state value of the debt/GDP ratio when we look at its trajectory starting from any point other than the steady-state point itself. This occurs when $r > g$.
- The intertemporal budget constraint is an accounting identity that does not recommend how adjustments to budget balance should be made.
- The intertemporal budget constraint is also silent about the timing of any required adjustment.
- A fiscal rule is a restraint on fiscal policy defined in terms of an indicator of overall fiscal performance. In India, the rule is defined in terms of the fiscal deficit (3 per cent of GDP by 2008–2009) and a public debt rule disqualifying borrowing from the central bank.
- Fiscal adjustment to meet a rule requires expenditure restructuring and the introduction of a productive tax structure.
- Fiscal adjustment must be accompanied by a monetary policy that keeps real interest rates low. Real exchange rates should also not be overvalued as competitiveness gets reduced which affects aggregate income. In high inflation economies, an incomes policy that indexes wages is also important to control inflation expectations.


## NOTES

1. See Shankar Acharya, "Macroeconomic Management in the Nineties," *Economic and Political Weekly* (20 April 2002): 1515–1538.
2. Note that the tax-payment term in Eq. (A9.1.9) in Appendix 9.1 of Chapter 9 that results in tax revenues varying with income is given by $t_1$. This appears with a positive sign in the denominator on the right-hand side. If there were no income taxes, $t_1 = 0$, the denominator would have been smaller, thus making the inverse relationship between $P$ and $Y$ steeper.
3. We are taking the position here that the intertemporal budget constraint is satisfied if the debt/GDP ratio converges without specifying its target value. Some authors (Buiter and Patel, 2006) take sustainability to be the case when the debt/GDP ratio returns to its initial level which is a stronger condition that requires that a government that has outstanding debt must run future surpluses on the primary budget sooner or later. See Willem Buiter and Urjit Patel, "Excessive Budget Deficits, a Government-Abused Financial System, and Fiscal Rules," in Suman Bery, Barry Bosworth, and Arvind Panagariya (eds.) *India Policy Forum*, No. 2 (New Delhi: Sage Publications, 2006).
4. Strictly speaking, the algebraic expression for debt should be estimated in a structural time series framework. The estimate from such a framework results in a policy conclusion that replicates the one from regression analysis and so we do not report those estimates here.
5. The actual estimate is $b_t = 0.0107 + 1.0126b_{t-1}$ but we have estimated the regression for an intercept of 0.0420 equal to the average primary deficit/GDP ratio for the data period. The coefficients are significant and $R^2 = 0.92$ which indicates a good fit.

6. We could also break up the debt into its domestic ($B$) and foreign ($B^*$) components so that with a spot exchange rate of $E_t$, we would have $(M_t - M_{t-1}) + (E_t B_t^* - E_{t-1} B_{t-1}^*) + (B_t - B_{t-1}) = rB_{t-1} + r^* E_{t-1} B_{t-1}^* + G - T$. In the data above, we have included the external debt liabilities of the government as part of the debt/GDP ratio.
7. M. J. Manohar Rao and Raj Nallari, *Macroeconomic Stabilization and Adjustment* (New Delhi: Oxford University Press, 2001).
8. However, borrowing from the RBI to cover temporary excess of cash disbursements over receipts—ways and means advances—is permitted.
9. Willem Buiter and Urjit Patel, op. cit.
10. See Parthasarathi Shome, "India: Tax Policy for the Ninth Five Year Plan," Report of the Working Group on Tax Policy of the Steering Group on Financial Resources, New Delhi, 1997; and the Kelkar Committee Report.
11. Rudiger Dornbush, "Policies to Move from Stabilization to Growth," Proceedings of the World Bank Annual Conference on Development Economics, 1990.
12. Many Latin American countries resorted to crawling pegs where the exchange rate is depreciated at a pace that maintains external competitiveness in order to tackle the real appreciation problem.

## TEST YOURSELF

1. Distinguish between automatic and discretionary fiscal policy.
2. Define the intertemporal budget constraint of the government. How is it related to the sustainability of the debt?
3. Why is the level of debt not as suitable a measure as the debt/GDP ratio of a country's debt obligations?
4. What is the difference between the fiscal deficit and the primary deficit? Which is a more appropriate measure of the deficit of the government and why?
5. Define a steady state. When is a steady state stable and when is it unstable? What condition must be satisfied to ensure that the debt/GDP ratio results in a stable steady state?
6. India is said to be experiencing a "demographic dividend" in that a large section of the population is in the labour force. What is the likely impact of this on future deficits?
7. In the case where the debt/GDP ratio is on an unstable steady-state path what should the government do to reduce the debt?
8. Discuss the deficit and debt scenario in India between 1980–1981 and 2002–2003. In what way would FRBMA alter the debt and deficit position of the government?
9. What are fiscal rules and how effective are they in restraining the growth of government debt?
10. What correlated macroeconomic policy is required to sustain a fiscal adjustment?

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
2. Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy"
3. Click on "Table 118: Combined Deficits of Central and State Governments" and insert data into a spreadsheet on the Gross Fiscal Deficit and the Gross Primary Deficit from 1980–1981 to 2005–2006.
4. Then click on "Table 1: Macro-economic Aggregates (at current prices)" and insert data on GDP at factor cost (Base 1999–2000) for the period 1980–1981 to 2005–2006.
5. Calculate the Primary Deficit and Fiscal Deficit as a percentage of GDP for each year. For the year 1980–1981 for instance, the primary deficit was INR 78.18 billion and the GDP was INR 1325.20 billion. The primary deficit as a per cent of GDP in 1980–1981 then was $(78.18/1325.20) \times 100 = 5.90$. In a similar way the fiscal deficit as a per cent of GDP in 1980–1981 was $(107.8/1325.20) \times 100 = 8.13$.
6. Create a graph of the primary deficit and fiscal deficit as a percentage of GDP for the years 1980–1981 to 2005–2006.
7. Insert a trendline for each of the two line graphs that you created in Step 6.
8. What do you notice about the trend of the primary deficit and the fiscal deficit over the period? Do the primary deficit and fiscal deficit diverge over time? If so, what is the root cause of the phenomenon that the difference between the fiscal deficit and the primary deficit increases over time?
9. The text demonstrated that the primary deficit in the recent past has been sustainable. Notice that the slope of the trend line of the fiscal deficit is mildly positive whereas the slope of the trend line of the primary deficit is negative. What policy decision(s) can result in a reduction of the fiscal deficit over time so that its trend has a negative slope?
10. If the policy objective is to attain the sustainability of deficits which definition of the deficit—fiscal or primary—is more meaningful and why? Why then do you think are targets for fiscal responsibility fixed in terms of the fiscal deficit and not the primary deficit?

# 13

# Open Economy II: The Current Account in an Intertemporal Framework

CRITICAL QUESTIONS

- *What is the current account balance in an intertemporal framework?*
- *What motivates nations to borrow and lend financial assets in the international capital market?*
- *In the medium run, where capital accumulation adds to capacity output, what impact do fiscal and current account deficits have on the national income and output?*
- *What is current account sustainability?*

# 13.1 Intertemporal Accounting and the Current Account Determination

In the last 15 years, a large literature on the dynamic optimizing (or intertemporal) approach to the current account has been developed. These models have emphasized the effects of real factors, such as productivity, terms of trade, government spending, and taxes on the current account balance. These operate through intertemporal substitution in consumption, production, and investment. These models were a response to a major shortcoming in the MUNDELL–FLEMING approach to an open economy.

› In the **Mundell–Fleming** model, a fiscal contraction leads to a contraction in the economic activity. In many economies, however, fiscal contractions have led to expansions in the economic activity. Intertemporal models of the current account can explain this anomaly.

In the Mundell–Fleming approach under a flexible exchange rate, a current transitory fiscal expansion shifts the *IS* schedule to the right and raises the level of output and the domestic interest rate. To maintain interest parity, the rise in the interest rate must result in the appreciation of the domestic currency. The current account deteriorates as output has risen and the domestic currency has appreciated. Under a fixed exchange rate, interest arbitrage ensures equality between the domestic and foreign interest rates. A fiscal expansion in this regime gains full potency in raising the level of output since there is no currency appreciation to offset it. As a result, the current account deteriorates.

However, the links between the fiscal deficit and the trade deficit, and between the trade deficit and the value of the domestic currency have been found to be empirically weak. Moreover, in many countries, contrary to the Mundell–Fleming prediction, large cuts in government spending carried out as part of stabilization programmes have led to *expansions* rather than contractions in economic activity and have resulted in improvements in the current account balance. For instance, in Denmark in the early 1980s and Ireland in the late 1980s, the government deficit was large and the public debt was growing. In the aftermath of large budget cuts following a stabilization programme, the current account improved.[1] Similarly, the mid-1985 disinflation programme in Israel included severe fiscal and monetary restraint with the public sector domestic deficit falling from about 12 per cent prior to stabilization to 2 per cent of GDP. Again fiscal consolidation resulted in output booms.[2] Such results are inconsistent with the predictions of income–expenditure models but are consistent with the predictions of the intertemporal models of the current account.

The intertemporal approach to macroeconomics differs from the income–expenditure one since it focuses on the intertemporal aspects of investment and savings (consumption) decisions. It also distinguishes between SHOCKS that are transitory in duration and those that are more permanent, and takes into account the fact that these different types of shocks drive the current account balance in distinctly different ways.[3] To grasp the basic insight of this approach we must first familiarize ourselves with the intertemporal accounts of a nation. We now turn our attention to this.

› A shock is an unexpected disturbance to the economy that affects the demand or supply side of the economy.

### 13.1.1 Intertemporal Accounts of a Nation

The sum of the trade balance and the three components of net receipts from services—non-factor services, factor incomes from abroad, and transfers—is the current account balance (see Chapter 1, Section 1.2.3). In this chapter, we will not consider the net receipts from services in the form of transfers from abroad. The current account balance is then the sum of the net exports of goods and services (or non-factor incomes) and the net factor incomes from abroad. As net exports equals output less domestic expenditure, $X - M = Y - (C + I + G)$, where $Y$ is the output in any period $t$, the current account balance is

$$\mathrm{CA}_t = (X_t - M_t) + \mathrm{NFI}_t$$

$$= Y_t - C_t - I_t - G_t + \mathrm{NFI}_t$$

Net factor incomes from abroad, $\mathrm{NFI}_t$, comprises the compensation of employees and investment income. In what follows we restrict our description of net factor incomes to just the investment income.

If $F_{t-1}$ stands for the net financial claims against residents abroad accumulated by a country at time $t$, then, $rF_{t-1}$ is the income received on this stock of financial assets, where $r$ is the interest rate.

Accordingly, $\mathrm{NFI}_t = rF_{t-1}$. We can then write

$$\mathrm{CA}_t = Y_t - C_t - I_t - G_t + rF_{t-1} \quad \textbf{(13.1)}$$

As is standard in many macroeconomic models, the government expenditure is purely consumption expenditure. So we can club the private consumption expenditure and government expenditure into a component that we label the consumption expenditure:

$$C_t^{\mathrm{pvt+G}} = C_t + G_t$$

The current account can then be written as

$$\mathrm{CA}_t = Y_t - C_t^{\mathrm{pvt+G}} - I_t + rF_{t-1}$$

where the superscript pvt+G stands for the private sector plus government sector.

› A current account surplus is an accumulation of financial assets by residents due to lending abroad.

› A current account deficit is financed through the accumulation of financial claims or liabilities against the rest of the world.

When the residents of a country lend more to foreigners than they borrow, and thus accumulate a net financial claim against the rest of the world, we say that the country has a CURRENT ACCOUNT SURPLUS. On the other hand, when the country is accumulating a net liability, or equivalently running down its net claims, against the rest of the world, the economy has a CURRENT ACCOUNT DEFICIT. This chapter stresses on the fact that the current account is the intertemporal budget constraint of the national economy.

If the economy runs a current account deficit today, the net debt to the rest of the world by its residents is escalating. At some time, the country will have to cut back on domestic expenditure in order to pay back the interest on this accumulated debt. When domestic expenditure is cut back, the national output that was used for consumption, for instance, gets to be increasingly used for net exports. The country's net exports then are its method of paying back the interest burden on the liabilities it accumulated when running current account deficits.

Recall that the gross national product (GNP) is given by (see Chapter 2)

**Gross National Product ($Y^n$) = Gross domestic product + Net factor incomes from abroad**

$$Y^n = Y + \mathrm{NFI}$$

GDP is given by $Y = C + I + G + X - M = C^{\mathrm{pvt+G}} + I + X - M$, and net factor incomes from abroad are given by $\mathrm{NFI}_t = rF_{t-1}$. Then, GNP is

$$Y^n = C^{\mathrm{pvt+G}} + I + (X - M) + rF_{t-1}$$

Thus,

$$Y^n - C^{\mathrm{pvt+G}} - I = (X - M) + rF_{t-1}$$

However, the term $Y^n - C^{pvt+G} = S$, or domestic savings. Similarly, the term on the right-hand side is the expression for the current account balance allowing us to write

$$S - I = CA$$

When the economy saves less than it invests, $(S - I) < 0$, it runs a current account deficit, which amounts to the country borrowing from the rest of the world and thereby building up financial claims or liabilities vis-à-vis residents in other countries. The current account deficit then is an accumulation of financial liabilities against the rest of the world. Similarly, if the country invests less than it saves, $(S - I) > 0$, and it runs a current account surplus and accumulates financial assets by lending abroad. We can then write the current account as

$$CA_t = F_t - F_{t-1} \tag{13.2}$$

where $F_t$ are the assets of national residents vis-à-vis foreigners *minus* the liabilities of residents vis-à-vis foreigners. In other words, $F_t$ is the *net foreign asset position* or the net international investment position of the country. When $F_t$ is positive, the country is a net creditor of the rest of the world. When $F_t$ is negative, a country is a net debtor of the rest of the world.

Obviously, the level of foreign assets $F$ in any period is a result of past current account surpluses and deficits. Starting from an initial year 0, the international investment position of a country in year $t$, $F_t$ will equal $F_0$ plus the sum of current accounts in the years between 0 and $t$

$$F_t = F_0 + CA_1 + CA_2 + \cdots + CA_t \tag{13.3}$$

For the purposes of analysis, it suffices to use a two-period model. We now consider the case of a two-period model.

### 13.1.2 Intertemporal Budget Constraint in a Two-Period Model

In a two-period model, $t = 1$ and 2, we assume the country starts with no foreign assets, $F_0 = 0$. Then, Eq. (13.2) may be written as

$$CA_1 = F_1 - F_0 = F_1$$

Substituting from Eq. (13.1) when $F_0 = 0$

$$CA_1 = F_1 = Y_1 - C_1^{pvt+G} - I_1 \tag{13.4}$$

Recall that $S = Y - C^{pvt+G}$, so that the above equivalently states $F_1 - F_0 = S_1 - I_1$, or the change in the financial asset holding of the nation is the difference between saving and investment.

In the second period, interest payments on the net foreign assets have to be included so that

$$CA_2 = Y_2 + rF_1 - C_2^{pvt+G} - I_2 \tag{13.5}$$

Using Eq. (13.2) we may rewrite this as

$$F_2 - F_1 = Y_2 + rF_1 - C_2^{pvt+G} - I_2$$

or,

$$F_2 = (1 + r)F_1 + Y_2 - C_2^{pvt}{+}G - I_2 \tag{13.6}$$

However, in a two-period model, in the second period the country must end with no net foreign assets ($F_2 = 0$), and it undertakes no investment in the second period ($I_2 = 0$).

Therefore,

$$0 = (1 + r)F_1 + Y_2 - C_2^{pvt+G} \tag{13.7}$$

We can rewrite this Eq. (13.7) as

$$C_2^{pvt+G} = (1 + r)\, F_1 + Y_2$$

or,

$$\frac{C_2^{pvt+G}}{1+r} - F_1 = \frac{Y_2}{1+r}$$

Combining this with Eq. (13.4), $F_1 = Y_1 - C_1^{pvt+G} - I_1$

$$\frac{C_2^{pvt+G}}{1+r} - F_1 = \frac{Y_2}{1+r}$$

$$C_1^{pvt+G} + F_1 = Y_1 - I_1$$

$$\overline{C_1^{pvt+G} + (C_2^{pvt+G}/1+r) = (Y_1 - I_1) + (Y_2/1+r)} \tag{13.8}$$

› The **national intertemporal budget constraint** states that the present value of the aggregate consumption expenditure equals the present value of the national production net of investment.

This is the NATIONAL INTERTEMPORAL BUDGET CONSTRAINT. The discounted value of the aggregate consumption expenditure (private and government) must equal the discounted value of national production net of investment.

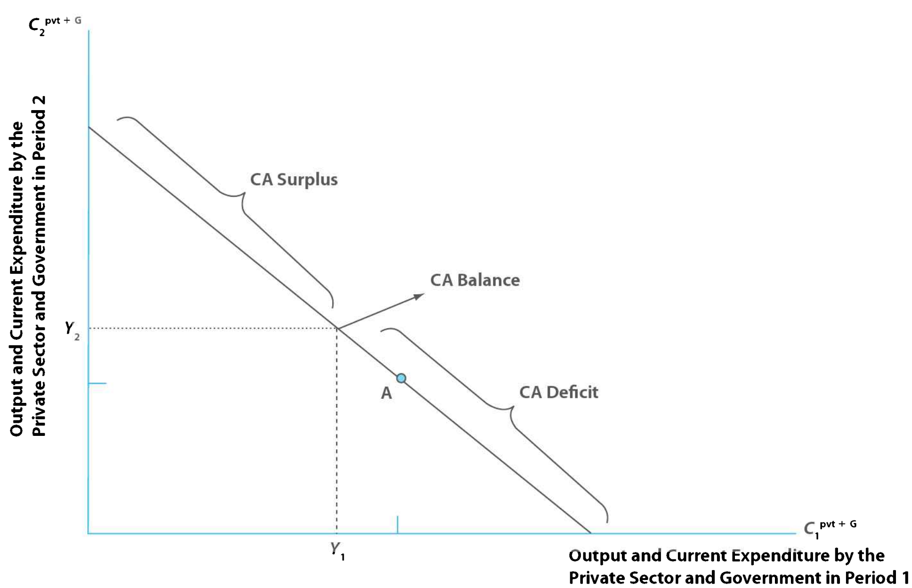
In the special case, where there are no attractive investment opportunities ($I_1 = 0$), the economy's only decision is to decide how much to consume today and how much to save. The budget constraint is then given by

$$C_1^{pvt+G} + \frac{C_2^{pvt+G}}{1+r} = Y_1 + \frac{Y_2}{1+r} \tag{13.9}$$

This budget constraint and the associated current account positions are given in Figure 13.1.

If $C_1^{pvt+G} > Y_1$, such as at point A, the economy is consuming more than it produces, which is only possible if it increases its indebtedness to the rest of the world and runs a current account deficit. In that case, in the future the economy must run a current account surplus, $C_2^{pvt+G} < Y_2$, so as to repay the indebtedness and in the process balance its intertemporal accounts.

That the discounted value of consumption must equal the discounted value of output net of investment is one way of writing a country's intertemporal




› **Figure 13.1**
The Current Account Position when All Output is Consumed ($I = 0$).
When no part of the output is invested, the discounted value of the current expenditure by the private sector and government equals the discounted value of national production. This is given by the downward-sloping national intertemporal budget line. If aggregate consumption then is exactly equal to the production in each period, the current account is in balance. If aggregate consumption is greater in period 1 than the output produced by the economy, as occurs at point A, the excess consumption is serviced by imports and a current account deficit characterizes the economy.

budget constraint. There are two other ways of writing the intertemporal budget constraint:

1. *Intertemporal budget constraint in terms of the trade balance*—A second way to write the budget constraint is to rearrange terms to describe it in terms of the trade balance of goods and non-factor services in the two periods. By definition,

$$CA_t = TB_t + \text{Net Factor Income from Abroad (NFI)}$$

$$CA_t = TB_t + rF_{t-1} \qquad \textbf{(13.10)}$$

and using this we can write the current account in the two time periods as

$$CA_1 = TB_1 + rF_0 = TB_1 \quad \text{as } F_0 = 0$$

$$CA_2 = TB_2 + rF_1$$

From Eq. (13.4)

$$CA_1 = TB_1 = Y_1 - C_1^{\text{pvt+G}} - I_1 \qquad \textbf{(13.11)}$$

From Eq. (13.5)

$$CA_2 = TB_2 + rF_1 = Y_2 + rF_1 - C_2^{\text{pvt+G}} - I_2$$

or

$$TB_2 = Y_2 - C_2^{\text{pvt+G}} \quad \text{as } I_2 = 0 \qquad \textbf{(13.12)}$$

We write Eq. (13.8) again as follows:

$$C_1^{\text{pvt+G}} + \frac{C_2^{\text{pvt+G}}}{1+r} = Y_1 - I_1 + \frac{Y_2}{1+r}$$

Taking all the terms to the left-hand side we can rewrite Eq. (13.8) as

$$(Y_1 - I_1 - C_1^{\text{pvt+G}}) + \frac{Y_2 - C_2^{\text{pvt+G}}}{1+r} = 0$$

Substituting from Eqs. (13.11) and (13.12) into the above equation we obtain:

$$TB_1 + \frac{TB_2}{1+r} = 0 \qquad \textbf{(13.13)}$$

This means that the TRADE DEFICIT in the first period must be balanced by a trade surplus in the second period of equal present value.

› A second expression for the intertemporal balance can be that it is the condition that a **trade deficit** in the present must be balanced by a trade surplus of equal present value in the future.

2. *Intertemporal budget constraint in terms of the current account*—The third way of expressing the country's intertemporal budget constraint is in terms of the CURRENT ACCOUNT. The current account is the accumulation of net foreign assets, or

$$CA_1 = F_1 - F_0$$

$$CA_2 = F_2 - F_1$$

› A third expression for intertemporal balance is that the sum of the present and future **current account balances** must add up to zero.

If a country starts with no net foreign assets, $F_0 = 0$, and ends with no net foreign assets, $F_2 = 0$, we have

$$\begin{array}{r} CA_1 = F_1 \\ CA_2 = -F_1 \\ \hline CA_1 + CA_2 = 0 \end{array} \qquad \textbf{(13.14)}$$

To sum up, the three alternative ways in which the current account position of a country may be written then are

$$C_1^{pvt+G} + \frac{C_2^{pvt+G}}{1+r} = (Y_1 - I_1) + \frac{Y_2}{1+r} \tag{13.8}$$

$$TB_1 + \frac{TB_2}{1+r} = 0 \tag{13.13}$$

$$CA_1 + CA_2 = 0 \tag{13.14}$$

A qualification that must be borne in mind is that the analysis assumes that a debtor country always honours its debts. In the international economy, where enforcement of contracts is more difficult, debtors may sometimes choose not to repay.

## 13.2 Current Account Determination

To understand this perspective of current account determination, let us specify a simple model. The economy is endowed with an initial stock of capital $K_0$ and a production function gives us the output that can be produced from the capital as $Y_t = F(K_{t-1})$. As is usual, the production function exhibits positive and diminishing marginal productivity. The capital stock at the end of period 1, $K_1$, is augmented by the investment that takes place so that

$$K_1 = K_0 + I_1$$

**MACROFOCUS 13.1**

### Global Imbalances

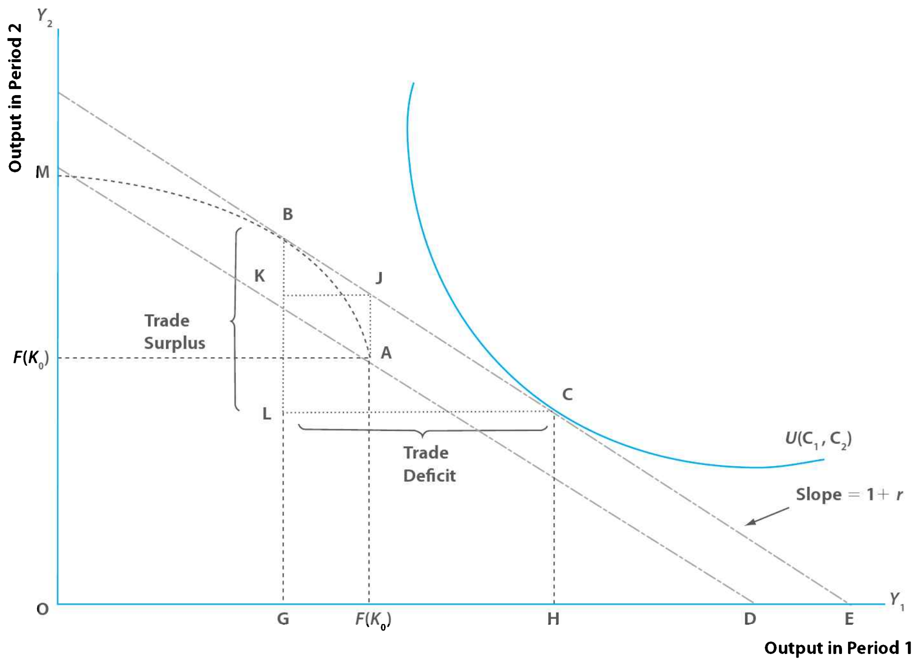
The large and persistent US current account deficit is considered to be a global imbalance and has led to concerns about the possibility of a major global crisis. The United States has a record of continued deficits since 1991 that totalled over $5.2 trillion by 2005. Doomsayers opine that once the financing from the rest of the world, which pays for such a wide United States external account gap dries up, perhaps due to satiation from owning large quantities of US debt, a chain of events would be set off. The dollar will collapse, interest rates will rise, and a global recession will ensue.

Calvo and Talvi* calculate that the deterioration in the current account (*CA*) position of the United States between 1997 and 2005 amounted to 5 per cent of the US GDP. A *CA* deficit in the United States must be balanced by a *CA* surplus in some other countries. Indeed, the *CA* position in emerging economies—Asia (excluding Japan), Middle Eastern countries, and Latin American countries—showed a sharp improvement in this period.

These emerging markets have moved on from being large net borrowers to net lenders in the world capital markets. Moreover, the holding of the US debt by the emerging markets was not by private agents but due to central bank purchases whose accumulation of reserves denominated in US dollars accelerated significantly after 2001. The accumulation of reserves denominated in US dollars by foreign central banks has risen from US$83 billion in 2001 to US$500 billion in 2004. There has been a savings glut in emerging markets and rather than invest their savings domestically, emerging markets have preferred to invest it in the United States. This savings from abroad has kept the US interest rates low despite its large borrowings to finance its deficit. The large *CA* imbalances in the United States are thus a result of a large expansion in the supply of savings in emerging markets, a factor exogenous to the US economy.

The likelihood of a contraction in the savings of emerging markets depends on whether countries like China, which today consume 40 per cent of the world's cement, a third of its coal, and a quarter of its steel, decide to stop accumulating reserves due to strong pressure from the United States and experience a deterioration in their *CA* following a sharp real appreciation of its currency. In the wake of this, other Asian countries would experience a similar outcome and the brunt of the adjustment would have to be borne by the United States or the non-Asian emerging markets. The decline in the availability of foreign savings could result in interest rates in the United States rising and an extensive credit rationing that sends the world economy into recession. Anticipating such an outcome of financial turbulence, investors would attempt a flight to quality. Investors choose to hold United States assets even when they typically pay a lower return because they perceive these assets to be safe. In fact, due to this feature the United States earns a higher return on its holdings of foreign assets than it pays to foreigners on its liabilities. The flight to quality will fill the US *CA* financing gap and can prevent a significant increase in US interest rates. The United States is then unlikely to be affected, while emerging markets may see dramatic reductions in output and a decline in domestic savings.

*G. Calvo and E. Talvi, "The Resolution of Global Imbalances: Soft Landing in the North, Sudden Stop in Emerging Markets?" *Journal of Policy Modeling* 28, no. 6 (2006): 605–613.




**› Figure 13.2**
Intertemporal Current Account Determination. **With no investment, the same amount is produced in each period as given by point A. The output can be transferred across periods at the interest rate (1 + *r*) and this is depicted by drawing a line of this slope through point A. Investment requires forgoing consumption today for higher future output, as given by the transformation curve ABM. The returns from investment are maximized at point B where the productivity of investment is equal to the returns from investment. With preferences for consumption today versus tomorrow being given by an indifference curve, the economy consumes OH in period 1 and runs up a trade deficit that is paid back in period 2 by the economy running up a trade surplus.**

where we are ignoring depreciation. Also, for simplicity, we are assuming a given labour force. Alternatively, we can think of the aggregate output and the capital stock as being expressed in per worker terms. Hence, output in period 2 will be

$$Y_2 = F(K_1) = F(I_1 + K_0)$$

Firms maximize the present value of profits or invest up to the point where the productivity of investment equals the cost of investment, as given by the interest rate.

What investment does is modify the intertemporal pattern of available outputs. In the absence of investment, the endowment sequence of the economy is

$$Y_1 = F(K_0), \quad Y_2 = F(K_0)$$

This is the point A in Figure 13.2. To find the present value of this sequence of output, draw a dashed line through point A with slope $(1 + r)$. The present value of the sequence of outputs is then given by point D. Investment in the diagram is measured in a leftward direction from A and is a curve that specifies how forgone consumption today is transformed into consumption tomorrow. As there are diminishing returns to additional investment, the slope of the transformation curve (curve ABM) diminishes gradually. At point M, the maximum level of investment is reached and no consumption occurs in the present. Investment will occur up to the point where the productivity of investment equals the returns or $F'(K) = 1 + r$. This is because at this point the marginal product of a unit of investment pays for the principal (1) and the interest $(r)$ cost of the project. At point B, this investment rule is satisfied and the amount of investment undertaken is the distance from G to $F(K_0)$ along the $Y_1$-axis.

We now turn to the demand side of the model. A representative individual's lifetime utility is given by $U(C_1, C_2)$, where the marginal utility of consumption in each period is positive and the marginal rate of substitution of consumption between two consecutive periods is diminishing along any indifference curve, a quasi-concave utility function. Utility is maximized where the slope of the budget line is tangential to an indifference curve—point C.

How is it that the preferred consumption and investment is possible? After all, the consumption in period 1 is OH, and investment is $GF(K_0)$ and consumption plus investment exceeds GDP as given by $OF(K_0)$. This is made possible by borrowing from abroad. In the current period, the economy runs a trade balance deficit equal to GH. That trade deficit is paid back in the subsequent period at interest rate $r$.

Distance GL represents the second-period consumption. Distance KL is the portion of the second-period output (BG) that is used to repay the first-period consumption [given by the distance $F(K_0)$H on the $X$-axis] that is over and above the consumption from first-period output [distance O$F(K_0)$]. This excess of consumption over the output in the first period was financed by INCREASING THE INDEBTEDNESS to the rest of the world, resulting in a trade deficit that is equal to distance GH (=LC). Distance BK is the portion of the second period output that is used to repay the borrowing for investment [$GF(K_0)$ = KJ] undertaken in the first period. Distance BK + KL = BL is the second-period trade surplus that is required intertemporally to finance the first-period trade deficit.

› A country can consume and invest more than its current output by increasing its indebtedness to the rest of the world via a trade deficit. This requires a future trade surplus that finances the debt incurred in the present.

The ratio of the distance BL to LC (=GH) is the ratio of the second-period trade balance surplus $TB_2$ to the first-period trade balance deficit $TB_1$. Also, this ratio BL/LC is the ratio of the vertical to the horizontal distance between points B and C and is, therefore, the slope of BC.

$$\frac{\text{Distance BL}}{\text{Distance LC}} = \text{slope BC}$$

or,

$$\frac{TB_2}{TB_1} = \text{slope BC}$$

However, we already know that BC has a negative slope equal to $(1 + r)$. Hence,

$$\frac{TB_2}{TB_1} = -(1 + r)$$

or,

$$\frac{TB_2}{1+r} = -TB_1$$

or,

$$TB_1 + \frac{TB_2}{1+r} = 0$$

This reiterates Eq. (13.13) that the trade deficit in the first period is balanced by a trade surplus in the second period of equal present value.

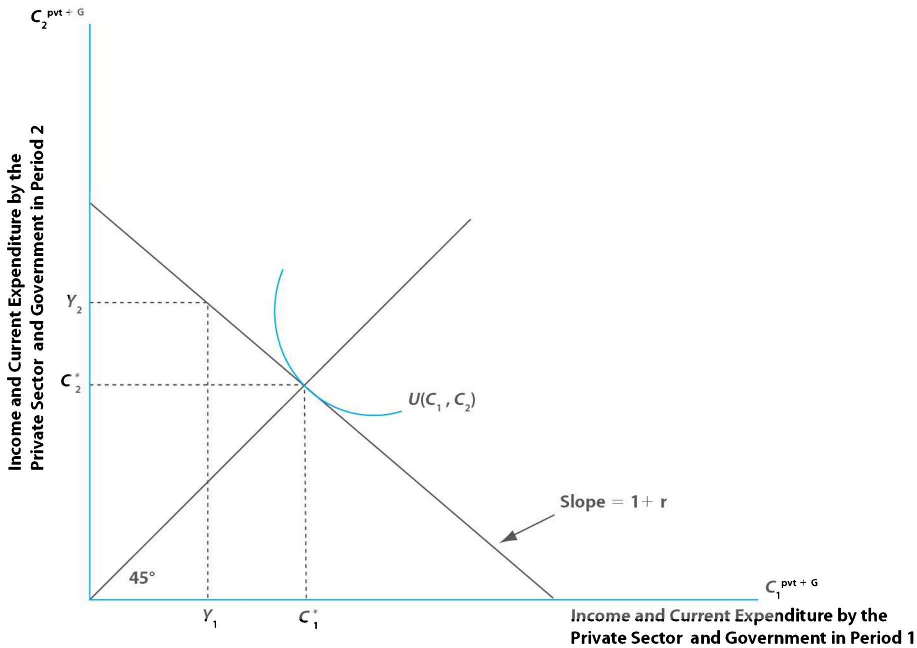
Point E measures the discounted sum of the current and future national output when investment [equal to $GF(K_0)$] financed by world capital markets is undertaken. Point D measures the discounted sum of current and future national product with no investment. The international capital market enables the economy to set aside some of its current resources for investment and yet increase the current consumption. It also provides a greater present value of the current and future consumption.

### *13.2.1* Three Determinants of Borrowing/Lending

The lack of synchronization between the time series of consumption and income is reconciled by resorting to the world capital markets. The RELIANCE ON CAPITAL MARKETS and the accumulation of debt is a function of three separate motives:

› Countries resort to international capital markets to finance expenditure due to three motives— consumption smoothing, consumption tilting, and consumption augmenting.

1. consumption smoothing
2. consumption tilting
3. consumption augmenting




**› Figure 13.3**
Consumption Smoothing. With consumption smoothing, even though income fluctuates over time, the economy borrows in period 1 to finance a consumption greater than income and repays the amount in period 2.

To isolate each factor separately we focus on special cases.

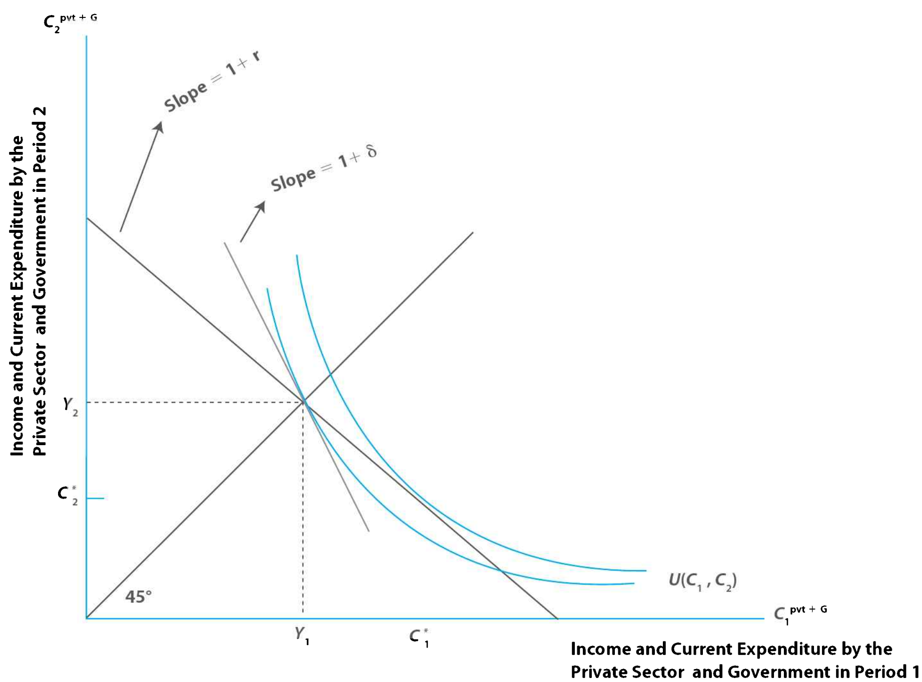
1. *Consumption Smoothing*—Assume that the subjective time preference rate δ from $U = U(C_1) + (1/1 + \delta)U(C_2)$ is equal to the market discount rate, $r$. Also, assume that no investment takes place. However, the periodic levels of income differ from each other so that $Y_1 \neq Y_2$, as in Figure 13.3. However, from the life-cycle theory we know that consumption will be given by permanent income and is along the 45-degree ray from the origin. As the idea is to SMOOTHEN CONSUMPTION, per period it will take a value between $Y_1$ and $Y_2$, where $C_1^* = C_2^*$. Consumption smoothing occurs through borrowing in period 1 and repaying the loan plus interest in the subsequent period.
2. *Consumption Tilting*—Assume that the subjective time preference rate and the market discount rate differ, $\delta \neq r$. Again, there is no investment, and the periodic levels of income are equal, $Y_1 = Y_2$, as in Figure 13.4. Then, equilibrium consumption does not lie along the 45-degree ray. Suppose the subjective rate of time preference exceeds the world rate of interest, $\delta > r$. As a result, individuals facing a FLAT TIME PROFILE OF INCOME wish to tilt the time profile of consumption towards the current period $(C_{1*} > Y_1)$. This motive is satisfied through world capital markets where individuals borrow in the current period and settle their debts in the next.
3. *Consumption Augmenting*—Again, assume equality between the subjective and the market discount factors, $\delta = r$. Also, period levels of income are equal, $Y_1 = Y_2$. However, now there is positive investment as the marginal productivity of investment is greater than the discount rate, $F'(K) > 1 + r$. The investment opportunity tilts the profile of income unlike in the consumption tilting case where the profile of consumption is tilted. The investment opportunity augments the level of consumption in each period without introducing variability to its time profile (the economy proceeds along the 45-degree ray). CONSUMPTION AUGMENTING is effected through world

› With fluctuating national income, **consumption smoothing** occurs when a nation resorts to the international capital market to stabilize consumption.

› Consumption tilting occurs when a nation with a flat-time profile of the national income has a time preference for greater (lesser) consumption in the present and lesser (greater) consumption in the future.

› Consumption augmenting is the outcome of resorting to the international capital markets to finance investment whose productivity enables a higher present as well as future consumption.

**› Figure 13.4**
Consumption Tilting. **With consumption tilting, even though the time path of income is smooth, the economy borrows in period 1 to finance a consumption greater than income and repays the amount in period 2.**




capital markets with individuals borrowing in the current period and repaying in the next. In the absence of the international capital market, the investment carried out in the current period would have crowded out private consumption in that period. Access to world capital markets facilitates the augmentation of consumption at a rate that is uniform over time. These different motives are summarized in Table 13.1.

**› Figure 13.5**
Consumption Augmentation. **Investment tilts income by increasing the future income relative to the current income. The economy can then borrow in period 1 to finance a consumption greater than current income and repay the amount in period 2. Not only current but also future consumption in period 2 is augmented because of the increase in the output and income due to the increased investment.**

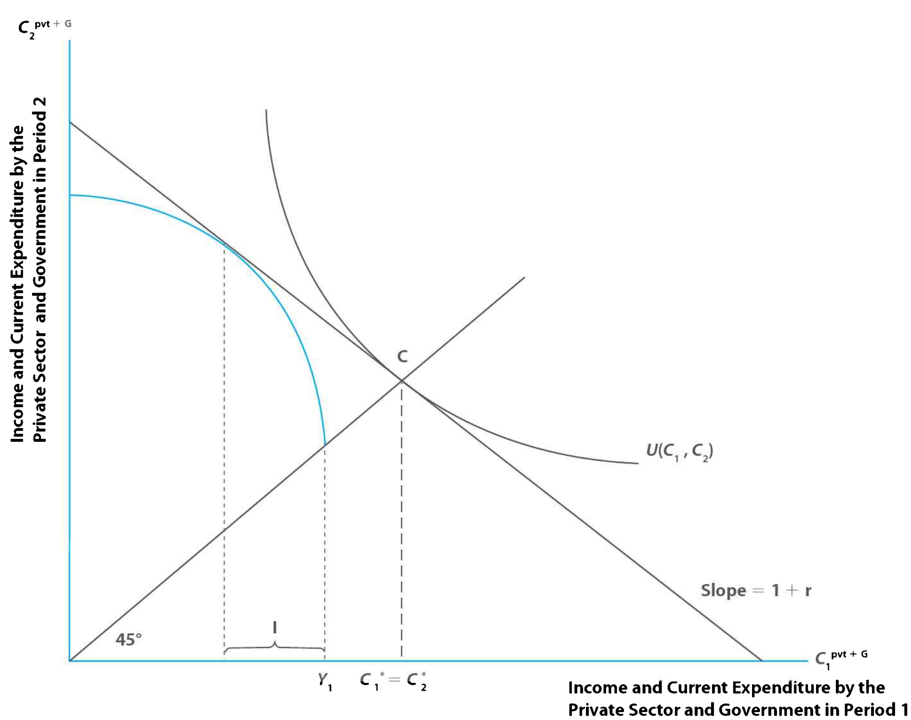


**Table 13.1**
Consumption Motives, Income Fluctuation, and Investment.

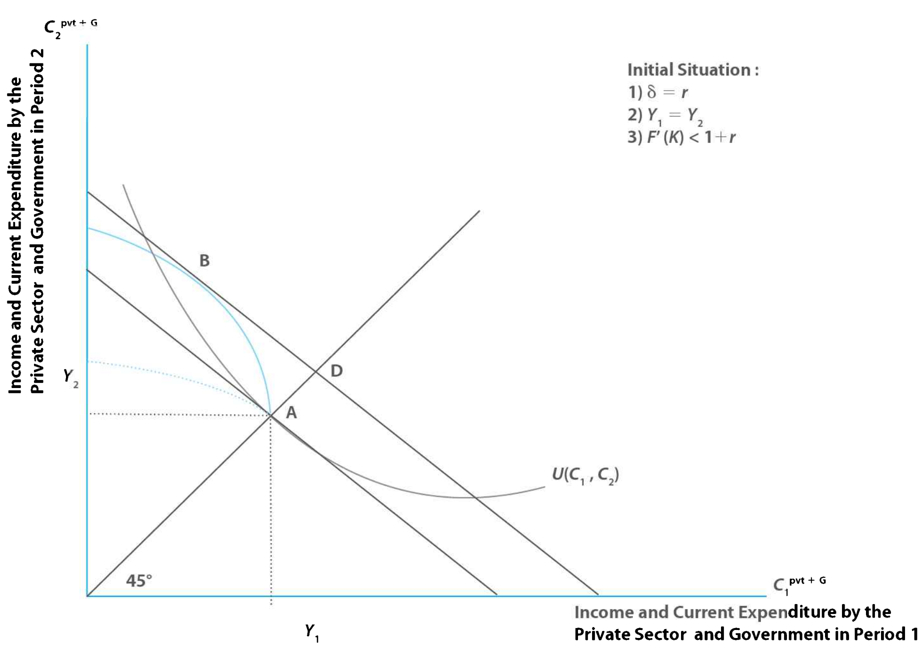
| Motive → | Consumption Smoothing | Consumption Tilting | Consumption Augmenting |
|---|---|---|---|
| Discount Factors | $\delta = r$ | $\delta \neq r$ | $\delta = r$ |
| Income Stream | $Y_1 \neq Y_2$ | $Y_1 = Y_2$ | $Y_1 = Y_2$ |
| Investment Profitability | $F'(K) < 1 + r$ | $F'(K) < 1 + r$ | $F'(K) > 1 + r$ |

### *13.2.2* Current Account and Shocks to the Output

In general, the three motives coexist and interact in generating the equilibrium pattern of consumption, investment, and debt accumulation. To take a case where both consumption smoothing and consumption augmenting effects operate, consider the situation when there is a positive supply shock stemming from a technological improvement in the process of investment.

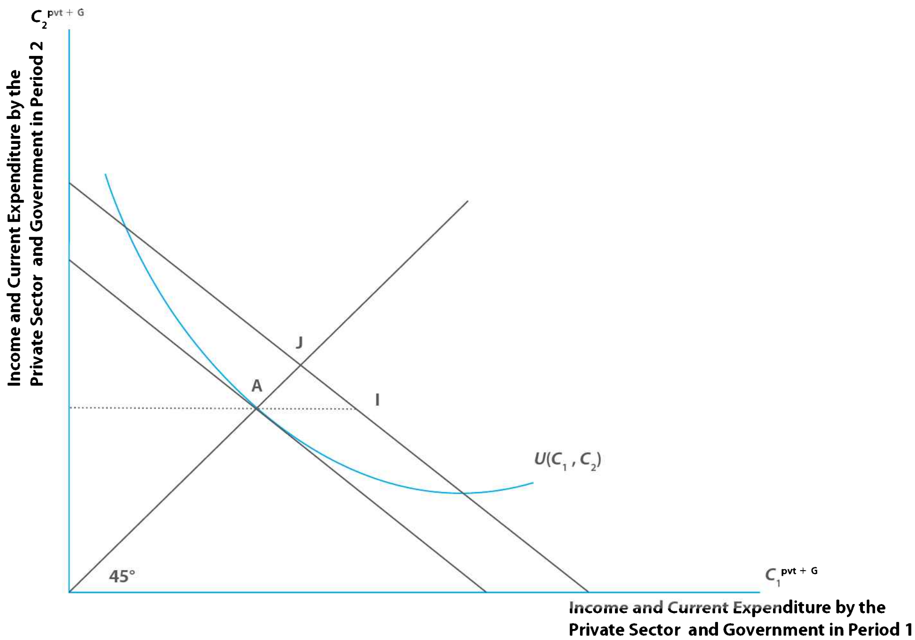
Under the initial technology, the investment opportunities schedule is the dashed curve from point A (Figure 13.6). As the marginal product of investment is less than the required return, no investment takes place. Then, if there is a technological improvement, the new investment opportunity curve from point A passes through point B. The level of production will now be given by point B and the level of consumption by point D. The current consumption rises even before the process of investment bears fruit. This represents both the consumption smoothing and the consumption augmenting effects. The economy runs a deficit in its balance of trade that is equal to the difference between the new consumption point in period 1, as given by the horizontal distance between points B and D in Figure 13.2. The counterpart of this trade deficit is a trade surplus in the subsequent period. Consumption smoothing occurs through the world capital market which allows the time profile of consumption to be smoothed relative to the variability in the time profile of GNP. The variability of GNP gets reflected in the time profile of the trade balance.

A *temporary* shock will also have a different impact from a *permanent* shock. Consider a temporary shock that raises the income in the current period




**› Figure 13.6**
Consumption Smoothing + Augmenting. Technological change increases the productivity of investment and results in the investment opportunity curve AB that increases the future income relative to the current income. The economy borrows some of the augmented future income to increase the current consumption. The increase in consumption in period 1 is serviced by a trade deficit (horizontal distance BD), which is repaid in period 2 through a trade surplus. Consumption is smooth across the two periods as well as augmented through borrowing in the world capital markets to finance profitable investment that augments output.

› **Figure 13.7**
Temporary Versus Permanent Shocks. A temporary increase in income takes the income profile of the economy from point A to point I. With consumption smoothing the optimal consumption is at point J. The economy runs a trade surplus equal to the horizontal distance JI and a corresponding deficit in period 2 with an increase in imports to service the higher consumption depicted by point J. If the increase in income is permanent, however, both the income and consumption move from point A to point J with no change in the balance of trade.




to point I (Figure 13.7). Equilibrium consumption is described by point J. To enable this pattern of consumption, the economy runs a surplus in its balance of trade that is equal to the horizontal distance between points J and I. In the subsequent period, this is balanced by a trade deficit with *consumption exceeding income*. If there is a *permanent* shock, on the other hand, the income in both periods is raised by the same proportion. The new pattern of consumption coincides with the new smooth income stream given by point J on the new higher budget line. The permanent shock results neither in a surplus nor a deficit in the balance of trade.

# 13.3 The Current Account in the Medium Run

› The *IS–LM–BP* **model** assumes an underutilized capacity so that the output adjusts to the changes in the aggregate demand. It does not satisfactorily incorporate the supply- side effect of investment expenditure, which adds to capital stock and capacity.

That additions to the capital stock augment the aggregate output on the supply side is not even considered in the open economy Mundell–Fleming approach to income determination. In fact, in that approach, an *IS–LM* type concept is employed that presumes like in all *IS–LM* approaches that the economy has underutilized capacity and so the national output adjusts endogenously to changes in aggregate demand. The effects of the supply side factors—additions to capital stock and accretion/depletion of foreign financial assets or capital flows—are not taken into account in the SHORT-RUN MUNDELL–FLEMING TYPE MODEL satisfactorily. We now recast the two-period intertemporal approach based on utility maximization, consumption, and investment behaviour in the framework of aggregate expenditure and aggregate output to facilitate an easy comparison of the difference in macroeconomic outcome when we explicitly allow for consumption and output augmentation.

As we know macroeconomic equilibrium in the market for goods is where the aggregate output equals aggregate expenditure. Aggregate expenditure, in turn, is the sum of domestic expenditure or absorption, $C + I + G$ and the current account balance, CAB = $(X - M)$ + NFI where NFI is the net factor incomes from abroad. We assume real interest rate parity[4] and perfect capital mobility. Moreover, we also assume static inflation and exchange rate expectations and ignore other risk factors that restrict capital mobility. Under these conditions[5] the exogenous world interest rate $r^*$ is equal to the domestic interest rate or

$$r = r^* \tag{13.15}$$

With the interest rate fixed exogenously, consumption does not respond to interest rate changes and is solely a function of the disposable income:[6]

$$C = c_0 + c_1(Y^n - T) \tag{13.16}$$

Similarly, with static exchange rate expectations, the competitive effect of exchange rates on exports and imports are temporarily downplayed. Imports respond to the aggregate domestic income whereas exports respond to the exogenous aggregate income in the rest of the world.[7] In an open economy, the relevant aggregate output and income variable is the GNP, which is defined as:

$$\text{GNP} = \text{GDP} + \text{Net Factor Incomes from Abroad}$$

The net factor incomes from abroad is simply the interest receipts on net foreign assets held or $r^*F$, where $F$ is the stock of foreign assets. Thus,

$$Y^n = Y + r^*F \tag{13.17}$$

We begin the analysis at the point of time where the initial preexisting stock of foreign assets is zero, $F = 0$. In that case, $Y^n = Y$, the national product equals the domestic product, and the current account balance equals the trade balance. That means when $F = 0$, NFI $= r^* F = 0$ and CAB $= X - M$.

From the balance of payments (see Chapter 2), if there is no initial stock of foreign assets held by the country, the overall balance of payments is zero, or the sum of the current account balance and capital account balance is zero.

$$\text{Current Account Balance} + \text{Capital Account Balance} = 0$$

$$\text{Current Account Balance} = -\text{Capital Account Balance}$$

In the initial period of analysis then, if the current account is in deficit, that is $(X - M) < 0$, the excess imports must be financed through a capital account surplus or a capital inflow—a net inflow of FDI, portfolio investment, and banking capital. The capital account balance is none other than the net foreign asset position of the nation's residents vis-à-vis the rest of the world. Hence, after the initial period, the capital account balance (KAB) is capital inflows net of outflows and this is the net accretion to the nation's stock of foreign assets:

$$-\text{KAB} = \Delta F$$

Hence,[8]

$$\text{CAB} = -\text{KAB} = \Delta F \tag{13.18}$$

In a medium-run model, the aggregate output produced is a function of capacity creation via additions to the capital stock. With a given labour force $\bar{N}$, the OUTPUT PRODUCED is a function of the capital stock or

› In a medium-run model, the aggregate output is a function of the capital stock, which includes the previous period additions to the capital stock.

$$\text{Output} = F(K_{t-1}, \bar{N}) = f(K_{t-1}) \tag{13.19}$$

Increments to the capital stock result in increments to the output but at a diminishing marginal rate. The marginal productivity of capital is subject to diminishing returns. Ignoring the depreciation of the capital stock, any additional investment adds to the capital stock:

$$K_t = K_{t-1} + I_t$$

The additional capital stock resulting from investment expenditure enlarges the output and income in a subsequent period. A marginal addition to the capital stock results in an increment to the product or the output, as given by the marginal product of capital, $\text{MP}_K = f'(K)$.

Hence, if the capital stock increases by $\Delta K = I$ units, the increment to the output is approximately $\text{MP}_K\,(I)$. We may then write

$$\textbf{Change in Output} = \text{MP}_K\,(I)$$

or,
$$\textbf{Output in Period } (t+1) - \textbf{Output in Period } (t) = \text{MP}_K(I_t)$$

or,
$$\textbf{Output in Period } (t+1) = \textbf{Output in Period } (t) + MP_K(I_t) \qquad \textbf{(13.20)}$$

Any increase in the capacity to produce output is only possible if there is an addition to the capital stock $(K_t - K_{t-1})$ or to the investment $(I_t)$ in the economy. Equation (13.20) states that the increased production from an increase in the capital stock or investment $(I_t)$ contributes to the increase in the output producing capacity of the economy.

The output produced results in income for the agent in the economy, which is either consumed or saved. For the private sector saving is disposable income $(Y^n - T)$ less consumption or

$$S_{\text{pvt}} = Y^n - T - C$$

Saving by the government is given by tax receipts less outlays by the government or

$$S_{\text{govt}} = T - G$$

Total national saving as we saw in Chapter 2 is then given by

$$\begin{aligned} S = S_{\text{pvt}} + S_{\text{govt}} &= (Y^n - T - C) + (T - G) \\ &= Y^n - C - G \qquad \textbf{(13.21)} \end{aligned}$$

The aggregate expenditure equation from Chapter 2 is given by

$$Y^n = C + I + G + (X - M + \text{NFI})$$

or,
$$Y^n - C - G = I + \text{CAB}$$

where the current account balance is $\text{CAB} = X - M + \text{NFI}$. Substituting for the left-hand side of the above from Eq. (13.21),

$$S = I + \text{CAB}$$

or,
$$S - \text{CAB} = I \qquad \textbf{(13.22)}$$

Thus, domestic savings and the savings of foreign citizens[9] finance the investment expenditures in the economy. This is nothing but the familiar condition that the market for output is in equilibrium when the aggregate savings equals the aggregate investment. We now explore the implication of allowing investment expenditure to influence output in a diagrammatic approach.

### *13.3.1* Determination of Expenditure–Output

We now bring all this together in Figure 13.8, where the horizontal axis measures the output of final goods and services produced domestically and made available for sale in the domestic market and abroad. This is because part of the output, unlike in a closed economy, is sold to meet the demand for exports. Therefore,

$$\text{Domestic Output} = \text{Domestic Sales} + \text{Exports}$$

$$= (C + I + G) + X$$

or,

$$\text{Domestic Output} = Y^n + M \tag{13.23}$$

where $Y^n = C + I + G + (X - M)$, when NFI $= 0$.

On the vertical axis we measure the aggregate domestic expenditure, which comprises the demand for domestic goods and the demand for foreign goods or imports:

$$\text{Aggregate Domestic Expenditure} = \text{Aggregate Demand for Domestic Goods} + \text{Imports}$$

or,

$$E = (C + I + G) + M \tag{13.24}$$

Now suppose that in the initial period where there is no preexisting stock of financial assets, aggregate domestic expenditure is less than output. Then, with reference to Eqs. (13.23) and (13.24),

$$C + I + G + M < Y^n + M$$

or,

$$0 < Y^n - (C + I + G)$$

or,

$$0 < (X - M)$$

or,

$$0 < \text{CAB}$$

The current balance is accordingly in surplus when the aggregate domestic expenditure is less than the output—any point below the 45-degree line. The current account balance will, by similar reasoning, be in deficit when the aggregate domestic expenditure is greater than output—any point above the 45-degree line.

We now draw in the consumption function Eq. (13.16) in Figure 13.8, which is a linear function of the output. Also, as not all the income (output) is consumed, $c_1 < 1$ and the slope of the consumption function is less steep than the 45-degree line.

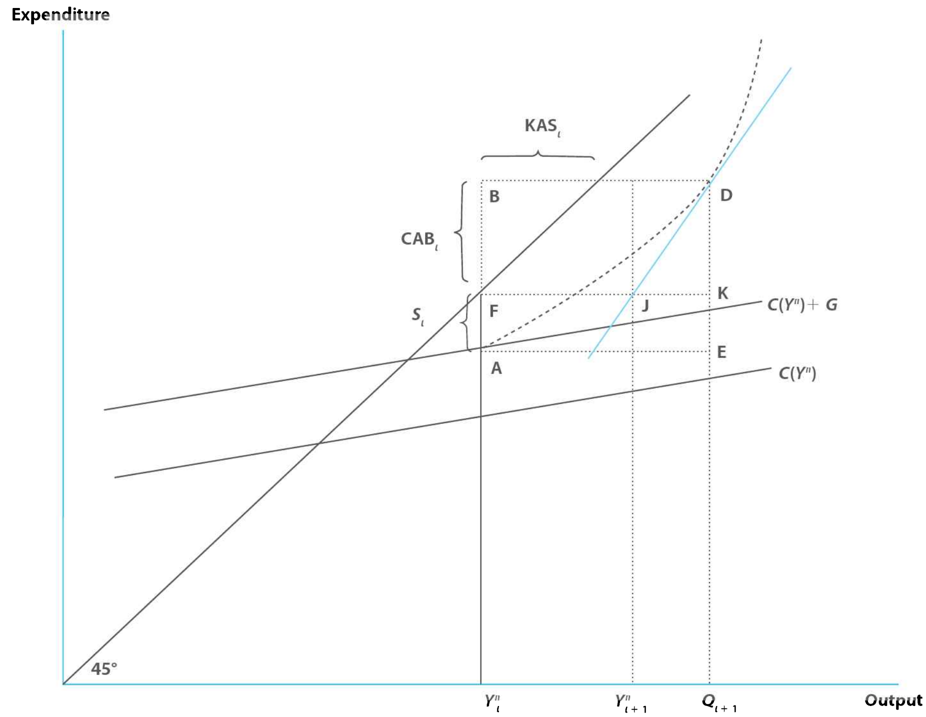
To the consumption function add another component of the aggregate domestic expenditure, which is the government expenditure, $G$. As government expenditure is exogenous and taken to be public consumption expenditure, the $C(Y^n)+G$ schedule is parallel to the $C$ schedule as the government expenditure is constant. If the national output produced in the initial time period is $Y_t^n$, then from Eq. (13.21) the national saving is given by $S_t = Y_t^n - C(Y_t^n) - G_t$. $S_t$ or the national saving is the difference between the value of $Y_t^n$ as measured on the vertical axis via the 45-degree line and the $C(Y_t^n) + G$ schedule.

We now map out the investment schedule. Recall from Eq. (13.20) that investment multiplied by the marginal productivity of capital results in the incremental output produced by the addition to capital stock

$$\text{Change in Output} = MP_K (I)$$

or,

$$\frac{I}{\text{Change in Output}} = \frac{1}{MP_K}$$




**› Figure 13.8**
Macroeconomic Equilibrium in an Open Economy. **Expenditure equals output from domestic sales and exports along the 45-degree line. When expenditure is less than the output, exports exceed imports, or there is a current account surplus—below the 45-degree line. Consumption increases with GNP, $Y^n$, but by less than the increase in GNP and is a straight line with a slope less than the 45-degree line. Investment expenditure increases the output of the economy as indicated by the dashed upward-sloping line from point A, indicates the sum of consumption and government expenditure. The benefit of investment, as given by the marginal productivity or slope of the output augmenting investment curve, equals the cost or interest rate at point D where the tangent indicates this equality. DE is the amount of investment expenditure undertaken that is financed by national savings $S_t$ and foreign savings CAB$_t$, or a current account deficit. The increase in foreign financial liabilities of the economy equal to distance DK requires to be serviced in a subsequent period to the extent of distance JK, thereby reducing the nation's income to $Y^n_{t+1}$ that is less than its domestic output $Q_{t+1}$.**

Initial levels of investment are associated with a high marginal productivity. As investment increases, the marginal productivity of capital diminishes. With investment $I$ being measured on the vertical axis and the change in output on the horizontal axis, the ratio of investment to the change in output is initially relatively flat as the marginal productivity of capital at low levels of investment is high—($1/MP_K$) low—and becomes steeper as investment and output increase. Thus, the investment opportunities schedule, which begins at point A in Figure 13.8, is initially somewhat flat, and becomes steeper at higher levels of output—the dashed line from A.

Unlike in the short-run, Mundell–Fleming type models where investment has only an aggregate demand effect, here the additional capital stock also has an effect on the aggregate supply side where it raises the output.

How much investment expenditure will result on the demand side? Investment should take place up to the point where the benefit from that investment equals the cost. The benefit from investment is the additional output it makes possible or the marginal productivity of capital, and the cost is the interest rate. Hence, investment occurs up to the point where

$$MP_K = r^* = r \tag{13.25}$$

The slope of the tangent to the investment opportunity schedule at point D has slope $1/r^* = 1/MP_K$ and determines the investment expenditure as distance DE = distance BA in Figure 13.8. Distance FA represents the amount of national savings $S_t$ that finances investment. The remainder of investment (distance BF) is financed by foreign saving and represents the current account deficit CAB$_t$. Thus, as Eq. (13.22) makes clear, investment is financed by national savings and by running a current account deficit. The demand for foreign currency to finance the current account deficit is met by the supply of currency through capital inflows resulting in a capital account surplus, KAS, as given in Figure 13.8.

The slope of the tangent to the investment opportunity schedule at point D is

$$\text{Slope of tangent at D} = \frac{1}{r^*} = \frac{\text{distance DK}}{\text{distance JK}}$$

Hence,

$$\text{distance JK} = r^* (\text{distance DK}) = r^* \, \text{CAB}_t$$

The current account deficit was financed by incurring financial liabilities to the rest of the world [see Eq. (13.18)] and those liabilities[10] equal to $F$ are serviced at the interest rate $r^*$ in the subsequent period. Thus,

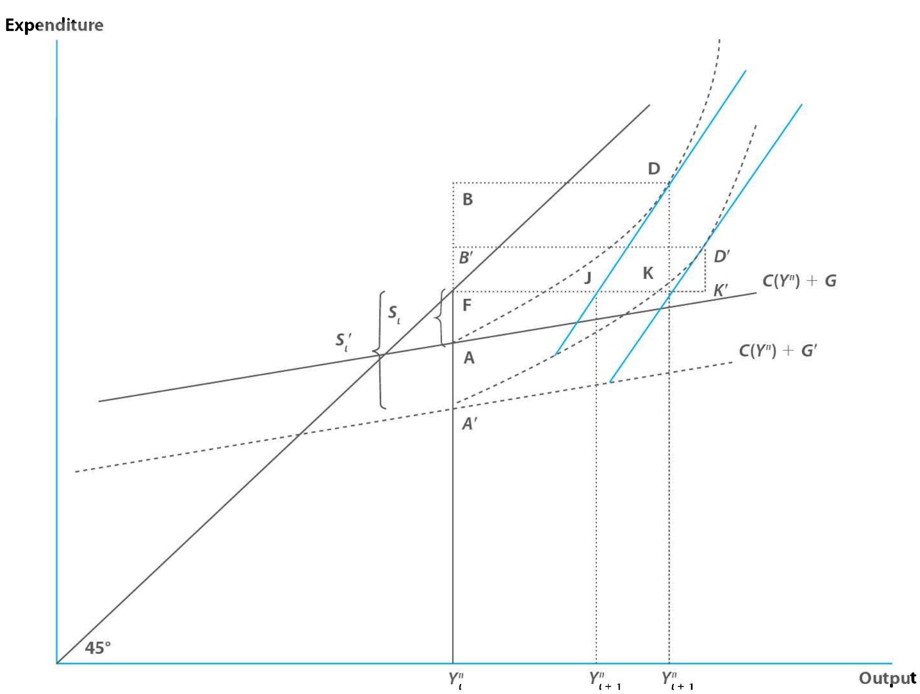
$$\text{distance JK} = r^* \, \text{CAB}_t = r^* F$$

The distance JK then is the interest payment on the foreign liabilities held and must be subtracted out of the domestic output $Q_{t+1}$ to arrive at the national income $Y^n_{t+1}$ [see Eq. (13.17)]. Unlike in short-run models where the capacity is given, here output in period $t + 1$ is above the output of period $t$ due to the additional investment that has augmented the capacity of the economy to produce. Also, the figure demonstrates that the RESORT TO FOREIGN SAVING as a source of financing investment can be maintained as long as the servicing of the debt is smaller than the increase in the output produced from the investment undertaken. External deficits facilitate higher investment and growth of output so long as the return on the investment funded by this means is beyond the cost of servicing the debt incurred.

› A country can resort to foreign savings to finance investment as long as the servicing of foreign debt is smaller than the increased output resulting from the investment undertaken.

### 13.3.2 Fiscal Adjustment and the Aggregate Income

What is the link between fiscal deficits and external sector deficits? Suppose the government reduces its expenditure from G to G′. Then public saving rises as $S_{\text{govt}} = T - G$ and so does national saving, $S = Y^n - C - G$. This reduction in the fiscal deficit results in a downward shift of the $C(Y^n) + G$ schedule to $C(Y^n) + G'$ in Figure 13.9.




**› Figure 13.9**
Effects of Reduction in Fiscal deficit. A reduction in the fiscal deficit reduces government expenditure to G′ and the economy moves from point A to A′, where national savings has increased to $S_t'$. There is a smaller requirement for funding investment expenditure through incurring financial liabilities abroad (D′K′ < DK) and subsequent reduced servicing of foreign liabilities, which raises income relative to output produced in subsequent periods or $Y^{n\prime}_{t+1} > Y^n_{t+1}$. A contraction in the fiscal deficit can thus increase the national income and product.

› A reduction in the fiscal deficit increases national saving. Greater domestic finance reduces the reliance on external finance for investment and the servicing of the external debt. This can lead to a higher subsequent national income depending on the productivity of investment expenditure.

The increase in saving to $S'_t$ reduces the expenditure in the current period as well as the current account deficit (from BF to B′F = D′K′). However, the higher national saving increases the sources of domestic finance for investment while the reduced current account from BF to B′F results in a decline in the servicing of external debt. The HIGHER NATIONAL SAVING and the lower external debt servicing together result in a higher national product in subsequent periods, $Y^{n\prime}_{t+1} > Y^{n}_{t+1}$. Thus, unlike in Mundell–Fleming type models fiscal contractions can actually be expansionary in the medium run as capacity gets augmented.

Of course, there are certain nuances that the exposition above still does not address. For instance, if the fiscal contraction is accompanied by a reduction in public investment and infrastructure spending instead of public consumption expenditure, it will diminish the output generating capacity of the economy and will result in a lower output and income in the future. This is particularly true in India where public investment is known to crowd in private investment.[11,12] The composition of deficit reduction is also as important as the size of the reduction.

Second, the reduction in government expenditure must be perceived to be permanent rather than temporary. If agents in the economy perceive the fiscal contraction as temporary, agents do not expect a change in the future tax obligations that is required to service the public debt. So private saving and consumption will not respond thereby reducing the growth promoting impact of a fiscal deficit reduction.

In an intertemporal set-up then lowering the fiscal deficit can be expansionary. Also, economies with low savings can sustain the large current account deficits and at the same time can experience good economic growth. If the investment opportunities are sufficiently productive and they result in an output that is capable of servicing the increased foreign indebtedness, there is no reason for foreign investors to be reluctant about channelling their funds and savings into such an economy. Economies with low national savings can thus undergo periods of high output growth.

MACROFOCUS 13.2

### Expansionary Fiscal Contractions

Standard *IS–LM* theory suggests that any reduction in the fiscal deficit would lead to a decline in economic activity. Recent experience, however, suggests that fiscal contractions may actually be expansionary. In a study that includes many OECD countries for the period 1960–1994, Alesina and Ardagna* find evidence of this phenomenon. Two unambiguously expansionary cases are that of Ireland during 1987–1989 and Australia in 1987. What explains the non-contractionary effect of fiscal consolidation?

One explanation that has been advanced is an expectation-based one. If public spending cuts are perceived to be permanent, consumers anticipate a reduction in the tax burden and a permanent increase in their lifetime-disposable income. Similarly, if the public debt is high, as this makes solvency more difficult, it may be accompanied by an interest rate premium to cover a default risk. A fiscal contraction that is perceived to be successful can then reduce the interest rate. The efficiency of financial markets may impede such an outcome. If consumers have a constrained access to financial markets, for instance, they may save rather than spend the increase in their lifetime-disposable income.

In addition to the demand-side channel that emphasizes the expectations regarding long-run tax liabilities, fiscal policy also has a supply-side effect on labour markets that affects competitiveness. A fiscal correction that relies on tax hikes rather than expenditure compression results in reduced post-tax real wages. Workers then bargain for higher wages, which results in higher labour costs and a loss in competitiveness. If unions are present and able to press for their wage demands there could be large losses of competitiveness.

A devaluation of the currency that increases net exports can offset the negative impact on the aggregate demand due to fiscal consolidation. The impact, however, is different in that fiscal adjustments in the OECD countries have been associated with a worsening of the trade balance rather than an improvement as predicted by the Mundell–Fleming model.

A fiscal adjustment that is preceded by a devaluation is also found to result in increased economic activity.

Thus, a fiscal adjustment that compresses public expenditure and that is implemented together with wage moderation and devaluation is expansionary. Fiscal adjustments that increase taxes are not successful. Countries such as Greece in 1986–1987 and Italy in 1993–1995 had tax-based fiscal adjustments that were unsuccessful even though the adjustment was accompanied by a devaluation.

* A. Alesina and S. Ardagna, "Fiscal Adjustments—Why They can be Expansionary," *Economic Policy* (October 1998): 488–517.

The United States and Japan are two examples that come to mind in this context. In 2004, the United States had a current account deficit of 4 per cent of the GDP and the national saving was only 14 per cent of the GDP.[13] The high external deficit has resulted in a call for protectionist measures as well as in lobbying for getting a main trading partner, China, to revalue its currency. Meanwhile, Japan had a current account surplus of 2 per cent of GDP and a national saving rate of 22 per cent of GDP. Yet, the United States experienced an average annual growth rate between 2000 and 2004 of 2.6 per cent, whereas Japan with a current account surplus experienced a lower growth rate of 1.3 per cent. Many other excluded factors are surely important for explaining this difference in growth rates.

What the above analysis reveals, however, is that apart from low savings/high current account deficits or high savings/low current account deficits, it is the productivity of investment as depicted in the slope of the investment opportunity schedule that is often a key determinant of sustained economic growth. From an economic policy point of view, the United States may not need to control its high external deficit and Japan rather than being smug with a high savings rate may need to initiate supply-side measures that improve productivity growth.

## 13.4 The External Debt Burden and Sustainability

How do we know whether the debt generated due to current account deficits is sustainable? Defining sustainability in the context of current account imbalances is complex since the imbalances reflect the interaction between the saving and investment decisions of the government and domestic private agents, as well as the lending decisions of foreign investors. Sustainability based on solvency considerations is simpler for fiscal imbalances given that these can be associated with direct policy decisions on taxation and government expenditure. An alternative way of formulating the question of whether the current account imbalances are sustainable is to ask if the current policy stance can be continued into the future indefinitely. This is what we examine now: the issue of sustainability.

How do price changes, the level of interest rates, and the rate of economic growth interact with trade imbalances in shaping the evolution of the external debt burden—ratio of the external debt to the GNP?

We write the current account balance in nominal terms. In real terms, it is

$$CA = F_t - F_{t-1} = (Y^n - C - I - G) + rF_{t-1} \tag{13.26}$$

where $F$ is the stock of net foreign assets and $r$ is the world interest rate. In nominal terms, we can write this[14] as

$$P(Y^n - C - I - G) + (r + \pi) E_t P_t^* F_{t-1} = E_{t+1} P_{t+1}^* F_t - E_t P_t^* F_{t-1}$$

where $E_t$ is the nominal exchange rate and $P^*$ the price level abroad (the foreign GNP deflator). Dividing both sides by the nominal GNP and rearranging terms, we may write

$$\frac{P(Y^n - C - I - G)}{PY^n} + (r + \pi)\frac{E_t P_t^*}{P_t}\frac{F_{t-1}}{Y_t^n} = \frac{E_{t+1} P_{t+1}^*}{P_{t+1}}\frac{P_{t+1}}{P_t}\frac{F_t}{Y_{t+1}^n}\frac{Y_{t+1}^n}{Y_t^n} - \frac{E_t P_t^*}{P_t}\frac{F_{t-1}}{Y_t^n}$$

Define the real exchange rate as $q_t = E_t P_t^* / P_t$ and let the ratio of the foreign assets to GNP in real terms be $f_t = q_t F_{t-1} / Y_t^n$. Then,

$$\frac{P(Y^n - C - I - G)}{PY^n} + (r + \pi)f_t = \frac{P_{t+1}}{P_t}\frac{Y_{t+1}^n}{Y_t^n} f_{t+1} - f_t$$

Now, $[P(Y^n - C - I - G)/PY^n]$ is the trade balance to nominal GNP ratio. Hence,

$$tb + (1 + r + \pi)f_t = \frac{P_{t+1}}{P_t}\frac{Y_{t+1}^n}{Y_t^n} f_{t+1}$$

Now, the inflation rate domestically is $\pi = [(P_{t+1} - P_t)/P_t] = (P_{t+1}/P_t) - 1$. Hence, $1 + \pi = P_{t+1}/P_t$. Also, $Y_{t+1}^n / Y_t^n = 1 + \gamma$, where $\gamma$ is the growth rate of the economy (growth rate of GNP). Then,

$$tb + (1 + r + \pi)f_t = (1 + \pi)(1 + \gamma)f_{t+1} \qquad \textbf{(13.27)}$$

or subtracting $f_t$ from both sides,

$$f_{t+1} - f_t = \frac{1}{(1 + \pi)(1 + \gamma)}[tb + (1 + r + \pi)f_t - (1 + \pi)(1 + \gamma)f_t]$$

or, $$f_{t+1} - f_t = \frac{1}{(1 + \pi)(1 + \gamma)}[tb + f_t(1 + r + \pi - 1 - \pi - \gamma - \pi\gamma)]$$

or, $$f_{t+1} - f_t = \frac{1}{(1 + \pi)(1 + \gamma)}[tb + f_t(r - \gamma - \pi\gamma)] \qquad \textbf{(13.28)}$$

› The change in the ratio of foreign assets to GNP, which reflects the current account position of the economy is driven by the trade imbalance, which is financed through the acquisition of financial assets abroad, and a debt-dynamics term that rises as the differential between the interest rate and the growth rate of the economy $(r - \gamma)$ diverges.

This expression states that changes in the RATIO OF FOREIGN ASSETS TO GNP are driven by trade imbalances and by a "debt dynamics" term proportional to $f(r - \gamma)$, ignoring the term $\pi\gamma$ which is a discrete time residual.

Recall that, $f_t = q_t(F_{t-1} / Y_t^n)$. Hence, the "debt-dynamics" term rises with the world rate of interest $(r)$, and falls with the rate of growth of the domestic economy $(\gamma)$, and the rate of real exchange rate appreciation $(q)$.

Note that from $q = EP^*/P$, $\Delta q/q = (\Delta E/E) + (\Delta P^*/P^*) - \pi$, an increase in $\pi$, the inflation rate, thus gets reflected in an appreciation of the real exchange rate $q$.

To understand the current account imbalances we analyse the Eq. (13.27) for its stability properties:

$$tb + (1 + r + \pi)f_t = (1 + \pi)(1 + \gamma)f_{t+1} \qquad \textbf{(13.27)}$$

We can write this as

$$(1 + \pi)(1 + \gamma)f_{t+1} = tb + (1 + r + \pi)f_t$$

or, $$f_{t+1} = \frac{tb}{(1 + \pi)(1 + \gamma)} + \frac{(1 + r + \pi)}{(1 + \pi)(1 + \gamma)} f_t \qquad \textbf{(13.29)}$$ [15]

This is a difference equation which can be graphed. If we measure $f_{t+1}$ on the $Y$-axis and $f_t$ on the $X$-axis, the above equation is a straight line with intercept $(tb)/(1 + \pi)(1 + \gamma)$ and slope $(1 + r + \pi)/(1 + \pi)(1 + \gamma)$. This slope is less than unity when

$$\frac{1 + r + \pi}{(1 + \pi)(1 + \gamma)} < 1$$

or, $$1 + r + \pi < 1 + \pi + \gamma + \pi\gamma$$

or, $$r < \gamma$$

ignoring $\pi\gamma$ which is a discrete time residual and small.

Similarly, the slope of Eq. (13.29) is greater than unity when $r > \gamma$.

The steady-state solution is obtained by setting $f_{t+1} = f_t = \bar{f}$ in the above equation. Hence,

$$\bar{f} = \frac{tb}{(1+\pi)(1+\gamma)} + \frac{(1+r+\pi)}{(1+\pi)(1+\gamma)}\bar{f}$$

or,

$$\left[\frac{(1+\pi)(1+\gamma)-(1+r+\pi)}{(1+\pi)(1+\gamma)}\right]\bar{f} = \frac{tb}{(1+\pi)(1+\gamma)}$$

or,

$$\frac{1+\pi+\gamma+\pi\gamma-1-r-\pi}{(1+\pi)(1+\gamma)}\bar{f} = \frac{tb}{(1+\pi)(1+\gamma)}$$

or,

$$\bar{f} = \frac{1}{(\gamma - r)}tb, \qquad \text{when } \pi\gamma \text{ is a discrete time residual}$$

Suppose a country runs a trade balance deficit, which means that the country's current absorption $(C + I + G)$ is greater than the current income $(Y^n)$. From the above equation:

$$tb = (\gamma - r)\bar{f}$$

and so $tb < 0$ can result in a PERMANENT CURRENT ACCOUNT DEFICIT, if $\gamma > r \Rightarrow \bar{f} < 0$. Alternatively, if $\gamma < r$, the above equation implies that $\bar{f} > 0$ or a permanent current account surplus, when $tb < 0$.

› The permanent current account position of an economy occurs where there is no tendency for the stock of foreign assets to GNP to change. This is a steady state given by $f_{t+1} = f_t = \bar{f}$. With a growth rate of the economy, $\gamma$, the permanent current account is $\gamma\bar{f}$.

Thus, a country can be currently running a trade deficit and yet the permanent current account may be in deficit or surplus depending on whether $\gamma > r$ or whether $\gamma < r$. When $\bar{f} > 0$, the country is holding positive claims against the rest of the world and is a net creditor to the rest of the world. Alternatively, when the country runs a permanent current account deficit, $\bar{f} < 0$, and the country is a net debtor to the rest of the world. Recall from Eq. (13.10) that the current account is the trade balance plus interest payments on net foreign assets, or

$$ca = tb + rf$$

from which the permanent current account deficit $\overline{ca}$ is given by

$$\overline{ca} = (\gamma - r)\bar{f} + r\bar{f} = \gamma\bar{f}$$

The permanent current account deficits are thus directly proportional to the steady-state holding of foreign assets $\bar{f}$ and have the same sign as $\bar{f}$.

For the case where the TRADE BALANCE IS CURRENTLY IN DEFICIT, Figure 13.10 depicts the permanent current account position of the country for the two cases $\gamma > r$ and $\gamma < r$.

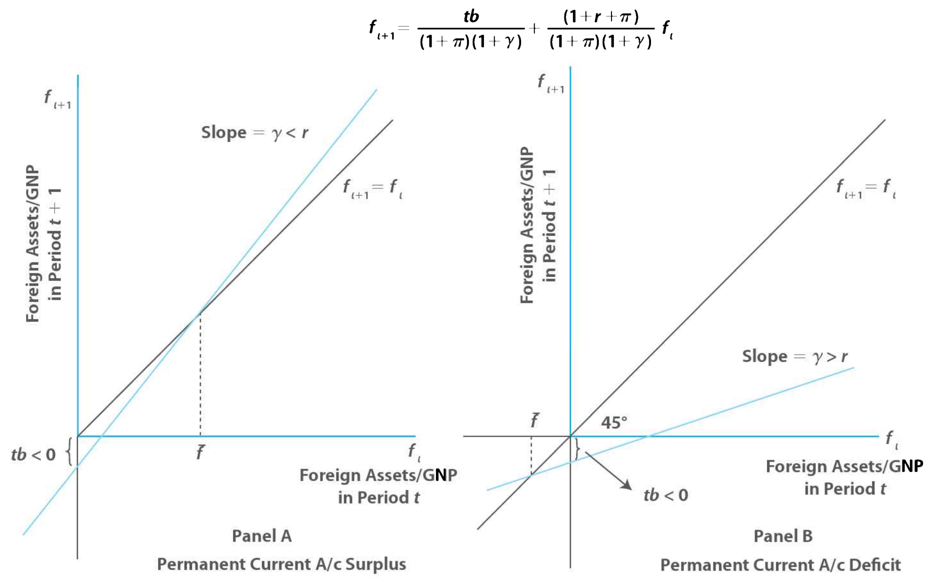
› A country that spends more than its output can absorb runs a trade deficit. For the trade deficit to be associated with a permanent current account surplus ($\gamma\bar{f} > 0$), the country must be a net creditor of the rest of the world ($\bar{f} > 0$), and the growth rate of the economy is less than the interest rate on its holding of foreign assets. For a trade deficit to be associated with a permanent current account deficit ($\gamma\bar{f} < 0$), the country must be a net debtor to the rest of the world ($\bar{f} < 0$), and the growth rate of the economy exceeds the interest payments on its foreign liabilities.

Similarly, when a country runs a trade balance surplus currently, the permanent current account position of a country can be a deficit or a surplus depending on whether $\gamma < r$ (deficit) or whether $\gamma > r$ (surplus). This situation is depicted in Figure 13.11.

Now consider an economy in a steady state in which consumption, investment, public expenditure and the stock of foreign assets are constant as a fraction of GNP.[16] What is the long-run net resource transfer (trade surplus) that an indebted country must undertake in order to keep the debt-to-output ratio constant? From Eq. (13.28) we get

$$tb = 1 - c - i - g = -f(r - \gamma) \qquad \textbf{(13.30)}$$

where $tb$ is the long-run trade balance and small-case letters are the capital letters to GDP ratios. This expression highlights the role played by the average

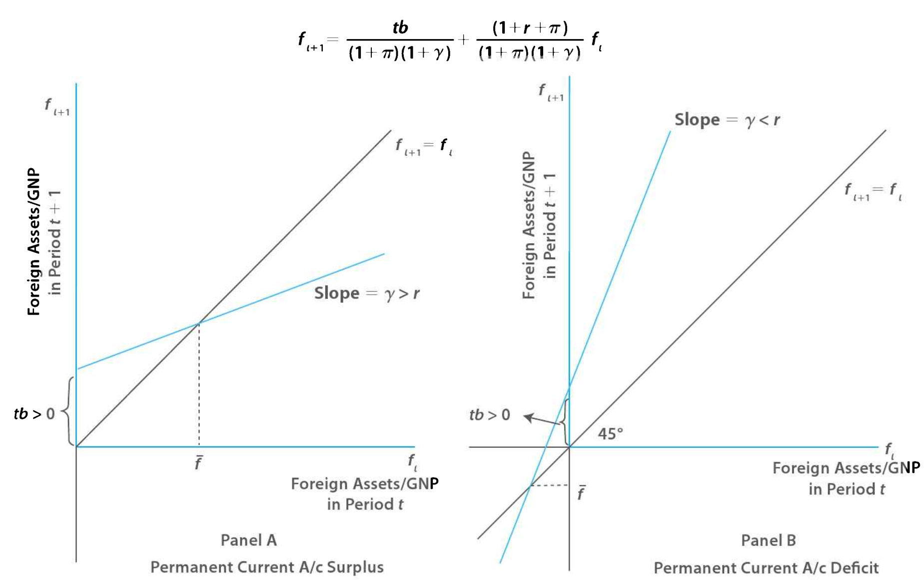



**› Figure13 .10**
Trade Balance Deficits and the Permanent Current Account. The ratio of foreign assets to GNP in a period is given by the trade imbalance (which is financed through the acquisition of financial assets abroad) and a debt-dynamics term that rises as the differential between the interest rate and the growth rate of the economy ($r - \gamma$) diverges. If the economy grows at a smaller rate than the interest rate (panel A), the trade deficit results in the economy becoming a net creditor to the rest of the world, $\bar{f} > 0$. If the economy grows at a faster rate than the interest rate (panel B), the trade deficit results in the economy becoming a net debtor to the rest of the world, $\bar{f} < 0$.

future value of world interest rates, domestic growth, and the long-run trend in the real exchange rate in determining the resource transfers necessary to keep the debt-to-GNP ratio from increasing. Consider the case in which the long-run real exchange rate is constant ($q$ = constant).

From Eq. (13.30), the country's long-run absorption can be higher than its income $c + i + g > 1$, only if the country is a net creditor. In this case, the country will run a trade deficit equal to $(r - \gamma)f$ but a current account surplus equal to $\gamma f$ due to the interest it earns on its foreign assets ($rf$). Thus, in the presence of economic growth, the permanent current account deficits can be consistent with solvency even when the growth rate is below the world interest rate, provided they are accompanied by sufficiently large trade surpluses.

If the long-run growth rate of the economy is zero, the current account must be balanced for the foreign assets (debt)-to-GDP ratio to be constant. In this case a country that is a debtor in the long run will have to run a trade surplus equal to $-rf$ to pay the interest on its external liabilities, while a country that is a long-run creditor will run a trade deficit.




**› Figure13 .11**
Trade Balance Surplus and the Permanent Current Account. The ratio of foreign assets to GNP in a period is given by the trade imbalance (which is financed through the acquisition of financial assets abroad) and a debt-dynamics term that rises as the differential between the interest rate and the growth rate of the economy ($r - \gamma$) diverges. If the economy grows at a faster rate than the interest rate (panel A), the trade surplus results in the economy becoming a net creditor to the rest of the world, $\bar{f} > 0$. If the economy grows at a slower rate than the interest rate (panel B), the trade surplus results in the economy becoming a net debtor to the rest of the world, $\bar{f} < 0$.

The *practical application* of this is to postulate a continuation into the indefinite future of the current policy stance and no change in the relevant features of the macroeconomic environment. The CURRENT POLICY STANCE is then deemed to be sustainable if its continuation in the indefinite future does not violate solvency (budget) constraints.

› The actual current account deficit of an economy is sustainable, if it is lower than the value of the permanent current account deficit.

Callen and Cashin[17] carry out exactly such an exercise. To the stock of external debt to GDP[18] of 23 per cent towards the end of 1998–1999; they add an estimate of the stock of external equity liabilities calculated by accumulating the flows of FDI and FII in the capital account of the balance of payments from 1970. They estimate an outstanding stock of equity liabilities of 8 per cent of GDP in the same period. The stock of foreign assets to GDP is thus estimated at 31 per cent of the GDP. Given a constant real exchange rate, assuming a real GDP growth of 5 per cent a year ($\gamma = 5$ per cent) and the cost of foreign liabilities to be 4 per cent a year (r = 4 per cent), from $tb = -f(r - \gamma)$ we get, 31 per cent × (4 per cent − 5 per cent) or 0.3 per cent GDP. This is what the deficit on goods, services, and transfers must average to maintain the stock of foreign assets to GDP ratio at 31 per cent. The cost of servicing the outstanding stock of external liabilities is $rf = 4$ per cent (31 per cent) = 1.2 per cent of GDP. Hence, the trade deficit of 0.3 per cent of GDP plus the debt servicing of 1.2 per cent of GDP gives a current account deficit of 1.5 per cent of GDP in 1998–1999 that can be sustained. By contrast the actual current account deficit in 1998–1999 at 1 per cent of GDP is lower than the sustainable current account deficit, which is 1.5 per cent of GDP. Using the same tools, they demonstrate that the current account deficits were not sustainable from the 1980s up to 1990–1991. In the 1990s, by contrast, India has been in a position to sustain a current account deficit from 1.5 to 2.5 per cent of the GDP depending on the economic and policy environment.

## SUMMARY

- In the Mundell–Fleming model, a fiscal contraction leads to a contraction in economic activity. In many economies, however, fiscal contractions have led to expansions in economic activity. Intertemporal models of the current account can explain this anomaly.
- A current account deficit is financed through the accumulation of financial claims or liabilities against the rest of the world. A current account surplus is an accumulation of financial assets by residents as a result of lending abroad.
- The national intertemporal budget constraint states that the present value of aggregate consumption equals the present value of national production net of investment expenditure.
- A second expression for intertemporal balance can be stated as the condition that a trade deficit in the present must be balanced by a trade surplus of equal present value in the future.
- A third expression for intertemporal balance is that the sum of present and future current account balances must add up to zero.
- A country can consume and invest more than its current output by increasing its indebtedness to the rest of the world via a trade deficit. This requires a future trade surplus that finances the debt incurred in the present.
- Countries resort to international capital markets to finance expenditure due to three motives: consumption smoothing, consumption tilting, and consumption augmenting.
- With a fluctuating national income, consumption smoothing occurs when a nation resorts to the international capital market to enable it to enjoy stable consumption.
- Consumption tilting occurs when a nation with a flat-time profile of national income has a time preference for greater (lesser) consumption in the present and lesser (greater) consumption in the future.
- Consumption augmenting is the outcome of resorting to international capital markets to finance investment whose productivity enables higher present as well as future consumption.
- The three motives of consumption smoothing, tilting, and augmenting, coexist and interact in generating an equilibrium pattern of consumption, investment, and debt accumulation.
- The *IS–LM* model assumes underutilized capacity so that the output adjusts to changes in aggregate demand. It does not satisfactorily incorporate the supply-side effect of investment expenditure, which adds to capital stock and capacity. In a medium-run model, the aggregate output is a function of the capital stock, which includes the previous period additions to the capital stock
- Output $= f(K_{t-1}) = f(K_{t-2} + I_{t-1})$
- A country can resort to foreign savings to finance investment as long as the servicing of the foreign debt is smaller than the increased output resulting from the investment undertaken.

- A reduction in the fiscal deficit increases government and national saving. The greater availability of domestic financial capital reduces the reliance on external finance for investment and the servicing of external debt. This can lead to higher subsequent national income depending on the productivity of investment expenditure.
- The change in the ratio of foreign assets to GNP, which reflects the current account position of the economy, is driven by the trade imbalance (which is financed through the acquisition of financial assets abroad) and a debt dynamics term that rises as the differential between the interest rate and the growth rate of the economy ($r - \gamma$) diverges.
- The permanent current account position of an economy occurs where there is no tendency for the stock of foreign assets to GNP to change. This is a steady state given by $f_{t+1} = f_t = \bar{f}$ With a growth rate of the economy, $\gamma$, the permanent current account is $\gamma\bar{f}$.
- A country that spends more than its output can absorb runs a trade deficit. For the trade deficit to be associated with a permanent current account surplus ($\gamma\bar{f} > 0$), the country must be a net creditor of the rest of the world ($\bar{f} > 0$), and the growth rate of the economy is less than the interest rate on its holding of foreign assets.
- For a trade deficit to be associated with a permanent current account deficit ($\gamma\bar{f} < 0$), the country must be a net debtor to the rest of the world ($\bar{f} < 0$), and the growth rate of the economy exceeds the interest payments on its foreign liabilities.
- The actual current deficit of an economy is sustainable if it is lower than the value of the permanent current account deficit.

## NOTES

1. Francesco Giavazzi and Marco Pagano, "Can Fiscal Contractions be Expansionary? Tales of Two Small European Countries," in Olivier Blanchard and Stanley S. Fischer (eds.) *NBER Macroeconomics Annual* (MIT Press, 1990).
2. Assaf Razin and Efraim Sadka, "Fiscal Balance During Inflation, Disinflation and Immigration: Policy Lessons," in Mario Blejer and Teresa Ter-Minassian (eds.) *Macroeconomic Dimensions of Public Finance: Essays in Honor of Vito Tanzi* (IMF, 1996).
3. This chapter is inspired heavily by Jacob A. Frenkel, Assaf Razin, and Chi-Wa Yuen, *Fiscal Policies and Growth in the World Economy* (MIT, MA: MIT Press, 1996).
4. This assumption is also made in Mundell–Fleming and intertemporal two-period current account models. The assumption of perfect capital mobility can be relaxed but it would not affect the main results.
5. See Eq. (6.8) in Chapter 6. With $\pi_t$ constant $i = i^*$ implies $r = r^*$, where $r$ is the real market interest rate.
6. See Eq. (8.11) in Chapter 8.
7. See Eqs. (A5.1.1) and (A5.1.2) in Appendix 5.1 of Chapter 5.
8. Alternatively, we know that $-\text{KAB} = \Delta F$ from Eq. (13.2).
9. A current account deficit must be financed by capital inflows, which is the savings of foreigners deployed in the domestic market.
10. In this case as there is no initial stock of foreign assets $\Delta F = F - 0 = F$.
11. Luis Serven, "Does Public Capital Crowd out Private Capital?—Evidence from India," The World Bank, Policy Research Working Paper No. 1613, 1996.
12. K. N. Murty and Alamuru Soumya, "Macroeconomic Effects of Public Investment in Infrastructure in India," IGIDR, Working Paper No. 2006–003, Mumbai, 2006.
13. "World Development Report", Table 3, World Bank, Oxford University Press, 2006.
14. The yield $(1 + r)$ in nominal terms is actually $(1 + r)(1 + \pi) = 1 + r + \pi + r\pi$, so that in nominal terms we should write $(r + \pi + r\pi)E_tP_t^*F_{t-1}$. We are assuming $r\pi$ to be small and ignoring it.
15. $(1+ r + \pi) = (1 + r)(1 + \pi)$ for small $r\pi$ and so the slope may also be considered to be $[(1 + r)(1 + \pi)]/[(1 + \pi)(1 + \gamma)] = (1 + r)/(1 + \gamma)$.
16. With $f_{t+1} = f_t = f$ in a steady state, we can solve for $tb$ from Eq. (13.22). Also, recall that the nominal trade balance-GNP ratio is $P(Y^n - C - I - G)/PY^n = 1 - c - i - g$.
17. Tim Callen and Paul Cashin, "Assessing India's External Position," in T. Callen, P. Reynolds, and C. Towe (eds.) *India at the Crossroads-Sustaining Growth and Reducing Poverty* (Washington, DC: IMF, 2001).
18. Callen and Cashin write the stock of foreign assets as a ratio of GDP rather than GNP. For empirical purposes, the difference arising due to measuring aggregate income as GDP rather than GNP in the denominator is marginal. Tim Callen and Paul Cashin, op. cit.
19. Martin Feldstein and Charles Horioka, "Domestic Saving and International Capital Flows," *The Economic Journal* 90, no. 358 (June 1980): 314–329.

## TEST YOURSELF

1. What is the advantage of the intertemporal over the Mundell–Fleming approach to the current account?
2. Demonstrate how the intertemporal accounts of a nation can be written as: (a) the discounted value of the aggregate consumption equal to the discounted value of the national output net of investment, (b) the sum of the present value

of the trade balance over time equal to zero, and (c) the sum of the current account balance over time equal to zero.

3. What motives result in a country running up trade imbalances and resorting to international capital markets?
4. How do the productivity of investment and the nature of shocks to the economy (temporary or permanent) affect the trade balance of an economy?
5. In the medium run how does investment expenditure affect the economy's supply of output?
6. An economy is running a current account deficit. How does a reduction in the fiscal deficit affect the current account and the gross national product (GNP) in the medium run?
7. An economy is running a current account deficit. The fiscal deficit is reduced by reducing public investment expenditure. With the aid of a diagram demonstrate what impact this can have on the current account and on GNP.
8. How does the net foreign asset investment position of the economy evolve over time?
9. When is the trade deficit associated with a permanent current account surplus or a permanent current account deficit?
10. What is the value of evaluating the permanent current account position of an economy from a policy perspective?

## ONLINE APPLICATION

In a closed economy, the current account is zero and national saving equals domestic investment. Accordingly any increase in savings in a closed economy will be accompanied by an equal rise in domestic investment. In contrast in an open economy with capital mobility, saving and investment can diverge due to the opportunities available to an economy to gain from intertemporal trade. With perfect capital mobility there should be no relation between the domestic saving and investment as saving in each country responds to the worldwide opportunities for investment, and investment in a country is financed by the international pool of capital. Feldstein and Horioka[19] found that for a sample of 16 OECD countries, which typically have high rates of capital mobility, changes in domestic saving rates ultimately change domestic investment rates by the same amount over the period 1960–1974. This phenomenon has been termed the Feldstein–Horioka Puzzle.

1. Go to the home page of the United Nations Statistics Division (UNSD)
   Internet (URL: http://unstats.un.org/unsd/default.htm).
2. Click on the "Statistical Databases" icon.
3. The statistical databases section has an unrestricted access subsection. In this category, click on "National Accounts Main Aggregate Databases".
4. Under the Downloads icon click on "File List".
5. You now enter a section of the Web site where Excel files can be downloaded for free. Go to the section titled "GDP and its breakdown at constant 1990 prices in US dollars". Under this section click on the icon "All regions/subregions (totals) for all years—sorted alphabetically".
6. We need a selection of countries that would typically have high capital mobility. A good sample in this regard would be the countries of Western Europe that includes among others France, Germany, the Netherlands, and Switzerland. Most of these countries are OECD countries just like the case in the Feldstein–Horioka original sample. The two countries that are not—Liechtenstein and Monaco—are classified as high-income economies (high-income economies have a per capita income of US$9,386 or more). The Western Europe economies statistics are given in rows 344–360 of the Excel file.
7. Aggregate Investment is given by the data in the row Gross Capital Formation.
8. Aggregate national savings is not reported in the data. Estimate aggregate savings as a residual by subtracting the Final Consumption Expenditure from GDP.
9. Divide national savings by GDP to get $S/Y$ or the savings ratio. Similarly, divide domestic investment by GDP to get the investment ratio $I/Y$.
10. Create a graph of the data for the period 1970–2006 with $S/Y$ on the horizontal axis and $I/Y$ on the vertical axis. Draw a linear trendline for this graph.
11. If the slope of your graph is close to 1, which is the slope of a 45-degree line, it would indicate that most of the increment to savings in this group of countries has remained there. If the slope of your graph is flatter—close to zero—it would indicate that an increase in savings is associated with responses to international opportunities for investment and that the domestic investment is financed by the worldwide supply of capital. What does the graph you have created depict? Do these countries show a relationship between domestic savings and investment? Can you explain the relationship as shown in the graph?
12. The apparent lack of capital mobility that the Feldstein–Horioka Puzzle reports is sometimes explained by claiming that developed economies usually have sufficient provision of capital and are near the steady-state levels of their external debt. They then do not have much to gain from intertemporal trade that enables them to borrow for investment purposes. Yet another explanation is that when there is a pecking order to finance then an increase in firm saving (retained profits) can relax the financial constraint and cause higher investment even when international capital markets are as accessible as domestic market.

# 14

# Financial Sector Reforms

CRITICAL QUESTIONS

» *What is information asymmetry and how do financial markets attempt to deal with it?*

» *In what way is the market for loans different from bonds (debt markets) and stock markets?*

» *What is financial repression?*

» *What were the major banking sector reforms in India?*

» *How does the presence of risk affect a bank's decision to give loans and affect its composition of assets?*

» *What was the impact of economic deregulation on the banking sector?*

# 14.1 Features of Financial Markets

At one level of abstraction economics is concerned with commodities, money, and securities. Goods such as rice and cars, services such as transportation and labour, and assets such as land and machines are commodities. Money, such as dollars and rupees, is not a commodity, but a means of exchange of commodities. Securities, such as bonds and stocks, are neither commodities nor money but contracts for the exchange of commodities or money.

Financial economics is a specialized branch of economics because securities or financial products are different from commodities and money—FINANCIAL PRODUCTS are contracts that allow investors to delegate the management of their assets to others. An example of such delegation is a mutual fund or unit trust, where investors retain full legal title to their assets, and in the case of an open-ended fund, can also get a pro-rata share of the fund back upon demand. In the case of equity, debt, and bank-mediated instruments, investors surrender their assets in return for control rights that constrain managers and directors and motivate them to pay dividend and interest payments.

› Financialp roducts are contracts that allow investors to delegate the management of their assets to others.

Markets, as Arrow and Debreu[1] demonstrated, work best in allocating resources when: (1) there is full information, (2) agents are atomistic—too small to influence prices, and (3) there are no externalities—unpriced side effects, such as environmental pollution. Transactions in such markets, when all agents have the same piece of information, result in prices that fully reflect the values of the goods transacted. However, the moment this INFORMATION IS PRIVATE in the sense that it is only accessible to individual or some of their close associates, individuals have an incentive to hide their information from the rest of the market to make gains. For instance, in the market for insurance, clients will typically know more about their own risk characteristics and have an incentive to hide this so that they are asked to pay smaller premiums on their insurance policy.

› The delegation of assets in the financial markets is circumscribed by the prevalence of asymmetric information.

This ASYMMETRIC-INFORMATION situation, in which one side of the market knows more about the properties of the good being traded than the other, results in inefficient trading and in some cases may not result in trade taking place at all. The essential feature of a financial market is one where information asymmetries prevail. Every one does not have access to this information—it is not exogenous, public information—but it is endogenous, and to some extent, is possible for an individual or agent to improve on the available information through research via screening, monitoring, or auditing activity. However, as the information generated through research of various kinds is a PUBLIC GOOD, there will be little incentive for spending private resources to acquire information that becomes publicly available. An example of this is when agents have some information about firms, and by purchasing or selling stocks—the prices of which are public information—they publicly reveal the information about the firms.

› Asymmetricinf ormation is prevalent when buyers and sellers are not equally well informed about either the characteristics of investment projects or the goods and services for sale, or about the effort expended to efficiently utilize the delegated resources.

› Eliciting information can be a public good in the sense that once the information is known by an agent, other agents can access it at negligible cost.

The two types of asymmetric information emphasized in the literature on financial markets are adverse selection and moral hazard.

### *14.1.1* Adverse Selection

This tends to occur before a transaction takes place and arises due to HIDDEN INFORMATION about the type of individual or good. In the case of insurance transactions, for instance, buyers know more about themselves and the risks they are exposed to than the provider of insurance. In the case of transactions in a used or second-hand car market, for example, the seller knows more about whether the car runs smoothly and reliably than does the buyer. In both cases

› Adverse selection occurs when the information possessed by an individual (or firm or institution) is concealed from others and the welfare (or pay-offs) of others depends on this information.

one side of the transacting parties—buyer or seller—has some information that is hidden from the other party.

› The effect of hidden information arising from adverse selection is minimized by screening agents or through using contracts to reveal the information.

The effect of this HIDDEN INFORMATION can be minimized by using screening devices. Insurance agents insist that clients are subject to a check-up by a doctor, and cars can be subjected to professional examination and a road test by certified mechanics. Similarly, banks use information systems to identify clients with poor credit records, and stockbrokers employ company analysts to tell them about the financial health of companies.

In some cases, it is possible to get clients to reveal private information by offering a range of different contracts at different prices.[2] If one insurance contract, for instance, offers limited cover at a low price and the other more cover at a commensurately higher price, then the low-risk clients will select the former insurance contract, whereas the high-risk clients will select the latter. This two-tier contract separates the high-risk clients from the low-risk clients and reduces the dispersion of risk characteristics.

### *14.1.2* Moral Hazard

› Moral hazard is the hidden action arising after a contract is endorsed since it is extremely costly to monitor the effort expended towards meeting the objectives of the contract.

Moral hazard occurs when a contract changes the incentives and behaviour of the parties that enter into it. MORAL HAZARD takes the form of *hidden actions* during the period of the contract. The best examples are insurance contracts, which reduce the incentive for the client to take care. Once an insurance contract is in place, for instance, the outcome of a car crash depends on the behaviour (rash driving) of the parties involved as well as chance, and it is difficult to know which factor prevailed. Table 14.1 summarizes the asymmetric-information issues.

**› Table 14.1**
Asymmetric Information Issues

| Time Frame: | Transaction<br>Ex ante | Contract Period | Outcome<br>Ex post |
|---|---|---|---|
| Problem | Adverse Selection | Moral Hazard | State Verification |
| Solution: | | | |
| (i) Contract | Two Tier | Equity Participation | Debt Contract |
| (ii) Research | Screening | Monitoring | Accounting and Audit by<br>Debt: Bankruptcy<br>Equity: All States |

The actions may not be easily observable but in principle it may be possible for one party to find out what the others are doing by monitoring their behaviour. However, since monitoring needs to be done on a continuous basis it is costly and can also be intrusive.

› Moral hazard is mitigated through clauses in contracts that induce agents to expend the appropriate effort.

When monitoring is not possible, payments and contracts can only be specified in terms of outcomes rather than actions. However, typically, CONTRACTS, which aim to minimize the adverse incentives that arise, end up exposing individuals to more risk. Insurance companies, for instance, introduce no-claims bonus schemes to induce reduced adverse behaviour, such as rash driving, but in the process of doing so expose clients to some of the risk the contract was designed to share. Similarly, giving executives equity-based remuneration contracts induces them to work harder but also exposes them to more risk. (These are typical problems that are dealt with in the principal–agent literature.) In financial markets, such problems arise when shareholders (principals) delegate the management of a company to a board of directors (agents), or when depositors (principals) delegate the management of their assets to a bank and its shareholders (agents).

### 14.1.3 *Ex post* State Verification

When the outcome of a venture is not apparent to all the parties to a financial contract, the uninformed may have to expend resources confirming this rather than believing what they are told by the informed party. These are known as *audit costs*. Thus, for example, an equity contract allows outside investors to share in the outcome of an investment. However, only insiders may know the true outcome, and they have an INCENTIVE TO UNDERSTATE this in order to minimize the distribution to outsiders.

Debt contracts offer a return that is independent of the outcome of a project. If the return on the project is insufficient to repay the principal and interest on the loan, the creditor takes over the management and ownership of the venture and keeps the residual value with the equity participants getting nothing in this case. In such a contract, if there is an attempt to cheat the lender (creditor) by declaring that the outcome was insufficient to repay the debt when it was not, the owners and managers of the firm lose ownership and get nothing at all. The debt contract thus is a mechanism for dealing with the veracity problem as it gets managers and equity holders to tell the truth about the outcome. However, it distorts the incentives of owners and managers who are encouraged to switch high-risk for low-risk projects as the gain from the upside accrues to them whereas in the downside they lose control over the project anyway. This EXCESSIVE RISK TAKING is a form of moral hazard known as asset substitution.

› Bond holders have the first claim on cash flows but receive fixed amounts if the firm makes enough income to meet its debt obligations. Equity investors have a claim on the residual cash flows and can declare bankruptcy if cash flows are insufficient to meet financial obligations. Thus, bondholders do not participate in the upside returns if a project succeeds, and bear a significant part of the cost if it fails. The nature of cash flow claims is a source for the **conflict of interest** between bondholders and shareholders.

› Once a debt issue is incurred through a bond, managers have an incentive to substitute high-risk projects for low-risk ones as they discount the negative effect on the bondholder value. This is asset substitution or **excessive risk taking**.

### 14.1.4 Banks and Debt Markets

There are various ways in which lenders and other creditors protect themselves from moral hazard. In the case of a public bond issue, the investment bank sponsoring the issue will usually attempt to protect its bond investors from any value transfer by writing COVENANTS into the prospectus that limit the scope for asset substitution and the issuance of new debt. Standard covenants prevent the firm from paying a dividend that is not covered by the company's earnings. Cash flow covenants require the firm to trade at a profit. Net value covenants PREVENT THE FIRM FROM TRADING unless its assets exceed its fixed liabilities. In the absence of these injunctions, the firm and its shareholders effectively have an option which allows them to trade on in the hope of recovery until the situation gets so bad that it is no longer worth putting equity into the company in a reorganization. It is the bond investor's responsibility to ensure that these covenants are adhered to. This requires monitoring, which is only possible if an objective set of accounts that cannot be fudged are available. Thus, though accounting systems are not necessary for bond investors to overcome the verification problem, it is necessary to buttress any covenants used to minimize moral hazard.

› Since bondholders are aware of the potential of shareholders to take actions inimical to their interest, they protect themselves by writing in **covenants** or restrictions on what stockholders can do.

› **Bond agreements can** restrict firms' dividend policy **by tying dividend policy to earnings. They can also restrict firms' investments and how much risk they can take on new investments.**

Banks typically do not use covenants to protect their investments from bad borrower behaviour. A bank will protect a loan by COLLATERAL or by securing it against specific assets which is difficult to organize in support of a bond issue. A bank also typically relies on an informal client agreement or understanding that can be monitored effectively from its POSITION AS BOOKKEEPER. The long-term benefits of the customer relationship are such that the necessity of repeat purchases in the form of future loans play an important role in restricting the borrower from bad behaviour. This is particularly the case for high-growth companies with intangible assets that cannot be pledged as collateral.

› Banks protect loans by **collateral** or assets (like machinery or building) backing the security of the loan.

› **Banks also manage the payments system through accounts held at the bank. This enables them to** monitor **the financial viability of a client for a loan. Borrowers who consider future loan requirements are careful to restrict excessive risk taking.**

Bond investors also tend to be smaller than banks. The problem of coordinating small lenders makes public debt a much tighter discipline upon company managers than bank debt. Thus, bank loans can be rescheduled or renegotiated more flexibly than public bond issues. Thus, small companies that need a sympathetic lender tend to rely upon bank finance despite the cost while larger companies would tend to use the bond market.

### *14.1.5* Banks and Capital Markets

The key features of banking and capital markets are as shown in Table 14.2.

› **Table 14.2**
Features of Capital and Banking Markets

| | Capital Markets | Banking Markets |
|---|---|---|
| Transparency | Open | Confidential |
| Typical Investors | Small | Large |
| Contracts | Equity and Bond | Loan and Deposit |

› Information disclosure requirements in bond and equity markets are high. Banking transactions are more confidential.

In contrast to capital–market relationships which are open, BANKING RELATIONSHIPS are highly confidential. Banks are given access to inside information, which allows them to finance a company privately without revealing its trade secrets to the market and its rivals. The information disclosure requirements by capital markets are fairly extensive and there is extensive disclosure of commercial information so as to minimize mispricing and opportunistic behaviour. That's why companies in knowledge-based industries try to remain private as long as they can despite their need for risk capital. This is known as information dilution.[3] Venture capital offers an appropriate solution to this problem by allowing a company to access finance while making the venture capitalist an insider through managerial involvement. In this case, there are no outside investors to free ride on the public good that is the screening and monitoring effort of the investor.

As large institutions banks enjoy economies of scale and scope in monitoring borrowers and reducing risk through portfolio diversification, public debt also differs from bank debt in this respect. The company's bank account is an excellent indicator of cash flow and financial viability. Banks are able to centralize the research and monitoring cost functions thus avoiding the problem of decentralization found in the capital market.

In reality, banks and capital markets are complementary. In a capital market, small shareholders find it not worthwhile to scrutinize company management because of the high cost and public good nature of the problem. However, shareholders know that if a bank lends to a project it must perceive little downside risk as it lends on fixed terms without any returns from the upside. Shareholders can free ride on this information of a loan approval and be secured in the knowledge that the bank's accountants and experts have scrutinized the project properly, and will bid the share price up. The bank does not mind as the extra equity provides more of a cushion for its own investment. As a result, blue chip companies often borrow from banks despite their credit ratings, which could help them borrow cheaper in the bond market.

Banks' inside information help mitigate the adverse selection problem through initial screening of loan applications. Banks become insiders in an informational sense, but unlike venture capital firms, they do not have day-to-day managerial involvement. However, the bank handles the borrower's transactions account, and this puts it in a position to monitor the borrower's financial behaviour in addition to screening new applications. This is an economy of scope, a synergy between borrowing and lending business, which provides a rationale for the joint provision of these facilities by the bank.

Often a debt instrument is considered safe because of the collateral offered. However, even after a bank efficiently screens customers when deciding loan applications, it faces the problem of having to monitor the behaviour of borrowers. If a lender monitors the borrower, competitors in the banking industry such as equity investors can free ride on this effort. If several lenders make loans to the same borrower, they may duplicate each other's monitoring effort or rely on each other's monitoring effort, so that no one actually supervises the borrower. Diamond[4] shows how a large bank solves these informational problems at a stroke by taking the firm off-market and becoming the sole provider of capital. This takes the form of a loan that promises a fixed return, financed by deposits with a similar structure. This resolves the problem of duplication of effort of monitoring lenders but shifts the problem to making it necessary for depositors to be sure that the bank is being run properly. Provided the bank's loan book is sufficiently diversified and the project returns from various loans are not highly correlated, depositors can be reasonably sure that the bank will not default in its obligations to them. Alternatively, depositors may require the owners of banks to put up more equity, so that with more of their capital at risk they would take more care with respect to the lending of depositors funds. These issues that ensure prudent behaviour by bankers is the subject of prudential regulation.

### *14.1.6* Functions of Financial Markets

As discussed by Levine,[5] the FUNCTIONS OF THE FINANCIAL SYSTEM can be classified as follows:

› Financial markets perform three main functions: allocate resources, monitor managers, and facilitate the trading, hedging, diversifying, and pooling of risks.

1. Allocating resources—The high information costs of evaluating firms, managers, and market conditions may keep capital from flowing to its highest value use. Given the fixed cost of acquiring and processing information, a financial intermediary that does it for a group of individuals can economize on the cost. This facilitates screening of projects and improves resource allocation.
2. Monitoring managers—Given verification costs, it is socially inefficient for outsiders to a firm to monitor it in all circumstances. It is thus efficient to have a contract whereby insider managers exercise control when the firm is solvent, and outsider bondholders exercise control in adverse circumstances when the firm is insolvent. If managers fail to make specified payments on bonds, which is easily verifiable, the bondholders have the right to seize the asset. Alternatively, a financial intermediary can mobilize the savings of many individuals and lend these. This intermediary is the delegated monitor of the borrower by individual savers, which reduces their duplication in monitoring. In turn, savers do not have to monitor the intermediary if it holds a diversified portfolio. Finally, stock markets can promote better corporate control by easing the takeover of poorly managed firms, which helps align managerial incentives with those of owners.
3. Facilitating the trading, hedging, diversifying, and pooling of risk—Financial markets and institutions arise to ease the trading, hedging, and pooling of risk. A HEDGE is a future transaction agreeing to purchase or sell a given quantity of an asset at a future date at a price agreed today in order to offset (hedge) the existing risk. For instance, if a firm takes a loan from a foreign source, then if the national currency, say INR, depreciates, the principal and interest payments to be paid out by the firm is devalued in local currency terms as the cash flows have to be converted into foreign currency to affect repayments. The foreign currency liabilities in domestic currency terms rise faster than the situation where the liability is measured at an unchanged exchange rate. To offset this risk, the firm can forward purchase the foreign currency liability amount and hedge the exchange rate risk.

› A hedge is a future transaction agreeing to purchase or sell a given quantity of an asset at a future date at a price agreed today in order to offset the existing risk.

› Diversification is the process of reducing risk by holding a number of financial securities in a portfolio.

DIVERSIFICATION refers to the ability of an agent to reduce risk by holding a number of financial securities in a portfolio. As the returns on different investments are not perfectly correlated, an institution can diversify away significant portfolio risk by exploiting the benefits of size. When a portfolio is allocated to a large number of stocks and bonds, then even if some financial asset experiences a decline in value, there is a good chance that other assets will rise in value with the gains offsetting the losses.

› Risk pooling is the process by which financial intermediaries protect savers from excessive risk through sharing the risks of investment projects.

RISK POOLING occurs because financial intermediaries that extend credit using funds raised from savers help savers to share the risks of investment projects. By sharing risks individual savers are protected from bearing excessive risk, and it enables savings to be directed to projects that are risky but are potentially high-return projects. Mobilizing savings—pooling risk—involves the agglomeration of capital from disparate savers for investment. Without access to multiple savers, many production processes would be constrained to economically inefficient scales. Further, mobilizing involves creating small-denomination financial instruments that enable savers to hold diversified portfolios and increase asset liquidity. However, there are transactions costs attached with collecting savings from different individuals, and information costs of convincing them to relinquish control of their savings, and financial arrangements address these in various ways while facilitating pooling.

Two types of risks, specifically important from the point of view of trading, hedging, and diversifying, are: liquidity and idiosyncratic risk.

› Liquidity is the ease and speed with which agents can convert assets into purchasing power.

LIQUIDITY is the ease and speed with which agents can convert assets into purchasing power at agreed prices. As an example, real estate is typically less liquid than equities, and equities traded on the New York Stock Exchange are more liquid than those traded on the Bombay Stock Exchange. With liquid capital, markets savers can hold assets (bonds, equity, or demand deposits) that they can sell easily and quickly if they seek access to their savings. Capital markets transform these liquid financial instruments into long-term capital investments in illiquid production processes.

› Idiosyncraticr isk is the risk that is specific to a financial instrument or project that cannot be diversified away. This comprises default risk, legal risk, and operational risk.

IDIOSYNCRATIC RISK is the risk specific to a specific financial instrument or project that cannot be diversified away. To assume this risk, a premium is usually charged for extending credit to such projects. Idiosyncratic risk, which is dealt with further in Section 14.3, comprises the following group of risks:

a. Default risk—The risk that a counterparty will default on its commitment, including the risk that any collateral will be inadequate if relied upon.
b. Legal risk—The risk that a contract may not be enforceable, usually applying to a default.
c. Operational risk—The risk of loss due to operational shortcomings—regulatory sanctions for inappropriate accounting or sales practice, the non-availability of specific human capital skills are examples of these.

In Section 14.2, we discuss the reform of one type of financial intermediation that is carried out in banking markets. These are financial markets where the transparency of loans is not high as confidentiality dominates the contract unlike in equity and bond markets. As a result of the idiosyncratic nature of their loans, banks' assets are illiquid (secondary trading in the loan is only attempted sometimes if the project does not perform and pay back), whereas their liabilities—deposits—are liquid. This makes banking markets different from capital markets. Table 14.3 sums up the financial markets in India.

› **Table 14.3**
Financial Markets in India

| Financial Intermediaries | |
|---|---|
| Banking Intermediaries | Public Sector Banks, Private Banks, Foreign Banks, Cooperative Banks |
| | Development Financial Institutions |
| Non-Bank: Financial Companies | Equipment Leasing, Hire Purchase Finance |
| | Mutual Benefit Companies, such as *nidhi* and *chit* Funds |
| | Mutual Funds (include UTI) |
| | Insurance Funds (include LIC) |
| | Pension Funds (include Employees Provident Fund Organization) |
| **Capital Markets** | |
| | Fixed Income |
| | Equity |
| | Derivatives |

### *14.1.7* Types of Banking Reforms

The main aim of the reforms in banking markets has been to relax credit and interest rate controls and remove restrictions on market entry and diversification. The stance of public policy towards banks may be classified according to their intent as being directed towards economic, prudential, or structural regulation.

ECONOMIC REGULATION deals with the control of credit allocation (reserve requirements, directed credit) and the pricing of credit (interest rate controls). PRUDENTIAL REGULATION refers to the set of laws and rules designed to minimize the risks banks assume and to ensure the safety and soundness of both individual institutions and the system as a whole. Examples include lending limits, minimum CAPITAL ADEQUACY guidelines, liquidity ratios. STRUCTURAL REGULATION deals with laws that specify the existence or absence of a separation between commercial and investment banks, between banks and insurance companies.

› Economic regulation in banks deals with control of credit allocation and with the pricing of credit.

› Prudential regulation refers to the set of laws or rules designed to minimize the risks to the financial system.

› Bank capital includes items such as paid-in capital, reserves, and retained profit, and is considered adequate if it is sufficient to provide for the possibilities of losses that may arise.

› The range of financial products that may be issued by a bank is the purview of structural regulation.

The first generation of *banking reforms in India* addressed the issue of economic deregulation and the second phase was an exercise in enhancing prudential regulation. STRUCTURAL REGULATION issues regarding whether banks should be allowed to deal in insurance products and the benefits/costs of universal banking is the third phase of the deregulation of the banking sector. A universal financial intermediary structure allows individual financial service organizations to offer a broad range of banking, insurance, securities, and other financial products. In most countries regulatory barriers prohibit the ability of a financial intermediary operating in one area of the financial service industry to expand its product set into other areas. In what follows we concentrate on the economic deregulation that has taken place since the Narasimham Committee report on the financial system. The aim of economic deregulation was to remove financial repression.

## 14.2 Financial Repression and Major Financial Reforms in India

A way used by governments in developing countries to finance expenditures in excess of tax revenues is to force private sector agents to buy government securities at rates below the market rates of interest. For instance, insurance companies and pension funds are often required to hold larger proportions of assets, in the form of government securities, than they would choose voluntarily. The largest captive buyers of government securities, however, in developing countries are commercial banks. By setting high-liquid asset ratios (known as the statutory liquidity ratio, SLR) and ensuring that government securities are the only eligible asset that satisfies this requirement, governments can borrow substantial amounts at below market rates of interest.

A second way the government finances its deficits is by setting high-reserve requirements. In this way the government can borrow from the banking system at a zero rate of interest.

Finally, governments in developing countries set ceilings on institutional interest rates to limit the competition from the private sector for loanable funds. Interest rate ceilings are imposed to check the competition to public sector fund-raising from the private sector. As private assets compete with public ones in individual portfolios and if private assets bear a market-determined interest rate, they would be preferred over public assets. One way to limit the appeal of these competing assets is by restricting the rates of interest they may legally bear.[6]

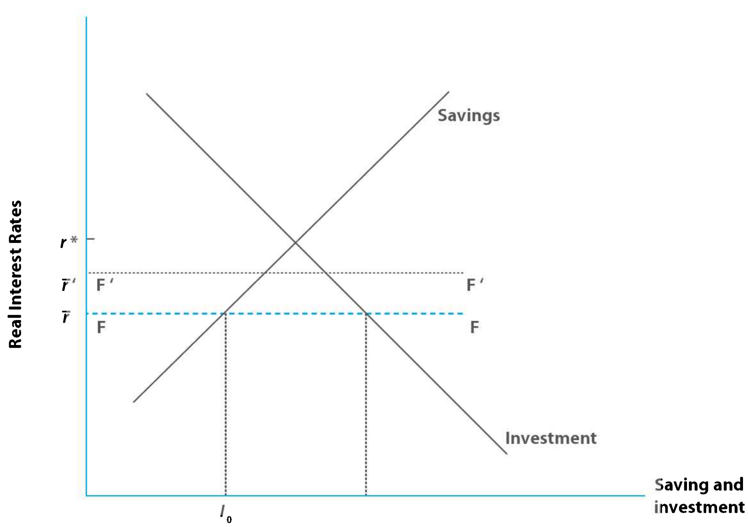
› Financial repression refers to the prevalence of credit controls, interest rate controls, entry barriers into banking, operational restrictions on banks, the predominance of state-owned banks, and restrictions on international financial transactions.

McKinnon[7] and Shaw[8] were the first to argue that FINANCIAL REPRESSION that involves distortions of interest rates reduces the size of the financial system and reduces the growth rate of the economy. The essence of their approach to financial repression is presented in Figure 14.1.

Saving is a positive function of the rate of interest—the rate of interest that they refer to is the real rate of interest. Investment as usual is a negative function of the real rate of interest as higher rates of interest reduce the profitability of investment projects and reduce the investment expenditure. Financial repression is represented by the line FF which represents the administratively fixed interest rate $\bar{r}$ that is below the equilibrium interest rate $r^*$. Actual investment is then LIMITED to $I_0$, the amount of saving forthcoming at the real rate of interest $\bar{r}$.

› Financial repression **limits** the economy's investment expenditure below the potential for investment in the economy.

Non-price rationing of loans must then occur. Credit is allocated not according to the expected productivity of the investment projects but according to transactions costs and perceived risk of default. As Fry[9] comments, "Quality of collateral, political pressures, 'name', loan size and covert benefits to loan officers may also influence allocation". The average efficiency of investment is also reduced as those who would have been dissuaded from applying for bank




**› Figure 14.1**
Financial Repression. Savings increases with the return on savings—the interest rate—and investment decreases with the cost of investment or the interest rate. In a financially repressed economy, the interest rate is administered at $\bar{r}$ and the actual investment at $I_0$ is limited by the amount of savings at this rate of interest.

loans at interest rates of $r^*$ would now find that their projects are profitable at lower rates of interest. Loan rate ceilings also deter financial institutions from taking any voluntary risk as risk premia cannot be charged when ceilings are binding and operative.

Interest rate ceilings interfere with the functioning of the economy. As low interest rates produce a bias against future consumption they reduce saving below what is possible. Then again, as borrowers are able to obtain the funds they require at low-loan rates, they would be induced to choose relatively capital-intensive projects. Finally, they deter banks from spending resources on loan assessments and the pool of potential borrowers will always contain low-yielding projects that would not have been presented for loan assessments if the higher market clearing interest rate had prevailed.

McKinnon and Shaw recommend that the policy prescription for the financially repressed economy is to raise INSTITUTIONAL RATES OF INTEREST and, if possible, to abolish interest rate ceilings altogether so that the average efficiency of investment goes up at the market determined rate of interest. Raising the interest rate to $\bar{r}'$, for instance—the financial repression line F′F′—increases saving and investment. It also deters firms from undertaking low-yielding investment projects that are no longer profitable at the higher interest rate.

› Financial repression theorists advocate the removal of interest rate ceilings and portfolio controls on banks to enable higher savings that gets directed to high yielding investments.

**MACROFOCUS 14.1**

### Financial Repression—Revenues and Liberalization

Financial repression is indicated by the prevalence of the following six policy dimensions:

1. Credit controls such as directed credit towards favoured sectors (agriculture) or industries, ceilings on credit towards other sectors, and excessively high-reserve requirements.
2. Interest rate controls with the government controlling interest rates directly.
3. Entry barriers in the banking sector such as licensing requirements, limits on the participation of foreign banks, and restrictions relating to bank specialization or the establishment of universal banks.
4. Operational restrictions such as on staffing, branching, and advertising.
5. Restrictions on privatization and extent of state-owned banks in the financial sector.
6. Restrictions on international financial transactions, such as on capital and current account convertibility, and the use of multiple exchange rates.

Giovannini and De Melo* argue that financial repression allows governments to finance themselves at artificially low interest rates. Taking the interest that governments pay in international capital markets to reflect the opportunity cost of funds, they measure the revenue directly accruing to government from financial repression as the difference between foreign and domestic interest rates times the stock of domestic government liabilities held outside the central bank. Their estimates of revenue from financial repression are reported in Table 1.

Financial intermediaries hold a large fraction of the government domestic debt and offset the artificially low rates on this asset by paying low rates on their liabilities such as savings and time deposits. With financial liberalization, domestic residents begin to access higher yielding portfolios, which require comparable increases in interest rates on government securities in order to reduce the risk of financial intermediaries.

Countries with highly repressed financial systems tend to stay that way for long periods of time. Reforms tend to be triggered by shocks or discrete events.** Balance of payments crises increase the likelihood of reforms whereas banking crises reduce their likelihood. Reforms are found to be more likely during the first year of an incumbent government in office and during recessions and high inflation. Also, left-wing governments are no less reform oriented than right-wing governments. The ideology of governments does not affect financial reforms.

**› Table 1**
Revenues from Financial Liberalization

| Country (time period) | Revenue from Financial Repression: % of GDP | Revenue from Financial Repression: % of Tax Revenue |
|---|---|---|
| Algeria (1974–1987) | 4.30 | 11.42 |
| Greece (1974–1985) | 2.53 | 7.76 |
| India (1980–1985) | 2.86 | 22.38 |
| Indonesia (1976–1986) | 0.00 | 0.00 |
| Malaysia (1974–1981) | 0.12 | 0.31 |
| Mexico (1984–1987) | 5.77 | 39.65 |
| Pakistan (1982–1983) | 3.23 | 20.50 |
| Sri Lanka (1981–1983) | 3.40 | 19.24 |
| Zimbabwe (1981–1986) | 5.50 | 19.13 |

*A. Giovannini and M. De Melo, "Government Revenue from Financial Repressions," *American Economic Review* 83, no. 4 (1993): 953–963.
**A. Abiad and A. Mody, "Financial Reform: What Shakes It? What Shapes It?" *American Economic Review* 95, no. 1 (2005): 66–88.

Since the advance of the financial repression thesis over the last two decades, development practitioners as well as academics have increasingly come to hold the view that finance matters and that financial systems that are market oriented have beneficial effects on the economy. Financial reform, which assigns a greater role to market forces and does away with controls and restrictions on the financial system, is considered to be a necessary ingredient for improving the efficiency of the financial sector as well as enabling it to better meet the growing and diverse needs of the real sector. In the real sector of the economy, controls were deemed to distort investment patterns and lead to inordinate delays and corruption. In the financial sector, controls in the form of regulated interest rates and directed credit programmes led to inefficiencies in credit allocation and adversely affected the viability of the banking system. Influenced by this thinking many countries have embarked on a wide ranging programme of reform and deregulation.

In the Indian context, reforms which began in the early 1980s were gradual till mid-1991 when there was a marked increase in the pace of implementation. These reforms were guided by the orthodox recommendation that real sector reforms should precede financial sector reform so that the future pattern of real returns is clearer. Prior to liberalization, the instruments of control in the financial sector included various interest rates on deposits, and since lending was fixed by the central bank, high-reserve requirements, quantitative credit allocations, concessional interest rates for specified sectors along with cross-subsidization, and restrictions on the scope of activities of particular financial institutions. Controlled interest rates resulted in very low real rates of interest in the financial sector due to inflation and were deemed to affect the deepening of the financial system and the savings rate adversely. Credit allocations and priority sector lending did deploy credit in accordance with social goals but eroded bank profitability and reduced the effectiveness of monetary policy as non-bank financial institutions were developed to get around regulations that were imposed on banks. The financial sector was dominated by nationalized banks which led to a loss in the autonomy of operations of these banks and political interference in the use of bank resources without regard to prudent commercial principles. The inefficiency in the deployment of credit and deteriorating bank profitability also went hand in hand with inadequate capitalization and insufficient provision for bad debts by the banks.[10]

Reforms of the financial sector were taken up by the Government of India as part of the structural reform programme that began in July 1991. The government appointed a committee under the chairmanship of M. Narasimham, a former governor of the Reserve Bank of India (RBI), in August 1991, to make recommendations as to how to go about reforming the financial system.[11] The committee recommended a gradual deregulation of interest rates and a reduction in the SLR to 25 per cent from the then existing level of 38.5 per cent, and higher interest rates to be paid by the RBI on cash reserves whose ratio was also to be brought down. The committee also made recommendations about allowing nationalized banks to raise equity from the capital market, entry into the market by private commercial banks, government borrowing at market-related interest rates, the implementation of a prudential regulatory framework, and recapitalization of banks over a period so as to attain the Bank of International Settlements Basle norms on capital adequacy standards. The reforms carried out by the government have been on the lines recommended by the Narasimham Committee.

So far, important financial sector reforms have included the following:

1. Moderation of the automatic monetization of the budget deficit (ad hocs have been phased out and a system of ways and means advances instituted) and a move towards market rates of interest on government borrowings. Banks in the financial-repression era were obliged to fund large fiscal deficits through the imposition of an SLR and when this proved insufficient the remainder was accommodated by the RBI through the medium of ad hoc Treasury bills that the government issued at will. A Treasury bill is the simplest form of borrowing. The government raises money by selling bills to banks, which buy the bills at a discount from the stated maturity value. At the bill's maturity, the bank receives from the government a payment equal to the face value of the bill. The difference between the purchase price and the ultimate maturity value constitutes the earnings of the bank. The RBI used to accept these bills and credit the central government account. Thus, there was automatic monetization of the deficit. A former Deputy Governor of the RBI stated in his Second Schumacher memorial lecture that when the definitive history of the financial sector reforms is written, this will be recognized as the *keystone* of the reforms.[12]
2. Reduction in pre-emptions through SLRs from an effective 37.4 per cent in March 1992. The SLR specifies the required holding of government securities (typically at sub-market rates of interest) as a fraction of the total assets as set by the central bank. It was brought down to 25 per cent in November 1997 and has remained unchanged since then. Despite this, right up to March 2005, the banks maintained the SLR at 38 per cent of net demand and time liabilities. In March 2006, this reduced to 31.3 per cent and in March 2007, to 28 per cent—still above the statutory requirements.

   Other fundamental reforms intended to strengthen the institutional structure for the secondary market to make the internal debt-management policy more effective include (a) setting up of the primary dealers system; (b) liquidity support by the RBI to mutual funds exclusively dedicated to government securities to the extent of 20 per cent of holding of government securities; (c) movement to mark to market system for valuation of government securities; and (d) the delivery versus payments system to ensure secure transactions.
3. Reduction in the cash reserve ratio (CRR) from 16.5 per cent in March 1992 to 7.5 per cent in late 2007. Cash reserves are the cash balances maintained with the RBI by commercial banks. As per the Reserve Bank of India Act of 1934 these cannot be less than 3 per cent of their net demand and time deposits. The CRR in the 1980s had resulted in the interest rate on cash reserve balances progressively rising to 10.5 per cent. Subsequently, the interest paid on CRR above 3 per cent and up to 5 per cent of net demand and time liabilities—known as eligible cash balances—was paid an interest rate of 3.5 per cent. The RBI has amended the RBI Act of 1934, and sought to not make interest payments on CRR balances. As the official gazette notifications on these provisions have not been made, the RBI continues to pay interest on eligible balances that are consistent with the monetary policy stance at relevant periods of time.
4. Introduction of the Basle norms on capital adequacy, asset classification and provisioning, and income recognition. The primary means of protection against the risk of insolvency and failure of a financial institution is its capital. This is the extent of a financial institution's owners' equity stake in the financial institution as measured by the difference between the market value of its assets and liabilities or what is sometimes called its net worth.[13] Capital adequacy norms are intended to protect a financial institution and

› Financial reforms also aimed to reduce pre-emption of bank assets through a high-statutory liquidity ratio and cash reserve ratio, and to introduce norms of income recognition, asset classification, provisioning, and capital adequacy.

allow it to absorb unanticipated losses. The higher the amount of capital the more protected are uninsured depositors in the event of liquidation of the institution and the more it inspires confidence about the financial institutions' ability to continue as a going concern.

Note that an asset is regarded as non-performing when it ceases to generate income for the bank. A non-performing asset (NPA) is defined as a credit facility in respect of which the interest and/or instalment of the principal has remained past due for a specified period of time.

Banks are required to classify NPAs further into the following three categories based on the period for which the asset has remained non-performing and the realizability of the dues:

a. Sub-standard Assets—This refers to cases where the current net worth of a borrower is not enough to ensure recovery of the dues to the bank in full. An NPA for a period less than a year is characterized as a sub-standard asset.
b. Doubtful Assets—A loan classified as doubtful has all the weaknesses inherent in assets that were classified as sub-standard, with the added characteristic that the weaknesses make collection or liquidation full, highly questionable and improbable. An asset is classified as doubtful if it remained in the sub-standard category for 12 months.
c. Loss Assets—A loss asset is one where the loss has been identified by the bank but the amount has not been written off wholly. Such an asset is considered uncollectible and of such little value that its continuance as a bankable asset is not warranted although there may be some salvage or recovery value.

Income-recognition norms are norms that recognize income on assets only if it is received within a specified time period after it is past due—the international norm is 90 days. The idea is to keep the confidence of the public from eroding due to the balance sheet impact. The policy of income recognition is based on the record of recovery. Internationally income from NPAs is not recognized on an accrual basis but is booked as income only when it is actually received. Internationally banks do not charge and take to the income account interest on any NPA.

Provisioning norms refer to provisions made by bank managements on the expectations of losses on the current loan portfolio. It is the recognition of a bank management of bad loans expected during a time period that causes them to make provisions for these in the balance sheet.

The NPAs of public sector banks were approximately 25 per cent of the total advances in March 1994 and are currently approximately 3.5 per cent of the total advances. The government has contributed over INR 125 billion by way of capital contribution towards bank restructuring.

A couple of years after the introduction of the financial sector reforms, it is increasingly being perceived that the response of the banking sector to these reforms has been contaminated by the impact of the adjustment on the portfolio of banks and the extent of the government's intervention in the financial market. However, despite this perception most analysis proceeds as if the banking sector is a black box. The banking sector is looked on as being in the business of intermediation and once investors are convinced of the viability of a project, banks just respond to this by placing the desired amount of finance with investors. As a consequence the banking sector is characterized by the absence of agents who make decisions on the basis of their preferences and given the incentives prevalent in the environment in which they operate. We need to try open the black box, that is the banking sector, so as to be able to sensibly analyse the financial sector reforms and that is what we now proceed to do.

# 14.3 The Supply of Loans and the Demand for Deposits

The supply of credit by financial intermediaries depends on the cost of raising funds (deposits) and the risks involved in the supply of credit services. The TWO TYPES OF FINANCIAL RISKS that a financial intermediary faces are borrower-specific risk or idiosyncratic risk and systemic risk. We examine the implications of these for the supply of loans by banks.

› Financial intermediaries face idiosyncratic and systemic risk while supplying credit services.

## 14.3.1 Risk and the Supply of Credit

SYSTEMIC RISK is that risk which is not diversifiable within a given domestic jurisdiction.[14] These risks typically stem from three sources:

› Systemic risk stems from macroeconomic uncertainty, weakness in the contractual and informational environment, and geographical limitations of small jurisdictions.

1. Macroeconomic uncertainty—This results in volatility in the rate of inflation, the real interest rate, and the real exchange rate. An increase in macroeconomic volatility as measured by the increased fluctuation in inflation and GDP growth makes business planning more difficult, and increases the resources that are devoted to managing inflation risks. When output is highly variable, firms and households are confronted with uncertainty about the future course of the economy and spend resources in back-up plans.
2. Weakness in the contractual and informational environment—It arises from poorly defined creditor rights that are difficult to enforce, deficient accounting and disclosure practices, and lack of information about credit histories due to the absence of a credit bureau.
3. Geographical limitations—These are important in small jurisdictions, which are prone to natural disasters such as floods or earthquakes.

SYSTEMIC RISKS impede the supply of credit as it raises the default probability on loans and the loss given default for all loan contracts written in a given jurisdiction. This results in a higher cost of funds and a higher floor to the interest rate required by a creditor before a loan is sanctioned.

› Systemic risk is the default risk associated with general economy-wide or macroeconomic conditions that affect all borrowers and lenders.

IDIOSYNCRATIC RISKS are specific to individual borrowers and/or projects that are not correlated with systemic risk. These result due to agency problems as well as due to the limits on the diversification of risk and are priced in loan contracts as a spread over the interest rate floor set by the extent of systemic risk. Agency problems arise due to information asymmetry whereby the debtor is privy to relevant information about the project that a creditor may not be able to secure without incurring a prohibitive cost. This is a source of credit risk arising from adverse selection.

› Idiosyncratic risks are risks arising from asymmetric information about the default risk associated with a specific borrower.

Under adverse selection, the creditor does not have sufficient information to sort out good borrowers, who are less likely to default, from the bad borrowers. This is an agency-related idiosyncratic risk, and a creditor would require compensation for it in terms of a risk premium charged on the loan. However, raising the interest rate charged to cover this attracts riskier borrowers, who are more likely to default in case of a bad outcome where they are unable to meet their debit obligations.

Individual borrowers differ in their probabilities of repayment. Consider the situation where there are two types of borrowers. A good borrower repays the loan with probability $p_g$ and a bad borrower repays with probability $p_b < p_g$. If lenders can observe borrower types each borrower will be charged a different interest rate to reflect the differing repayment probabilities. If the marginal cost of raising funds for a bank is $i_{mc}$, then, in order to recover this amount a risk-neutral lender in a competitive banking market would lend to good borrowers at the interest rate $i_g$ such that the expected return from lending equals the opportunity cost of funds. That means

$$p_g i_g = i_{mc}$$

or,
$$i_g = \frac{i_{mc}}{p_g}$$

Similarly, if lender is able to identify a bad borrower, they will lend to borrower at an interest rate given by

$$i_b = \frac{i_{mc}}{p_b}$$

As $p_b < p_g$, the interest rate charged on the bad borrower is greater than that at which the good borrower may borrow

$$\frac{i_{mc}}{p_b} > \frac{i_{mc}}{p_g}$$

In the presence of asymmetric information, the lender will not be in a position to observe the borrower's type—good or bad. The lender has to, therefore, perforce charge a uniform rate of interest $r$ on the given loans. The lender, however, knows general things about the population of borrowers—the distribution of borrower types is known. It is known that a fraction $\theta$ of all borrowers are good borrowers. Then, a lender will charge an interest rate that in a competitive loan market will allow the lender to break even. The interest rate charged $i$ will then be given by

$$\theta p_g i + (1 - \theta) p_b i = i_{mc}$$

or,
$$i = \frac{i_{mc}}{\theta p_g + (1 - \theta) p_b}$$

At this interest rate on loans, the lender expects to earn the required rate of return $i_{mc}$. Note, however, that this interest rate lies between the interest rate the lender would have charged good and bad borrowers if they were in a position to identify borrower types. Since the lender is charging an average interest rate this rate must lie between the rate charged to a good borrower when identified as one and the rate charged to a bad borrower when the lender is able to identify such a borrower. To see this first note that the uniform interest rate due to asymmetric information is greater than the interest rate a good borrower would have been charged if it were possible to identify borrower types,

$$i_g = \frac{i_{mc}}{p_g} < \frac{i_{mc}}{\theta p_g + (1 - \theta) p_b} = i$$

That the above inequality holds can be quickly established by cross multiplying the denominators

$$\left[\theta p_g + (1 - \theta) p_b\right] i_{mc} < p_g i_{mc}$$

Cancelling the $i_{mc}$ term from both sides and taking terms in $p_g$ to the right-hand side, we obtain:

$$(1 - \theta) p_b < (1 - \theta) p_g$$

or,
$$p_b < p_g$$

which is true by assumption. Hence, the inequality holds. Moreover, in a similar manner, we can establish that the uniform interest rate due to asymmetric information is less than the interest rate a bad borrower would have been charged if it were possible to identify borrower types,

$$i < (i_{mc}/p_b)$$

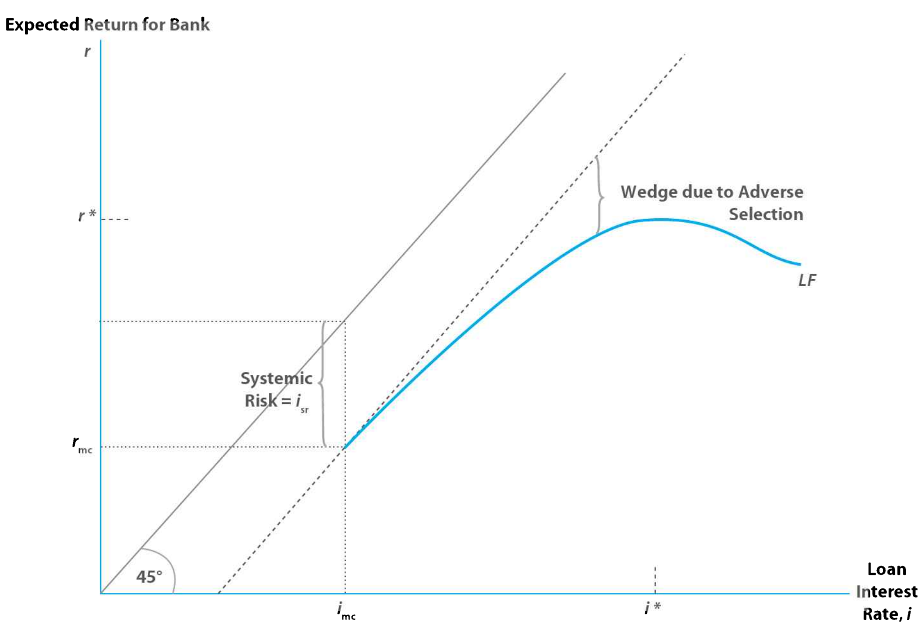
Hence, $$\frac{i_{mc}}{p_g} = i_g < i < i_b = \frac{i_{mc}}{p_b}$$

The uniform interest rate, therefore, overcharges good borrowers ($i_g < i$) and undercharges bad borrowers ($i < i_b$). The lender will then deter good borrowers from taking loans and will attract bad borrowers. As a result, the lender's pool of borrowers will not be drawn randomly from the population distribution of borrowers and with more than a fraction $(1 - \theta)$ of bad borrowers obtaining loans the return to the lender will turn out to be less than $i$. As the rate of interest charged on loans, $i$, is raised, good borrowers will drop out of the market for loans and bad borrowers will increasingly stay in.

As the RATE OF INTEREST charged on loans rises, the repayments received from those who repay loans increase. However, higher interest rates attract bad borrowers with a low probability of repayment—higher default rate—and this reduces the repayments received by the lender. At some rate of interest the default rate will be sufficiently high to outweigh the receipts from those charged high interest, and the expected return to the lender will begin to decline. A lender in an asymmetric information situation thus faces a situation where the expected returns from lending eventually decline as the interest charged on loans increases as bad borrowers with a higher default rate accumulate on the loan book of the lender. A lender would attempt to mitigate this adverse selection problem of attracting riskier borrowers by non-price screening devices such as requiring collateral, or, simply denying credit to applicants they are not able to screen effectively.

› As the interest rate charged on loans is raised, the financial intermediary attracts bad borrowers. At some threshold rate of interest, this results in higher default rates that offset repayments due to the bank and causes the expected return to the bank to decline. The loan return frontier for a bank is thus an inverted U–shape.

We illustrate the way in which systemic and idiosyncratic risk influence the cost and availability of credit supply in Figure 14.2. This figure measures the expected return to the lender, $r$, on the vertical axis and the interest rate at which funds are lent, $i$, on the horizontal axis. If there were no risk prevalent and transactions costs were negligible, we would have a situation where $r = i$, as represented by the 45-degree line from the origin. Due to the presence of risks a wedge is driven between the interest rate charged by the creditor and the expected return.




› **Figure 14.2**
The Credit Supply in Presence of Systemic and Idiosyncratic Risks. Without risk the expected return from lending equals the interest rate charged along the 45-degree line. The bank acquires funds at a marginal cost of $i_{mc}$. Subtracting from this the risk premium associated with the systemic risk that cannot be reduced through diversification gives the minimum expected return to lending or $r_{mc}$. From this floor, as the interest rate on lending is raised, the returns to the bank increases but good borrowers with a high probability of repayment drop out, and the bank attracts more and more bad borrowers. This raises the default rate on loans, which results in the expected return to the bank to decline after $i$*.

Let the marginal cost of acquiring funds be $i_{mc}$[15] for the bank lender. From this we deduct the country-level systemic risk premium $i_{sr}$ representing the premium accounting for deficiencies in contractual and informational frameworks and macroeconomic uncertainty, to get the floor to the expected return to the bank when lending $r_{mc} = i_{mc} - i_{sr}$. Next we introduce agency issues. This results in a nonlinear wedge between the 45-degree dashed line and the inverted U-shaped curve labelled *LF*. That implies that as the lending interest rate *i* rises, the expected return to the bank rises with repayments on loans rising with the interest rate charged. However, beyond a threshold given by ($i^*$, $r^*$), the likelihood of default rises and the expected returns begin to decrease.

### 14.3.2 The Determination of Banks' Asset Portfolio

The asset portfolio of the bank involves making a choice between different types of assets, one of which is loans that are subject to idiosyncratic and systemic risks.

We consider a simplified banking system where the liabilities of a bank are its deposits, *D*, and its assets comprise required reserves, *R*, advances in the form of loans, *A*, and government paper in the form of T-bills, *T* (see Table 14.4). If $\tau$ is the required reserve ratio, then, $R = \tau D$. Hence, we can write $A + T + R = D$ as

$$A + T = D(1 - \tau)$$

or,

$$D = (A + T)(1 - \tau)^{-1}$$

**› Table 14.4**
A Simplified Bank Balance Sheet

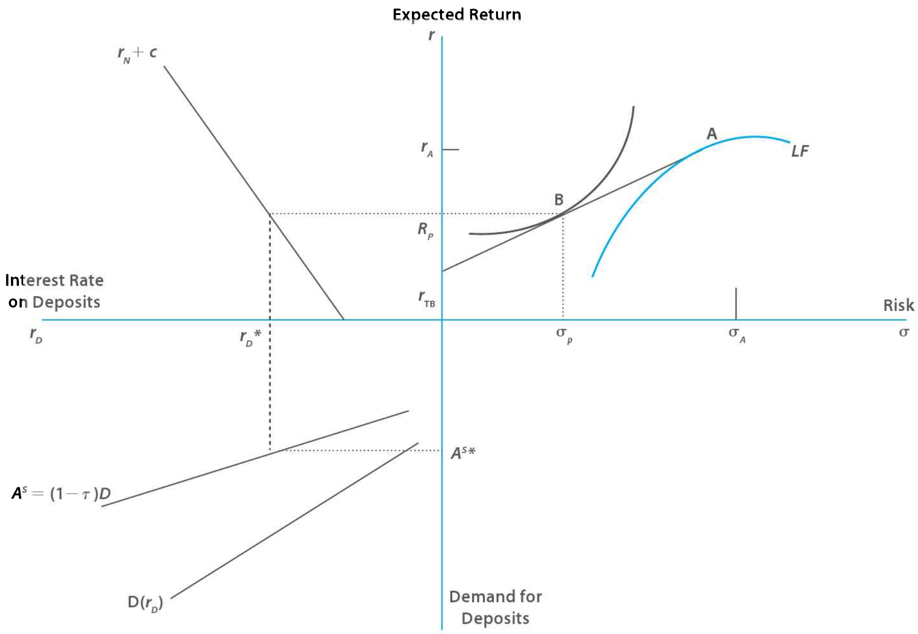
| Assets | Liabilities |
|---|---|
| Reserves *R* | Deposits *D* |
| Loans/Advances *A* | |
| Government Securities *T* | |

Thus in order to raise one INR of finance for lending or putting into T-bills, the bank must raise INR $(1 - \tau)^{-1}$ in the money market. After raising this amount it pays INR $r_D(1 - \tau)^{-1}$ to depositors, where $r_D$ is the interest on deposits. It puts a fraction $\tau$ of the amount it raises, that is INR $\tau(1 - \tau)^{-1}$ into reserves and earns an interest of INR $r_R\,\tau(1 - \tau)^{-1}$ where $r_R$ is the interest rate on reserves. Hence, the net marginal cost of accepting deposits is $r_N = r_D(1 - \tau)^{-1} - r_R\tau(1 - \tau)^{-1}$ or, $r_N = (r_D - r_R\tau)/(1 - \tau)$.

Apart from the cost of accepting deposits, the bank faces a cost of intermediation, *c*. The cost of intermediation of banks is a function of the high-establishment costs and low productivity. By international standards our banks are overmanned and technologically antiquated. We take the cost of intermediation, *c*, to be given.

› After accounting for the required reserves, which the bank must maintain as a ratio of deposits, the bank is free to allocate its portfolio amongst riskless government securities that give a return $r_{TB}$ and loans that are risky.

After ACCOUNTING FOR RESERVES, bankers face a choice of holding risky loans or riskless treasury bills that pay a return $r_{TB}$ without fail. Following Stiglitz and Weiss[16] adverse selection occurs in the credit market, as the rate of interest rises the quality of borrowers diminishes and the effective returns to banks falls. Recall that in Figure 14.2 the expected return on loans *r* differed from the interest charged on loans, *i*, due to the presence of the idiosyncratic risk associated with adverse selection. The higher the interest rate *i* charged, the higher was the risk of default due to good borrowers being driven away from the market. A rise in the interest rate charged is then equivalent to a rise in risk associated with a loan. Hence, we may plot the loan return frontier (LF) of banks with expected returns




**› Figure 14.3**
Banks' Equilibrium Asset Choice and Deposits. The loan return frontier, LF, gives the expected return to the bank from lending. This eventually declines due to asymmetric information. The other asset that banks hold are T-bills that being sovereign instruments give a return $r_{TB}$ without any risk. The line $r_{TB}$ A depicts the various portfolios of T-bills and loans the bank can hold. The bank selects point B where its subjective trade-off between risk and return, as given by the slope of the upward-sloping indifference curve, equals the slope of the portfolio line. The expected return on this portfolio is $R_p$ and this equals the net cost of accepting deposits and intermediation given by $r_N + c$. The corresponding equilibrium interest rate that a competitive bank will offer on deposits is given by $r_D^*$, which results in deposits to the bank D($r_D^*$), and a supply of investible resources with the bank of $A^{s*}$.

on loans, $r$, and the risk associated with loan returns, $\sigma$, on the two axes as depicted in Figure 14.3. Bankers are assumed to be risk averse and their indifference curves are upward sloping in ($r$, $\sigma$) space. This means that bankers are willing to assume higher risk provided it is accompanied by a higher return.

The banks' asset choice is to select a combination of risky loans from the loan return frontier and the riskless T-bill asset issued by the government, which earns a return $r_{TB}$ without fail as it has sovereign backing. The straight line drawn as a tangent from the point $r_{TB}$ to the loan return frontier, LF, represents combinations of such EFFICIENT INVESTMENT PORTFOLIOS. The efficient investment portfolio line given by $r_{TB}$ A represents linear combinations of risk and the expected returns of the risky loan asset and the riskless T-bill. The fraction of bank assets going to loans and T-bills is determined at B, the point of tangency between an indifference curve and the efficient investment portfolio line $r_{TB}$ A. The investment rate on loans is given by point $r_A$. The expected return on the portfolio at POINT B when a proportion $x$ of the portfolio is held in the riskless T-bill asset, which earns a rate of return $r_{TB}$, and a proportion $(1 - x)$ is held in the form of risky loans that give a return $r_A$, is given by $R_p = x r_{TB} + (1 - x)\, r_A$.

Note that $x$ is the ratio of the distance BA to $r_{TB}$ A and $(1 - x)$ is the ratio of the distance $r_{TB}$B to $r_{TB}$ A. The risk of this portfolio[17] is given by $\sigma_p$.

In equilibrium the banking sector will have the expected return on the portfolio combination of T-bills and advances as loans set equal to the net cost of accepting deposits and the costs of intermediation. Thus, in equilibrium, $r_N + c = R_p$. In Figure 14.3, this determines the equilibrium interest rate on deposits given by the point $r_D^*$. The demand for deposits, D($r_D$), is an increasing function of the rate of interest on deposits as drawn in the figure, and the supply of investible resources with the bank is then $A^s = (1 - \tau)\, D(r_D)$. Given the equilibrium return on deposits $r_D^*$ the supply of investible resources with the bank is $A^{s*}$. The amount invested by the bank in government paper is $xA^{s*}$ and the amount invested in risky loans is $(1 - x)A^{s*}$, where the fraction $x$ is determined at the point B in the diagram.[18]

› The efficient investment portfolio line is a straight line beginning at point $r_{TB}$ and is a tangent to the loan return frontier.

› Risk-averse bankers select a **point** on the efficient investment-portfolio line where an indifference curve representing points of indifference between the risk and the return is tangential to the efficient investment portfolio line.

## 14.4 Banks' Response to Financial Liberalization

This sketch of banking sector activity allows us to look at the impact of the financial sector reforms.[19] One major plank of the reforms, as was noted earlier, was that the interest rate on government paper was artificially repressed prior to the reform and subsequently there was a move toward market rates of interest. Thus, $r_{TB}$ the interest on government paper goes up to say $r^*_{TB}$. As shown in Figure 14.4, the efficient investment portfolio line also moves up and the banking sector equilibrium moves to $B^*$.

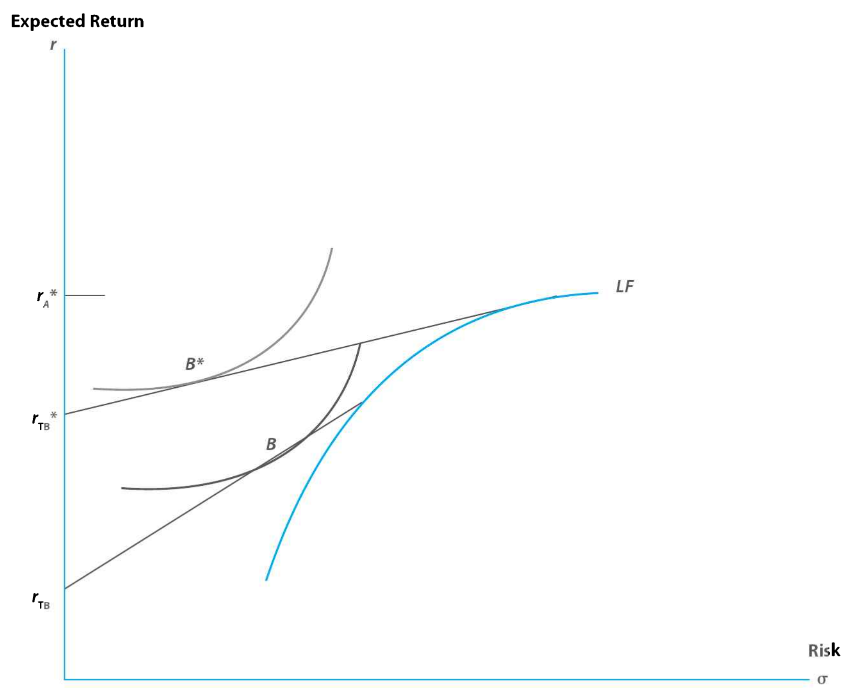
› In response to financial reforms where the interest rate on government securities becomes market determined, risk-averse banks prefer to hold a larger proportion of government securities in their portfolios and to substitute away from investing in loans.

The impact of the move towards MARKET DETERMINATION OF INTEREST RATES on government paper is that the interest rate on loans rise to $r^*_A$. If we trace through the effects (not shown in the diagram) we can see that the interest on deposits, $r_D$, also rises, and the supply of investible resources, $A^s$, also goes up. However, as can be seen in the diagram, the banks at $B^*$ choose to hold a larger proportion of the riskless asset and the proportion of investible resources given as advances falls despite more liquidity being available. There is a resulting combination of high-interest rates and lesser advances being made by the banks—credit rationing has exacerbated.

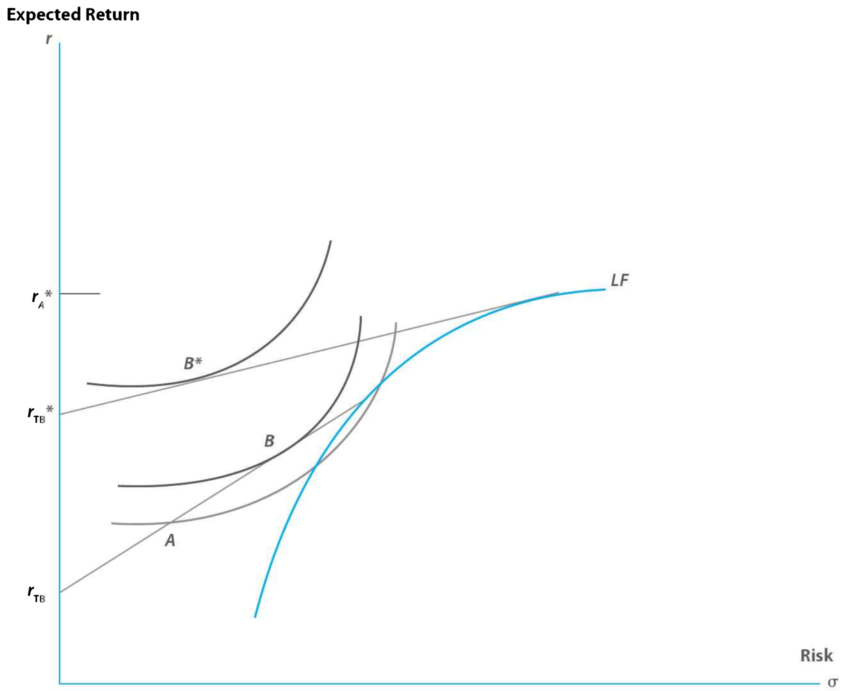
A second area of reform was that portfolio controls in the form of high SLR existed prior to the reforms and banks were forced to hold a high proportion of government paper prior to the reforms.

Let the portfolio controls be such that banks are forced to operate at point A in Figure 14.5.

Prior to the reforms there were interest rate ceilings on lending which implies an efficient portfolio line, such as that passing through points A and B, which is not tangential to the loan return frontier. Let us say that the financial sector reform removes control over portfolios and banks move from point A




› **Figure 14.4**
Interest Rate Liberalization on Government Securities. **The removal of financial repression resulted in market rates of interest on T-bills and the interest rate increased to $r^*_{TB}$. Equilibrium asset allocation between T-bills and loans for banks now takes place at $B^*$ where the indifference curve is tangential to the new portfolio line. The higher return on government securities attracts banks towards holding more T-bills than before and giving out less loans.**




**› Figure 14.5**
Removal of Portfolio Controls. **Under financial repression the interest rate on government securities was $r_{TB}$. Control over bank portfolios required banks to hold more government securities, as depicted by point A, than optimal (point B). The removal of portfolio restrictions induces banks to reduce their holding of government securities and give out more loans, as represented by point B. However, the increase in the return on government securities as a result of liberalization to $r^*_{TB}$ attracted banks to hold high- return and low-risk T-bills as given by point *B**. The net result was less loans in their portfolio of assets.**

to point B, thereby implying a GREATER AMOUNT OF INVESTIBLE RESOURCES available for lending as advances with the banks. However, this is not the way the story gets completed as the government at the same time has liberalized interest rates on government debt, which resulted in higher rates of interest on government paper such as $r^*_{TB}$. As a result, the banks optimal position is now at point $B^*$ and despite the cuts in SLR they disburse fewer loans and hold government paper instead.

› Banks tend to increase the proportion of loans in their portfolio as portfolio controls requiring banks to hold a larger fractions of government securities are withdrawn. However, the liberalization of interest rates on government securities makes loans less attractive and banks reduce their exposure to loans.

It is well known that in India banks have been holding SLR eligible securities well in excess of their requirements. This is due to the endeavour by the government to move towards market rates of interest on government debt. The thinking seems to have been that with the additional investible resources in the hands of the banks as a result of the removal of portfolio controls, the incremental credit-deposit ratios of banks would be extremely high and that banks have weak credit assessment capacities to handle such an expansion. The shift to market determination of government security rates would provide high yields on debt besides the advantage of a zero-risk weight assigned to government securities in assessing capital adequacy. At the margin this would lead to a shift in bank portfolios from advances to investment in government paper. The official thinking seems to have been concerned with preventing banks from a weakening of their capital position. It had the impact, however, of resulting in a regime of high-interest rates and credit rationing.

The Chakravarty Committee[20] had advocated a cautious liberalization keeping very much to the framework of an administered rate system.[21] The primary purpose of the increase in yields was to make government securities sufficiently attractive to induce voluntary holding of the government debt by the public at large, to reduce the monetization of the public debt, while at the same time enhancing bank profitability. The Narasimham Committee[22] noted that interest rates on government debt had been progressively increased but that they were still not at the levels to attract voluntary subscribers. It sounded caution about deregulation, however, until the health of financial intermediaries was restored and macroeconomic stability prevailed.

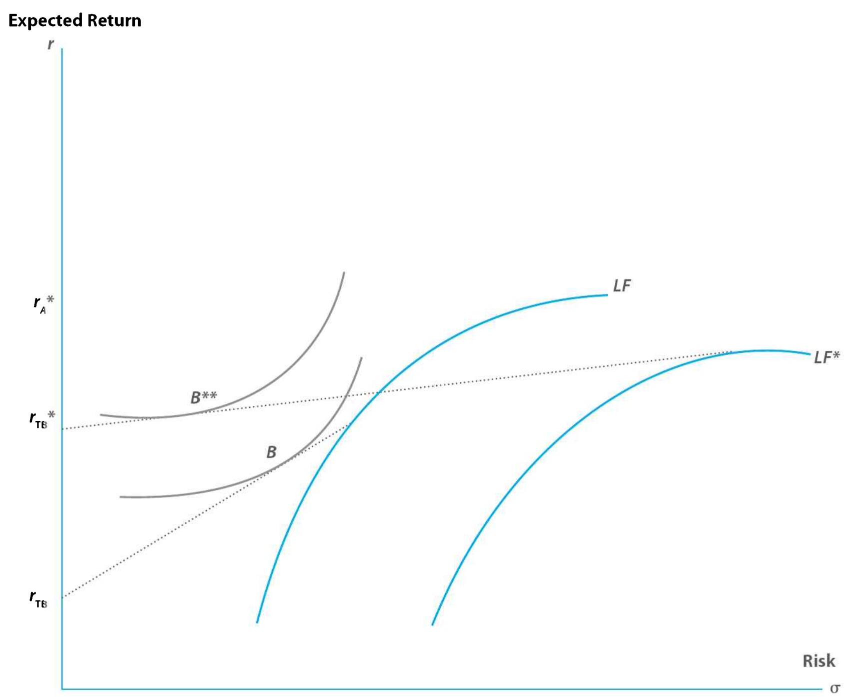
In short, mechanisms to strengthen supervision such that the quality of credit risks is adequately assessed should have been put in to place so that with a removal of portfolio controls the incremental credit-deposit ratio increased and banks made advances instead of holding government paper. A fast move towards market determination of interest rates has not been beneficial for adequate and cheap credit availability.

Prior to the reforms, bank incentive systems were geared to follow government instructions on credit allocation causing bankers to underinvest in risk assessment and monitoring skills. Financial reforms entail a portfolio shift of banks away from forced holding of government securities and directed credit. The less control that banks had over their assets prior to the reforms, the greater will be the learning problems associated with the reforms. With the reforms banks will increasingly be held responsible for any likely loan losses. Given the BANKS' STOCK OF HUMAN AND MANAGERIAL CAPITAL that has been shaped by the pre-reforms environment, bankers will behave as if they are facing a riskier loan portfolio. This can be viewed as a rightward shift of the loan return frontier, LF, to LF* in Figure 14.6.

› Prior to liberalization, bankers underinvested in credit-risk assessment and monitoring skills. Only when **bankers' skills and human capital** is upgraded to cope with the demands of liberalization will the credit-deposit ratios of banks improve.

This results in the equilibrium moving to $B^{**}$ where the banks hold a greater amount of riskless assets than before. Also, it can be traced out in the extended diagram (Figure 14.3) that as the interest rate on loans goes up, $R_p$ falls, and so does the supply of investible funds for loans from the banking sector.

Macroeconomic uncertainty and instability also affect bank behaviour. Suppose there is an economic downturn which means that aggregate income reduces. The profitability of firms gets adversely affected as a result and banks view all loans as being riskier than before—that is to say, the loan return frontier, LF, shifts to the right. Similarly, in an era of deflation the real value of debt goes up and debt servicing corners a higher proportion of debtors' income. The greater burden of debt results in a decline in the net market value of equities




**› Figure 14.6**
Increase in Perceived Riskiness. **Banks, which had underinvested in risk assessment and monitoring of loans, were given increasing control over their asset holding decisions as financial liberalization took place. This caused banks to perceive the risk associated with lending to have gone up and the loan return frontier, *LF*, shifts to the right to *LF**. As interest rates on T-bills become market determined at $r^*_{TB}$, the banks reacted to the increased perception of risk by holding more government securities, as at *B***.**

and hence a decline in the net worth of business, which precipitates defaults on loans and bankruptcies, threatening the solvency and liquidity of the financial system.[23,24] The banks respond to this by viewing all loans as becoming riskier—the *LF* shifts to the right—and they cut access to credit as the net worth of firms declines.

Weak bank supervision and lending to interrelated entities has often been witnessed concomitant with repressed financial systems. Bank solvency in such a situation reflects the weak capital bases, poor loan portfolio, and inefficient management that is a part of a repressed financial system. When financial liberalization occurs there are inadequate knowledge and skills and institutional capacities among financial institutions to assess and control risks. A minimal system of bank supervision and regulation in such a situation will go a long way in improving credit allocation processes.

An immediate focus on capital adequacy, lending to interrelated enterprises, large loan concentrations, loan classification and provisioning procedures when liberalization is attempted can help reduce the risk of financial crisis. It is important that these microeconomic problems of the institutional weaknesses of the financial system be dealt with prior to liberalization of interest rates on government paper so that with a reduction in reserve requirements credit expansion can be diverted to projects that are not too risky. Also, the fiscal deficit will not be under pressure due to higher interest rates on government debt, transfer to the budget of the cost of subsidized lending, and the costs of restructuring the banking system. Once financial markets are well-regulated and in a position to function effectively and to assess risks, they will be able to improve overall savings mobilization and so mitigate the adverse consequences of interest rate liberalization.

**MACROFOCUS 14.2**

## Liberalization, Growth, and Collateral Benefits

There are two divergent opinions about financial liberalization. One view is that it strengthens financial development and contributes to the long-term growth. The other view emphasizes that liberalization induces excessive risk taking and increases macroeconomic volatility and the frequency of crises.

As is to be expected the crisis view emphasizes the gravity of output declines during financial crises but ignores the growth benefits during tranquil times. The growth view, on the other hand, highlights only the average growth effect across the booms and busts generated by financial liberalization.

Ranciere, Tornell, and Westermann* find for 60 countries over the period 1980–2002 that financial liberalization has a direct positive effect on per capita GDP growth. Financial liberlization also increases the probability that a banking crisis or a currency crisis will occur. However, the direct positive effect of financial liberalization on growth is found to far outweigh its indirect effect through a higher propensity for crises. The contrasting experiences of Thailand and India illustrate this well. In India (a partially liberalized economy) GDP per capita grew by only 99 per cent between 1980 and 2001, whereas Thailand's GDP per capita grew by 148 per cent despite experiencing a major crisis.

More importantly, financial liberalization which may not increase private savings or have a significant impact on the quantity of investment has nevertheless been found to have a positive effect on growth through its effect on the efficiency with which investment is allocated across firms and sectors, and not mainly through the quantity of resources mobilized. Galindo, Schiantarelli and Weiss** find that liberalization increases the share of investment going to firms with a higher marginal return to capital.

As stressed by Kose, Prasad, Rogoff and Wei,† far more important than the direct growth effects of access to more capital is how capital flows generate collateral benefits. The development of the financial sector through better quality of financial services and greater efficiency of financial intermediation can enhance total factor productivity growth. Moreover, financial reforms have induced countries to improve their government and corporate governance structures and well-regulated and well-supervised financial systems deliver the collateral benefit of enhancing the efficiency of financial intermediation.

*R. Ranciere, A. Tornell, and F. Westermann "Decomposing the Effects of Financial Liberalization: Crises vs. Growth," *Journal of Banking & Finance* 30 (2006): 3331–3348.
** A. Galindo, F. Schiantarelli, and A. Weiss, "Does Financial Liberalization Improve the Allocation of Investment? Micro-evidence from Developing Countries," *Journal of Development Economics* 83 (2007): 562–587.
† M. A. Kose, E. Prasad, K. Rogoff, and Shang-Jen Wei, "Financial Globalization: A Reappraisal," IMF, Working Paper No. 06/189, 2007.

## 14.5 Human Capital: A Keystone of Financial Reforms

Banking by its nature is an information (intensive) and human-capital-intensive industry. One of the surrogate measures of output of a bank is the value of deposits and loans that the bank is able to generate. The efficiency of the employees of a bank on this notion can be summarized by the ratio of the deposits plus advances (the turnover) per employee.

On this count, the trends in the efficiency of the various types of banks in India can be gauged. While the turnover per employee increased 3.1 times from INR 3.85 million per employee in 1990–1991 to INR 12.21 million per employee in 1999–2000 for the public sector State Bank of India group of banks, it increased 3.1 times for the nationalized banks, 9.3 times for the private banks, and 4.4 times for the foreign banks during the decade. In 1990–1991, the turnover per employee for the private and foreign banks combined was 1.28 times that in the public sector banks. By 1994–1995, the multiple was 2.1 times and by 1999–2000 it was 2.56 times; the turnover per employee in the private and foreign banks doubled relative to the public sector banks during the decade and it has remained at a multiple of 2.07 times in 2004–2005. It is true that the public sector banks have a large presence in the rural and semi-urban areas where the foreign banks, for instance, do not even have a branch. However, even when we compare the turnover per employee amongst banks in urban/metropolitan areas, foreign and private banks are still doing more than twice better than the public sector banks in their ability to mobilize deposits and disburse advances per employee.[25]

It would be expected that deregulation and increased competition acts as a disciplining force on the operating efficiency of nationalized banks. The adopted approach of reducing portfolio controls and encouraging the growth of the private and foreign banks was in order to create a more competitive environment with the expectation that this would lead to efficiency improvements. In the current scenario, while banks have considerable discretion over setting of interest rates and on deciding on the product mix, the public sector banks that dominate the banking industry do not, however, take full responsibility for their decisions. If they are undercapitalized and run into losses, they face no risk of bankruptcy as the government has given an implicit insurance that the stability of the financial system will be maintained—the government's guarantee has in effect imported confidence that less care needs to be taken about the non-marketable, idiosyncratic loans made by banks.[26] The outcome of this is that public sector banks have come to expect support in financial difficulties, which reduces their sensitivity to efficiency and shields them from the disciplining force of competition from the non-public sector banks. As a result, even if operational freedom is fully with the public sector banks, and even if the government's stake in the capital of the banks falls to 33 per cent, the management may not operate as efficiently as they possibly could. In India, the situation was reinforced in the Budget Speech of the Finance Minister in 2000 when he stated that the "public nature of the nationalized banks will continue even if the government stake drops to 33 per cent".

A related aspect of human capital issues in banking is with regard to the compensation policy in public sector banks where the pay is compressed relative to the private sector due to the method of pay fixation with the pegging of earnings to the Pay Commission awards. First, there is a pecking order in the financial sector with the RBI, which is the regulator and financial institutions that hitherto performed a developmental function being at the top of the pecking order. The State Bank of India group comes next due to the special function it performs of transacting the government business of both the central and the state governments. Last in the hierarchy are the other nationalized banks. The salary of the chairman and managing director of the banks is fixed by the government keeping in mind parity with the Indian Administrative Service and this then becomes the ceiling on the salary that can be earned in the banks. The subordinate staff, knowing this upper limit, as well as the managerial grades that are operational, seven in all, go into the bargain of their unions, such as the All India Bank Employees Association and National Confederation of Bank Employees, with the Indian Banks Association. This bargain is done with the objective of minimizing the distance between their earnings and that set by the government for the chairman and managing director. As the settlement with the subordinate staff sets the floor for determining officers' salaries, the pay scales for the officers leaves little room for maneuverability and is sandwiched between the subordinate staff award settlement and the salary of the chairman and managing director.

The pay structure that emerges in the public sector banks is, therefore, compressed and even though the relative responsibility increases significantly as one moves up the organizational hierarchy, the relative increase in compensation does not match up to this increased responsibility. As a result, the use of promotions as they take place currently amongst the public sector banks, is not an effective mechanism for sorting out the higher ability individuals and neither do they provide incentives for increased effort. It is not surprising accordingly that many employees by-pass the opportunity for promotion and that with the emergence of the new private banks there is an incentive for shifting jobs amongst those employees who are less risk-averse and are willing to forego job security.

This has had the effect that public sector banks have found it very difficult to retain motivated employees and their talent pool has been subjected to raiding by the private banks. Talented and not so risk-averse individuals prefer to be rewarded on the basis of performance rather than independent of it. As a result, when the private banks pay more on average and offer stronger pay-for-performance rewards, it is to be expected that talent will migrate to them. Accordingly, low ability and risk-averse individuals will continue in public sector banks with bureaucratic compensation systems that ignore performance. As the rewards do not vary with the value of the organization they are not effective in motivating the top management from acting in the interests of maximizing the value of the organization and that affects performance.

A banking system that is dominated by public sector banks that have compensation systems that do not motivate their employees to perform in line with market-oriented systems will find that its employees under-invest in risk management and credit monitoring. This affects the portfolio of the banks and the commercial efficiency of the banks. The essence of financial intermediation is the value created in the process of reducing information asymmetry and that is made possible by the skills and motivation that bank employees invest in and deploy in the assessment of the risks that affect the efficiency of intermediation. When attempting banking reforms it is imperative that incentives be created for bankers to upgrade their information and human capital.

## SUMMARY

» Financial products are contracts that allow investors to delegate the management of their assets to others.
» Delegation of assets in financial markets is circumscribed by the prevalence of asymmetric information.
» Asymmetric information is prevalent when buyers and sellers are not equally well informed about either the characteristics of investment projects and goods and services for sale in the marketplace, or about the effort expended to efficiently utilize delegated resources.
» Eliciting information can be a public good since once the information is known by an agent, other agents can access it at negligible cost.
» Adverse selection occurs when the information possessed by an individual (or firm or institution) is concealed from others and the welfare (or pay-offs) of others depends on this information.
» The effect of hidden information arising from adverse selection is minimized by screening agents, or using contracts that reveal the information.
» Moral hazard is the hidden action arising after a contract is endorsed, as it is extremely costly to monitor the effort expended towards meeting the objectives of the contract.
» Moral hazard is mitigated through clauses in contracts that induce agents to expend the appropriate effort.
» Bondholders have first claim on cash flows but receive fixed amounts if the firm makes enough income to meet its debt obligations. Equity investors have a claim on residual cash flows and can declare bankruptcy, if the cash flows are insufficient to meet financial obligations. Thus, bondholders do not participate in the upside returns if a project succeeds, and bear a significant part of the cost if they fail. The nature of cash-flow claims is a source of conflict of interest between bondholders and stockholders.
» Once a debt is incurred through a bond issue, managers have an incentive to substitute high-risk projects for low-risk ones as they discount the negative effect on bondholder value. This is asset substitution or excessive risk taking.
» Bondholders, being aware of the potential of stockholders to take actions inimical to their interests, protect themselves by including covenants or restrictions on what stockholders can do.
» Bond agreements can restrict firms' dividend policy by tying dividend payments to earnings.
» Bond agreements can also restrict where firms can invest and how much risk they can take on new investments.
» Banks protect loans by collateral or assets (like machinery or building) backing the security of the loan.
» Banks also manage the payments system through the accounts held at the bank. This enables them to monitor the financial viability of a client for a loan. Borrowers who consider future loan requirements are careful to restrict excessive risk taking.
» Information disclosure requirements in bond and equity markets are high. Banking transactions are more confidential.
» Financial markets perform three main functions—allocate resources, monitor managers, and facilitate the trading, hedging, diversifying, and pooling of risks.
» A hedge is a futures transaction agreeing to purchase or sell a given quantity of an asset at a future date at a price agreed today in order to offset (hedge) existing risk.
» Diversification refers to the ability of an agent to reduce risk by holding a number of financial securities in a portfolio.
» Risk pooling is the process by which financial intermediaries protect savers from excessive risk through sharing the risks of investment projects.
» Liquidity is the ease and speed with which agents can convert assets into purchasing power. With liquid capital, markets savers can hold assets (bonds, equity, or demand deposits) that they can sell easily and quickly if they seek access to their savings.
» Idiosyncratic risk is the risk specific to a financial instrument or project that cannot be diversified away. This comprises default risk, legal risk, and operational risk.
» Default risk is the risk that a counterparty will default on its commitment, and includes the risk that any collateral will be inadequate if relied upon. Legal risk is the risk that a contract may not be enforceable. Operational risk is the risk of loss due to operational shortcomings.
» Economic regulation in banks deals with the control of credit allocation and with the pricing of credit.
» Prudential regulation refers to the set of laws and rules designed to minimize the risks to the financial system.
» The range of financial products that may be issued by a bank is the purview of structural regulation.
» Financial repression refers to the prevalence of credit controls, interest rate controls, entry barriers into banking, operational restrictions on banks, the predominance of state-owned banks, and restrictions on international financial transactions.
» Financial repression limits the economy's investment expenditure to below the potential for investment in the economy.
» Financial repression theorists advocate the removal of interest rate ceilings and portfolio controls on banks to enable higher savings that gets directed to high-yielding investments.
» In India the Narasimham Committee made recommendations on removing the financial repression. The keystone of the reforms was the move towards market determination of interest rates on government securities.
» Financial reforms are also aimed to reduce pre-emption of bank assets through a high SLR and CRR, and to introduce norms of income recognition, asset classification, provisioning, and capital adequacy.
» Financial intermediaries face idiosyncratic risk and systemic risk while supplying credit services. Systemic risk stems from macroeconomic uncertainty, weakness in the contractual and informational environment, and geographical limitations of small jurisdictions.

- Systemic risk is the risk of the default associated with general economy-wide or macroeconomic conditions that affect all borrowers.
- Idiosyncratic risks are risks arising from asymmetric information about the default risk associated with a specific borrower.
- As the interest rate charged on loans is raised, the financial intermediary attracts bad borrowers. At some threshold rate of interest, this results in higher default rates that offset the repayments due to the bank and causes the expected return to the bank to decline. The loan return frontier (LF) for a bank thus has an inverted U shape.
- After accounting for required reserves which the bank must maintain as a ratio of deposits, the bank is free to allocate its portfolio amongst riskless government securities that give a return $r_{TB}$ and loans that are risky.
- The efficient investment portfolio line is a straight line beginning at point $r_{TB}$ and is a tangent to the LF.
- Risk-averse bankers select a point on the efficient investment portfolio line where an indifference curve representing points of indifference between risks and returns is tangential to the efficient investment portfolio line.
- In response to financial reforms where the interest rate on government securities becomes market determined, risk-averse banks prefer to hold a larger proportion of government securities in their portfolios and to substitute away from investing in loans.
- As portfolio controls requiring banks to hold a large fraction of government securities are withdrawn, banks tend to increase the proportion of loans in their portfolio. However, the liberalization of interest rates on government securities makes loans less attractive and banks reduce their exposure to loans.
- Prior to liberalization, bankers underinvested in credit-risk assessment and monitoring skills. Only when bankers' skills and human capital is upgraded to cope with the demands of liberalization will the credit-deposit ratios of banks improve.

## NOTES

1. Kenneth J. Arrow and Gerard Debreu, "Existence of Equilibrium for a Competitive Economy," *Econometrica* 22 (1954): 265–290.
2. Michael Rothschild and Joseph E. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90 (1976): 629–649.
3. Patrick Bolton and Xavier Freixas, "A Dilution Cost Approach to Financial Intermediation and Securities Markets," Working Paper, Universitat Pompeu Fabra (1998).
4. Douglas W. Diamond, "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies* 59 (1984): 393.
5. Ross Levine, "Financial Development and Economic Growth: Views and Agenda," *Journal of Economic Literature* 35, no. 2 (1997): 688–726.
6. Curtailing the availability of good substitutes would make the demand for money more inelastic.
7. Ronald I. McKinnon, *Money and Capital in Economic Development* (Washington, D.C.: Brookings Institution, 1973).
8. Edward S. Shaw, *Financial Deepening in Economic Development* (New York: Oxford University Press, 1973).
9. Maxwell J. Fry, *Emancipating the Banking System and Developing Markets for Government Debt* (London and New York: Routledge, 1997).
10. GOI, "Report of the Committee on the Financial System," (New Delhi: Government of India), 1991.
11. Ibid.
12. S. S. Tarapore, "Financial Sector Reform: Retrospect and Prospect," RBI Bulletin (May 1996): 299–306.
13. In reality a banks capital is divided into Tier-I and Tier-II capital. Tier-I capital is closely linked to a banks book value of equity reflecting the concept of the core capital contribution of a bank's owners. Tier-II capital is a broad array of secondary capital resources including a bank's loan loss reserves plus various convertible and subordinated debt instruments.
14. Note that systemic risk may be diversified by resorting to international capital markets.
15. Because of arbitrage the marginal cost of funds to the bank will be equal to the interest rate paid on short-term government debt securities.
16. Joseph E. Stiglitz and Andrew Weiss, "Credit Rationing in Markets with Incomplete Information," *American Economic Review* (1981): 393–410.
17. Since the T-bill pays a return without any risk, the risk of the portfolio $\sigma_p$ is the risk associated with the loan $\sigma_A$ times the weight of the loan in the portfolio, $(1 - x)$. Thus, $\sigma_p = (1 - x)\,\sigma_A$.
18. It is quite possible that the loan demand at and below rA exceeds the lending capacity of banks, $(1 - x)A^{s*}$, in which case there is credit rationing.
19. Gerard Caprio Jr., "Banking on Financial Reform? A Case of Sensitive Dependence on Initial Conditions," in Gerard Caprio Jr., Ijak Atiyas, and James A. Hanson (eds.) *Financial Reform—Theory and Experience* (Cambridge, MA: Cambridge University Press, 1996).
20. Reserve Bank of India, "Report of the Committee to Review the Working of the Monetary System," Bombay (1985).
21. The Committee recommended that the discount rate on short-term T-bills (91 days) be set to provide a yield marginally positive in real terms, and that the rest of the yield structure be adjusted accordingly to return expected real yields of 1 per cent to 3 per cent per annum, depending on maturity.
22. GOI, op. cit.

23. Irving Fisher, "The Debt Deflation Theory of Great Depressions," *Econometrica* (October 1933).
24. James Tobin, "Asset Accumulation and Economic Activity," Yrjo Jahnsson Lectures, Basil Blackwell, Oxford (1980).
25. Errol D'Souza, "Employment and Human Resource Practices in Public Sector Banks in the 1990s", in Shuji Uchikawa (ed.) *Labour Market and Institution in India–1990s and Beyond* (New Delhi: Manohar Publishers, 2003), pp. 95–120.
26. Errol D'Souza, "Prudential Regulation in Indian Banking," in S. Arumugam (ed.) Indian Capital Markets—Modern Perspectives and Empirical Evidence (Mumbai: Allied Publishers Ltd., 2001).

## TEST YOURSELF

1. In what way are financial markets different from the market for goods and the market for money?
2. What are adverse selection and moral hazard and how do financial markets attempt to mitigate the consequences of their prevalence?
3. In what way is the market for bank loans different from the bond markets and stock markets?
4. What are the functions of financial markets? What are the types of risk that banks face in the loan market?
5. What is financial repression, and how does it impact aggregate savings and investment?
6. What are the major financial reforms that have been initiated in India and what was the expected outcome of these reforms?
7. How do banks decide on the proportion of assets to hold in government securities and in loans?
8. Explain how banks decide on the equilibrium supply of bank loans and demand for deposits?
9. What is the initial response of banks to the scaling down of preemptions of their portfolio and liberalization of interest rates on government securities?
10. To what extent are credit-risk assessment skills and human capital responsible for the preference for government securities in bank portfolios?

## ONLINE APPLICATION

1. (a) Go to the home page of the Reserve Bank of India (URL: www.rbi.org.in/home.aspx).
   (b) Click on the "Database" icon.
   (c) Then under the classification of Annual data, click on "Handbook of Statistics on Indian Economy".
2. Click on "Table 46: Major Monetary Policy Measures—Bank Rate, CRR & SLR". Create a spreadsheet of data on the CRR for the period 1991 till currently.
3. Table 46 gives data only up to 2004. The subsequent CRR rates affected by the Reserve Bank of India are given in the table below:

Cash Reserve Ratio

| Rate | Effective Date |
|---|---|
| 5.25 | 23.12.2006 |
| 5.50 | 06.01.2007 |
| 5.75 | 17.02.2007 |
| 6.00 | 03.03.2007 |
| 6.25 | 14.04.2007 |
| 6.50 | 28.04.2007 |
| 7.00 | 04.08.2007 |
| 7.50 | 10.11.2007 |

4. Again from Table 46 insert data in your spreadsheet on the SLR as prescribed by the Reserve Bank of India from 1990 till date. Note from Table 46 that from October 1997 there have been no changes in the mandated SLR. In fact, till date (30 August, 2007) the SLR has remained unchanged.
5. Apart from CRR and SLR the central bank uses the Liquidity Adjustment Facility (LAF) that was introduced in June 2000 to inject liquidity into the system (repo) and absorb liquidity from the banking system (reverse repo). This has been the RBI's primary instrument for managing day-to-day liquidity and transmitting interest rate signals to the market. The repo and reverse repo auctions set a corridor for the short-term interest rate consistent with policy objectives. Call money rates generally stay within the reverse repo and repo rate corridor. The movements in the LAF rates are given in the table below:

Movement in LAF Rates

| Repo (Injection) | | Reverse Repo (Absorption) | |
|---|---|---|---|
| Date | Rate | Date | Rate |
| March 2002 | 8.50 | March 2002 | 6.00 |
| May 2002 | 8.00 | June 2002 | 5.75 |
| November 2002 | 7.50 | October 2002 | 5.50 |
| March 2003 | 7.00 | March 2003 | 5.00 |
| August 2003 | 6.50 | August 2003 | 4.50 |
| March 2004 | 6.00 | October 2004 | 4.75 |
| October 2005 | 6.25 | April 2005 | 5.00 |
| January 2006 | 6.50 | October 2005 | 5.25 |
| June 2006 | 6.75 | January 2006 | 5.50 |
| July 2006 | 7.00 | June 2006 | 5.75 |
| | | July 2006 | 6.00 |
| | | October 2006 | 7.25 |
| | | January 2007 | 7.50 |
| | | March 2007 | 7.75 |

Insert the data in the table above in your spreadsheet.

6. Create a graph of the CRR rates and LAF rates till date. (Note that the CRR is the major direct instrument for monetary control with the RBI and the LAF is the principal indirect instrument for managing the supply of funds. The introduction of the LAF is therefore the introduction of a signal by the RBI of its monetary policy stance to the financial markets which are characterized by asymmetric information.)
7. Go back to the Database on the Indian Economy and click on "Table 183: Scheduled Commercial Banks—Select Aggregates". From 1991 onwards find out the "Investments" (Column 10)—that is, investment in government and other approved securities—as a proportion of demand and time liabilities of banks (Column 4 + Column 7), for the month of March in each year. For instance, in March 1991, Investments by banks were INR 750.65 billion. Aggregate deposits (Column 4) plus Other demand and time liabilities (Column 7) were INR 1,925.41 billion + INR 125.89 billion, or, INR 2,051.3 billion. Then, the investment by banks in government securities was (750.65/2,051.3)*100 = 36.6 per cent of demand and time liabilities in March 1991.
8. Create a graph of the proportion of liabilities that banks have invested in government securities. In the same graph insert a graph of the mandated SLR for each year that was downloaded by you in Step 6.
9. What do the graphs you created in Steps 8 and 10 tell you about the nature of financial reforms in the economy?

# 15

# Monetary Policy Objectives and Targets

CRITICAL QUESTIONS

- *What are the goals of monetary policy?*
- *Why do central banks use intermediate targets while attempting to achieve their ultimate objectives?*
- *Which intermediate target—interest rate or money stock—is effective when there are real aggregate demand shocks and money demand shocks?*
- *Which intermediate target is effective in the presence of supply shocks to the economy?*
- *What is time inconsistency and do monetary policymakers have an incentive to deviate from a pre-announced policy?*
- *How do individuals with rational expectations react to announcements of monetary policy when policymakers have the discretion to change policy?*
- *How does a central bank achieve credibility where the announcement of a policy is believed by agents in the economy?*

## 15.1 Goals of Monetary Policy and Intermediate Targets

Apart from attempting to understand how the aggregate economy functions, macroeconomics is also concerned with the goals that policymakers may appropriately engage with. The goals of macroeconomic policymaking are usually the following:

1. Output: Several theories point out that there is little scope for monetary and fiscal policy to have long-run effects. However, in the short-run, unexpected changes in the growth rate of the money stock can have an effect on real output. This makes one of the goals of macroeconomic policy, that is, the prevention of spikes in output fluctuations relative to its natural, full-information level. Such a POLICY GOAL also alleviates business cycles.
2. Inflation: Inflation redistributes resources from creditors to debtors by reducing the real value of debt and redistributes incomes from the private sector to the government as it pushes people into higher nominal tax brackets. Moreover, inflation variability induces uncertainty about the future gains from sinking resources into a new project and results in reductions in investment and economic growth. These costs of inflation provide good reasons for policymakers to attempt to keep inflation rates low and to limit the variability in inflation rates.

› The **goals of monetary policy** are to keep output and inflation as stable as possible. The output is stabilized around the potential or natural rate of output and the inflation is kept stable at a low rate.

Some policymakers are also burdened with demands to prevent significant unemployment in the economy, especially when workers are a significant vote bank. However, in the short run with a trade-off between inflation and unemployment, as depicted by the Phillips curve, there could be conflicts among macroeconomic objectives. How should monetary and fiscal policies be conducted in the face of conflicting goals? Moreover, because of the time lags associated with policymaking, as pointed out by Friedman, a policymaker focusing on ultimate objectives may not react with the instruments at its disposal until a disturbance has already occurred.

As a way around this problem, economists suggest an indirect management of the ultimate target by means of an INTERMEDIATE TARGET. In fact, typically, most central banks have sought to achieve intermediate targets of monetary policy. An intermediate target is a macroeconomic variable that a central bank attempts to control because of the view that doing so is consistent with its ultimate objectives. Intermediate targets serve as proxies for ultimate objectives.

› An **intermediate target** is a macroeconomic variable that the central bank can influence with some predictability and is related to its ultimate objectives.

› **Intermediate targets** are used in monetary policy as a compromise in the absence of complete agreement among policymakers about how best to achieve its ultimate objectives and also because policymakers possess limited current information about their ultimate policy goals such as output and inflation.

There are two reasons for using INTERMEDIATE TARGETS when conducting monetary policy. The first is that there are a variety of theories about how monetary policy affects inflation and output. Given that individual policymakers may subscribe to different theoretical outlooks on the economy, it can be difficult to reach a consensus on the best way to attain the policy objectives. In such a situation pursuing an intermediate policy target is a compromise, when it is difficult to agree about how to achieve ultimate objectives directly. For instance, even though there is disagreement about the extent of the effect of changes in the money stock on output, most theories point out that a change in the nominal quantity of money has an effect on the price level. In that case, central bank officials might resolve to pursue a money stock growth rate objective as it is an approximation to the ultimate goal on which a consensus exists.

Second, even if policymakers can agree on how policies influence economic activity, they typically possess limited information about the economy. Current information about ultimate policy goals—inflation and real output—is available only with a lag and even then the estimates of these variables are revised in the period following their initial release. By contrast, interest rates, money stock, and credit data are in hand for inspection and for use at any given moment. By pursuing an intermediate target defined around such financial variables, the central bank can infer whether the eventual outcome for the economy is in the direction of its eventual policy objectives rather than having to wait to make such an assessment.

The choice of a macroeconomic variable, that is, to serve as an intermediate target is usually made on the basis of the following considerations:[1]

1. Frequent observability: It must be readily observable with little or no lag which is the rationale for preferring to use intermediate targets rather than focusing on ultimate goal variables. For instance, the price level as measured by the WPI is observable at fortnightly intervals and real GDP at quarterly intervals but with a lag in the availability of the data. Hence, the central bank must choose an intermediate target variable that can be observed more frequently, say weekly, if not daily.
2. Consistent with ultimate goals: The policies resulting in the target variables taking on certain values must in turn result in the ultimate goal variables taking on certain values. A central bank successfully meeting values for the INTERMEDIATE TARGET should at the same time meet its goals for inflation and output.
3. Controllable: An intermediate target variable should be one whose value the central bank can rapidly influence, so that on attaining the intermediate target the ultimate objective is also eventually achieved.

› Macroeconomic variables selected as intermediate targets must be frequently observable, consistent with ultimate goals, and controllable.

› Central banks have a range of options from which to select intermediate targets: monetary or credit aggregates, interest rates, and nominal GDP.

The ALTERNATIVE INTERMEDIATE TARGET VARIABLES that central banks have adopted include the following:

1. Monetary or credit aggregates: The rationale for targeting monetary aggregates such as $M_1$ or $M_3$ has been that the quantity of money determines aggregate demand and influences the price level. Also, monetary aggregates are controllable by central banks and frequently observable as well as their values are known weekly or at most fortnightly. Many countries including Japan (1975–1994), the UK (1976–1994), and Germany (1975–2001) have used monetary aggregates as intermediate target variables. Financial innovations, however, began to affect the stability of money demand in the 1980s as new forms of money-like assets have emerged, and a given stock of money has been able to service a larger volume of transactions. The 1999 Bank of England survey of monetary frameworks reported 43 of 50 central banks as viewing monetary aggregates as relevant intermediate targets while 7 central banks preferred the interest rate as intermediate targets.

   Another financial aggregate that has been emphasized in China and Russia is a credit aggregate or the volume of lending. In India, the RBI had traditionally used credit—aggregate as well as sectoral—as an intermediate target of monetary policy. This adherence to credit targeting went hand in glove with the concept of development banking that accompanied the planning process in India. As inflation was considered to be structural selective credit controls were used to regulate bank advances to sensitive commodities so as to influence production outlays and curb speculation. Over time as deficit financing of public investment leaked out into inflation, the focus of monetary management shifted to the management of demand rather than the regulation of supply. The

Chakravarty Committee recommended a monetary targeting framework to target inflation in line with output growth.[2] Monetary targeting was successful as a tool of monetary policy in the late 1980s and early 1990s as inflation was driven by excessive monetary expansion.

2. Interest rates: The price of credit or the nominal interest rate is another intermediate monetary policy target. Central bank policy actions have direct effects on nominal interest rates and interest rates can be observed daily. However, the relationship between interest rates and economic activity is not necessarily directly obvious. Although lower interest rates stimulate capital investment, it is also the case that the increases in aggregate income raise the demand for credit and push up interest rates. While the operating procedures of central banks in developed economies can be described as interest rate targeting, there are significant differences in implementation. Many countries target the interest rate with refinancing facilities[3] but some countries have a preference for open-market operations. The operating target can also relate to the rate for overnight money or the money market rate for longer maturities such as the 3-month rate. There is also a difference regarding the precision with which the money market rate is controlled—whether a specific interest rate level is targeted or whether fluctuation within a given interest rate band in permissible. Interest rate targeting does not involve keeping the money market rate constant over long periods but rather aims to exert an effect over economic events via a money market rate. The Federal Reserve deployment of this operating procedure in the United States demonstrates that there can be phases when the money market rate is noticeably stable as in the second half of the 1990s as well as phases involving rapid and abrupt interest rate changes as occurred during 1972–1975.
3. The nominal GDP: In recent times, the nominal GDP has been advanced as an intermediate target. As the nominal GDP is equal to the real GDP times the GDP price deflator, minimizing the nominal GDP helps contain the volatility in either of the ultimate goal variables—real output and the price level.

**MACROFOCUS 15.1**

### Evolution of the Operating Instruments of the Monetary Policy in India

In an emerging market such as India, the availability of credit is often an important dimension of policy because of the reliance of small and medium enterprises, which dominate the economy, on bank credit. In the 1990s with the opening up of the economy, the linkages between money, credit, and output began to weaken. Until the mid-1990s, banks surrendered the foreign currency they raised via non-resident deposits and current account transactions and thus there was a direct link between the process of money and credit creation. As the external account got liberalized, banks could increasingly invest their foreign currencies abroad and as domestic credit demand was weak, they resorted to this measure. Between 1995–1996 and 2003–2004, bank credit to the commercial sector grew at a trend rate of 14.2 per cent whereas their foreign currency assets grew[4] at 20 per cent. The Reddy Working group accordingly re-emphasized the importance of monitoring credit aggregates again as there was a substantial accretion to net foreign assets of the banking sector.* In the absence of a clear connection between intermediate targets and final objectives, the RBI has turned to a multiple-indicator approach. The RBI has shifted from direct instruments of monetary control to indirect instruments in consonance with macroeconomic outcomes. It now modulates market liquidity to steer monetary conditions to the desired trajectory. This is achieved via a mix of policy instruments including changes in reserve requirements and standing facilities and open-market (including repo) operations, which affect the quantum of liquidity, and changes in policy rates, such as the bank rate and repo rate, which influence the price of liquidity.** The liquidity adjustment facility is used by the RBI to manage liquidity on a daily basis and to transmit interest rate signals to the market and has emerged as a principal indirect operating instrument of monetary policy.

*Y. V. Reddy, "Parameters of Monetary Policy," *RBI Bulletin*, February 2003.
**RBI, "Report of the Working Group on Money Supply: Analytics and Methodology of Compilation," Chairman: Dr. Y. N. Reddy (June 1998), Mumbai.

# 15.2 Choosing Intermediate Targets in the Case of Demand Shocks

The analysis of the merits of interest rate targeting and monetary targeting goes back to Poole.[5] In the context of a deterministic macro model, Poole argued that the "policy prescription may be in terms of either the interest rate or the money stock; it makes no difference which instrument is selected."

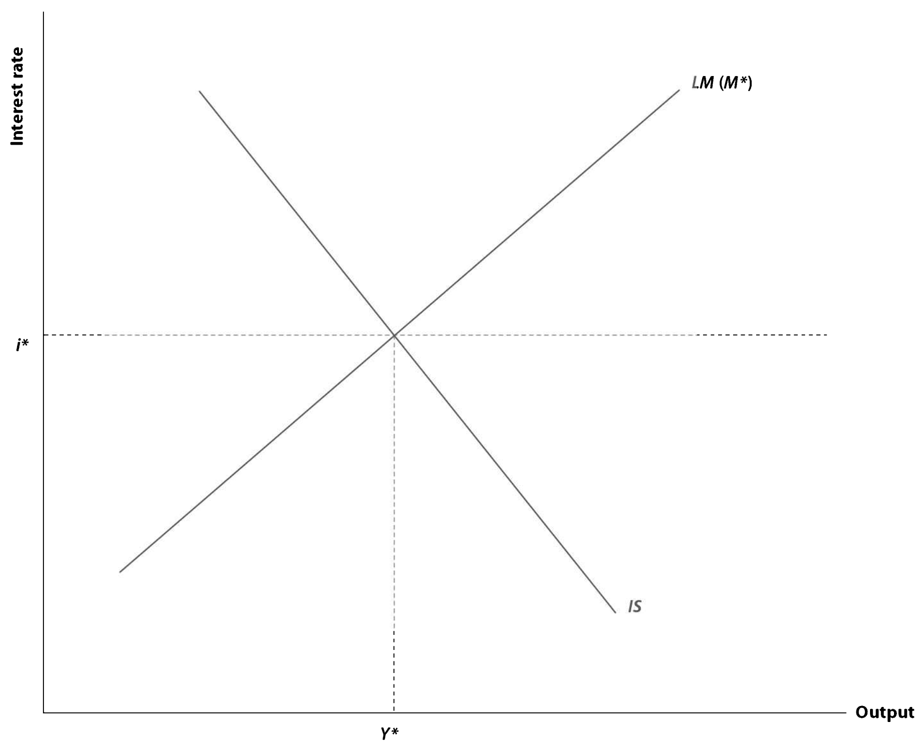
In Figure 15.1, the equilibrium income level $Y^*$ is reached either by the monetary authorities setting the interest rate of $i^*$ or by setting the money stock at some level $M^*$ that makes the $LM$ function intercept the $IS$ function at $Y^*$. From a policy prescription viewpoint, the CHOICES are equally serviceable.

› In a **deterministic environment**, it is irrelevant whether it is the interest rate or the money stock that is targeted.

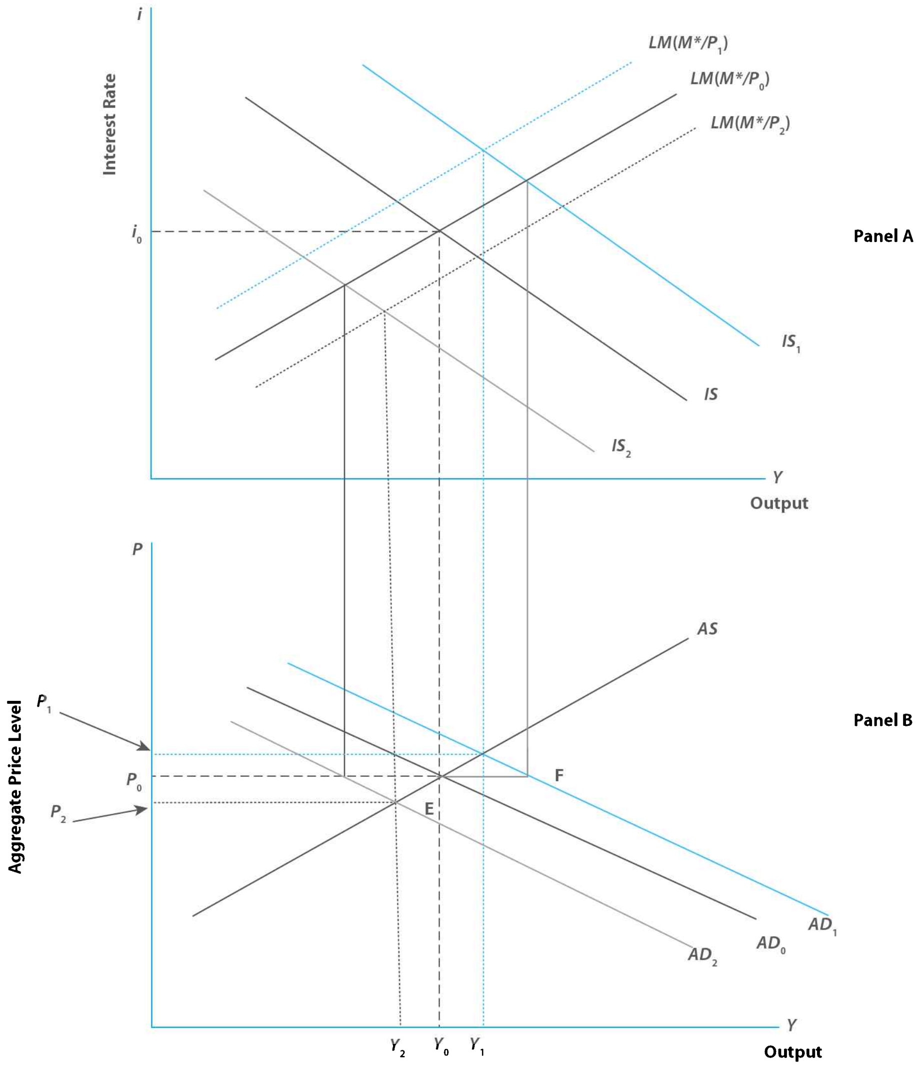
Thereafter, Poole went on to show that this is not the case when the $IS$ and $LM$ functions are subject to random shocks. All the stochastic shocks studied here are variations of the additive error terms that have to be added to the money demand and aggregate expenditure equations ($IS$ equation) of the model. We abstract from shocks that affect the slopes of the curves in the macroeconomic model. Poole's analysis neglected the impact of shocks on the price level. We extend his approach to an aggregate demand–supply model with flexible prices and also examine the effects of the supply shocks, which are important in a development context.

### 15.2.1 *IS* Curve Shocks

We begin with shocks such as changes in autonomous consumption, investment, government spending, export or import spending, or net taxes that cause a multiple change in the equilibrium real income and result in a shift of the $IS$ curve. A shift in the $IS$ curve to $IS_1$ causes the aggregate demand curve




› **Figure 15.1**
Equivalence of Instruments in a Deterministic Environment. **The equilibrium income level *Y** can be achieved by the monetary authorities setting an interest rate of *i**, or controlling the money supply at a level *M** that ensures that the *LM* curve intersects the *IS* curve at *Y**. Both instruments give an identical result.**



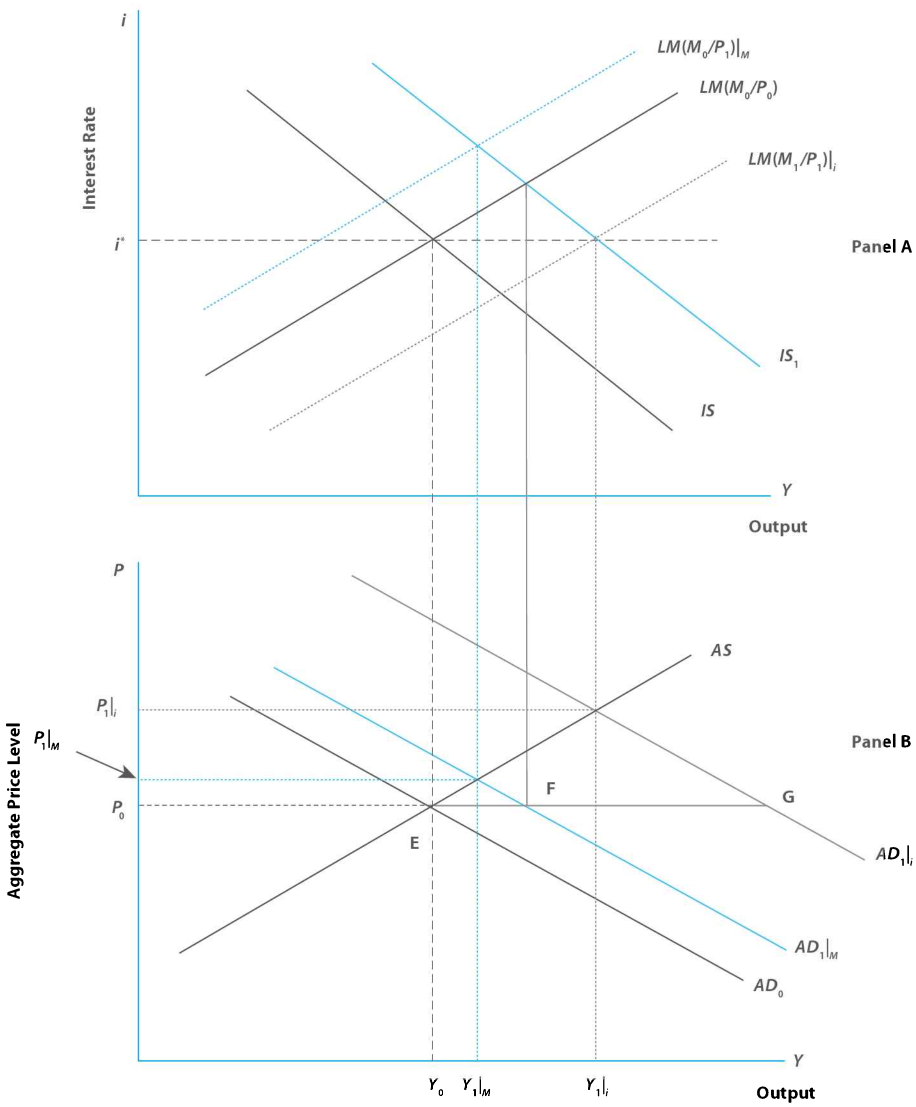

› **Figure 15.2**
Shocks to *IS* Curve. A positive shock to the *IS* curve shifts it rightwards to $IS_1$ in panel A and the aggregate demand curve to $AD_1$ in panel B. The excess demand for goods given by distance EF increases the price to $P_1$ in panel B and reduces the real money supply, which shifts the *LM* curve in panel A to $LM(M^*/P_1)$. Similarly, a negative shock to the *IS* curve results in a price $P_2$ and a rightward shift of the *LM* curve to $LM(M^*/P_2)$. Output fluctuates between $Y_1$ and $Y_2$ and the price fluctuates between $P_1$ and $P_2$.

to shift to $AD_1$ [Figure 15.2(B)]. At the initial price level $P_0$, there is an excess demand for goods (distance EF) that results in a rise in the price level. In turn, the increase in the price reduces the real money supply to $M^*/P_1 < M^*/P_0$ and the *LM* curve shifts to the left in Figure 15.2(A). The price rises till the excess demand is eliminated, $P = P_1$ and $AD_1 = AS$. Similarly, a leftward shift of the *IS* curve to $IS_2$ will result in an excess supply of output at the initial price $P_0$ causing the price level to decline to $P_2$ and the *LM* curve to shift rightward to $LM(M^*/P_2)$. The excess supply is eliminated at $P = P_2$ and $AD_2 = AS$. Under monetary targeting, the central bank keeps the money supply fixed at $M^*$. In such a situation, fluctuations in the *IS* curve result in the real output fluctuating between $Y_1$ and $Y_2$ and the price level varying between $P_1$ and $P_2$.

If the central bank targeted the interest rate, it would adjust the money supply in such a way that the *LM* curve intersects the *IS* curve at the desired interest rate $i^*$. When there is an occurrence of a positive shock to the *IS* curve that shifts right to $IS_1$, then the $IS_1$ and *LM* curve intersect at an interest rate greater than the targeted $i^*$. Now, the central bank has to resort to expanding the nominal money supply so as to shift the *LM* curve rightwards and achieve $i^*$.

In Figure 15.3, the increase in the money supply to $M_1 > M_0$ reinforces the expansionary effect of the shift in the *IS* curve and increases the pressure on the

**› Figure 15.3**
Monetary and Interest Rate Targeting under *IS* Shock.
A positive shock to the *IS* curve shifts it to $IS_1$ in panel A. The central bank targeting an interest rate of $i^*$ increases money supply to $M_1$ that shifts the *LM* curve to $LM\,(M_1/P_1)|_i$. The rightward shift of the *IS* and *LM* curves causes the *AD* curve to shift to $AD_1|_i$ in panel *B* and the price to rise to $P_1|_i$. Compared to money supply targeting, which was dealt with in Figure 15.2, the price and the output variation is higher under interest rate targeting.

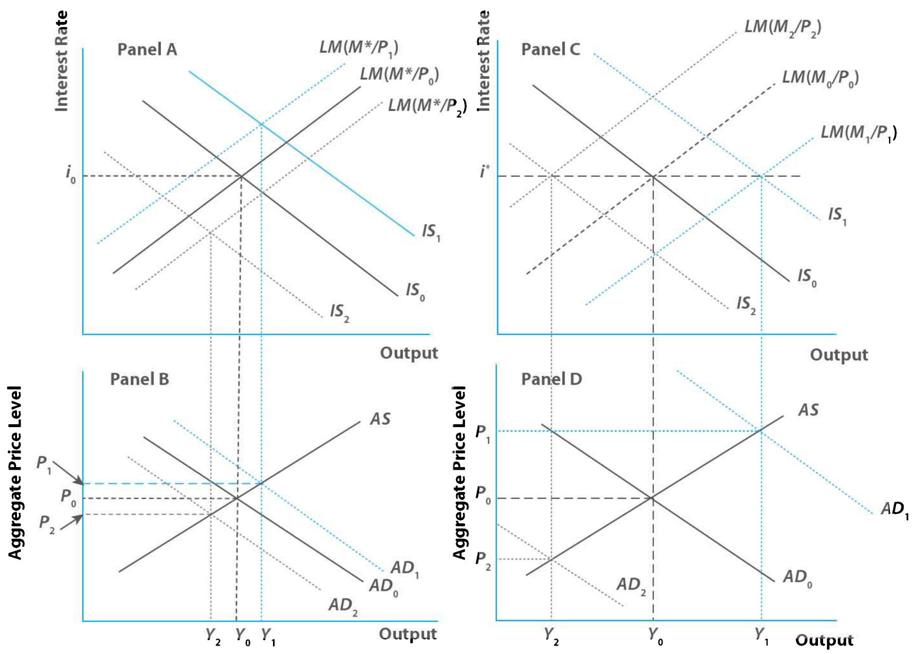



price level as excess demand at the original price level $P_0$ is now even greater at EG. As a result, the aggregate demand curve under interest rate targeting $AD_1|_i$ is even further to the right of the aggregate demand curve under money supply targeting, $AD_1|_M$. As depicted in Figure 15.3, the real output and price level variation is higher under interest rate targeting than monetary targeting with $Y_1|_i - Y_0 > Y_1|_M - Y_0$ and $P_1|_i - P_0 > P_1|_M - P_0$.

If a restrictive *IS* shock occurs, which pushes the *IS* curve leftwards to $IS_2$, then the monetary authority in an attempt to achieve $i^*$ will have to restrict nominal money supply that in turn reinforces the contraction because of negative *IS* shock [see Figure 15.4(C and D)].

› If the economy is subject to real shocks that shift the *IS* curve, then money stock targeting results in less output and price variability than interest rate targeting.

If the central bank selects an intermediate target[6] on the basis of minimizing the distance of the actual income, such as $Y_1$ and $Y_2$, from a desired level of income such as $Y_0$, then if the economy is subject to real shocks to the *IS* curve, it is clear that MONETARY TARGETING [Figure 15.4(A and B)] is superior to interest rate targeting [Figure 15.4(C and D)]. This is independent of whether the ultimate objective is to stabilize prices or output. Using the money stock as an intermediate target is preferable to targeting the nominal interest rate, which could result in a much wider variation in real output and prices.




**› Figure 15.4**
Variability under Monetary and Interest Rate Targeting for *IS* Curve Shocks. **Under money supply targeting, which was analysed in Figure 15.2, the shocks to the *IS* curve cause it to vary between $IS_1$ and $IS_2$ as depicted in panel A. This results in the aggregate demand curve varying between $AD_1$ and $AD_2$ in panel B. Under money supply targeting the output varies between $Y_1$ and $Y_2$ and the price level varies between $P_1$ and $P_2$ in panel B. Under interest rate targeting, which was analysed in Figure 15.3, the shocks to the *IS* curve are accompanied by corresponding shifts to the *LM* curve as shown in panel C, where the monetary authority targets an interest rate of *i**. The combined effect of the shifts of *IS* and *LM* results in the *AD* curve shifting left to $AD_2$ or right to $AD_1$ in panel D. Under interest rate targeting, the output variation and price variation in panel D is much higher than the variation in panel B.**

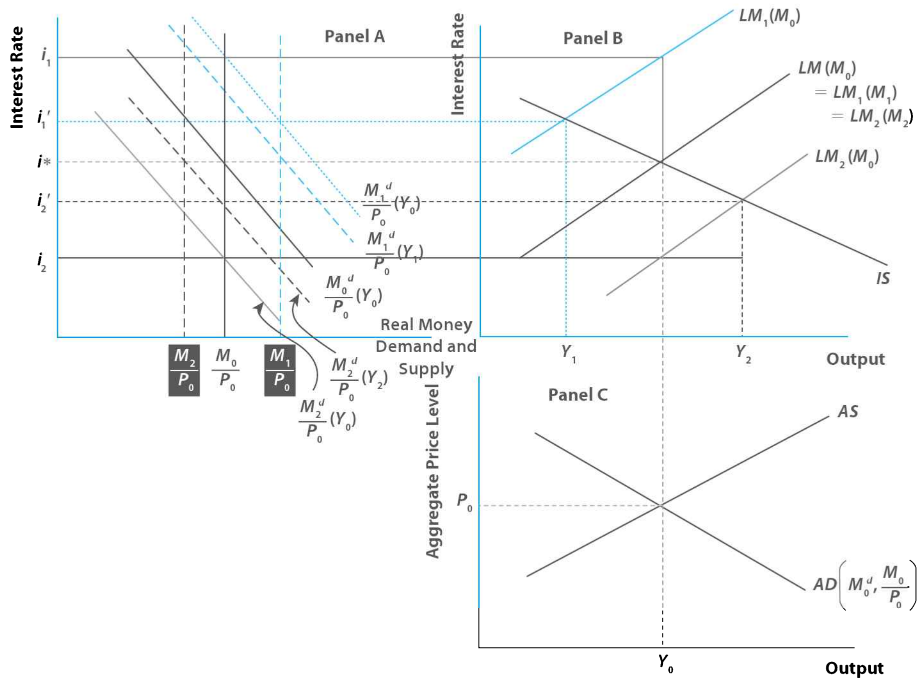
### *15.2.2 LM* Curve Shocks

Shocks to the *LM* curve can occur through variations in the demand for money that arise from changes in the technology by which people make payments, and not from any change in real income. In Figure 15.5, we examine the case where the central bank maintains a nominal interest rate target $i^*$.

To achieve this intermediate target,[7] the central bank will have to make sure that the supply of real money balance $(M_0/P_0)$ intersects the demand curve for real money balances $(M_0^d/P_0)(Y_0)$ at the interest rate $i^*$ in panel A of Figure 15.5. This results in the $LM(M_0)$ curve in panel B that intersects with the *IS* curve and results in a real income equal to $Y_0$. In turn, the accompanying aggregate demand curve $AD[M_0^d, (M_0/P_0)]$ and aggregate supply curve *AS* intersect at $(P_0, Y_0)$ in panel C. An increase in the demand for real money balances due to financial innovation to $(M_1^d/P_0)(Y_0)$ causes the equilibrium nominal interest rate in the money market to rise to $i_1$ in panel A. The *LM* schedule will accordingly shift upward by the vertical distance $i_1 - i^*$ in panel B. The rise in interest rate leads to a decline in real investment in the goods market (movement upwards along the *IS* curve) and the real income will fall to $Y_1$.

As the real income declines, it is accompanied by a reduction in the demand for money for transactions purposes and the interest rate settles at $i_1'$, where the *IS* and $LM_1(M_0)$ curves intersect.[8] The interest rate is now above the targeted interest rate, $i_1' > i$, and the central bank will respond by raising the quantity of money to $M_1$. This increase in the supply of money in panel A will shift the $LM_1(M_0)$ curve in panel B to $LM_1(M_1)$ and the original *IS*–*LM* equilibrium will be re-attained. This monetary disturbance is offset by an adjustment in the money supply, the *LM* curve does not change its position, and real output and the price level remain unchanged at $P_0$, $Y_0$.

If, instead, the money demand falls from $(M_0^d/P_0)(Y_0)$ to $(M_2^d/P_0)(Y_0)$, there is a reduction in the interest rate to $i_2'$ (up from the initial decline to $i_2$ because of the subsequent increase in the money demand for transaction as the investment expenditure and real income increase). As the interest rate has declined relative to the target rate, the monetary authority will respond by reducing the money




**› Figure 15.5**
Interest Rate Targeting under Shock to Money Demand.
When money demand is given by $M_0^d/P_0$ the central bank that targets the interest rate of $i^*$ must fix the money supply at $M_0/P_0$ in panel A. This results in the *LM* curve given by $LM(M_0)$ in panel B that intersects the *IS* curve to give the aggregate demand curve depicted in panel C that results in output $Y_0$ and price level $P_0$. An increase in the demand for money due to a shock shifts the money demand curve to $M_1^d/P_0$ and the interest rate to rise to $i_1$ in panel A. The *LM* curve shifts left to $LM_1(M_0)$ in panel B and income declines to $Y_1$ whereas the interest rate moves to $i_1'$ at the intersection of the *IS* curve with $LM_1(M_0)$. The demand for money for transactions reduces to $M_1^d/P_0\,(Y_1)$ in panel A. The monetary authorities now increase money supply to $M_1/P_0$ in panel A to intersect the demand for money at the target rate of interest $i^*$. The new *LM* curve is given by $LM_1(M_1)$ in panel B, which takes the same position as the original *LM* curve before the shock. The level of output and price are, therefore, unaffected in the goods market in panel C. A similar outcome occurs for a negative shock to money demand.

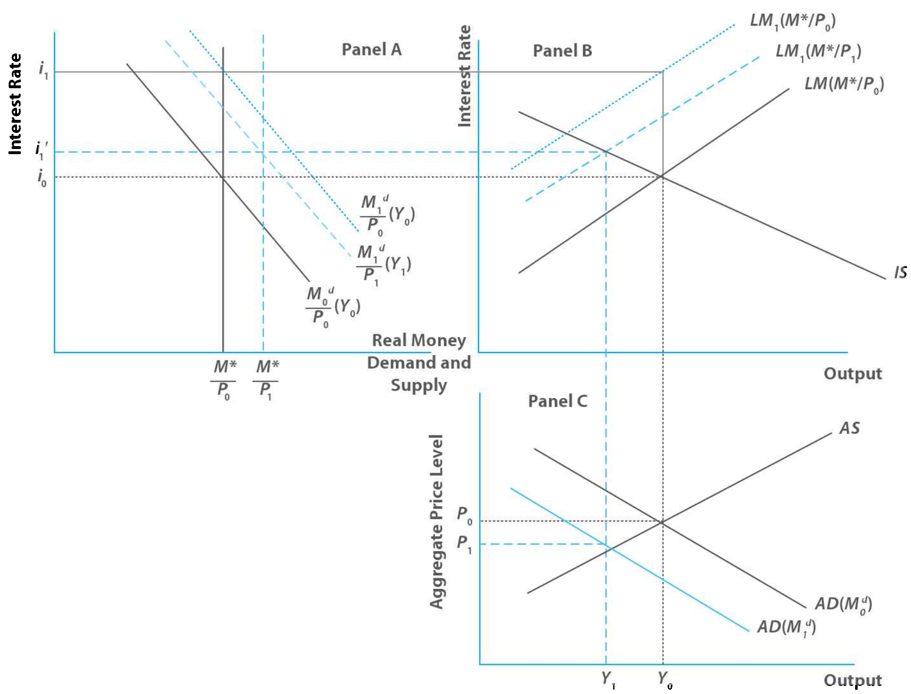
supply to $M_2$, shifting the *LM* curve back to the left in panel B to $LM_2(M_2)$ until the original equilibrium is re-attained. As a result, the real output and the price level remain unchanged at $P_0$, $Y_0$.

What if the central bank targets the nominal money supply at $M^*$? In Figure 15.6, the initial equilibrium where the demand and supply of real money balances is equal occurs at interest rate $i_0$, which results in the price and output level $P_0$ and $Y_0$, respectively. Suppose that now there is an increase in the demand for money that is not caused by a change in the real income, the increase in money demand to $(M_1^d/P_0)(Y_0)$ results in a rise in the interest rate to $i_1$ in Figure 15.6(A) and shifts the $LM(M^*)$ curve leftwards to $LM_1(M^*/P_0)$. The aggregate demand curve also shifts leftward to $AD(M_1^d)$ in Figure 15.6(C) and at price $P_0$, there is now an excess supply of goods, which pushes the price level down.

The rise in the interest rate reduces investment expenditure and real income, which shifts $(M_1^d/P_0)(Y_0)$ to the left in panel A. At the same time, the reduction in the price shifts the money supply curve $M^*/P_0$ to the right. The intersection of the new money demand $(M_1^d/P_1)(Y_0)$ and money supply curve $M^*/P_1$ results in interest rate $i_1'$ and this is accompanied by a price-output combination $P_1$, $Y_1$ in panel C.

Suppose there is a decrease in the demand for money, that is, unrelated to the level of income. In this case, the decline in real money demand to $(M_2^d/P_0)\,(Y_0)$ results in a decline in the interest rate to $i_2$ in Figure 15.7(A) and shifts the $LM(M^*)$ curve rightwards to $LM_2(M^*/P_0)$. The aggregate demand curve shifts rightwards to $AD(M_2^d)$ and the excess demand for goods causes the price to rise. The decline in the interest rate spurs investment expenditure and increases real income that shifts the money demand curve in panel A to the right. The rise in price reduces the real money supply and shifts the $M^*/P_0$ curve to the left in panel A. The final equilibrium occurs where the money demand and supply intersect at interest rate $i_2'$, which is accompanied by a price output combination given by $P_2$, $Y_2$ in panel C.

When the central bank pursues a policy of monetary targeting, the aggregate demand curve fluctuates between $AD(M_1^d)$ in Figure 15.6 and $AD(M_2^d)$




**› Figure 15.6**
Positive Shock to Money Demand when Money Stock is Targeted.
With the central bank targeting the nominal money supply at $M^*$ the intersection of money demand and supply results in the interest rate $i_0$ in panel A. The *LM* curve is given by $LM(M^*/P_0)$ in panel B and the aggregate demand curve representing *IS* and *LM* equilibrium is $AD(M_0^d)$ in panel C. Aggregate demand and supply intersect to give output $P_0$ and price level $Y_0$. A positive shock to money demand shifts the demand for money curve to the right in panel A and raises the interest rate to $i_1$. The *LM* curve shifts to the left in panel B to $LM_1(M^*/P_0)$ and results in the aggregate demand curve shifting to the left in panel C to $AD(M_1^d)$. The price level declines to $P_1$ and the income also declines to $Y_1$. This is associated with a shift to the left of the money demand curve in panel A to the dashed downward-sloping curve, and a rightward shift of the *LM* curve in panel B to $LM_1(M^*/P_1)$. Both price and output have declined.

in Figure 15.7. The result is a fluctuation in the output between $Y_1$ and $Y_2$ and in the price level between $P_1$ and $P_2$. As a result, an economy confronted with SHOCKS IN MONEY DEMAND will face less variation in output and prices if it targets the interest rate (Figure 15.5) rather than if it targets the nominal money supply (Figures 15.6 and 15.7). Monetary targeting is fruitful only when there is a stable money demand.

› If the economy is subject to shocks to the demand for money, interest rate targeting results in less variation in output and prices than money stock targeting.

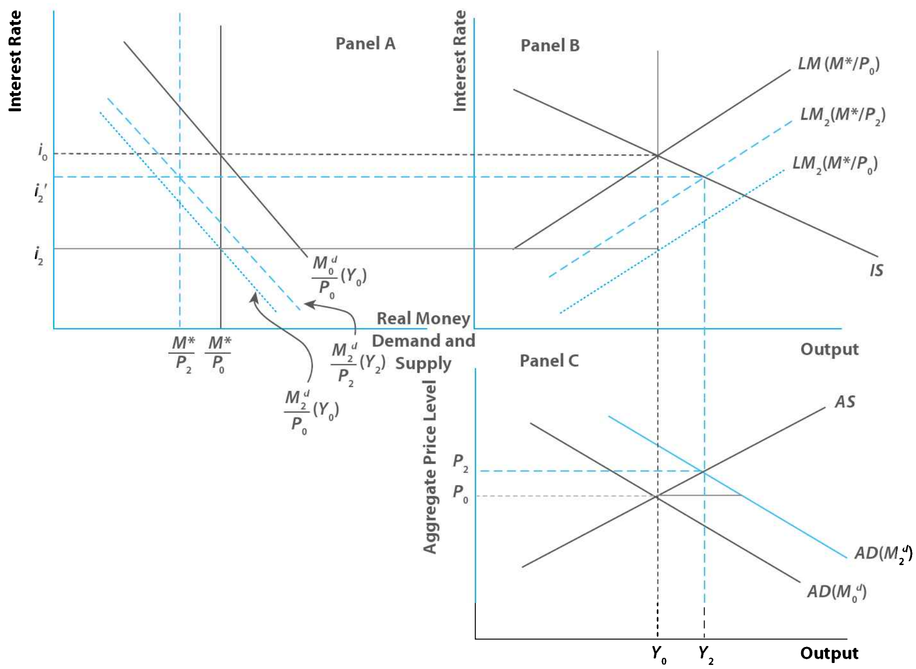
# 15.3 Choosing Intermediate Targets in the Case of Supply Shocks

So far the shocks considered were because of monetary disturbances or shifts in the demand for goods. Poole did not consider supply shocks that could arise, say, because of a rise in nominal wages, arising from trade union activity that causes an upward shift in the aggregate supply curve. Similarly, a rise in oil prices or a drought could result in the aggregate supply curve shifting up. Fluctuations in the supply curve lead to fluctuations in both the output level and the price level in the economy.

If the aggregate supply fluctuates between $AS_1$ and $AS_2$, as depicted in Figure 15.8(B), the intersection of the aggregate supply and demand curves determines the equilibrium incomes $Y_1$ and $Y_2$ and price levels $P_1$ and $P_2$. The change in the price level changes the real money stock and in the context of monetary targeting with nominal money supply $M^*$ shifts the *LM* curve to $LM(M^*/P_1)$ when the supply shock is adverse and results in $AS_1$. It shifts the *LM* curve to $LM(M^*/P_2)$ when the supply shock is favourable and results in $AS_2$.

Under a policy of interest rate targeting, the monetary authority must counteract the interest rate changes that accompany supply shocks through changes in the nominal money supply. If the supply shock shifts the aggregate

**› Figure 15.7**
Negative Shock to Money Demand when Money Stock is Targeted. With the central bank targeting the nominal money supply at $M^*$, the intersection of money demand and supply results in the interest rate $i_0$ in panel A. The $LM$ curve is given by $LM(M^*/P_0)$ in panel B and the aggregate demand curve representing $IS$ and $LM$ equilibrium is $AD(M_0^d)$ in panel C. Aggregate demand and supply intersect to give output $Y_0$ and price level $P_0$. A negative shock to money demand shifts the demand for money curve to the left in panel A and reduces the interest rate to $i_2$. The $LM$ curve shifts to the right in panel B to $LM_2(M^*/P_0)$ and results in the aggregate demand curve shifting to the right in panel C to $AD(M_2^d)$. The price level increases to $P_2$ and income also rises to $Y_2$. This is associated with a shift to the right of the money demand curve in panel A to the dashed downward-sloping curve and a leftward shift of the $LM$ curve in panel B to $LM_2(M^*/P_2)$. Both price and output have risen.




supply curve to $AS_1$, there is a rise in the price to $P_1$ and a decline in the real output to $Y_1$ [see Figure 15.9(B)].

With the interest rate increasing above $i^*$, the central bank increases the money supply to $M_1$, and this results in an outward shift of the aggregate demand curve to $AD_1$. The money supply is increased until the target rate of interest $i^*$ is achieved that occurs when the $LM$ curve is back at $LM(M_0/P_0)$. This requires that the real income to be back at $Y_0$ that occurs when the shift in the $AD$ curve due to the expansion in money supply results in an intersection with the $AS_1$ curve at real income $Y_0$. Corresponding to the output $Y_0$, there will be a price level $P_3$ that results in a real money supply of $M_1/P_3$ that corresponds to the $LM$ curve $LM(M_1/P_3)$, which is identical to $LM(M_0/P_0)$. In a similar vein if the aggregate supply shock shifts the supply curve to $AS_2$, in an attempt to get back to its interest target of $i^*$ the central bank contracts the money supply till the real money supply $M_2/P_4$ results in the $LM$ curve given by $LM(M_2/P_4) = LM(M_0/P_0)$.

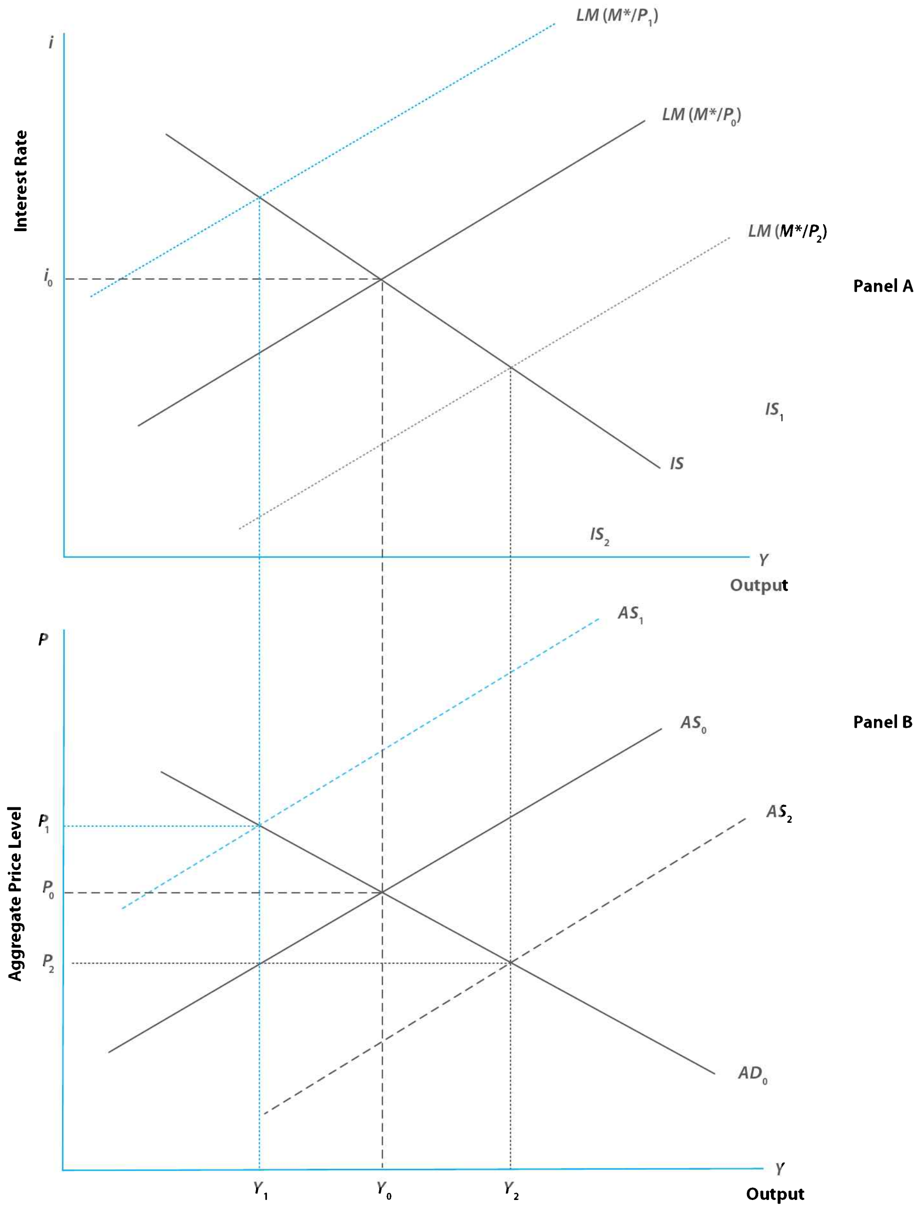
› If the economy is subject to a supply shock then money stock targeting results in less variation in prices than interest rate targeting. However, interest rate targeting results in less variation in the output compared to money stock targeting.

Comparing Figures 15.8 and 15.9, we can see that monetary targeting results in a smaller variability of prices ($P_1$ to $P_2$) compared with INTEREST RATE TARGETING ($P_3$ to $P_4$). On the other hand, interest rate targeting results in a smaller variability of the real output ($Y_0$) compared with monetary targeting ($Y_1$ to $Y_2$).

The effectiveness of monetary versus interest rate targeting can be summarized in Table 15.1.

**› Table 15.1**
Effectiveness of Intermediate Targets

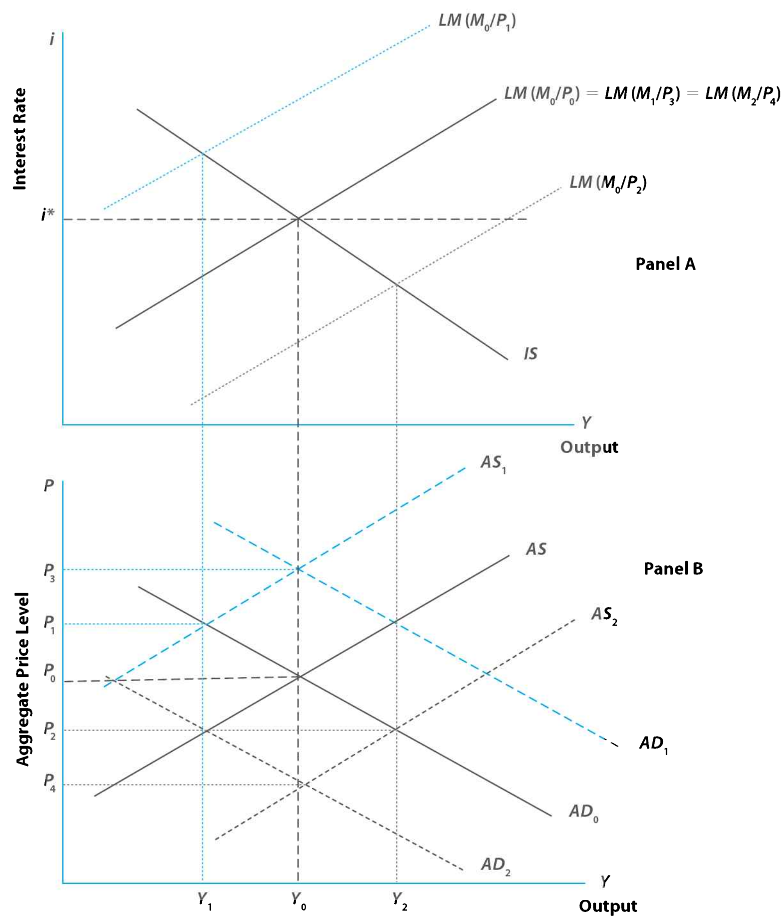
| | Effective Intermediate Target | |
|---|---|---|
| Type of Shock | Interest Rate | Nominal Money Stock |
| Demand Shock | Money Demand Shock | Real Aggregate Demand Shock |
| Supply Shock | If Real Output Target is the Ultimate Objective of the Central Bank | If Price Level Target is the Ultimate Objective of the Central Bank |




**› Figure 15.8**
Supply Shock under Monetary Targeting. An adverse supply shock shifts the aggregate supply curve upwards to $AS_1$ in panel B and raises the price and reduces output. With the central bank targeting the money supply at $M^*$ the higher price and lower output shifts the *LM* curve to the left to $LM(M^*/P_1)$ and the interest rate increases as depicted in panel A. A favourable supply shock shifts the aggregate supply curve downwards to $AS_2$ in panel B and reduces price whilst increasing output. With the central bank targeting the money supply the *LM* curve shifts to the right to $LM(M^*/P_2)$ in panel A. Under monetary targeting supply shocks result in prices varying between $P_1$ and $P_2$ and output varying between $Y_1$ and $Y_2$.

## 15.4 Targeting the Nominal GDP

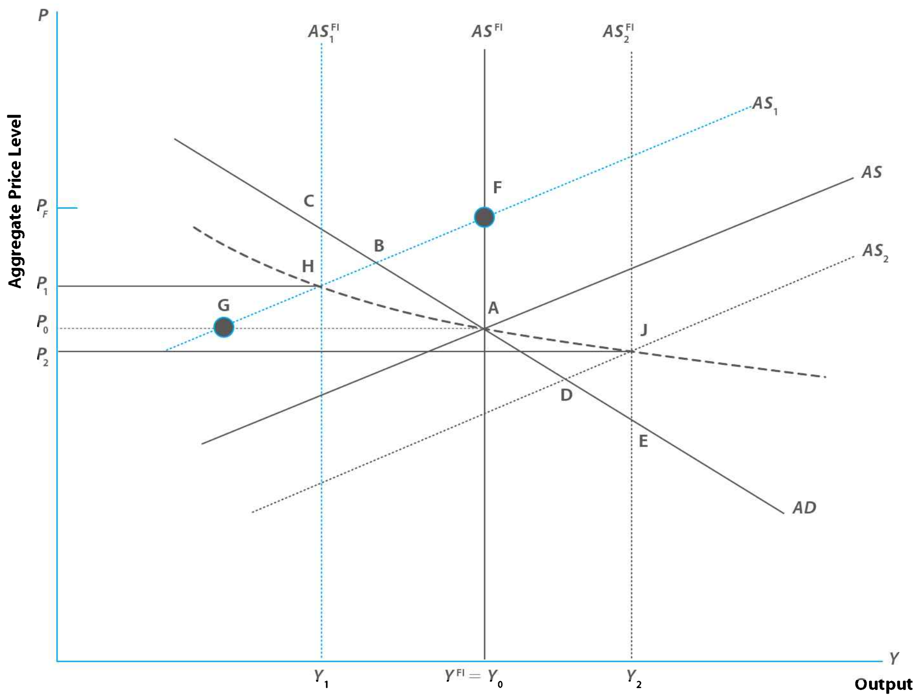
An alternative policy to money stock or interest rate targeting is to target the nominal GDP. Interest rate and monetary targeting are effective when the source of variability is on the demand side of the economy, either shocks to *IS* or shocks to money demand and *LM*. This is because demand shocks move prices and output in the same direction and so any policy that offsets price changes will also offset the real output simultaneously. This does not occur with supply shocks as we saw. Under a supply shock, interest targeting stabilizes output but prices vary considerably. Also, for monetary targeting a supply shock stabilizes prices better but at the cost of output variability. This is because under supply shocks, prices and output change in opposite directions. An adverse supply shock, for instance, reduces the output and results in a rise in the price level.




**› Figure 15.9**
Supply Shock under Interest Rate Targeting. An adverse supply shock shifts the *AS* curve up to $AS_1$ in panel B. The price rises to $P_1$ and the output declines to $Y_1$ resulting in the *LM* curve shifting leftwards to $LM(M_0/P_1)$ in panel A. With the central bank targeting the interest rate at $i^*$ it increases the supply of money to $M_1$ which causes the *LM* curve to shift to the right in panel A till it intersects the *IS* curve at $i^*$. The rightward shift of the *LM* curve is accompanied by a rightward shift of the aggregate demand curve in panel B to $AD_1$ that results in the price rising further to $P_3$ and the output returning to the pre shock level of $Y_0$. Similarly, a favourable supply curve shock shifts the *AS* curve to $AS_2$ and results ultimately in a reduction of the price to $P_4$ and the output returning to the pre shock level of $Y_0$.

When the position of an economy's aggregate supply schedule varies from time to time due to labour militancy, energy price changes, or drought, the nominal GDP can be an attractive intermediate target. To understand the theoretical case for nominal GDP targeting, we need to flesh out the macroeconomic framework we have adopted up to now. So far we have drawn the aggregate supply curve as upward sloping, that is a reflection of the Keynesian argument that nominal wages are set in a staggered manner, or set by agreement for a fixed period and the amount of employment is determined by the employer. Alternatively, the aggregate supply curve could be upward sloping because of mistaken expectations on the part of workers. In both scenarios, nominal wages adjust slowly to unexpected changes in the economic environment. The aggregate supply curve that takes into account expected changes in the environment when nominal wage accords are made can be labeled the full-information aggregate supply curve $AS^{FI}$, which takes into account changes in the economic environment that are expected to occur when wage agreements are endorsed by workers and employers. For expected changes, the applicable supply curve as we know is vertical at the output level, $Y^{FI}$, indicating no direct relationship between the price and output as the expected changes regarding the economic milieu have been factored into the decision making regarding wages and employment and thereby output. Thus, expected changes in the price level will be compensated for by adjustments in nominal wages when workers and firms negotiate wage agreements.

The initial macroeconomic equilibrium, when we include expected changes in the analysis is at point A in Figure 15.10 where *AD* and *AS* intersect the $AS^{FI}$




› **Figure 15.10**
Nominal GDP Targeting. Initial equilibrium is at point A where aggregate demand intersects the short-run aggregate supply curve *AS* and the long-run full information aggregate supply curve $AS^{FI}$. An adverse supply shock shifts the short-run *AS* curve to $AS_1$ and the long-run curve to $AS_1^{FI}$. In the short-run equilibrium is at point B where as prices have risen workers renegotiate contracts and the short-run curve shifts upwards further till it intersects the long-run curve at point C. Similarly, a favourable supply shock will result in short-run equilibrium at point D where the aggregate demand curve intersects the $AS_2$ curve and long-run equilibrium at point E. The initial short-run curve after the shock and the corresponding long-run $AS^{FI}$ curves result in the same level of nominal spending as at point A. Joining these points of intersection such as points H, A, and J, results in the dashed downward-sloping curve giving the locus of points along which nominal spending is the same. A central bank that targets output after an adverse supply shock that results in $AS_1$ will have to increase money supply and shift the aggregate demand curve to pass through point F where the price rises further to $P_F$. If it targets prices then it will have to reduce aggregate demand so that it passes through point G and output declines further than $Y_1$. Nominal GDP targeting at point H avoids these extremes and the effect of the supply shock is distributed between the objectives of price and output stabilization.

curve. Supply shocks, which are unexpected changes in supply, are different from demand shocks as they shift both the short-run and the full-information aggregate supply curves. For example, a negative supply shock shifts both the long-run and short-run aggregate supply curves to the left to $AS_1^{FI}$ and $AS_1$. In the short-run aggregate demand and supply $AS_1$ intersect at point B. As prices have risen, workers renegotiate their nominal wages upward that shifts the short-run supply curve leftward again until it intersects the full-information aggregate supply curve $AS_1^{FI}$ at point C in Figure 15.10. The FULL-INFORMATION LEVEL OF OUTPUT has declined because the adverse supply shock had reduced the productive capacity of the economy, and by definition, the full-information aggregate supply is that output that is produced when the economy has completely adjusted to any shock so that there are no unexpected changes in the economic environment affecting economic decisions that impinge on output. The reduction in the full-information aggregate supply can be temporary in which case the full-information output level will return to $AS^{FI}$ after the shock disperses. In such a case, the decline in potential output is transitory.

› The full-information aggregate supply curve is vertical and established at the level of output where expectations are realized and the economy has completely adjusted to any shocks.

The situation for a positive supply shock is symmetric with the full-information level of output and the short-run aggregate supply curve shifting to the right. As the shock is unexpected, the initial equilibrium is at point D and after wage adjustments have been made, the economy produces at the full-information level of output at point E. The intersection of the short-run aggregate supply and full-information output curves occurs at the same level of nominal spending (see the proof in Appendix 15.1). This implies that the value of $P_0Y_0$ at point A, $P_1Y_1$ at point H, and $P_2Y_2$ at point J are identical. The dashed line is the locus of all the points of intersection of the short-run aggregate supply and full-information output after a supply shock, with the nominal wage held constant. As these points of intersections are points of identical nominal spending, this dashed line is a rectangular hyperbola. With the product of *P* and *Y* along this dashed line equal to a constant, *k*, the equation of the dashed line is

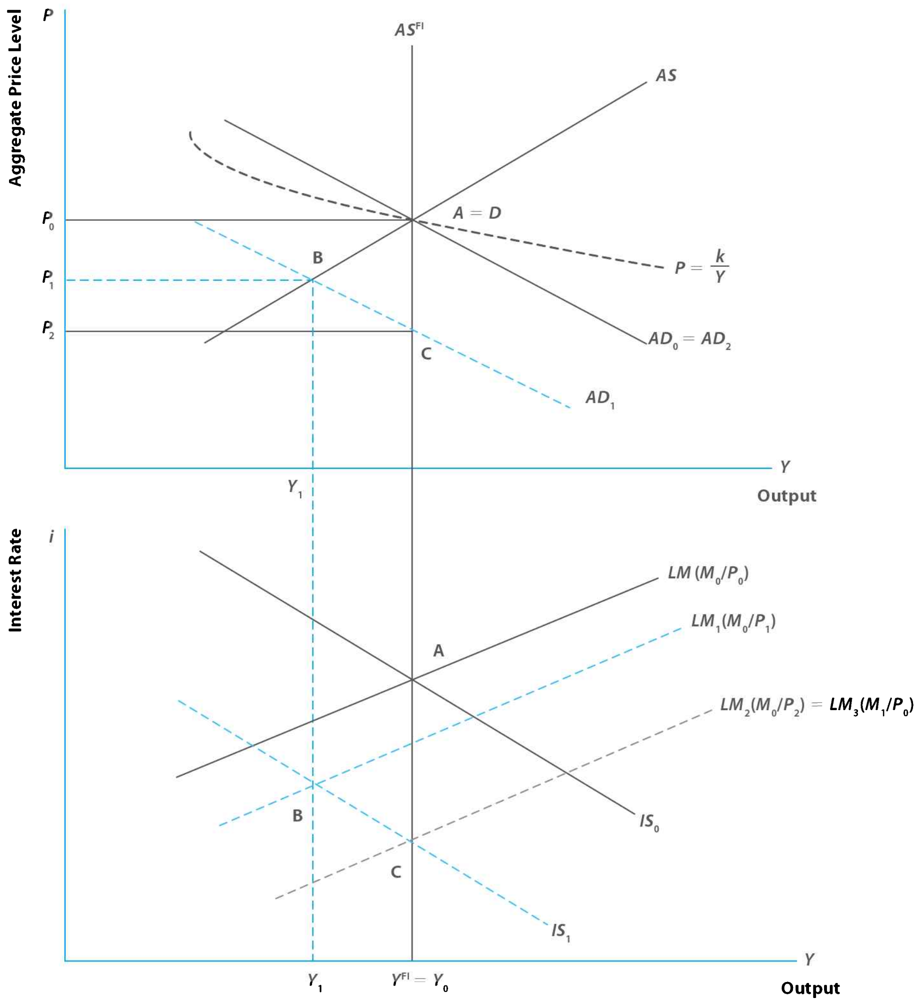
$$P = \frac{k}{Y}$$

Now, consider the case of a negative supply shock that in the absence of a policy response would shift the configuration of the economy from point A to B in the short run and then to point C in the long run. Both ultimate objectives for price and output have been negatively affected. If the central bank attempts to return to the targets on one of these aggregates—either $P_0$ or $Y_0$—it will find that this is possible only at the cost of an even greater target deviation on the other aggregate. An expansionary monetary policy, for instance, may restore the output to $Y_0$ by shifting the aggregate demand curve to the right to pass through point F, but this will cause the price level to rise sharply to $P_F$. As the rise in prices is a surprise to workers, they will bargain for a nominal wage hike and the rise in input prices will cause the short-run aggregate supply curve to shift to the left again. If monetary policy perseveres with getting the economy back to the real output level $Y_0$, prices will keep spiraling upwards. If the monetary authorities pursue a price level target instead, then they will pursue a restrictive monetary policy and reduce the demand curve to pass through point G to keep the price level at $P_0$. However, this leads to a further decline in output from that due to initial supply shock and even more undershooting from the target output $Y_0$. Going further, the decline in output results in a rise in unemployment that pushes down nominal wages and shifts the short-run aggregate supply curve to the right, where it intersects with the full-information aggregate supply curve at an even lower price level. The economy may be pushed towards deflation.

› Nominal GDP targeting in response to variations in the aggregate supply stabilizes output and minimizes the inflationary effects of such variations.

Nominal GDP targeting represents a compromise between these two extremes as it distributes the negative effects of the supply shock equally between the two ultimate objectives. With nominal GDP targeting, the intersection of short-run aggregate supply and full-information output after a supply shock results in the product of price and output, $P_1Y_1$, being at the same level of nominal spending as their intersection before the occurrence of the shock, $P_0Y_0 = P_1Y_1$. Thus, nominal GDP targeting requires the economy after a shock to be on the dashed locus of points given by HAJ. This means the central bank must pursue a slightly restrictive monetary policy that is enough to move the economy from point B after the occurrence of the supply shock to point H. As a consequence, the rise in the price level is less sharp and the fall in output ensures that it is at the new full-information long-run output level that is achievable by the economy after the shock has dissipated. The advantage of the nominal income rule can be visualized by noting that the negative supply shock results in the nominal wages that had been agreed before the shock had been realized having become too high. If the central bank allows a temporary increase in the price level, a partial reduction in real wages results so that the decline in output is mitigated. If the central bank held the price level constant, the correction to nominal wages would then take place in the next round of wage negotiations and that means a higher real output loss in the present. If policy is aimed at keeping the real output at the full-information level, nominal GDP targeting is superior in responding to supply shocks in comparison to price level or output targeting.

What about shocks to aggregate demand? A negative demand shock, say, due to falling autonomous investment leads to an unexpected decline in the aggregate demand and lower prices and output on the short-run aggregate supply curve to point B in Figure 15.11. Because of the decline of the price level from $P_0$ to $P_1$, the real money supply increases and shifts the $LM$ curve to $LM_1(M_0/P_1)$. In this situation, if there is no policy response workers and firms will react to the unexpected decline in prices by adjusting nominal wages downwards until the full-information level of output point C is achieved and




**› Figure 15.11**
Demand Shock and Nominal Income Targeting. The economy is initially in equilibrium at point A in the upper panel where aggregate demand intersects the short-run and long-run full information aggregate supply curves. An adverse demand shock shifts the $AD_0$ curve to the left to $AD_1$. The price declines to $P_1$ and output declines to $Y_1$ and this results in the $LM(M_0/P_0)$ curve shifting to the right to $LM_1(M_0/P_1)$ in the lower panel. In response to lower prices at point B workers will revise downwards their nominal wage demands and the short-run aggregate supply curve will shift downwards over time to eventually intersect the $AS^{FI}$ curve at point C in the upper panel. The lower price will cause the *LM* curve to eventually settle at $LM_2(M_0/P_2)$. Alternatively the central bank pursuing a nominal GDP target expands the nominal money supply to $M_1$ and shifts the *LM* curve to $LM_3(M_1/P_0)$ in the lower panel. This shifts the aggregate demand curve in the upper panel to $AD_2$ where it passes through point D which is identical to point A which was the pre shock equilibrium.

the relevant *LM* curve is $LM_2\,(M_0/\,P_2)$. On the other hand, if the central bank pursues a nominal GDP target, then it must expand the money supply and cause the aggregate demand curve to shift upwards until it again intersects the supply curve at the supply curve's point of intersection with the nominal GDP hyperbola given by $P = k/\,Y$. This requires an expansion in the money supply till the *LM* curve shifts to $LM_3(M_1/\,P_0) = LM_2(M_0/\,P_2)$, the aggregate demand curve shifts out to $AD_2 = AD_0$ at point D, and there is a restoration of the original price and output level $P_0$ and $Y_0$, respectively.

If the central bank followed a price level target or an output level target, monetary policy would have been required to respond in an identical manner as in the case of a nominal GDP target. Thus, a central bank that is obliged to target, say, the price level in the short run is automatically obliged to target the real output and also the nominal income. There would be no conflict between the output and employment objectives and the goal of price level stability when an economy is subject to shocks to demand unlike the situation where supply shocks occur. Under a supply shock, however, price level targeting leads to further deviations away from the pre-shock equilibrium position, and nominal income targeting is superior in meeting the ultimate objectives. Thus, for an economy subject to DEMAND AND SUPPLY SHOCKS, nominal income targeting demonstrates a clear advantage over a price-level target. It results in the same performance in the situation of a demand shock and allows a more flexible reaction in the situation of a supply shock.

› Nominal GDP targeting is superior to price level targeting when an economy is subject to variations in aggregate demand and/or supply.

Given its advantage it seems attractive to prescribe a stable nominal income as the short-run goal of monetary policy and continuing to have price stability as a long-term goal. Despite its theoretical advantage, there are implementation difficulties, however. First, a short-run nominal income target can be met with precision if monetary policy is able to react, which, in turn, requires SUFFICIENTLY PRECISE FORECASTING values of nominal GDP. The technical skills of the central bank would have to be of a high order to provide the information at an early stage so that the monetary authority can act on it. Otherwise, instead of fine-tuning the nominal GDP, the central bank will end up aggravating cyclical fluctuations. Second, monetary policy is liable to be held responsible for mistakes in other areas of economic policy that cause a decline in the growth of the real GDP. For the case of a negative supply shock caused by negotiated wage increases that are not in consonance with the underlying fundamentals of the economy, MONETARY POLICY would have to take on some of the blame for fluctuations in output and employment, while workers and firms would be discharged of their responsibility that led to such an outcome in the first place.

› Nominal GDP targeting is difficult to achieve due to the absence of good forecasts for nominal GDP.

› Shocks to the economy, say, emanating from expectational errors, could wrongly be attributed to a monetary policy that targets nominal GDP.

In practice, therefore, most central banks do not explicitly announce a short-term nominal income target but act to accommodate supply shocks in the short run. Rather than being tied to an explicit nominal income target, it gives a central bank the discretion to decide on the short-run orientation of its monetary policy. However, this results in the risk that a central bank may use its discretion to cause surprises in inflation in an attempt to grow the economy. Economic agents realizing this incentive of the central bank will expect an even higher rate of inflation, which is welfare reducing. We now turn to this issue of whether it is better to have rules or to give discretion to monetary authorities.

## 15.5 Rules Versus Discretion in the Monetary Policy

If any rule can be adopted by discretion it seems natural to think that discretion should be superior to rules. After all, with discretion there is flexibility available to policymakers to choose and one such choice is to choose to do what a rule would oblige policymakers to do. The rule is therefore a subset of the policy choices under discretion. The concept of dynamic inconsistency that was introduced into macroeconomics in the rules versus discretion context by Finn Kydland and Edward Prescott[9] and further elaborated by Robert J. Barro and David B. Gordon,[10] however, showed that a monetary authority that operates under a discretionary regime can be trapped by its own good intentions and generate excessive inflation, which has no offsetting benefit.

› A policy that is optimal from the vantage point of the announcement date may not subsequently be optimal from the vantage point of a later date and the policymaker then has an incentive to deviate from it. This is dynamic or time inconsistency.

DYNAMIC OR TIME INCONSISTENCY occurs when a future policy decision that forms part of an optimal plan formulated at an initial date is no longer optimal from the viewpoint of a later date, even though no new information or event has appeared in the meantime. An academic example provides some initial insight into the nature of the problem.

As is well known in academic institutions, the purpose of an examination apart from ranking students' abilities is to induce students to learn and study. At the beginning of a term it is, therefore, optimal to announce that there will be a final examination. In anticipation of this examination, students will study and learn more. On the examination day, students arrive to the examination having put in effort towards absorbing the study material. The purpose of pre-announcing the examination that was to get students to study has been achieved. Everyone is now better off if the professor cancels the examination and gives each student some satisfactory grade—students are spared the anxiety of giving an examination and can use the examination time for something else and the professor is spared the hassle of evaluating examination papers. With students having studied, the time-inconsistent behaviour by the professor in canceling the examination is optimal. However, students with rational expectations who know that a professor has discretion will realize that they have an incentive to cancel the examination at the end of the term and so may not put in adequate effort into their studies. Discretion then is a trap and it may be better if the announced examination must take place—as a rule, no examination can be cancelled.

### *15.5.1* The Social Welfare Function of Policymakers

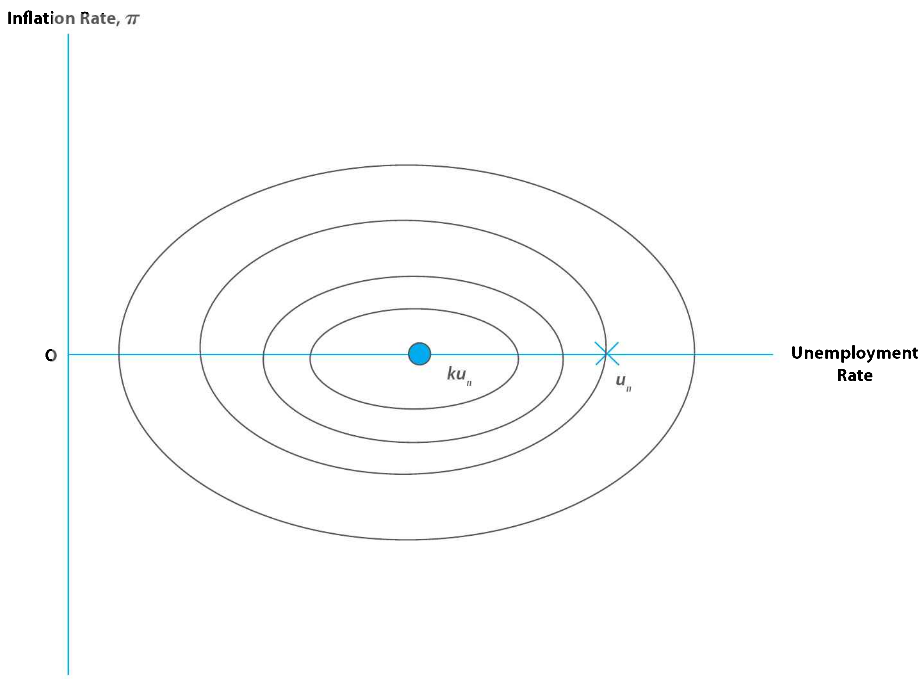
To demonstrate the advantage of rules, we follow the Barro and Gordon[11] approach that describes the strategic interaction between the central bank and the private sector. The model assumes that the central bank has the instruments to perfectly manage the economic process and its decisions are based on a social welfare function that corresponds to the preferences of private individuals. The social welfare function is formulated as a cost or loss function. Social loss, $L$, is considered to be a weighted average of the deviation of inflation ($\pi$) from the target level of inflation, which is taken to be zero, and the deviation of unemployment from the target unemployment rate. Thus,

$$L = \pi^2 + b(u - ku_n)^2, \qquad b > 0, 0 < k < 1 \tag{15.1}$$

Social costs are due to two factors:

1. The unemployment rate deviates from a target unemployment rate $ku_n$, where $u_n$ is the natural rate of unemployment. This gives rise to social costs irrespective of whether the deviation is positive or negative. The deviation is squared in Eq. (15.1) because both positive and negative deviations are costs. Also, squaring of the deviation implies that large deviations are weighed more heavily than small ones. When the unemployment rate is too high it implies that output is too low[12] and this results in a loss in social welfare. Too low an unemployment rate is also evaluated negatively because this implies mistaken expectations about prices. As the natural rate of unemployment can be too high due to allocative distortions, it is not considered as a target value for policy to achieve. $k$ is referred to as an "efficiency criterion" by Barro and Gordon.[13] They argue that the imposition of taxes or payment of unemployment benefits makes labour supply more expensive relative to consumption goods and individuals will respond by consuming more leisure. The aggregate supply of labour is lower than it would be without the corresponding state intervention in the labour market.[14] Alternatively, as argued by Canzoneri,[15] the labour supply curve includes only union members and wage setters' behaviour and systematically excludes other workers. In such a situation, the real wages are too high because of the power of trade unions and therefore employment is too low compared to the level regarded as required by all workers who are represented in the loss function.

**› Figure 15.12**
Loss Function of Policymaker/ Individual Agents. **Policymakers target an unemployment rate given by $kU_n$ and an inflation rate of zero, which is the bullseye point on the horizontal axis. The ellipses around this point are social-indifference curves. Along any one ellipse or social indifference curve the social loss of trading off inflation and unemployment that is not on target is constant. Social-indifference curves further away from the bullseye represent higher social losses.**




2. The actual inflation rate ($\pi$) deviates from the target inflation rate assumed to be zero. Positive inflation is an inflation tax on money balances and negative inflation impairs the unit of account function of money that requires adjustments to all nominal magnitudes. Wage and salary settlements have to be renegotiated, and firms have to post new prices (menu costs).[16]

Social costs will be minimal and welfare maximal when neither of the two deviations with respect to target unemployment and inflation occurs. However, deviations from the unemployment target are weighted differently from those associated with deviations from the inflation target, $b > 0$. When $b$ is large, a high priority is given to the unemployment target, and low values of $b$ indicates a high preference for price stability.

The Eq. (15.1) in graphical terms results in ellipsoids in ($\pi$, $u$) space (see Figure 15.12). The target unemployment and inflation rate is measured by the "bull's-eye" point. Ellipses around this point $(0, ku_n)$ represent social indifference curves with curves further away from the target $(0, ku_n)$ representing higher social losses or costs. Social costs increase with the distance from $(0, ku_n)$. The weighting factor $b$ will determine the actual shape of the indifference curve. The more the weight that is given to deviations from the desired unemployment rate ($ku_n$)—the larger is $b$—the closer to the origin will the points of intersection of the indifference curves with the $x$-axis lie.[17]

### *15.5.2* The Private Sector and Central Bank Policy

The conditions of interaction between policymakers and private agents may now be specified. Private individuals form their expectations regarding inflation and on the basis of this nominal wage settlements are made, which they cannot renegotiate till the end of the game of strategic interaction. The monetary authority, however, has discretion at all times. Due to expectational errors of wage setters, there are deviations of unemployment from the natural rate as given by an expectations-augmented Phillips curve[18]

$$\pi - \pi^e = \alpha(u_n - u) \tag{15.2}$$

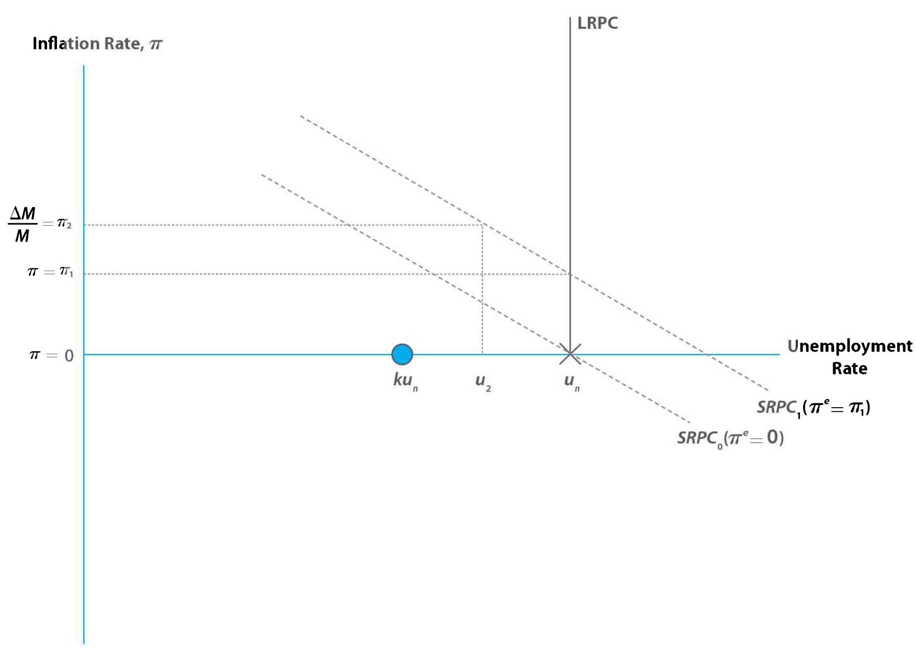
Once individual agents form their inflation expectations and commit to a nominal wage, on the basis of this, the policymaker deploys the policy instrument, which is the growth rate of the money supply. Under a simplifying assumption that there are no lags in the effect of money supply growth on inflation, the policymakers have direct and continuous control over the inflation rate.[19] This correspondence between the inflation rate and the growth rate in money supply implies:

$$\frac{\Delta M}{M} = \pi$$

The ultimate target ($\pi$) being identical to the intermediate target ($\Delta M/M$) implies no distinction between a rule geared to the ultimate target and a rule geared to an intermediate target. Thus, we can treat the inflation rate itself as an instrument.

The Phillips curve $\pi - \pi^e = \alpha(u_n - u)$ is depicted in Figure 15.13. When $\pi = \pi^e$, we have $u = u_n$ in Eq. (15.2). That is, the natural rate of unemployment is that rate where inflation expectations are realized and there is no unexpected inflation. The condition $u = u_n$ when $\pi = \pi^e$ is also the description of the Long-run Phillips curve (LRPC), which is depicted in Figure 15.13 as the vertical curve passing through $u = u_n$. The expected rate of inflation, $\pi^e$, associated with any Short-run Phillips curve (SRPC) is then the $Y$-coordinate of the point at which the SRPC described by Eq. (15.2) crosses the vertical LRPC. Thus, when private agents expect the inflation rate to be $\pi^e = \pi_1$, the relevant short-run Phillips curve is $SRPC_1(\pi^e = \pi_1)$. When the expected inflation rate is zero, $\pi^e = 0$, the relevant short-run Phillips curve is $SRPC_0(\pi^e = 0)$.

Given the public's expectations of a particular rate of inflation, the monetary authority and the public realize, due to perfect information, which SRPC



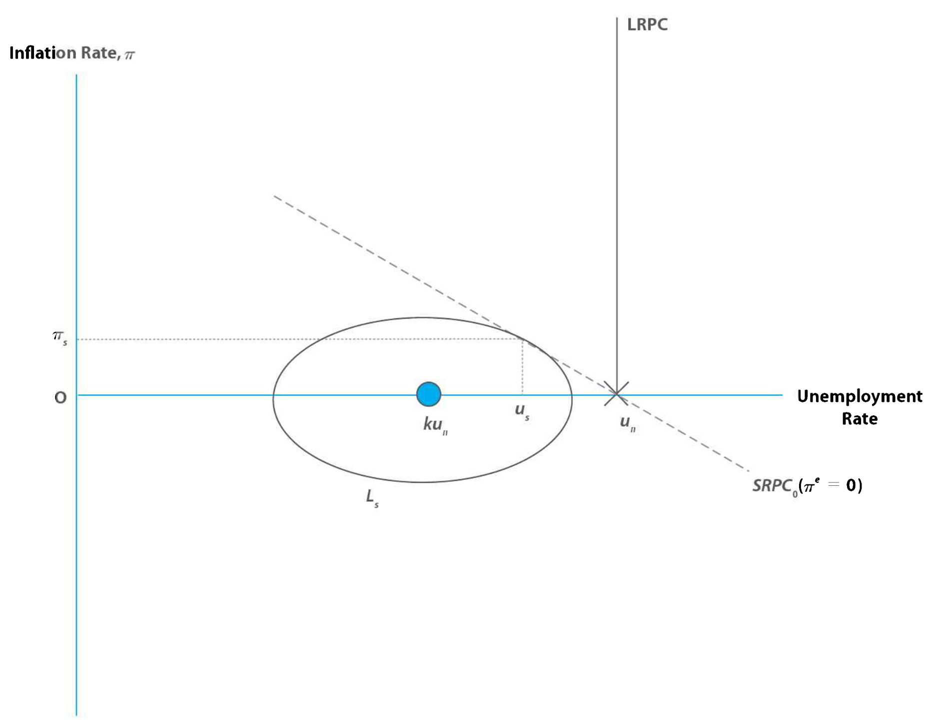

**› Figure 15.13**
Phillips curves. The short-run Phillips curve (SRPC) for a given expected inflation is an inverse relationship between the inflation and the unemployment rate. Higher expected inflation $\pi^e = \pi_1 > \pi^e = \pi_0$ shifts the SRPC to the right. In the long run, the Phillips curve is vertical at LRPC and passes through the natural rate of unemployment $u_n$. Policymakers control the inflation rate, which is equal to the growth in money supply. If the monetary authority selects a growth of money supply equal to an inflation rate of $\pi_2$ when the agents are on the short-run Phillips curve $SRPC_1$, then the unemployment rate is given by $u_2$.

the economy is operating on. The monetary authority can then choose an inflation rate $\pi$ (equal to money supply growth rate) and as a result the associated unemployment rate. For instance, if expected inflation is $\pi^e = \pi_1$ and the private sector is operating along $SRPC_1(\pi^e = \pi_1)$, the monetary authority can select an inflation rate of $\pi_2$ and by doing so will obtain an outcome for unemployment given by $u_2$ (see Figure 15.13).

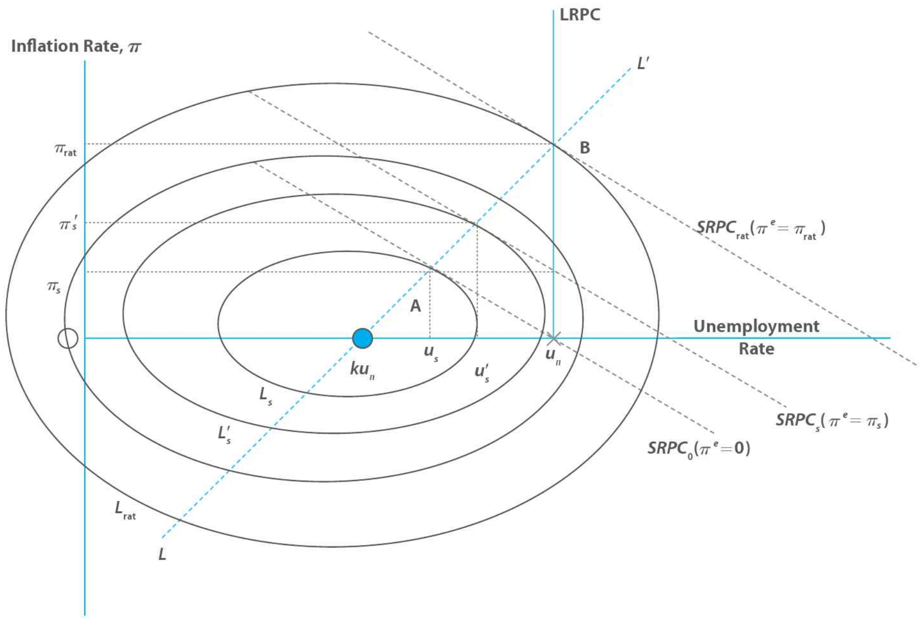
### *15.5.3* Discretion and Surprise (Unanticipated) Inflation

Now consider what happens in a discretionary regime. Suppose the central bank announces a policy target of zero inflation and private individuals go by this announcement and expect inflation to be zero, $\pi^e = 0$. They then set wages on the basis of an expected inflation of $\pi^e = 0$ and the economy is on the short-run Phillips curve $SRPC_0(\pi^e = 0)$. The central bank, however, has complete discretion at all times, it has a free hand to maximize social welfare (minimize $L$) at each point in time, while treating past events as given. Given that the private sector is locked in to operate along $SRPC_0(\pi^e = 0)$, the central bank will find that though it had originally announced to target $\pi = 0$, it can achieve a lower value of the loss function as given by the indifference curve $L_s$ that is tangential to $SRPC_0(\pi^e = 0)$ (see Figure 15.14).

The central bank exploits the lock-in as the private sector can only renegotiate wage contracts with a lag and so it surprises the private sector by increasing the money supply and inflation to $\Delta M/M = \pi_s > 0$ in order to generate an economic boom—the demand for labour increases with inflation as real wages decline—and reduce the unemployment rate to $u_s$. The central bank's original announcement of an inflation rate of zero is, therefore, time inconsistent and it has an incentive to surprise the economy by changing the inflation target to a higher level $\pi_s$ subsequently.




**› Figure 15.14**
Surprise Inflation. The central bank announces a policy of targeting zero inflation and agents set wages so that they are on the short-run Phillips curve $SRPC_0$. The central bank, which has discretion over policymaking, then finds that it can target an inflation rate $\pi_S > 0$, which puts it on a lower indifference curve representing lower social loss that is tangential to $SRPC_0$. This is where inflation expectations are unchanged. The higher inflation rate reduces real wages and unemployment declines to $u_S$. The central bank's original announcement of zero inflation as a target is time inconsistent.




› **Figure 15.15**
Rational Expectations Outcome. When the central bank announces an inflation target of zero, private agents will expect it to renege on this announcement and will expect instead an inflation rate of $\pi_S$ given by the tangency between the short-run Phillips curve $SRPC_0$ and the central bank's indifference curve representing a social loss of $L_S$. With inflation expectations $\pi^e = \pi_S$ the new short-run Phillips curve will be $SRPC_s$. At this expectation of inflation, private agents can expect the central bank to again surprise them and select an inflation rate of $\pi_s'$, given by the tangency between $SRPC_s$ and the central bank indifference curve representing a social loss of $L_s'$. Only at point B will the central bank have no incentive to deviate from the inflation rate given by $\pi_{rat}$ as the social-indifference curve is tangential to the SRPC at a point through which the LRPC also passes. At such a point, inflation expectations are correct and the central bank cannot surprise agents by choosing a different policy target. Point B represents the rational expectations equilibrium.

› A policy announcement is credible if the agents in the economy believe that the central bank will keep to its pre-announced policy promise of targeting low inflation, even if doing so may result in foregoing short-run economic benefits or imposing short-run costs.

However, the private sector has perfect information and knows the central bank's loss function as well as the unemployment-inflation trade-offs as given by the Phillips curve. They will then set their inflation expectations as *rational expectations* and equal to the central bank's optimal inflation rate of $\pi^e = \pi_s$ in the next time period despite an announcement of $\pi = 0$ now. This means that they will not take the central bank's announcement that it is targeting to achieve an inflation rate of zero as CREDIBLE. Having set $\pi^e = \pi_s$, the new short-run Phillips curve $SRPC_s(\pi^e = \pi_s)$ will be what the private sector will aim to operate on (Figure 15.15).

However, when it chooses to operate on $SRPC_s(\pi^e = \pi_s)$, the private sector will rationally expect the central bank to select the indifference curve $L_s'$ as with surprise inflation $\pi_s'$ the central bank can inflate the economy and achieve a lower unemployment rate $u_s'$ than with the natural rate. The private sector with rational expectations will, of course, foresee the incentive of the central bank to surprise the economy and so in order not to be caught in such a situation will set inflation expectations as rational expectations. They then set $\pi^e$ equal to the central bank's optimal inflation rate from which it has no incentive to deviate and surprise the economy. This occurs at $\pi^e = \pi_{rat}$ at which inflation rate the central bank has no incentive to deviate from the expected rate of inflation (and by consequence also from the natural rate of unemployment). Under discretion, then, the private sector will thwart the attempt by the central bank to renege on its initial announced target of zero inflation with a subsequent upward revision to get to point B by operating along $SRPC_{rat}(\pi^e = \pi_{rat})$. Point B is the only point that is consistent with:[20]

1. correct expectations by the private agents (along the LRPC where $\pi = \pi^e$) and
2. the social cost minimizing the welfare level that the central bank can achieve given that the private sector is operating for a given expectation of inflation along a SRPC (that is, along the dashed line $LL'$).

At any point in time the policymaker is motivated to generate unexpected inflation in order to stimulate the economy. But the private agents understand these incentives in advance and, therefore, form high expectations of inflation. The policymaker then must choose a high rate of inflation just to stay even, that is, in order for the unexpected inflation to be zero. Overall, then, two conditions must be satisfied in equilibrium. First, the private agents' expectations of inflation are correct on average—a rational expectations condition. This occurs along the vertical LRPC. Second, although monetary authorities retain the power to fool private agents via inflation surprises, they are not motivated to exercise this power. This occurs along the dashed $LL'$ line. Monetary authorities will not indulge in surprise inflation when inflation is high enough so that the marginal cost of inflation (a higher inflation rate than expected) equals the marginal benefit from inflation surprises (a lower unemployment rate than the natural rate of unemployment). Only at point B is the marginal cost equal to the marginal benefit of surprise inflation with no incentive for the monetary authority to deviate from this outcome. At this point the inflation rate is high ($\pi = \pi_{\text{rat}}$), but there is no surprise to this high inflation rate. The economy bears the cost of this high inflation and does not receive rewards, in the form of lower unemployment, which would arise from unexpected inflation. The inflation rate at $\pi_{\text{rat}}$ is higher than with surprise inflation $\pi_s$. Also with a correct anticipation by private agents of the inflation rate it is not possible to reduce the unemployment rate below the natural rate. Both the rational inflation rate and the natural unemployment rate will be higher at B than the case of surprise inflation, which is at point A. The social costs at B given by $L_{\text{rat}}$ are accordingly higher than the social costs at point A given by $L_s$.

### *15.5.4* Commitment to a Policy Rule

Barro and Gordon demonstrate that with rational expectations the social costs can be reduced by subjecting the monetary authority to a strict rule that the inflation rate (and also the growth rate in the nominal money supply) must equal the target value of zero

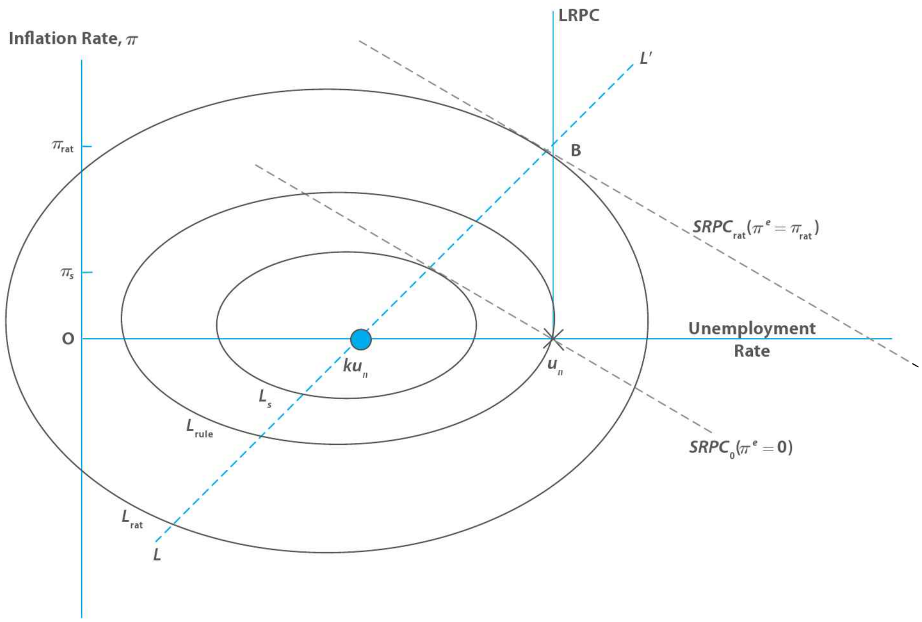
$$\pi_{\text{rule}} = 0$$

Thus, the policymaker must commit *ex ante* to low inflation. The rule must be formulated in such a way that it is absolutely binding—there is some mechanism that prevents violations *ex post*. There must be a mechanism for tying the monetary authorities' hands in advance so that unexpected higher inflation cannot be generated later even if such a choice appears attractive *ex post*. In that case, the commitment is credible and people also anticipate low inflation. Consequently, there are no inflation surprises and so no accompanying social costs to inflation. However, commitment to such a rule deprives the central bank from bringing unemployment down to below the natural rate as is the case with surprise inflation. The social cost then is lower than in the case of rational expectations, or $L_{\text{rule}} < L_{\text{rat}}$ (see Figure 15.16).

The social costs are lowest in the case of surprise inflation and highest for rational expectations, with rules being a midway solution:

$$L_s < L_{\text{rule}} < L_{\text{rat}}$$

The Barro–Gordon model points to the conclusion that an exogenous monetary rule is a preferred policy arrangement even when in principle optimal conditions exist for a discretionary policy. As Barro and Gordon note inflation has




**› Figure 15.16**
Commitment to a Rule. **If the policymaker can commit to a policy rule of zero inflation and it is credible to agents in the economy that there will be no incentive to deviate from this announcement, then inflation expectations will be set equal to zero and the SRPC on which the economy will operate will be $SRPC_0$. The social loss from such a commitment to a rule is lower than the loss from rational expectations due to discretion with the policymaker, or $L_{rule} < L_{rat}$.**

historically been higher since central banks abandoned the rules of the gold standard, despite the reality that fiat money central banks have the means to produce the same or lower inflation.

In the description of central banking, Blinder[21] who was a former vice chairman of the Federal Reserve Board argues that the descriptions of policymaking in the Barro–Gordon model bears little resemblance to the process of formulating monetary policy in the United States and many other countries. He argues that the whole inflation bias depends on $0 < k < 1$ or an unemployment target of monetary policy $ku_n < u_n$. According to Blinder, who has been a practicing central banker as well as an academic, monetary policy is not viewed in terms of the temptation to move up the SRPC in order to lower the rate of unemployment below the natural rate. In fact, central banks are extremely concerned about the potential inflationary consequences of pushing unemployment too low. A more accurate description is that the central bank targets the natural rate of unemployment and so the problem of time inconsistency flowing from a target below the natural rate does not arise.

However, this is true for central banks that are insulated from the political process where there is no pressure for monetary policy to follow an activist employment policy. If myopic politicians who are bothered about voter behaviour during elections have a say in monetary policy, $k < 1$ is an effective description of economic reality. In fact, many central banks in emerging markets are under a very tight control by the government. In such a situation, the Barro–Gordon and Kydland–Prescott setups provide an important insight into why central bank independence is required in order to pursue time-consistent policies that keep inflation in check. A central bank is independent when it is better able to take a long-term view of the economy and need not pursue policies that may be temporarily beneficial, such as reducing unemployment below the natural rate. A central bank is considered to be independent if

1. its actions are not subject to frequent interference by the legislature or the executive branch of government,

2. it does not have the obligation to finance the government's deficit, and
3. central bankers are appointed for long terms and cannot be replaced at the whims of the executive or legislature.

### 15.5.5 Achieving Credibility

Another important aspect of the Barro–Gordon framework is its identification of credibility as an important determinant of the real economic effects of monetary policy. Monetary policy announcements are regarded as credible by private agents to the extent that they base their inflation expectations on them. If a central bank lacks credibility as we saw, the economy will move to point B in Figure 15.15 where there is high inflation and monetary policy is incapable of influencing the unemployment rate. This would be a case of stagflation, a phenomenon that affected many economies in the first half of the 1980s. A central bank can regain lost credibility by its own endeavour by pursuing a policy of disinflation over a long period. Thus, despite positive inflation expectations, $\pi^e > 0$, the central bank must engage in a policy of price stability,

**MACROFOCUS 15.2**

### Monetary Policy Rules

A good monetary policy is one that is successful in stabilizing inflation around a low average level as well as stabilizing output around potential output. This is referred to as flexible inflation targeting in the monetary policy literature. How can these goals be achieved? Many economists argue that they are best achieved when central banks follow a simple policy rule. Friedman had advocated following a rule prescribing a fixed growth rate for the money supply but the growing instability of money demand in the 1980s caused interest in this rule to wane. Since then there has been a search for alternatives. Some academics have advocated a policy that targets the growth of nominal GDP.* Yet others** have advocated a "Taylor rule" in which interest rates are adjusted in response to movements in output and inflation.

John Taylor of Stanford University, who served as the treasury undersecretary for International Affairs under President George W. Bush, focused on the federal funds rate that the interest rate banks charge each other when they loan funds they have on deposit at the Fed. Taylor's rule states that the monetary authority (the Fed) should raise the real interest rate, $i - \pi$ ($i$ is the nominal interest rate and $\pi$ is the rate of inflation), by 1/2 of the excess of the annual rate of inflation over 2 per cent ($\pi - 0.02$). However, the monetary authority should also lower the real rate by 1/2 of the GDP gap—the difference between potential and actual output expressed as a percentage of potential output ($Y^*$)—that is, by $1/2[(Y^* - Y)/Y^*]$. Alternatively, the monetary authority should lower the real rate of interest by 1/2 of the unemployment gap or the difference between the unemployment rate and the natural unemployment rate, that is, by $1/2\,(u - u_n)$. More precisely,

$$i - \pi = 0.02 + \frac{1}{2}(\pi - 0.02) - \frac{1}{2}(u - u_n)$$

$$\text{or} \quad i = 0.02 + \pi + \frac{1}{2}(\pi - 0.02) - \frac{1}{2}(u - u_n)$$

If inflation is $\pi = 0.02$ or 2 per cent, and unemployment is at the natural rate then the real interest rate should be $i - \pi = 0.02$ or 2 per cent and the nominal rate $i = 0.02 + \pi = 0.04$ or 4 per cent. If inflation rises to 3 per cent, that is, by a percentage point, the Taylor rule requires the monetary authority to put the brakes on the economy and to increase the real interest rate by $(1/2)(0.01) = 0.005$ or by 0.5 percentage points, and the nominal interest rate by $i = 0.02 + \pi + 1/2(\pi - 0.02)$ or 5.5 per cent.

Hence, it is not enough for monetary policy to raise the nominal interest rate when inflation threatens—it must raise the nominal rate by more than the rate of inflation so that the real cost of borrowing increases. This is because when inflation increases and the central bank wants to decrease spending and output, it must increase the real interest rate, which is the rate that matters for spending. The Taylor rule also prescribes that if unemployment has increased by 2 per cent but the rate of inflation remains unchanged, then the nominal interest rate should be lowered by 1 per cent.

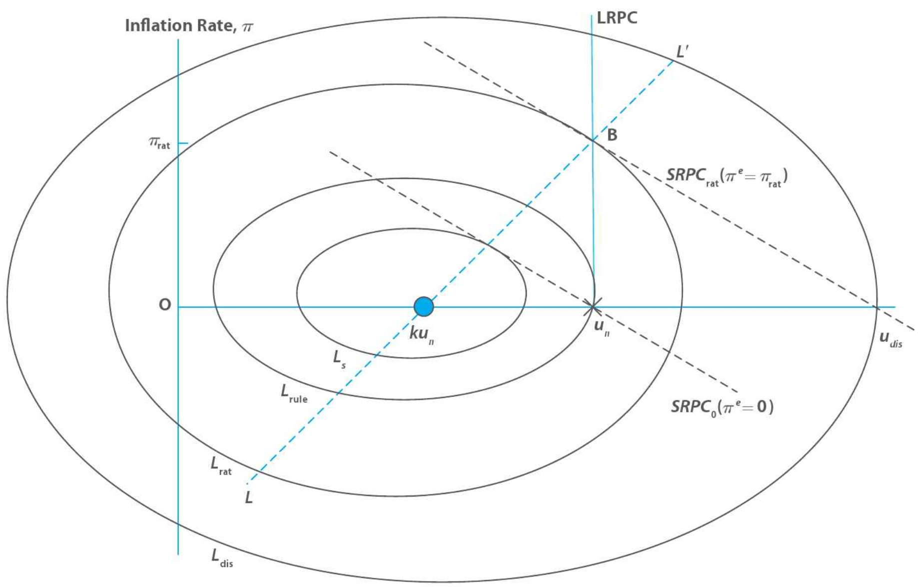
The Taylor rule describes what central banks have been doing for long as "leaning against the wind" or attempting to slow the economy when it is expanding too rapidly and stimulating it when it is chugging along at a slow pace.

In contrast to the expanding literature on rules, no central bank has explicitly adopted a rule and announced it as a guideline. Monetary reform in many countries has focused instead on formulating explicit and precise objectives for monetary policy and creating an institutional setting where the central bank is strongly committed to achieving those objectives. There has been a commitment to objectives rather than to rules.† Also, central banks in the pursuit of output and inflation stability collect and process large amounts of information such as on credit aggregates and asset prices, and exercise considerable judgement in analysing this large number of variables before making decisions on monetary policy. Yet the Taylor rule has been found to be a good description of monetary policy behaviour not only in the United States but also in many other countries. Thus, it has usefulness as a benchmark to assess policy decisions.

*R. E. Hall and N. G. Mankiw, "Nominal Income Targeting," in N. Gregory Mankiw (ed.) *Monetary Policy* (Chicago: University of Chicago Press, 1994); B. T. McCallum, "Specification and Analysis of a Monetary Policy Rule for Japan," *Bank of Japan, Monetary and Economic Studies* 11 (1993): 1–45.

** P. Krugman, "Stable Prices and Fast Growth: Just Say No," *The Economist* (31 August 1996): 15–18.

†L. E. O. Svensson, "What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy Through Targeting Rules," *Journal of Economic Literature* (June 2003), 426–477.




**› Figure 15.17**
Achievement of credibility.
To signal that it is serious about bringing down inflation in the economy the policymaker can select to operate at an exceedingly low rate of inflation that is associated with very high unemployment such as $U_{dis}$. Agents immediately see that by deliberately choosing a level of social loss that is higher than rational, $L_{dis} > L_{rat}$, the policymaker is attempting to regain credibility and that it is willing to accept a high social cost to signal that it intends to not indulge in surprises. This will result in agents to revise their expectations of inflation downwards.

$\pi = 0$. With $\pi < \pi^e$ from Eq. (15.2), $u_n < u$ or there is a decline in output and employment demanded as real wages have risen substantially.

This requires the central bank to provisionally consent to a rise in unemployment to $u_{dis}$ (see Figure 15.17) in order to convince private agents[22] that its announcement of a policy of price stability is to be taken seriously by them. By doing so the central bank adopts the indifference curve $L_{dis}$ and when private agents acknowledge that this high cost of $L_{dis} < L_{rat}$ is being done with a view to invest in credibility, the monetary authority regains its credibility.

Of course, the reinstatement of CREDIBILITY is a long drawn out process requiring the pursuit of a disinflationary policy, $\pi < \pi^e$, for many periods of time. There are other methods a central bank can adopt to convince private agents of its sincerity in controlling and targeting low inflation. One radical and extreme approach has been currency reform, where a currency that has lost its value through inflation is entirely withdrawn from circulation with the anticipation that a new currency will be subject to significantly lower inflation expectations. The German currency reforms of 1923 and 1948 showed that this is possible but similar reforms in Latin America and the Ukraine did not eliminate inflation expectations.[23]

› Credibility can be achieved by (1) withdrawing an existing currency and issuing new notes in circulation, (2) monetary union with countries whose central banks have credibility, and (3) making monetary policy transparent by allowing central bank independence and appointing central bankers who are hawkish about low inflation.

Another quicker way to achieve credibility is monetary union with a country or group of countries whose central bank has the benefit of a high degree of credibility. The former German Democratic Republic managed, for instance, to replace the unpopular East German mark with the globally revered Deutschmark almost in a matter of days. However, often there will be resistance by countries with a stable currency to form a common currency area with countries that have shown themselves to be incapable of controlling their domestic economy.

However, a less drastic way to achieve credibility is to make monetary policy more transparent for private agents so as to enable them to identify a policy of surprise inflation more rapidly and reduce the possible advantage a central bank can derive from such a policy. This is brought about by giving a central bank independence to set its own targets and appointing a central banker who is conservative, who places a great weight on the losses from deviations of inflation from the optimal rate.[24] Such central bankers have a reputation for pursuing price stability with reputation being an indicator to private agents of the expected future policy based on the perceived characteristic of the central banker and their actions observed in the past. As Barro[25] states somewhat tongue-in-cheek, "The ideal central banker should always appear somber in public never tell any jokes and complain continually about the dangers of inflation."

## SUMMARY

- The goals of macroeconomic policy are to keep output and inflation as stable as possible with the output stabilized around the potential or natural level of output and with a low rate of inflation.
- An intermediate target is a macroeconomic variable that the central bank can influence with some predictability, and which is related to its ultimate objectives.
- Intermediate targets are used in monetary policy as a compromise in the absence of complete agreement among policymakers about how best to achieve ultimate objectives and also because policymakers possess limited current information about their ultimate policy goals such as output and inflation.
- Macroeconomic variables selected as intermediate targets must be frequently observable, consistent with ultimate goals, and controllable.
- Central banks have a range of options from which to select intermediate targets: monetary or credit aggregates, interest rates, and nominal GDP.
- In a deterministic environment, it is irrelevant whether it is the interest rate or the money stock that is targeted.
- If the economy is subject to real shocks that shift the *IS* curve, money stock targeting results in less output and price variability than interest rate targeting.
- If the economy is subject to shocks to the demand for money, interest rate targeting results in less variation in output and prices than money stock targeting.
- If the economy is subject to a supply shock, money stock targeting results in less variation in prices than interest rate targeting. However, interest rate targeting results in less variation in the output than does money stock targeting.
- The full-information aggregate supply curve is vertical and established at the level of output where the expectations are realized and the economy has completely adjusted to any shocks.
- Nominal GDP targeting in response to variations in aggregate supply stabilizes output and minimizes the inflationary effects of such variations.
- Nominal GDP targeting is superior to price-level targeting when an economy is subject to variations in aggregate demand and/or supply.
- Nominal GDP targeting is difficult to achieve due to the absence of good forecasts for nominal GDP.
- Shocks to the economy, say, emanating from expectational errors could wrongly be attributed to a monetary policy that targets nominal GDP.
- A policy that is optimal from the vantage point of the announcement date may not subsequently be optimal from the vantage point of a later date, and the policymaker then has an incentive to deviate from it. This is dynamic or time inconsistency.
- A policy announcement is credible if the agents in the economy believe the central bank will keep to its pre-announced policy promise of targeting low inflation even if doing so may result in foregoing short-run economic benefits or imposing short-run costs.
- Credibility can be achieved by (1) withdrawing an existing currency and issuing new notes in circulation; (2) monetary union with countries whose central banks have credibility; and (3) making monetary policy transparent by allowing central bank independence, and appointing central bankers who are hawkish about low inflation.

## NOTES

1. Thomas R. Saving, "Monetary Policy Targets and Indicators," *Journal of Political Economy* 75, no. 4 (1967): 446–456.
2. RBI, "Report of the Committee to Review the Working of the Monetary System," Chairman: S. Chakravarty, Mumbai (1985).
3. A central bank refinancing facility can be thought of as an overdraft facility whereby it grants to each credit institution a generous credit line (at a given interest rate) that it can draw on as and when it needs to. Access to this facility depends on banks complying with supervisory regulations in force.
4. Calculated from *Handbook of Monetary Statistics of India* (Mumbai: RBI, 2006).
5. William Poole, "Optimal Choice of Monetary Policy Instrument in a Simple Stochastic Macro Model," *Quarterly Journal of Economics* 84, no. 2 (May 1970): 197–216.
6. The objective of the central bank can be described as

minimizing an expected loss function given by

$$L = E(Y - Y_0)^2 + E(P - P_0)^2$$

7. In Figure 15.5, the intermediate target $i^*$ is selected to achieve the ultimate policy objective of a GDP equal to $Y_0$.
8. There will also be price effects as the leftward shift of the *LM* curve will result in *AD* shifting left and the decline in the price increases real money supply that will shift $LM_1$ to the right again. We have ignored this detail and drawn $LM_1$ as the final curve after the intermediate price changes have occurred.
9. Finn E. Kydland and Edward C. Prescott, "Rules rather then Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 87 (June 1977): 473–492.
10. Robert J. Barro and David B. Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy* 91 (August 1983): 589–610.
11. Ibid.
12. Alternatively, we could write $L = a\pi^2 + (Y - kY^*)^2$, where $Y^*$ is the natural rate of output or can also be interpreted as full employment output.
13. Barro and Gordon, op. cit.
14. Here, the assumption is that while deciding on the supply of labour the individual does not take account of the related increase (in the case of income tax) or decrease (in the case of unemployment benefit) in the supply of public goods.
15. Matthew B. Canzoneri, "Monetary Policy Games and the Role of Private Information," *American Economic Review* 75 (December 1985): 1056–1070.
16. The term menu costs originates from the fact that in inflationary conditions restaurant menus need to be reprinted more often and as this entails fixed costs high inflation rates increase the incentive to change prices quickly.
17. The point of intersection of elliptical indifference curves for loss level $L_0$ with the $X$-axis is $kU_n + \sqrt{L_0/b}$. The indifference curve intersects the $Y$-axis at $\sqrt{L_0}$, which is independent of $b$. If $b = 1$, we get circular indifference curves.
18. See Eq. (7.25) of Chapter 7. Note that this is in another form the output supply function $Y = Y^* + b(\pi - \pi^e)$.
19. We are also abstracting from the temporary effect of a change in money growth on real income. The economy is always on its long-run money demand curve and the inflation rate equals its steady-state equilibrium value. These are useful abstractions that lay bare the essential ingredient of time inconsistent behaviour.
20. In game theory parlance point B is the Nash equilibrium. This is akin to a prisoner's dilemma game where the strategies that are optimal for the central bank and the private sector lead to an outcome detrimental to both of them.
21. Alan S. Blinder, *Central Banking in Theory and Practice* (Cambridge, MA: MIT Press, 1998).
22. The subscript in $U_{dis}$ stands for disinflation.
23. Michael Bruno, *Crisis, Stabilization and Economic Reform: Therapy by Consensus* (Oxford: Clarendon Press, 1993).
24. Kenneth Rogoff, "The Optimal Degree of Commitment to an Intermediate Monetary Target," *Quarterly Journal of Economics* 100, no. 4 (1985): 1169–1189.
25. Robert J. Barro, *Getting It Right-Markets and Choices in a Free Society* (Cambridge, MA: MIT Press, 1996), pp. 58.

## TEST YOURSELF

1. What are intermediate targets of monetary policy and why are they employed by monetary authorities?
2. What are the alternative intermediate targets of monetary policy that have been suggested and used by central banks?
3. Why are the choices of money stock and the interest rate irrelevant as intermediate targets in a deterministic environment, and which is superior in a stochastic environment subject to demand shocks?
4. When supply shocks prevail, is it better to target the interest rate or the money supply and why?
5. What is nominal GDP targeting and in what way does it smoothen out the conflict between the ultimate price level and output targets?
6. What is time inconsistency? In what way is it an outcome of discretionary policymaking?
7. What are the properties of the social welfare function that form the basis of the central bank policy?
8. In what way may a central bank attempt to inflate an economy once the private sector forms its expectations through discretion in monetary policy?
9. How do rational expectations account for surprise inflations engineered by a central bank?
10. In what way is commitment to a policymaking rule beneficial when rational expectations prevail? How can a policymaker achieve credibility that a rule will be adhered to?

## ONLINE APPLICATION

1. (a) Go to the home page of the European Central Bank—Internet (URL: http://www.ecb.int).
   (b) Click on the icon "Monetary Policy" on top of the page.
2. Under the Background heading, click on "Objective of Monetary Policy."
3. What are the objectives of monetary policy of the European

Central Bank? Can the objectives be broken up into a primary objective and secondary objectives?

4. Read the sections on "Benefits of Price Stability," "Role of Monetary Policy," and "Implementing Monetary Policy."
5. Now, go to the home page of the Federal Reserve of the United States—Internet (URL: http://www.federalreserve.gov/).
6. Click on the icon "About the Fed" on top of the page.
7. Click on "The Federal Reserve Board."
8. Click on "Purposes and Functions."
9. Click on "Monetary Policy and the Economy."
10. What are the objectives of monetary policy of the US Federal Reserve? Does the Federal Reserve have two primary goals as opposed to the European Central Bank's one primary objective? Does the presence of two goals affect the ability of a central bank to influence economic activity?
11. Read the sections on "How Monetary Policy Affects the Economy," "Limitations of Monetary Policy," and "Guides to Monetary Policy."
12. Now go to the home page of the Reserve Bank of New Zealand— Internet (URL: http://www.rbnz.govt.nz/).
13. Click on the icon "Monetary Policy" and read the Introduction that states the objective of monetary policy.
14. Click on the icon "About Monetary Policy" and then on "Explaining Monetary Policy."
15. Go to the section on "Monetary Policy Implementation in New Zealand" and read more about the objectives of the monetary policy in New Zealand that explains what is called the "Policy Targets Agreement" in New Zealand, which binds the Governor of the Reserve Bank to keep inflation between 1 per cent and 3 per cent per year on average. Also read the last section on "Other Roles for Monetary Policy."
16. Next click on the icon "Monetary Policy accountability and monitoring." Read the section titled "The Statutory Provisions." What does this state about the accountability of the Governor of the Reserve Bank of New Zealand? Can they be sacked for non-performance with regard to a target set down in the Policy Targets Agreement?
17. Put together the reading you have done on these three central banks. Are the goals of monetary policy different across these central banks? Why would this be the case? Are some central banks more independent of the legislatures of their countries and of political interference? How much discretion do central banks have and what are the means by which they seek to achieve their objectives?

# Appendix 15.1 Nominal Spending, the Short-Run and the Long-Run Full-Information Supply Curve

We follow Bean[1] in a case that firms are assumed to face a Cobb–Douglas technology

$$Y_t = L_t^{1-\alpha} e^{u_t}$$

where $u_t$ is a productivity shock,
$Y_t$ is output, and
$L_t$ is employment.
The capital stock has been normalized to unity. Taking logs of both sides,

$$\ln Y_t = (1 - \alpha) \ln L_t + u_t$$

or

$$y_t = (1 - \alpha) l_t + u_t$$

where the lower-case letters denote logarithms. The labour demand in a competitive market is given by the usual marginal productivity condition, or

$$\frac{\partial Y_t}{\partial L_t} = (1 - \alpha) L_t^{-\alpha} e^{u_t}$$

or

$$\ln\left(\frac{\partial Y_t}{\partial L_t}\right) = \ln(1 - \alpha) - \alpha \ln L_t + u_t$$

With the real wage equal to the marginal product of labour, we have

$$w_t - p_t = b - \alpha l_t^d + u_t$$

where $w_t = \ln W_t$, $p_t = \ln P_t$, $b = \ln(1 - \alpha)$. At $t - 1$, firms and workers negotiate a nominal wage, before the state of uncertainty about productivity and aggregate demand shocks is revealed, which they expect will equate the labour demand and supply in period $t$. The nominal wage for period $t$ is thereby set in period $t - 1$ and employment in period $t$ is determined by firms, $L_t = L_t^d$. The supply of labour is given by

$$w_t - p_t = c + d l_t^s$$

Equating demand and supply, we have

$$b - \alpha l_t^d + u_t = c + d l_t^s$$

or

$$(\alpha + d) l_t = b - c + u_t$$

or

$$l_t = \frac{b - c}{\alpha + d} + \frac{u_t}{\alpha + d}$$

Hence, market-clearing wages are given by

$$w^* = p_t + c + \frac{d(b - c)}{\alpha + d} + \frac{d u_t}{\alpha + d}$$

[1] Charles R. Bean, "Targeting Nominal Income: An Appraisal," *Economic Journal* (December 1983): 806–819.

Workers and firms choose $w_t$ to minimize the variance of $(w_t - w_t^*)$ where $w_t^*$ is the wage that would clear the market at time $t$ if it were a spot market.

$$\text{Min}E_{t-1}(w_t - w_t^*)^2$$

which implies that $2w_t(w_t - w_t^*) = 0$

or $$w_t = w_t^* = p_t + \frac{\alpha c + cd + bd - cd}{\alpha + d} + \frac{d}{\alpha + d}u_t$$

or $$w_t = w_t^* = p_t + \frac{\alpha c + bd}{\alpha + d} + \frac{d}{\alpha + d}u_t$$

Hence, the wage will be set equal to its expected market-clearing level,

$$w_t = E_{t-1}p_t + \phi_0 + \phi E_{t-1}u_t$$

where $\phi_0 = (\alpha c + bd)/(\alpha + d)$ and $\phi = d/(\alpha + d)$. Substituting this into the demand for labour equation,

$$E_{t-1}p_t + \phi_0 + \phi E_{t-1}u_t - p_t = b - \alpha l_t + u_t$$

From the production function, we have

$$l_t = \frac{y_t}{1-\alpha} - \frac{u_t}{1-\alpha}$$

or $$-\alpha l_t = \frac{-\alpha}{1-\alpha}y_t + \frac{\alpha}{1-\alpha}u_t$$

Substituting into the above expression,

$$E_{t-1}p_t + \phi_0 + \phi E_{t-1}u_t - p_t = b - \frac{\alpha}{1-\alpha}y_t + \frac{\alpha}{1-\alpha}u_t + u_t$$

or $$E_{t-1}p_t + \phi_0 + \phi E_{t-1}u_t - p_t = b - \frac{\alpha}{1-\alpha}y_t + \frac{\alpha + 1 - \alpha}{1-\alpha}u_t$$

or $$E_{t-1}p_t + \phi_0 + \phi E_{t-1}u_t - p_t = b - \frac{\alpha}{1-\alpha}y_t + \frac{1}{1-\alpha}u_t$$

Hence, $$y_t = \frac{1-\alpha}{\alpha}[(p_t - E_{t-1}p_t) + b - \phi_0 - \phi E_{t-1}u_t] + \frac{1}{\alpha}u_t$$

If we set $\beta = (1 - \alpha/\alpha)$, then $1 + \beta = (1 - \alpha/\alpha) + 1 = 1/\alpha$ and we may write

$$y_t = \beta[(p_t - E_{t-1}p_t) + b - \phi_0 - \phi E_{t-1}u_t] + (1+\beta)u_t$$

The full-information output $y_t^*$ is the output level that would occur if the nominal wage in period $t$ could adjust fully to the contemporaneous labour-market conditions. This wage $w_t^*$ equates $l_t^d$ and $l_t^s$. Hence, $y_t^*$ is obtained by setting $w_t = w_t^*$ in the labour demand equation and then substituting that result into the equation for the production function:

$$p_t + \phi_0 + \phi u_t - p_t = b - \alpha l + u_t$$

or $$\phi_0 + \phi u_t = b - \frac{\alpha}{1-\alpha}y_t + \frac{\alpha}{1-\alpha}u_t + u_t$$

or $$\frac{\alpha}{1-\alpha}y_t = b - \phi_0 - \phi u_t + \frac{\alpha}{1-\alpha}u_t + u_t$$

or $$\frac{\alpha}{1-\alpha}y_t = b - \phi_0 + \frac{1 - \alpha - (1-\alpha)\phi + \alpha}{1-\alpha}u_t$$

or $$\frac{\alpha}{1-\alpha}y_t = b - \phi_0 + \frac{1 - (1-\alpha)\phi}{1-\alpha}u_t$$

Hence, $$y^* = \frac{1-\alpha}{\alpha}(b - \phi_0) + \frac{1}{\alpha}[1 - (1-\alpha)\phi]u$$

or $$y^* = \beta(b - \phi_0) + (1 + \beta)\left[1 - \left(\frac{\beta}{1 + \beta}\right)\phi\right]u$$

or $$y^* = \beta(b - \phi_0) + (1 + \beta)\left[\frac{1 + \beta - \beta\phi}{1 + \beta}\phi\right]u$$

or $$y^* = \beta(b - \phi_0) + [1 + \beta(1 - \phi)]u$$

Thus, the deviation of output from its full-information level is given by

$$\begin{aligned} y_t - y^* &= \beta[(p_t - E_{t-1}p_t) + (b - \phi_0) - \phi E_{t-1}u_t] + (1 + \beta)u_t - \beta(b - \phi_0) - [1 + \beta(1 - \phi)]u_t \\ &= \beta(p_t - E_{t-1}p_t) - \beta\phi E_{t-1}u_t + (1 + \beta)u_t - (1 + \beta)u_t + \beta\phi u_t \\ &= \beta[(p_t - E_{t-1}p_t) + \phi u_t - \phi E_{t-1}u_t] \end{aligned}$$

The optimal policy requires the manipulation of demand so that the innovation to prices offsets the effects of the innovation in the productivity shock.

From $$y_t = \beta[(p_t - E_{t-1}p_t) + b - \phi_0 - \phi E_{t-1}u_t] + (1 + \beta)u_t$$

$$E_{t-1}y_t = \beta(b - \phi_0 - \phi E_{t-1}u_t) + (1 + \beta)E_{t-1}u_t$$

Therefore, $$y_t - E_{t-1}y_t = \beta(p_t - E_{t-1}p_t) + (1 + \beta)(u_t - E_{t-1}u_t)$$

or $$(1 + \beta)(u_t - E_{t-1}u_t) = y_t - E_{t-1}y_t - \beta(p_t - E_{t-1}p_t)$$

Now, let the nominal income be given by $X_t = P_tY_t$, or

$$x_t = p_t + y_t$$

Then $$y_t = x_t - p_t$$

and $$E_{t-1}y_t = E_{t-1}x_t - E_{t-1}p_t$$

Hence, $$y_t - E_{t-1}y_t = x_t - E_{t-1}x_t - (p_t - E_{t-1}p_t)$$

Substituting into the above expression,

$$(1 + \beta)(u_t - E_{t-1}u_t) = x_t - E_{t-1}x_t - (1 + \beta)(p_t - E_{t-1}p_t)$$

or $$u_t - E_{t-1}u_t = \frac{1}{1 + \beta}(x_t - E_{t-1}x_t) - (p_t - E_{t-1}p_t)$$

Substitute this into the expression for $y_t - y^*$,

$$\begin{aligned} y_t - y^* &= \beta\left[(p_t - E_{t-1}p_t) + \frac{\phi}{1+\beta}(x_t - E_{t-1}x_t) - \phi(p_t - E_{t-1}p_t)\right] \\ &= \frac{\beta}{1 + \beta}\phi(x_t - E_{t-1}x_t) + \beta(1 - \phi)(p_t - E_{t-1}p_t) \\ &= \beta[\alpha\phi(x_t - E_{t-1}x_t) + (1 - \phi)(p_t - E_{t-1}p_t)] \end{aligned}$$

When the short-run aggregate supply curve intersects the full-information aggregate supply curve, we must have $p_t = E_{t-1}p_t$ and $y_t = y^*$ as with prices as expected, there are no mistaken expectations and the short-run and full-information aggregate supply curves intersect. Then, the above equation can only be true if $x_t = E_{t-1}x_t$ that occurs when the nominal income targeting sets the value of $(x_t - E_{t-1}x_t)$ equal to zero. Hence, this proves the result that the intersection of the short-run aggregate supply and the full-information aggregate supply curves occurs at the same level of nominal spending.

# *Appendix 15.2* Derivations of the Equations in Section 15.5 on Rules Versus Discretion

Derivation of the Surprise Inflation and Discretion situation:

$$L = \pi^2 + b(u - ku_n)^2 \quad b > 0 \tag{A15.2.1}$$

$$\pi - \pi^e = \alpha(u_n - u) \tag{A15.2.2}$$

or $$\alpha u = -(\pi - \pi^e) + \alpha u_n$$

Hence, $$u = u_n - \frac{1}{\alpha}(\pi - \pi^e) = u_n - a(\pi - \pi^e)$$

where $$a = 1/\alpha$$

Under discretion, the policymakers sets $\pi$ and the expected inflation is taken as given. Substituting Eq. (A15.2.2) into Eq. (A15.2.1)

$$L = \pi^2 + b[u_n - a[\pi - \pi^e) - ku_n]^2$$

$$= \pi^2 + b[(1 - k)u_n - a(\pi - \pi^e)]^2$$

The central bank would calculate the optimal inflation rate by differentiating the above loss function in terms of $\pi$ and setting the first-order condition equal to zero.

$$L = \pi^2 + b(1 - k)^2u_n^2 + ba^2(\pi - \pi^e) - 2ab(1 - k)u_n(\pi - \pi^e)$$

Hence, $$\frac{\partial L}{\partial \pi} = 2\pi + 2a^2b(\pi - \pi^e) - 2ab(1 - k)u_n$$

Setting this equal to zero,

$$2\pi + 2a^2b\pi = 2a^2b\pi^e + 2ab(1 - k)u_n$$

or $$\pi^* = \frac{a^2b}{1 + a^2b}\pi^e + \frac{ab(1 - k)}{1 + a^2b}u_n$$

If the central bank announces an inflation rate of zero and individual agents consider this to be credible ($\pi^e = 0$), the actual inflation rate will be

$$\pi_s = \frac{ab(1 - k)}{1 + a^2b}u_n$$

This is depicted in Figure 15.14.

The optimal inflation thereby exceeds the expected value of zero as the central bank seeks to achieve a positive employment effect by means of a positive inflation rate.

Derivation of the Rational Expectations, Rules, and Discretion Outcomes: Under perfect information, the private sector knows the Phillips curve and the central bank's loss function. They will accordingly set their inflation expectations as the rational expectations equal to the central bank's optimal inflation rate, $\pi^e = \pi^*$. Then,

$$\pi_{\text{rat}} = \frac{a^2b}{1 + a^2b}\pi_{\text{rat}} + \frac{ab(1 - k)}{1 + a^2b}u_n$$

or $$\pi_{\text{rat}}(1 + a^2b - a^2b) = ab(1 - k)u_n$$

or $$\pi_{\text{rat}} = ab(1 - k)u_n$$

This is depicted in Figure 15.15.

The inflation rate is higher compared to the case where the central bank deviated from its pre-announced inflation rate of zero in order to inflate the economy, $ab(1 - k)u_n > [ab(1 - k)/1 + a^2b]u_n$. The social cost of discretion is then given by

$$L_{\text{rat}} = \pi^2 + b[(1 - k)u_n - a(\pi - \pi^e)]^2$$

$$= \pi^2_{\text{rat}} + b(1 - k)^2u_n^2$$

as $\pi = \pi^e$ under discretion that corresponds to $u = u_n$. Therefore,

$$L_{\text{rat}} = a^2b^2(1 - k)^2u_n^2 + b(1 - k)^2u_n^2$$

$$= b(1 + a^2b)[(1 - k)u_n]^2$$

This is the indifference curve as depicted in Figure 15.15. With rational expectations, the social costs can be reduced by requiring the central bank to adhere to the strict rule that the inflation rate, or the growth rate in the nominal money supply, equals the target value of zero:

$$\pi_{\text{rule}} = 0$$

Then $$L_{\text{rule}} = \pi^2_{\text{rule}} + b[(1 - k)u_n - a(\pi_{\text{rule}} - \pi^e)]^2$$

$$= b\{(1 - k)u_n\}^2$$

as there is commitment to the rule and $\pi^e = \pi_{\text{rule}} = 0$.

It is obvious that the social costs are lower with a rule,

$$L_{\text{rule}} < L_{\text{rat}}$$

This is depicted in Figure 15.16.

Similarly, if regardless of rational expectations the central bank is able to achieve a stable price level ($\pi = 0$), the disinflation has social cost given by

$$L_{\text{dis}} = b(1 + a^2b)^2[(1 - k)u_n]^2$$

This is depicted in Figure 15.17.

16

# Economic Growth

CRITICAL QUESTIONS

- *What is economic growth?*
- *What are the sources of economic growth?*
- *What are the determinants of the long-run (steady-state) growth?*
- *What is the golden rule of capital accumulation?*
- *Is there a tendency for economic growth to converge or diverge across nations?*
- *Why do economies get caught in poverty traps?*
- *How does the endogenous growth approach explain the growth process?*

# 16.1 Sources of Growth

The father of economics, Adam Smith, was preoccupied with the economic growth of nations and even titled his famous treatise, *An Inquiry Into the Nature and Causes of the Wealth of Nations*. More than 200 years since the publishing of this treatise, we are still grappling with the question of what makes some countries rich and others poor. Recall that economists use per capita income as an indicator of welfare (Chapter 1). The per capita income today varies across countries much more than it used to. Differences in per capita incomes across countries were small until the 19th century. They started to widen with the Industrial Revolution and have got enlarged during the last 100 years.[1] ECONOMIC GROWTH is measured as the rate of change of real per capita income. A country with a growth rate of 1 per cent per annum doubles its living standard every 70 years, whereas a country with a growth rate of 3.86 per cent, as was the case in India during 1990–1991 to 2003–2004, will double its living standard every 18 years. If India can push its growth rate up to 6 per cent, then in a course of 11.5 years, it can double its living standard. This means that differences in growth rates that persist can result in dramatic differences in the living standards.

› Economic growth is the rate of change of the real income (or output) per capita.

So what causes the growth rates to be different across countries? Macroeconomists emphasize the accumulation of physical and human capital as having the major influence on economic growth. Examples of physical capital comprise the stock of machines and equipment, whereas examples of human capital comprise the stock of education and training embodied in the labour force.

Robert Solow, however, demonstrated in an influential article in 1956 that accumulation of these factors in response to economic incentives would not lead to continued growth in per capita income. Instead, Solow recognized that growth was due to the invention of new technologies that made labour more productive. As the source of these innovations was unexplained in the model developed by Solow, his theory came to be known as the exogenous-growth theory.

Solow demonstrated that, in the US economy, seven-eighths of the growth in the nation's output per worker is accounted for by changes in technology and one-eighth by increases in the tangible capital per worker. Further research did depict a reduced contribution of technological change but has been unable to challenge Solow's basic result that such change accounts for more than half the growth in output per worker.[2] Moreover, Solow[3] showed that besides the fact that the real output per worker grows at a more or less constant rate on average, there are two other STYLIZED FACTS about long-run economic growth:

› The stylized facts about long-run growth are (1) output per capita grows at a more or less constant rate, (2) the share of labour and capital in income is constant, and (3) the share of consumption in aggregate income is constant.

1. The share of labour in income is more or less constant.
2. The share of consumption in aggregate income has remained approximately constant.

Growth theory begins with the assumption that output is related to the aggregate capital and labour through a production function:

$$Y = AF(K,L) \tag{16.1}$$

where, $Y$ is aggregate output, $K$ is capital, $L$ is labour, and $A$ is a technology parameter.[4] In applied work, a particular functional form for the production function is usually specified. A function that is frequently used is the Cobb–Douglas function:

$$Y = AK^{\alpha} L^{1-\alpha} \tag{16.2}$$

› The neoclassical Solow growth theory assumes that the economy is characterized by perfect competition with firms as price takers, and output given by a constant return to scale production function such as $Y = AK^{\alpha} L^{1-\alpha}$ where $A$ is a technology parameter.

The NEOCLASSICAL THEORY assumes that output is produced by a large number of competitive firms, each of which uses the same production function. Normalizing the price of output to unity,[5] profit-maximizing firms solve the following problem:

$$\begin{aligned} \underset{\{K,L\}}{Max}\, \Pi &= Y - rK - wL \\ &= AK^{\alpha} L^{1-\alpha} - rK - wL \end{aligned}$$

where, $w$ is the wage per unit of labour and $r$ the rental per unit of capital.

The first-order profit-maximizing conditions are the usual ones, the firm should rent capital until the marginal product of capital is equal to the rental price, and hire labour till the marginal product of labour is equal to the wage. Take the profit-maximizing condition with respect to labour hiring,[6]

$$w = (1-\alpha) A \left(\frac{K}{L}\right)^{\alpha} = MP_L$$

Dividing throughout by $L$, Eq. (16.2) gives us $Y/L = A(K/L)^{\alpha}$, which we use to write the above labour-market condition as,

$$w = (1-\alpha)\frac{Y}{L} = MP_L$$

where, $MP_L$ is the marginal product of labour. In this case, the share of labour in aggregate output will be given by the wage bill divided by the aggregate output:

$$\text{Share of Wages} = \frac{wL}{Y} = \frac{MP_L(L)}{Y} = (1-\alpha)$$

The Cobb–Douglas production function, accordingly, reproduces one of Solow's stylized facts that the share of labour in aggregate income is constant implicitly.

Similarly, the profit-maximizing condition for the capital stock is where the marginal cost of a unit of capital, $r$ , equals the marginal benefit of employing that capital given by the additional output that it produces—the marginal product of capital.

$$r = MP_K = \alpha A \left(\frac{K}{L}\right)^{\alpha-1}$$

Again, dividing Eq. (16.2) throughout by $K$ gives us $Y/K = AK^{\alpha-1} L^{1-\alpha} = A(K/L)^{\alpha-1}$. The share of capital in aggregate output is, then, given by

$$\text{Share of Capital} = \frac{rK}{Y} = \frac{MP_K(K)}{Y} = \alpha$$

Solow used these results emanating from the properties of the production function to get a hold on estimating the contribution of technological progress to output growth. The contribution of technological improvements to economic growth was hard to quantify before Solow's pioneering work. Output can be measured by GDP, and we can measure the hours worked in a year and the stock of capital as well. But how does one measure the increased productivity in the economy due to the introduction of Ford-type assembly lines and, more recently, the computer revolution?

› Solow used the production function to argue that the change in output is due to the growth of the labour force, the accumulation of capital stock, and technological progress.

Solow argued that the CHANGE IN OUTPUT is the sum of three components: the increased output due to technological advance, the increased output due to the growth of labour employed, and the increased output due to the availability of additional capital equipment. From Eq. (16.2), this may be written as[7]

$$\frac{\Delta Y}{Y} = \frac{\Delta A}{A} + \alpha\frac{\Delta K}{K} + (1-\alpha)\frac{\Delta L}{L} \quad \textbf{(16.3)}$$

where,

$\frac{\Delta Y}{Y}$ is rate of growth of output;

$\frac{\Delta K}{K}$ is rate of growth of the capital stock;

$\frac{\Delta L}{L}$ is rate of growth of the labour force; and

$\frac{\Delta A}{A}$ is rate of technological progress.

Solow recognized that the only unobservable variable in Eq. (16.3) is the rate of technological progress, $\Delta A/A$. This could be calculated by taking it to the left-hand side of the equation as follows:

$$\frac{\Delta A}{A} = \frac{\Delta Y}{Y} - \alpha\frac{\Delta K}{K} - (1-\alpha)\frac{\Delta L}{L} \quad \textbf{(16.4)}$$

This has come to be known as the SOLOW RESIDUAL—the estimate of technological progress derived as a residual after the contributions of labour and capital growth are subtracted from the rate of growth of output.

› The Solow residual is the estimate of the technological progress derived as a residual after accounting for the rate of growth of output due to labour force and capital stock growth.

MACROFOCUS 16.1

## Growth Accounting—China and India

The Solow residual is also called the total factor productivity (TFP) of an economy. One interpretation that is given is that it is a measure of the increased efficiency with which inputs are used, and the measure captures the technological progress that is taking place in the economy. As it is calculated as a residual by subtracting out from the output growth the contributions of the capital stock and labour, the measure is sensitive to the accuracy with which the factor inputs are measured. Labour input, for instance, is measured by the number of hours worked in a year in the aggregate. However, a person enters the labour force after accumulating the years of schooling and informal education. The more the accumulated years of schooling that are present in the working-age population, the more educated is the labour force, and the greater the efficiency of labour in contributing to output.

In the light of this, the measure of labour input requires to be adjusted for improvements in educational attainment ($H$ representing human capital). Bosworth and Collins* accordingly write the production function as follows:

$$Y = AK^{\alpha}(LH)^{1-\alpha}$$

This leads them to decompose the economic growth into the contribution of growth in capital stock per worker ($k = K/L$), increases in education per worker ($h = H/L$), and the Solow residual.

$$\underbrace{\left(\frac{\Delta Y}{Y} - \frac{\Delta L}{L}\right)}_{\text{Growth rate of per capita output}} = \underbrace{\frac{\Delta A}{A}}_{\text{Total fator productivity growth}} + \alpha\underbrace{\left(\frac{\Delta K}{K} - \frac{\Delta L}{L}\right)}_{\text{Growth rate of capital stock per worker}} + (1-\alpha)\underbrace{\left(\frac{\Delta H}{H} - \frac{\Delta L}{L}\right)}_{\text{Growth in educational attainment per worker}}$$

or, $$\frac{\Delta y}{y} = \frac{\Delta A}{A} + \alpha\frac{\Delta k}{k} + (1-\alpha)\left(\frac{\Delta h}{h}\right)$$

The estimated sources of growth for the period 1978–2004 for China and India are given in Table 1.

**› Table 1**
Sources of Economic Growth

| | Growth Rate of Output per Worker | Contributors | | |
|---|---|---|---|---|
| | | Capital Stock per Worker | Education per Worker | Total Factor Productivity |
| China | 7.3 | 3.2 | 0.2 | 3.8 |
| India | 3.3 | 1.3 | 0.4 | 1.6 |

The sources of China's economic growth, as depicted in Table 1, come from a high contribution of the large rate of capital accumulation per worker (44 per cent of per capita output growth) and from gains in TFP (52 per cent). As Bosworth and Collins point out, there are historical examples of countries achieving growth rates of GDP comparable to China's such as Germany and Japan post–World War II, and Taiwan and Korea more recently. However, none of these economies were able to sustain such rapid growth for long, whereas China's high growth has continued for more than 25 years. In contrast to China, India achieved its growth with relatively more emphasis on education per worker, contributing 12 per cent to growth in contrast to China's 3 per cent, and the TFP and capital accumulation contributed much less—48 per cent and 40 per cent, respectively—to the economic growth in India as compared to that in China.

* B. Bosworth and S.M. Collins, "Accounting for Growth: Comparing China and India," National Bureau of Economic Research, Working Paper no. 12943, Cambridge, MA (2007).

Solow[8] applied Eq. (16.4) to the annual data on the growth rate of output ($\Delta Y/Y$), the share of capital in output $(1-\alpha)$, the growth rate of the labour force ($\Delta L/L$), and the growth rate of the capital stock ($\Delta K/K$). He reported that the rate of technological progress was about 1.2 per cent per annum from 1909 to 1929 and about 1.9 per cent per annum from 1929 to 1949. As economic growth is the rate of change of the output per capita, we can write Eq. (16.3) as follows:

$$\underbrace{\left(\frac{\Delta Y}{Y}-\frac{\Delta L}{L}\right)}_{\substack{\text{Growth Rate}\\\text{of Per Capita}\\\text{Output}}} = \underbrace{\frac{\Delta A}{A}}_{\substack{\text{Rate of}\\\text{Technological}\\\text{Progress}}} + \underbrace{\alpha\left(\frac{\Delta K}{K}-\frac{\Delta L}{L}\right)}_{\substack{\text{Growth Rate of}\\\text{Capital–Labour}\\\text{Ratio}}} \tag{16.5}$$

› Solow demonstrated that the Solow residual or the total factor productivity growth contributes the major share to per capita income growth.

The left-hand side is the growth rate of per capita output,[9] and the term $[(\Delta K/K)-(\Delta L/L)]$ is the growth rate of capital intensity, or the capital–labour ratio, $K/L$. Solow estimated that about seven-eighths of the increase in output per worker over the 40-year period, 1909 to 1949, was due to TECHNOLOGICAL IMPROVEMENT and only one-eighth to an increase in the capital–labour ratio. For the period 1960–1990 for the United States, it is estimated that the contribution of technological progress to the increase in output per worker was 78.6 per cent, representing a drop of about 9 per cent since the time period of Solow's estimate.[10]

## **16.2 Potential and Feasible Output

The aggregate output that is referred to so far is the *potential output*—the output that would result if all the resources, capital, and labour are fully utilized. Strictly speaking, the variable $L$ refers to full employment, and $w$ is the full-employment wage rate,

$$N = L$$

and

$$w = w_f$$

where, $L$ is the active labour force, and $w_f$ is the full-employment real wage. We take the active labour force (those in the age group of 15–62 years) to be equal to the size of the population to make the exposition less cumbersome. Then,[11] $N = L$ implies that everyone in the population is employed—full employment. With full employment, the wage paid, which is equal to the marginal product of labour as in Eq. (16.4), is the full-employment wage. When the labour market clears ($N = L$), the wage associated with that is the full-employment wage.

When labour is not fully employed, $N < L$, and the economy will produce an output that is less than the potential output. As $N < L$, there is some unemployment prevailing in the economy, and the output that is produced in this situation can be called the *feasible output*. It is the output resulting because of the underutilization of the available labour force due to the institutional conditions prevailing in the labour market.

The reason for unemployment in the labour market (as we have seen in Chapter 7) is that the wage demand of workers does not match the wage offer of firms. This is either due to mistaken expectations by workers or because wages are somewhat sticky as workers seek to keep their wages at levels they think to be appropriate relative to other occupations. Alternatively, wages higher than the market-clearing wages could be paid by firms as an efficiency wage to motivate employees or to discourage them from quitting.

Here we pursue the simple approach, which states that if there is an excess supply of labour, pressure is generated by the unemployed to reduce money wages as they are willing to undercut those who are already employed [see the discussion following Eq. (7.8) in Chapter 7]. This relationship between the wages workers settle for and the unemployment rate is referred to as the bargained real-wage curve, or following Blanchflower and Oswald,[12] as the *wage curve*. The bargaining power of workers then depends directly on the rate of employment $(N/L)$, and the real wage negotiated by workers is the following function of the rate of employment:

$$w = \left(\frac{N}{L}\right)^{\beta} = \varepsilon^{\beta} \qquad \text{for} \qquad N < L \tag{16.6}$$

where, $\varepsilon = (N/L) < 1$ is the rate of employment, and $\beta$ denotes the elasticity of negotiated wages with respect to the employment rate. The production function when the employment rate is $\varepsilon$ is given by,

$$Y = AK^{\alpha}\left[L\left(\frac{N}{L}\right)\right]^{1-\alpha} = AK^{\alpha}(L\varepsilon)^{1-\alpha} \tag{16.7}$$

Equilibrium in the labour market is where the wages firms are willing to pay, given by the marginal product of employed labour, equal the wages bargained and sought for by workers as stated in Eq. (16.6). Hence,

$$(1-\alpha)AK^{\alpha}(L\varepsilon)^{-\alpha} = \varepsilon^{\beta} \tag{16.8}$$

or,
$$(1-\alpha)A\left(\frac{K}{L}\right)^{\alpha} = \varepsilon^{\alpha+\beta}$$

or,
$$\varepsilon = \left[(1-\alpha)A\left(\frac{K}{L}\right)^{\alpha}\right]^{1/(\alpha+\beta)}$$

or,
$$\varepsilon = [(1-\alpha)Ak^{\alpha}]^{1/(\alpha+\beta)} \tag{16.9}$$

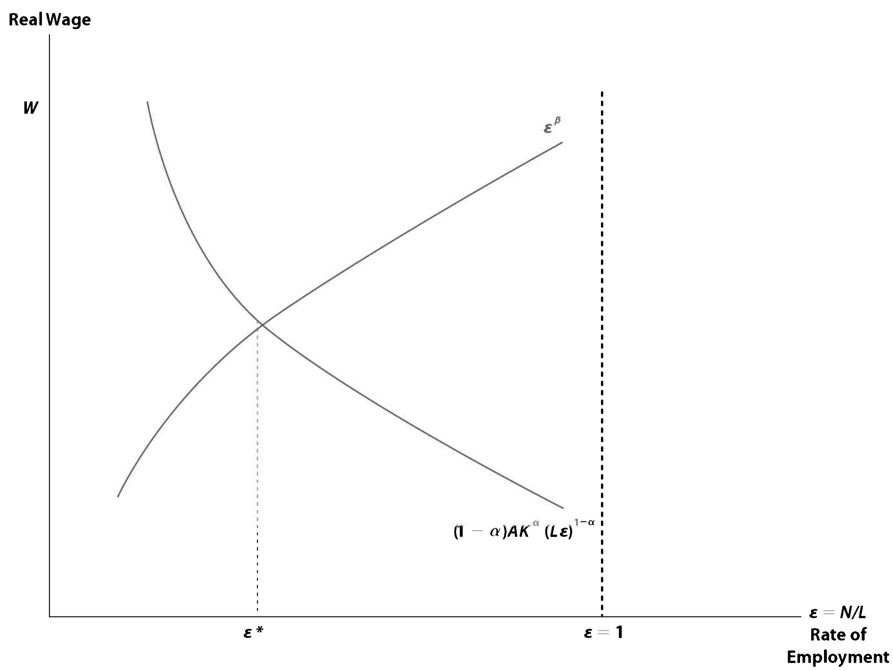
where, $k$ is the capital per worker. Note that $k = K/L$ is not equal to capital per employed worker, $K/N$. The equilibrium employment rate, therefore, depends on the intensity of the capital stock per capita, $k$, the extent to which workers demand for wages responds to the employment rate $\beta$, the technology parameter, $A$, and the share of wages in output $(1-\alpha)$.

The two sides of Eq. (16.8) are depicted in Figure 16.1. The desired wage of the workers equals the wage that is profitable to firms at $\varepsilon^* < 1$, and the unemployment rate in equilibrium is equal to $(1-\varepsilon^*)$. A reduction in the workers' bargaining power—an increase[13] in $\beta$—reduces the ability of workers to negotiate wage rates (the $\varepsilon^{\beta}$ curve shifts down), and firms find it profitable to employ more workers at the reduced wages, thereby increasing employment.

We now substitute the equilibrium level of employment as given by Eq. (16.9) into the production function [Eq. (16.7)] and write it in per capita terms as[14] follows:

$$y = \theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} \tag{16.10}$$

where, $y = Y/L$ is the per capita output, and $\theta = (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}$. Equation (16.10) is the *feasible production function*—the output that is attainable given the amount of capital and labour available as a result of the interaction between the aspirations of workers and the profit-maximizing behaviour of firms. The difference between the potential production function [Eq. (16.2)] and the feasible production function [Eq. (16.10)] is that the former incorporates purely




› **Figure 16.1**
The Equilibrium Employment.
The real wage bargained for by workers is increasing in the employment rate. The real wage paid by firms equals the marginal product of employed labour and is downward sloping. The wage bargain by workers equals the wage that is profitable to the firm at $\varepsilon^*$. The unemployment rate is $1 - \varepsilon^*$.

technical relations between inputs and output—additional input results in additions to output but at a diminishing rate. In addition, the feasible production function incorporates an institutional datum as given by the wage-setting behaviour of workers and the wage-offer behaviour of firms in the labour market. The feasible output of the economy increases with $\beta$—a reduction in the bargaining power and aspiration of workers—which increases employment at the same level of capital and population in the economy. Due to the resources of labour not being fully utilized, the feasible production function will always lie below the potential production function, which fully employs labour: Eq. (16.10) is less than the intensive form of Eq. (16.2).[15]

We restrict our attention to a production function with full employment of resources—the potential production function. Our interest is in the determinants of long-run growth. The unemployment arising from cyclical factors may be overlooked over a long time period where the effect of capital accumulation on the supply of output is explicitly accounted for. Our qualitative results, which give us insights into the process of economic growth, are not affected whether we employ the potential production function or the feasible production function. As expected, the rigidity in the labour market, which was accounted for in the feasible production function, results in the economy producing below its potential in every peiod. This will slow down the time taken by the economy to reach a steady state.

Up to now, one of our central quests in macroeconomics has been to understand why unemployment arises. In growth theory, this quest is sidestepped not because it is unimportant, but because giving centrality to unemployment does not provide additional insight. It does not alter the main conclusions of growth theory about the impact of savings, labour force growth, and technological progress on the growth rate of the economy. Readers who think of unemployment as so central an issue as to require its consideration in a theory of economic growth can read every subsequent result derived for a potential production function, and the corresponding result for the feasible production function in Appendix 16.1.[16] Solow[17] himself had made the following observation in defence of his approach to economic growth, which utilized a potential production function:

> All theory depends on assumptions, which are not quite true. That is what makes it theory. The art of successful theorizing is to make the inevitable simplifying assumptions in such a way that the final results are not very sensitive.

# 16.3 The Determinants of Long-Run Growth

To facilitate a diagrammatic approach to economic growth, we proceed by drawing a graph of the production function in per capita form as economic growth is about the growth of the per capita output. In per capita terms, Eq. (16.2) may be written as follows:

$$y = \frac{Y}{L} = AK^{\alpha} L^{-\alpha} = A\left(\frac{K}{L}\right)^{\alpha} = Ak^{\alpha} \tag{16.11}$$

where, $y = Y/L$ is per capita output, and $k = K/L$ is the capital intensity of production. For given parameters $\alpha$ and $A$, how does $y$ respond to the changes in $k$? The answer to this is obtained by finding out the response of the per capita output ($\Delta y$) to a marginal change in the capital intensity ($\Delta k$), or[18]

$$\frac{\partial y}{\partial k} = \alpha A k^{\alpha-1}$$

This expression is nothing but the expression for the marginal product of capital, $MP_K$, as can be witnessed by writing Eq. (16.11) in extensive form and finding the expression for $MP_K$. In extensive form, Eq. (16.11) is Eq. (16.2):

$$Y = AK^{\alpha} L^{1-\alpha}$$

The marginal productivity of capital from this is[19]

$$MP_K = \frac{\partial Y}{\partial K} = \alpha AK^{\alpha-1} L^{1-\alpha} = \alpha A\left(\frac{K}{L}\right)^{\alpha-1} = \alpha Ak^{\alpha-1}$$

The slope of Eq. (16.11) is, therefore, nothing other than the marginal product of capital. Given that the marginal product of capital is subject to diminishing returns, this translates into the slope of Eq. (16.11) diminishing as $k$ increases. We can, therefore, depict the graph of Eq. (16.11) in Figure 16.2.

Having dealt with the labour market, which is the supply side of the economy, it is now time to turn to the demand side of the economy, which is the familiar product-market condition that, in equilibrium, saving equals investment:

$$S_t = I_t \tag{16.12}$$

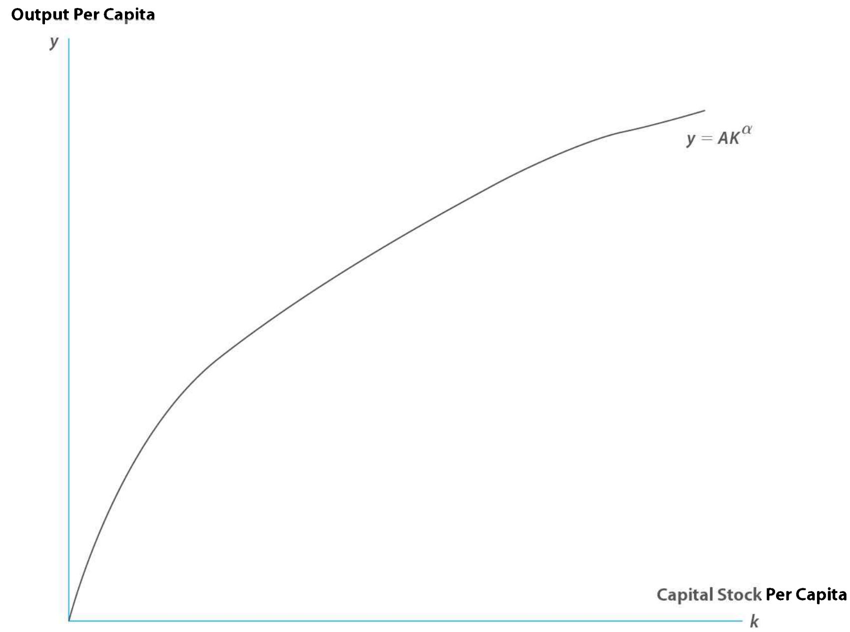
We depict savings behaviour by appealing to the other Solow-stylized fact, which is that the share of consumption and, therefore, saving in aggregate output has been more or less constant:

$$S_t = sY_t \tag{16.13}$$

where, $0 < s < 1$ is the marginal propensity to save.

Finally, we incorporate an accounting identity that investment adds to the capital stock:

$$K_{t+1} = (1 - \delta)K_t + I_t \tag{16.14}$$




**› Figure 16.2**
The Output Per Capita as a Function of Capital Stock Per Capita. **The output per capita increases as the capital stock per capita increases but at a diminishing rate due to the diminishing marginal productivity of the capital.**

Equation (16.14) states that next year's capital stock is equal to the fraction of this year's capital stock that is left over after depreciation plus any new gross investment in capital. $\delta$ represents the rate of depreciation. Finally, the population, which is equivalent to the labour force by assumption[20] is assumed to grow at a constant growth rate, $n$:

$$L_{t+1} = (1 + n)L_t \tag{16.15}$$

Equations (16.11) to (16.15) constitute the Solow growth model, which we, for convenience, state again in Table 16.1.

Recall that growth theory attempts to explain what drives the per capita output growth. Figure 16.2 depicts that changes in per capita output are due to changes in capital intensity ($k$) for the given parameters $\alpha, A$.

The neoclassical growth equation is, therefore, an equation describing the relationship between the capital stock per capita, $k$, in any two successive years. We now go through the steps that enable us to derive this growth equation. First, restate the capital-stock accounting identity in per capita terms by dividing Eq. (16.14) throughout by the population

$$\frac{K_{t+1}}{L_t} = (1 - \delta)\frac{K_t}{L_t} + \frac{I_t}{L_t}$$

or,

$$\frac{K_{t+1}}{L_{t+1}}\frac{L_{t+1}}{L_t} = (1 - \delta)\frac{K_t}{L_t} + \frac{I_t}{L_t}$$

**› Table 16.1**
The Solow Model

| | | | |
|---|---|---|---|
| SOLOW MODEL | Feasible Production Function | $y_t = Ak_t^{\alpha}$ | (16.11) |
| | Product Market Accounting Identity | $S_t = I_t$ | (16.12) |
| | Savings Behaviour | $S_t = sY_t$ | (16.13) |
| | Capital Stock Accounting Identity | $K_{t+1} = (1 - \delta)K_t + I_t$ | (16.14) |
| | Population Growth Behaviour | $L_{t+1} = (1 + n)L_t$ | (16.15) |

Substituting for the population-growth rate from Eq. (16.15) and the savings-investment identity and savings-behavioural Eqs. (16.12) and (16.13),

$$\frac{K_{t+1}}{L_{t+1}}(1+n) = (1-\delta)\frac{K_t}{L_t} + \frac{sY_t}{L_t}$$

The per capita output $Y_t/L_t$ is given by the production function [Eq. (16.11)]. Substituting Eq. (16.11) into the above and writing $K/L = k$ wherever it occurs,[21]

$$(1+n)k_{t+1} = (1-\delta)k_t + sAk_t^{\alpha}$$

or,

$$k_{t+1} = \frac{1-\delta}{1+n}k_t + \frac{s}{1+n}Ak_t^{\alpha} \tag{16.16}$$

Subtracting $k_t$ from both sides of Eq. (16.16),

$$k_{t+1} - k_t = \frac{1-\delta-1-n}{1+n}k_t + \frac{s}{1+n}\left(Ak_t^{\alpha}\right)$$

or,

$$k_{t+1} - k_t = \frac{s}{1+n}\left(Ak_t^{\alpha}\right) - \frac{n+\delta}{1+n}k_t$$

Multiplying throughout by $(1+n)$,

$$(1+n)(k_{t+1} - k_t) = sAk_t^{\alpha} - (n+\delta)k_t \tag{16.17}$$

This equation[22] states that the change in capital per worker each period—the left-hand side[23]—is dependent on two terms. The investment per worker, $sAk_t^{\alpha}$, increases the capital per worker, $k$, while the depreciation per worker $\delta k$ reduces $k$. In addition, in each period, there are $nL$ new workers around, who were not there during the previous period. The capital per worker would decline because of this increase in the labour force if there were no new investments and no depreciation. Note that no new investments implies that the first term on the right-hand side of Eq. (16.17) is zero, and no depreciation requires $\delta = 0$. Then,

$$(1+n)(k_{t+1} - k_t) = -nk_t$$

or,

$$k_{t+1} - k_t = \frac{-n}{1+n}k_t$$

or,

$$k_{t+1} = \frac{1}{1+n}k_t$$

which implies,

$$k_{t+1} < k_t$$

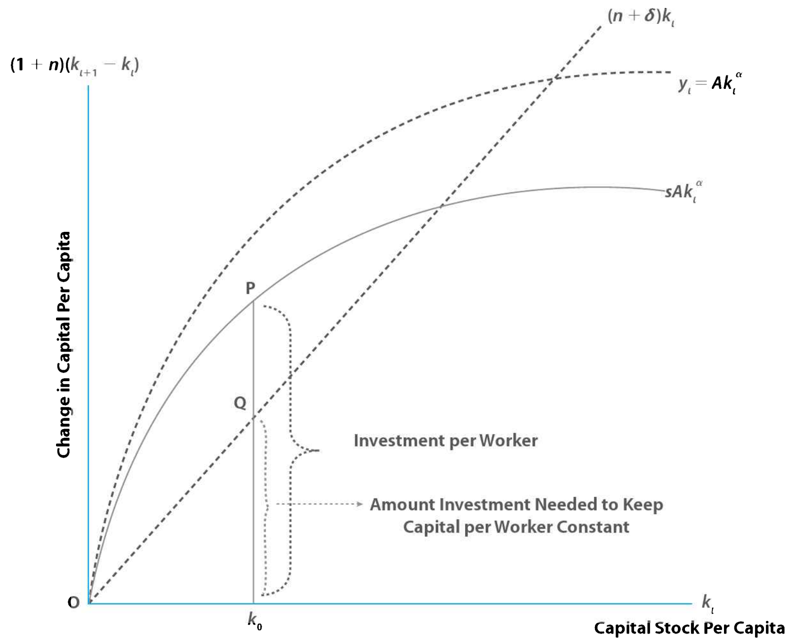
The capital per worker in $t + 1$, therefore, declines due to the increase in the labour force when no investment occurs to offset this decline.

# 16.4 The Solow Diagram

We now plot the graph of Eq. (16.17) in Figure 16.3. Note that the first term on the right-hand side is the graph of the production function, $y = Ak^{\alpha}$, which is depicted in Figure 16.3, but now that graph is translated down by the factor $s$.

The graph of the second term on the right-hand side of Eq. (16.17) is a straight line through the origin with slope, $n + \delta$, that represents the amount

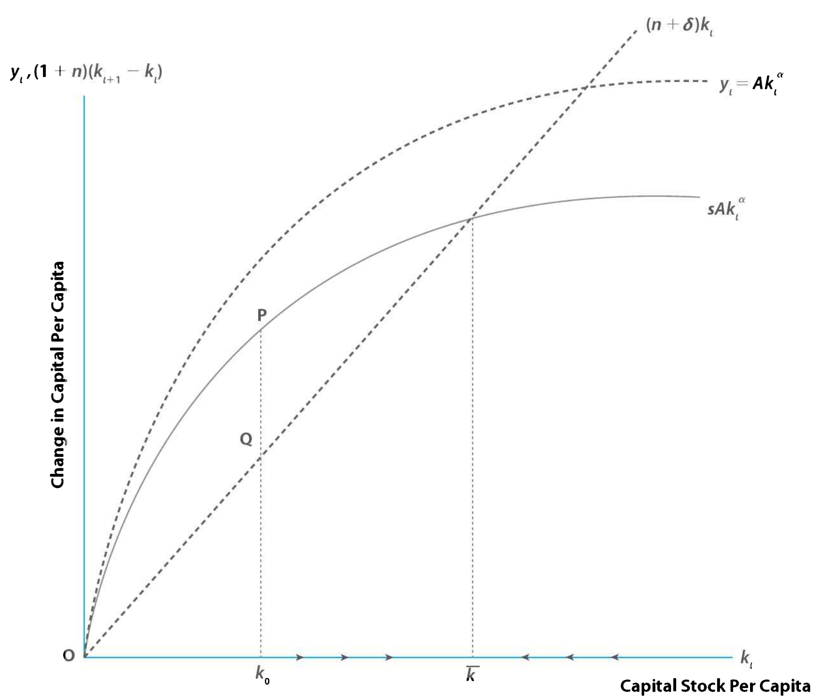
› **Figure 16.3**
The Solow Diagram. The economy has a capital stock per capita given by $k_0$ initially. Then the investment per worker (distance $Pk_0 = sAk_0^{\alpha}$) is greater than the investment required to maintain the capital stock per worker as constant due to the increase in the labour force and depreciation (distance $Qk_0 = (n + \delta)k_0$).




of the new investment per person, which is required to keep the amount of capital per worker constant as both the depreciation and the growth of the workforce reduce the amount of capital per person in the economy. The difference between these two curves is the change in the amount of the capital per worker $(1 + n)(k_{t+1} - k_t)$, given by the distance PQ for an initial capital stock per capita of $k_0$.

Suppose an economy has a capital stock per capita given by $k_0$ today, as in Figure 16.4. What occurs over time? At $k_0$, the amount of investment per worker—distance $Pk_0$—exceeds the amount needed to keep the capital per worker constant—distance $Qk_0$. Accordingly, $k$ increases over time as

› **Figure 16.4**
Stability in the Solow Diagram. The initial capital stock per capita is $k_0$. As the investment per worker is greater than the investment required for the capital stock per worker to be constant, $Pk_0 > Qk_0$, the capital stock per capita increases as depicted by the arrows till $\bar{k}$. At $\bar{k}$, the investment per worker is exactly equal to that required to keep the capital stock per worker constant and the economy is in a steady state.



$(1 + n)(k_{t+1} - k_t) > 0$ or $k_{t+1} > k_t$. The capital per worker increases, as depicted by the arrows pointing rightwards from $k_0$ until $k = \bar{k}$, at which point, the investment per worker equals the replacement of worn-out capital per worker inclusive of the additions to the workforce. Then, the right-hand side of Eq. (16.17) is zero so that,

$$(1 + n)(k_{t+1} - k_t) = 0$$

or,

$$k_{t+1} = k_t = \bar{k} \tag{16.18}$$

A point described by Eq. (16.18), where, $k_{t+1} = k_t$ is called a STEADY STATE.

› A steady state is characterized by a point where the amount of capital stock per capita remains constant over time. $\bar{k} = k_{t+1} = k_t$.

Suppose that the economy began with a capital stock per worker that is greater than $k = \bar{k}$. At any point to the right of $\bar{k}$, the amount of investment per worker is less than the amount required to keep the capital-labour ratio constant, as can be verified in Figure 16.4. The right-hand side of Eq. (16.17) is negative, which implies that $(1 + n)(k_{t+1} - k_t) < 0$, or $k_{t+1} < k_t$, which means that the amount of capital per worker begins to decline in the economy. The decline will occur until the amount of capital per worker reduces to $\bar{k}$.

The steady state in the Solow model is, therefore, a *stable steady state*. If the economy has been working for a long time, we would expect it to reach the steady state as any per capita capital stock other than $\bar{k}$ pushes the economy towards $\bar{k}$. What determines the size of the capital per person in the steady state?

In the steady state, $k_{t+1} = k_t = \bar{k}$, and slipping this condition into Eq. (16.17) gives us,

$$0 = sA\bar{k}^{\alpha} - (n + \delta)\bar{k}$$

or,[24]

$$\bar{k} = \left[\frac{sA}{n + \delta}\right]^{1/(1-\alpha)} \tag{16.19}$$

Recall that a STEADY STATE is one where the new investment per person is just sufficient to replace the worn-out capital and add to capital stock in line with population growth so that the capital stock per capita does not decline. If, after replacing the depreciated capital and providing for capital to new workers at par with the initial amount of capital per worker that existed before this new investment was undertaken there is still some savings (equal to investment) left over, the economy will see an increment in the capital stock per person. However, as the stock of capital per capita grows, the economy will need to devote a greater amount of investment to replace the worn-out capital, thereby leaving less left over for further growth. In the steady state, savings and investment is just sufficient to keep the capital per worker constant, and, at this point, the growth of output per person comes to a halt. Increased saving and investment then cannot cause increased growth, but it does result in high levels of capital and output per person.

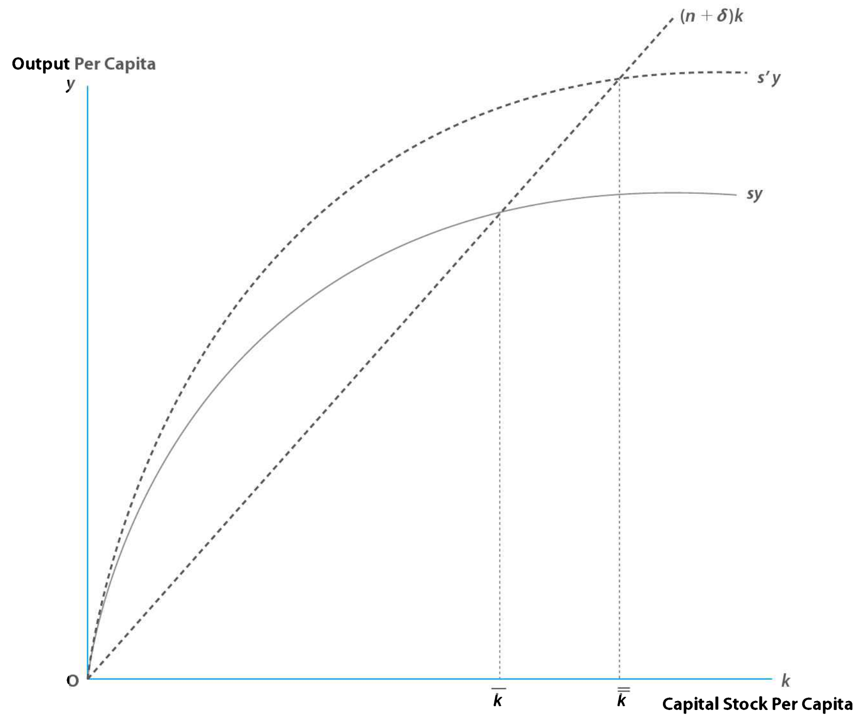
› In a steady state, investment per capita is just sufficient to replace the depreciated capital stock and add to the capital stock in pace with the population growth so that the capital stock per capita is constant.

## 16.5 Comparative Statics

Substituting the steady-state value of the capital stock per capita into the production function (16.11) gives us the steady-state quantity of the output per capita $\bar{y}$

$$\bar{y} = A\bar{k}^{\alpha}$$

**› Figure 16.5**
Increase in Investment Rate. A higher savings/investment rate, $s'$, increases the investment per worker. The higher investment exceeds that required to keep capital per worker constant, and the capital stock per capita increases till the new steady state, $\bar{\bar{k}}$, is achieved.




Substituting the value of $\bar{k}$ from (16.19):

$$\bar{y} = A\left\{\frac{sA}{n+\delta}\right\}^{\alpha/(1-\alpha)}$$

or,[25]

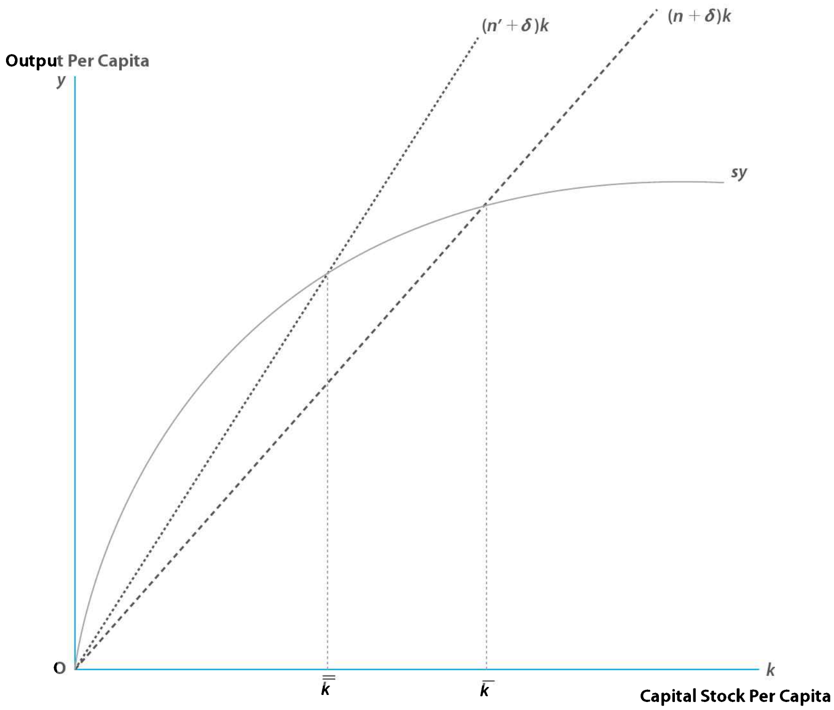
$$\bar{y} = A^{1/(1-\alpha)}\left[\frac{s}{n+\delta}\right]^{\alpha/(1-\alpha)} \qquad \textbf{(16.20)}$$

This equation is the Solow model's response to the question: "Why are some countries rich and some poor?" Countries that have a high savings-investment rate will tend to be richer, *ceteris paribus*. This is because such countries accumulate more capital per worker, and more capital per worker results in more output per worker. In contrast, countries with high population growth rates—high $n$—will tend to be poorer. This is because a higher fraction of savings in these economies goes towards keeping the capital–labour ratio constant in the face of a growing population, and they find it tougher to accumulate more capital per capita.

› An increase in the investment rate leads to a higher level of output per capita.

If an economy increases the INVESTMENT RATE from $s$ to $s'$ permanently, the $sy$ curve shifts up to $s'y$ (see Figure 16.5). At the current steady-state value of the capital stock, $\bar{k}$, the investment per worker exceeds the amount required to keep the capital per worker constant. As a result, there is an increase in the capital per worker, which continues till the capital stock per capita reaches the new steady state, $\bar{\bar{k}}$. As higher capital stock per capita is associated with higher output per capita, the economy is more well off than before.

However, the higher savings rate has no effect on the long-run growth rate of output per capita. Recall that, in a steady state, the capital stock per worker is constant. Accordingly, the output per worker, which is dependent on the capital stock per worker, as $y = Ak^{\alpha}$, is also constant when there is no technological progress ($A$ is a constant). If the per capita income is constant, the growth rate of per capita income must be zero. In a long-run steady state, the savings rate has no effect on the growth rate of the output per capita, which is equal to zero.




**› Figure 16.6**
Increase in Population Growth Rate. **An increase in the population growth rate to $n'$ reduces the capital stock per worker, which declines to the new steady-state level of $\bar{\bar{k}}$.**

Of course, when the savings rate increases initially at capital stock per capita of $\bar{k}$, there is a TEMPORARY INCREASE in the growth rate of the output per capita as the resultant investment is higher than that required to replace the depreciated capital stock and the growth in population that enters the labour force. The capital stock per capita increases until the new steady state, $\bar{\bar{k}}$, is reached where the period of growth in per capita income comes to an end.

› The increase in savings temporarily leads to a higher growth of output per capita but savings has no effect on the long-run growth rate of output per capita, which equals zero.

If an economy that has reached its steady-state witnesses an increase in the POPULATION GROWTH RATE from $n$ to $n'$, say, due to immigration, then, the $(n + \delta)k$ curve rotates counterclockwise to the new $(n' + \delta)k$ curve (see Figure 16.6). The investment per worker is now no longer high enough to keep the capital–labour ratio constant in the face of the rising population, and the capital–labour ratio begins to fall until the point at which $sy = (n' + \delta)k$, where $k = \bar{\bar{k}}$. The economy now has lesser capital per worker, and the per capita output is also lower.

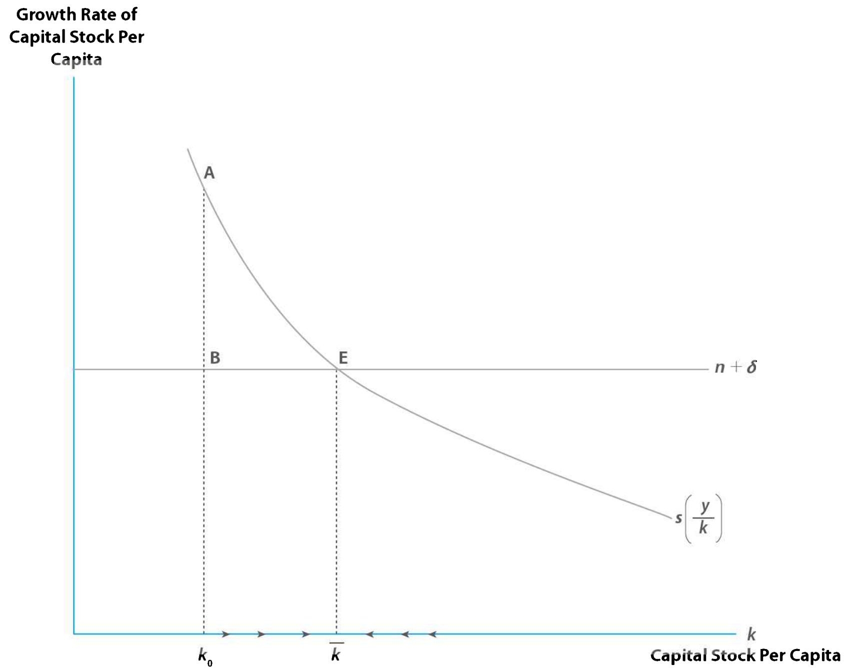
› An increase in the population growth rate reduces the level of output per capita.

What can we say about economic growth in the steady state? From Eq. (16.20), there is no growth of per capita output in the steady state as the output per worker is a constant equal to the expression of constant parameters on the right-hand side. However, an economy that begins with a stock of capital per worker below its steady-state value, $k < \bar{k}$, will experience growth in $k$ and $y$ along the path to the steady state. Over time, however, growth slows down as the steady state is approached, and, in due course, growth stops altogether. To see this, note that the growth rate of $y$ for the given parameters $A$ and $\alpha$ is proportional to the growth rate of $k$,[26] as $y$ is a function of $k$ given by $y = Ak^{\alpha}$. Now,

$$\text{Growth rate of } k = \frac{\text{Change in Capital Stock Per Capita}}{\text{Capital Stock Per Capita}}$$

$$= \frac{(1 + n)(k_{t+1} - k_t)}{k_t}$$

**› Figure 16.7**
Growth Rate of the Capital Stock Per Capita. The savings or investment from the average product due to the capital stock, $s(y/k)$, declines with an increase in the capital stock per capita due to diminishing productivity. If this is greater than the growth in population and depreciation of the capital stock, $n + \delta$, as occurs to the left of point E, there is a positive growth in the capital stock per capita. To the right of point E, the capital stock per capita diminishes over time. The point E, thus, represents a steady state of capital stock per capita, $\bar{k}$.




The numerator is the left-hand side of Eq. (16.17). Hence, on dividing Eq. (16.17) throughout by $k_t$,

$$\text{Growth Rate of } k = \frac{sAk_t^{\alpha}}{k_t} - \frac{(n+\delta)k_t}{k_t}$$
$$= \frac{sy_t}{k_t} - (n+\delta)$$

The first term on the right-hand side is the saving rate into the average product of capital, $y/k$. As the production function is subject to diminishing marginal returns, and because the average product declines when the marginal product declines, the graph of $sy/k$ slopes downwards as $k$ increases in Figure 16.7. The second term, $(n + \delta)$, is independent of $k$, and a horizontal line.

The difference between the two graphs in Figure 16.7 is the growth rate of the capital stock per capita. If the economy is at a point below its steady-state value of $k$, such as at $k_0$, the growth rate of the capital stock per capita is the distance AB. As the economy approaches its steady-state per capita capital stock, the diagram indicates that the growth rate of capital stock per capita, and indeed of per capita output, which is proportional to it, declines. At the steady state, given by point E, the two graphs are equal, and growth has ceased altogether. Of course, with $y = Y/L$, this implies that the aggregate output is growing in the steady state at the rate of population growth, but per capita output growth is nil.

That economies may grow for a while and then experience stagnation, as predicted by the Solow model, does not fit an important stylized fact that economies exhibit a sustained per capita income growth. To get around this problem, Solow assumed that technological progress occurs with parameter $A$ increasing over time. In the Solow setup, technological progress is exogenous: it occurs autonomously and regardless of whatever else is going on in the economy. With technological progress, the parameter $A$ increases with time, shifting the $sy/k$ curve to the right as the capital stock per capita increases. In such a case, growth need not slow down and, eventually, stop altogether.

# 16.6 The Golden Rule

As a higher capital stock per capita results in a higher output per capita should an economy keep increasing its investment rates? A rise in the saving rate increases the steady-state capital intensity (Figure 16.5) so that we may write,

$$\bar{k} = \bar{k}(s), \quad \text{where,} \quad \frac{\partial \bar{k}}{\partial s} > 0$$

The economy selects the steady-state per capita capital stock as well as the output per capita by choosing a saving rate. However, as saving is future consumption, we would be interested in that level of saving which maximizes sustainable consumption. Put another way, increasing the saving rate amounts to a reduction in consumption today. However, the increased saving results in a higher consumption in the future. Thus, sustainable consumption is that which provides the same amount of consumption to members of each present and future generation. To find this level of consumption, write the per capita consumption as the per capita output less per capita saving:

$$\begin{aligned} c &= (1-s)y \\ &= (1-s)Ak^{\alpha} \\ &= Ak^{\alpha} - sAk^{\alpha} \end{aligned}$$

However, in the steady state,[27]

$$(n+\delta)k = sAk^{\alpha}$$

Substituting this into the equation for the per capita consumption,

$$c = Ak^{\alpha} - (n+\delta)k$$

As $k$ is a function of the saving rate, $k = k(s)$, we write,

$$c = A[k(s)]^{\alpha} - (n+\delta)k(s)$$

Hence, the maximum consumption sustainable from a given saving rate is[28]

$$\frac{\partial c}{\partial s} = \left[\alpha Ak^{\alpha-1} - (n+\delta)\right]\frac{\partial k}{\partial s} = 0$$

Since $\partial k/\partial s > 0$, the term in the square brackets must be zero. The first term in the square brackets is none other than the slope of the production function, $\partial y/\partial k$, that we derived earlier[29] in Section 16.3. We can then write that, for a maximum,

$$\frac{\partial y}{\partial k} - (n+\delta) = 0$$

or

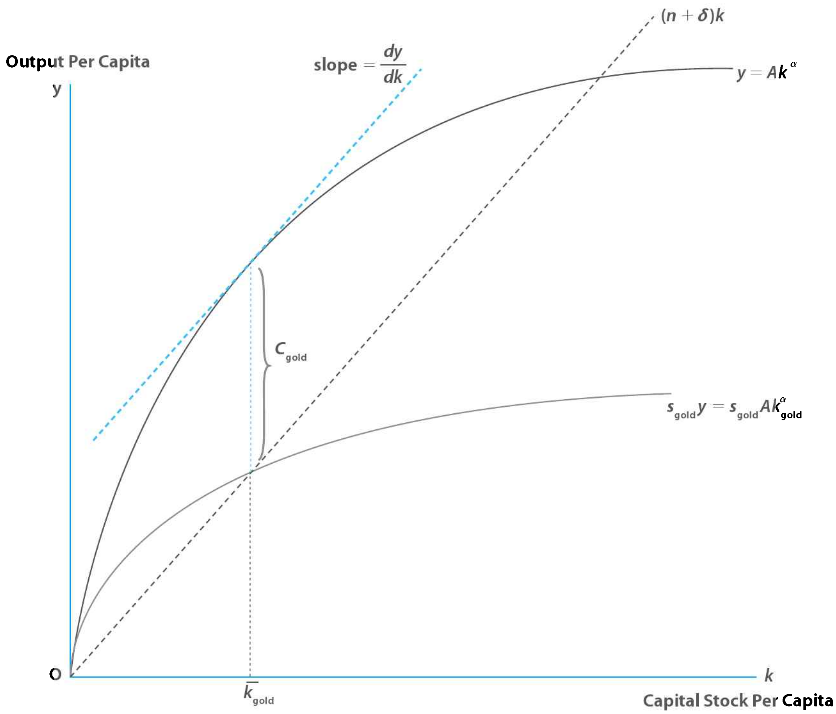
$$\frac{\partial y}{\partial k} = (n+\delta) \qquad \textbf{(16.21)}$$

This CONDITION FOR MAXIMUM SUSTAINABLE CONSUMPTION requires that the slope of the production function equals the slope of the line $(n+\delta)k$, which depicts the capital needed to keep the capital per worker constant.

This condition is depicted in Figure 16.8. If we do not provide less to future generations than to ourselves,[30] then the consumption per capita should be fixed at $c_{\text{gold}}$.

› The golden rule condition for the maximum sustainable consumption requires that the slope of the production function equals the amount of the capital stock required to keep the capital stock per capita constant. $\partial y/\partial k = n + \delta$.

**› Figure 16.8**
The Golden Rule. The golden rule consumption is the maximum consumption that is possible across steady states. This is derived in three steps. First is the condition that the slope of the per capita production function or the marginal productivity of capital equals the growth of population plus the depreciation rate, $\partial y/\partial k = \alpha A k^{\alpha-1} = n + \delta$. The solution to this equation gives the golden rule capital stock per capita or $\bar{k}_{gold}$. The golden rule saving rate then is the solution for $s = s_{gold}$ in the equation $sA\bar{k}^{\alpha}_{gold} = (n + \delta)\bar{k}_{gold}$. Finally, the golden rule consumption is given by $c_{gold} = y(1 - s_{gold})$ where the output per capita is that produced from $\bar{k}_{gold}$.




To derive the value of the golden rule savings rate, we first solve for $k = \bar{k}_{gold}$ in the equation for the maximum sustainable consumption, where, $\partial y/\partial k = n + \delta$,

$$\alpha A\left[\bar{k}_{gold}\right]^{\alpha-1} = n + \delta$$

This solution for $k = \bar{k}_{gold}$ is then substituted into Eq. (16.17): where as the capital per worker is constant in the steady state at $\bar{k}_{gold}$, the left-hand side is zero, which implies,

$$sA\bar{k}^{\alpha}_{gold} = (n + \delta)\bar{k}_{gold} \tag{16.22}$$

**MACROFOCUS 16.2**

## Calculating the Golden Rule Saving Rate

Suppose that the economy is characterized by the following parameters: the population growth, $n$; the depreciation rate, $\delta$; technology, $A$; and the share of capital in output, $\alpha$; Let

$$n = 0.02 \quad \delta = 0.06 \quad A = 145 \quad \alpha = 0.4$$

To solve for the golden rule saving rate, we need to first solve for the capital stock per capita that sustains the maximum possible consumption. This requires us to solve for $k = \bar{k}_{gold}$ in the equation

$$\alpha A[\bar{k}_{gold}]^{\alpha-1} = n + \delta$$

or, $0.4(145)[\bar{k}_{gold}]^{0.4-1} = 0.02 + 0.06$

or, $[\bar{k}_{gold}]^{0.4-1} = \dfrac{0.02 + .06}{0.4(145)} = 0.001379$

or, $[\bar{k}_{gold}] = (0.001379)^{-1/0.6} = (0.001379)^{-5/3}$

or, $\bar{k}_{gold} = 58509.99$

We then substitute this value of $\bar{k}_{gold}$ into the following equation to arrive at the golden rule saving rate, $s = s_{gold}$:

$$sA\bar{k}^{\alpha}_{gold} = (n + \delta)\,\bar{k}_{gold}$$

or, $s(145)(58509.99)^{0.4} = (0.02 + 0.06)58509.99$

or, $s = \dfrac{(0.02 + 0.06)}{145}\left(\dfrac{58509.99}{(58509.99)^{0.4}}\right)$

or, $s = \dfrac{0.08}{145}(58509.99)^{0.6} = \dfrac{0.08}{145}(725)$

or, $s = (0.000552)(725)$

or, $s = s_{gold} = 0.4$

In other words, if an economy described by these parameters wants to maximize the consumption enjoyed by the current and future generations, without sacrificing any generation's consumption for the other, it must save 40 per cent of its aggregate income.

The above parameter values are approximately the values that may be said to characterize the Indian economy.[31] According to the golden rule, it appears that the Indian economy is saving less than it should be. Of course, one reason that may not be true is that it is an outcome derived from the presumption that there is no technological progress.

Solving this equation for $s$ gives us the GOLDEN RULE SAVING RATE as $s = s_{gold}$. At $s_{gold}$ saving rate, the required amount of saving is generated that will sustain the golden rule steady-state capital stock per capita. This capital stock, $\bar{k}_{gold}$, in turn, will sustain the per capita consumption, $c_{gold}$, that maximizes the consumption, which is equally enjoyed by members of each current and future generation.

› The savings rate that sustains the golden rule capital stock per capita is the golden rule savings rate. $sA\bar{k}^{\alpha}_{gold} = (n + \delta)\bar{k}_{gold}$

The golden rule is a reminder that pushing for higher saving beyond a point does not pay off because each increment to capital stock is accompanied by a diminishing marginal productivity and, thus, a diminishing incremental output. At the same time, each addition to the capital stock requires more output to be set aside to replace the depreciated capital. A diminishing incremental output and increasing requirements for the depreciated capital to be replaced leaves a *net* incremental output, which is available for the consumption that is declining. That is why additions to saving and investment beyond a point give a diminished future consumption. An economy that saves and invests beyond $s_{gold}$ is over saving and per capita consumption now and in the future could, in fact, be raised by lowering the economy's saving rate. Such an economy is said to be DYNAMICALLY INEFFICIENT as the path of per capita consumption is below a feasible alternative path that is attainable. Of course, this result is valid for a given technology.

› An economy that saves more than the golden rule saving rate is dynamically inefficient as the consumption per capita over time is lower than a feasible alternative that is available to the economy.

If there are technological improvements that enhance the productivity of capital, improvements in the saving and investment rates beyond the $s_{gold}$ rate (that is associated with a given technology) can sustain higher present and future consumption. Just as we saw when looking at the intertemporal approach to the current account, it is not just thrift or saving, but productivity improvements that are important for sustaining higher incomes and consumption, and higher living standards improve welfare. From a policy perspective, improving saving in an economy helps up to a point: a technological change is required to sustain the growth of output over longer periods of time.

# 16.7 Convergence

An important feature of the Solow growth model then is that with technological progress constant, the long-run growth rate of the economy does not depend on economic policy variables such as the saving rate. An INCREASE IN THE SAVING RATE, for instance, TEMPORARILY raises the growth of the output per capita, but, eventually, the economy reaches a new steady state where the growth rate of the output returns to the level in the earlier steady state, a growth rate of $(n + \delta)$. The level of the income per capita, however, depends on policy variables, such as the saving rate, even in the long run. The second important feature of the model is that the growth rates vary with capital intensity: the growth rate of the output per capita is higher, the lower is the per capita capital stock (Figure 16.7). This has the important implication that the growth rate of a country declines over time as its capital stock per worker increases. Also, in a comparison across economies, countries with higher capital stock per capita will grow more slowly. This implies that there is a tendency for CONVERGENCE across economies.[32]

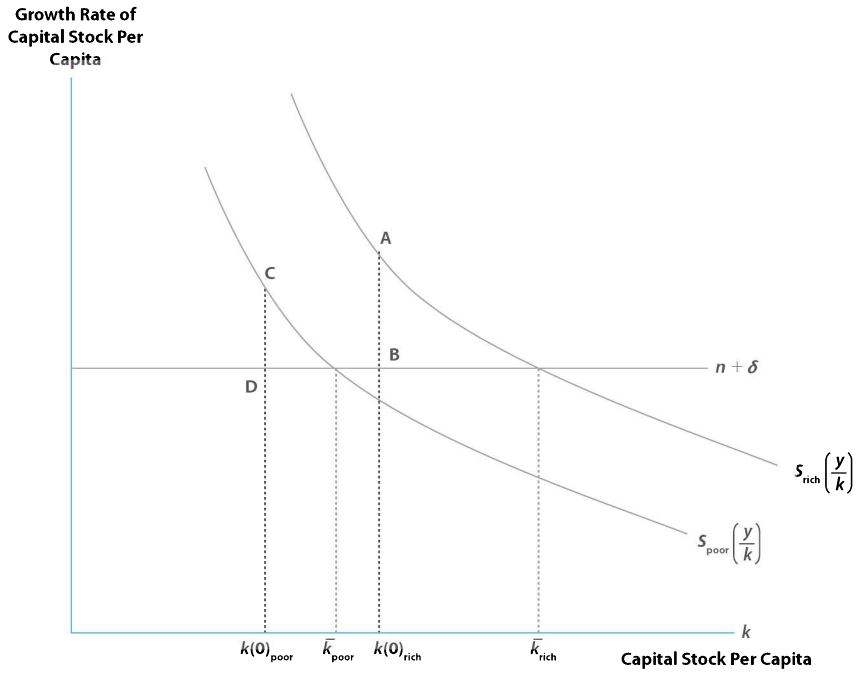
› Policy changes in the Solow model have no long-run growth effect and have only level effects on the level of the output per capita.

› Economies with lower per capita incomes grow faster than richer economies and the gap in the per capita income between the two economies reduces. This catching up is called convergence.

The hypothesis that poor economies tend to grow faster per capita than rich ones, without conditioning on any other characteristics of economies, is referred to as ABSOLUTE CONVERGENCE. This hypothesis receives little or no support except when the test of the hypothesis is limited to a relatively homogeneous

› Absolute convergence occurs when in economies with the same rates of investment, population growth, and levels of technology, the one with a smaller initial per capita income grows temporarily faster than the one with an initially higher per capita income.

**› Figure 16.9**
Conditional Convergence. A poor economy has less capital stock per capita than a rich one, $k(0)_{\text{poor}} < k(0)_{\text{rich}}$, and the saving rate in the poor economy is also smaller, $s_{\text{poor}} < s_{\text{rich}}$. In the case depicted, the growth rate of the capital stock per capita of the rich economy, as given by the vertical distance AB, is larger than the growth rate for the poor economy as given by the distance CD. Both economies converge to their respective steady-state capital stocks per capita conditionally.




group of countries.[33] The theory gets better empirical backing if we allow for heterogeneity across economies in the sense that economies do not have the same steady states. If steady states differ, then, an economy grows faster, the further it is from its own steady-state value, a concept known as CONDITIONAL CONVERGENCE. This concept is illustrated in Figure 16.9, where two economies differ in only two respects—they have different initial stocks of capital per person, $k(0)_{\text{poor}} < k(0)_{\text{rich}}$, and, secondly, they have different saving rates, $s_{\text{poor}} < s_{\text{rich}}$.

› Economies converge to their own steady states. Conditional convergence refers to the convergence of economies after we control for, or condition on, differences in steady states.

Then, it need not be the case that the poor economy will grow faster than the rich one. For the case illustrated in Figure 16.9, the growth rate of the per capita capital stock of the rich economy given by distance AB is greater than the growth rate of the poor economy given by distance CD. However, the prediction is that each economy will converge to its own steady state, $\bar{k}$, and the speed of its convergence relates inversely to the distance from its steady state.

Empirically, this requires us to look at the relation between the growth in the per capita capital stock (or the growth in per capita output) and the initial capital stock per capita (or the output per capita) after holding fixed variables that account for the differences in the steady-state position. To compare economies, of course, they must have something in common. One specification is that the economies share common technological knowledge. Economies could have different values of $K$ and $L$, but the same value of $A$ in the production function, $Y = AF(K, L)$.

A justification for such an assumption is that technological handbooks are easily available everywhere so that the basic commonality across economies is the knowledge of the potential production function.[34] Apart from this, the saving rates, the growth rate of employment, the rate of depreciation, and the initial per capita output would differ across economies. CONVERGENCE in the per capita output is fundamentally due to the convergence in per capita stocks of the physical capital and the diminishing returns to capital. Depreciation offsets capital accumulation and deters the onset of diminishing returns to capital: the convergence rate is an increasing function of the depreciation rate.

› Convergence arises due to the diminishing returns to capital. The convergence rate is an increasing function of the depreciation rate and an increasing function of the population growth rate.

Population growth reduces the growth rate in the capital stock per capita and delays the onset of diminishing returns: the convergence rate is an increasing function of the population growth rate.

Cashin and Sahay[35] estimated an absolute convergence model for Indian states using 1961–1991 data and found that convergence was taking place in the Indian states. However, their findings did not go through when the share of manufacturing in each state's net domestic product, which was included as a control variable in the growth equation, was omitted from the regression model. In that case, the estimated convergence rates decreased in magnitude and became statistically insignificant. Rao and Sen[36] indicate that the share of manufacturing in net domestic product may be proxying for the steady-state differences in growth rates across states in which case the Cashin and Sahay findings suggest the possibility of conditional convergence and not absolute convergence. Abler and Das[37] also estimate for convergence for the period 1961–1990 using state-level annual data for 17 largest Indian states. They find no tendency for convergence in India, a result inconsistent with the Solow model of growth.

## 16.8 Poverty Traps

Solow himself is circumspect about certain empirical applications of his growth model. As he states, "In my view growth theory was conceived as a model of the growth of an industrial economy.... I have never applied such a model to a developing economy, because I thought the underlying machinery would apply mainly to a planned economy or a well-developed market economy".[38]

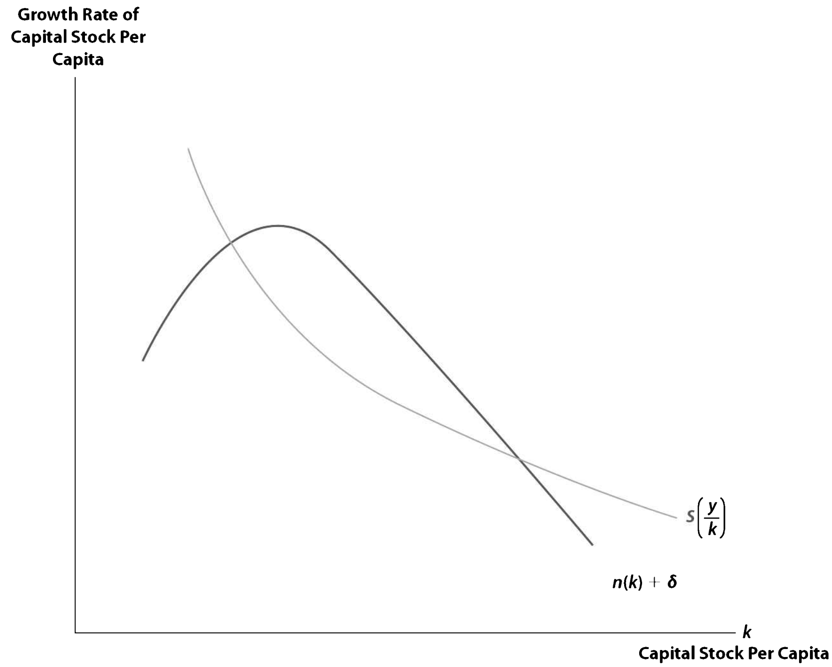
In developing countries, the composition of output has implications for the labour and product markets, which can result in outcomes that are different from the ones associated with the Solow model. Take the population growth, for instance, which is presumed to grow at a constant rate in the Solow model. In India, the current population growth rate is lower than the growth rate that applied around 1947.[39] As the income per worker rises with development, the GROWTH RATE OF THE LABOUR FORCE first rises and then diminishes. With the capital stock per worker an indicator of the level of development, the labour force growth rate, $n$, is a function of $k : n = n(k)$. The rate of population growth increases as agriculture is replaced by industry and slows down as the growth of industry is accompanied by urbanization and the improved levels of education that accompany fertility decline. This implies that, with a constant depreciation rate, the $(n + \delta)k$ line is as drawn in Figure 16.10. This conforms with the line of reasoning just advanced rather than with the horizontal line as stated in the Solow model.

› In a developing economy the rate of the population growth first increases and then declines as industrialization, urbanization, and improved education occur.

Again, in the Solow model, the average product of capital $y/k$ declines with $k$ due to diminishing returns. In the process of development, however, at low levels of development, the agriculture sector dominates the economy, and this is a sector where diminishing returns prevail. As the ECONOMY DEVELOPS, the industrial and service sectors begin to emerge. The industrial sector exploits the scale economies resulting from the division of labour and learning-by-doing more easily, and the economy experiences increasing returns with $y/k$ increasing with $k$. Finally, as the services sector emerges, productivity growth diminishes as this sector's productivity[40] is lower than manufacturing.[41] The economy again runs into a phase of diminishing returns.

› The diminishing returns to capital are offset as an industrial sector that exploits scale economies and learning by doing supplants the agricultural sector.

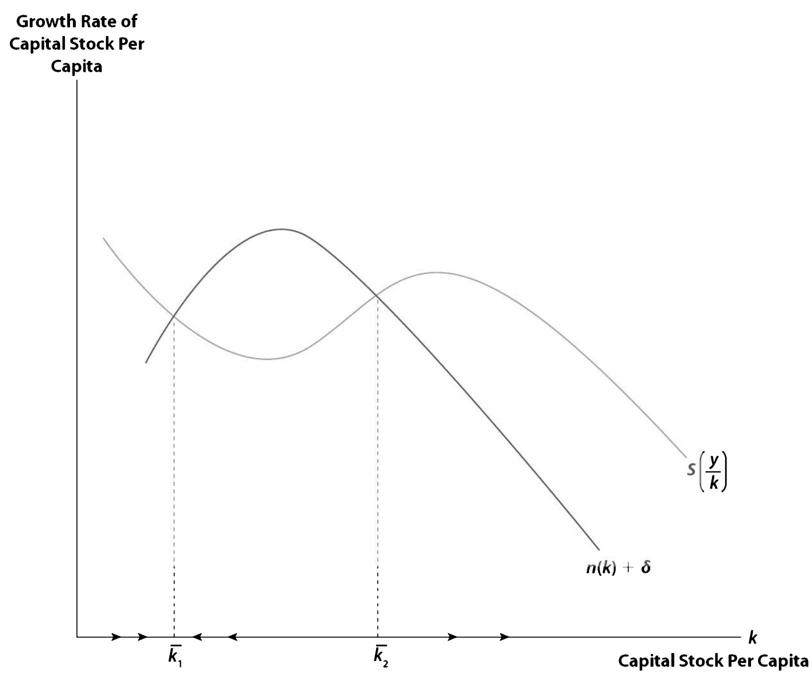
**› Figure 16.10**
Population Growth and Development. **With the capital stock per capita an indicator of the development phase of the economy, the rate of population growth increases in the early stages of development when an economy is primarily agricultural, and declines as the economy becomes industrialized giving rise to the inverted U-shaped $n(k)+\delta$ curve.**




This pattern of productivity changes as the economy develops, as depicted in Figure 16.11 in the shape of the $s(y/k)$ curve, which has a negative slope at low levels of $k$, followed by a range with a positive slope, and a negative slope again at very high levels of $k$.[42]

Recall from the discussion associated with Figure 16.4 that a stable steady state involves a positive growth rate of capital stock per capita for values of $k < \bar{k}$, and a negative growth rate of the capital stock per capita for $k > \bar{k}$. From Eq. (16.21), this implies that for any given steady state, $\bar{k}$, if the $s(y/k)$ curve lies above the $n(k) + \delta$ curve to the left of $\bar{k}$ and $s(y/k)$ lies below the $n(k) + \delta$ curve to the right of $\bar{k}$, $\bar{k}$ is a stable steady state. Accordingly, $\bar{k}_1$ is a stable

**› Figure 16.11**
Development and Productivity. **In the initial stages of development, with agriculture the predominant sector, the economy is subject to diminishing returns. As industrialization occurs, the economy enters an increasing returns phase, and the average productivity of capital $y/k$ rises. As the economy becomes a services economy, the productivity of the economy again declines. Multiple equilibriums then exist with the steady-state $\bar{k}_1$ being a stable steady-state equilibrium and $\bar{k}_2$ an unstable steady-state equilibrium. An economy with a value of $k < \bar{k}_2$ is caught in a poverty trap.**



steady state, but $\bar{k}_2$ is an unstable steady state. If the economy starts out with a value of $k$ such that $\bar{k}_1 < k < \bar{k}_2$, it has an innate tendency for the value of $k$ to reduce to $\bar{k}_1$.

If, somehow, the capital stock per capita gets to $k > \bar{k}_2$, then, the economy will have a tendency to grow further. Hence, for any $k < \bar{k}_2$, the economy is caught in a POVERTY TRAP. Any attempt to raise the per capita output and capital stock has the propensity to return the economy to the low-level steady state, $\bar{k}_1$.

› A poverty trap is the tendency of an economy to return to a low level steady state despite attempts to raise the per capita output.

There are three possible ways of overcoming the barrier to development and getting out of the poverty trap:

1. A domestic policy of raising the savings rate.[43] This results in an upward shift of the entire $s(y/k)$ curve. If the new curve with a higher savings rate still crosses the $n(k) + \delta$ line, the economy can attain a somewhat higher steady-state level of the per capita output and capital stock, but is unable to get away from the poverty trap. However, as the $s(y/k)$ curve gradually rises at some time, the $s(y/k)$ and $n(k) + \delta$ curves will become tangential to one another. Then, suddenly, the economy would find itself on an inevitable growth path towards $\bar{k}_2$. In this case, a large dose of saving and investment is required before a developing country can overcome the existing barriers to growth.
2. A REDUCTION in the population growth rate. This lowers the $n(k) + \delta$ line enough so that it no longer intersects the $s(y/k)$ curve at $\bar{k}_1$, and the economy escapes from the poverty trap.
3. An exogenous injection of additional capital, such as a foreign loan, a gift, or the discovery of new resources. If the World Bank, for instance, makes a contribution of some quantity of capital to a country that begins at, say $\bar{k}_1$, this bestowed capital raises $k$ to a level below $\bar{k}_2$. The economy will enjoy

› An economy can escape a poverty trap with sufficiently large increases in the savings rate, a decline in the population growth rate, or an exogenous increase in the stock of capital.

**MACROFOCUS 16.3**

## Lawyers—Do They Reduce Economic Growth?

Solow's theory informs us that the output per worker is determined by the rate of investment and the labour force growth, and by the efficiency with which these inputs are utilized. However, why do some countries invest less than others, and why are the factors of production less productive in some countries than others?

Economists have come to realize that the social infrastructure of an economy—its policies, laws and regulations, and the institutions of governance that enforce them—is a major determinant of the extent to which individuals have the incentive to make investments in the capital, the technology, and the skills that boost the long-run performance of an economy. In economies where the institutions of governance favour diversion of resources away from the production of goods and services—through theft, red tapism and bureaucratic regulation that encourages corruption, the lobbying of government by special interests, defrauding investors, and frivolous litigation—individuals will be less willing to invest in new technology or to augment the capital stock. When higher returns are available in wealth redistribution rather than wealth creation, individual attention is diverted towards rent seeking rather than profit seeking, with a detrimental impact on the dynamism of the economy. The search for those socio-political characteristics that support economic growth was initiated by Robert Barro.*

In this context, there has been a lively debate on whether professions such as law or the armed services promote or reduce economic growth. Murphy, Shleifer, and Vishny** argued that data on college-law enrolments is an indicator of the talent allocated to diversion of resources or rent seeking. Engineering enrolments, on the other hand, reflect entrepreneurial endeavour and are argued to be good for economic growth. They found that engineering enrolments were good for growth, whereas law enrolments were not. A 10 per cent increase in law enrolments reduced future economic growth by 0.3 per cent.

Revisiting this debate, Cameron and Thorpe† use data on the number of lawyers in 53 countries from the Web site of the International Centre for Commercial Law, www.Legal500.com, among others. In this data set, the United States has the highest number of lawyers among all the countries listed. China is a far second with 12 per cent of the number of lawyers in the United States, and India is third. They are unable to find support for the idea that lawyers result in lower economic growth. This is because not all lawyers are involved in rent seeking through strategic litigation intended to restrict competition and impose costs on competitors. Some lawyers contribute to wealth creation through reducing the transactions and compliance costs involved in meeting regulatory authority's requirements. Finally, some lawyers, such as those involved in human rights cases, are involved in socially useful activities that may be difficult to put a monetary value on. Not all legal cases are about junk lawsuits.

* R. J. Barro, "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics* 106 (May 1991): 407–443.
** K. M. Murphy, A. Shleifer, and R. W. Vishny, "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics* 106 (1991): 503–531.
† S. Cameron and A. Thorpe, "Legislating for Economic Sclerosis: Are Lawyers a Baleful Influence on Growth Rates?" *Kyklos* 57 (2004): 67–86.

temporary higher levels of income per capita but would still return eventually to $\bar{k}_1$ and not escape the poverty trap. A sufficiently large provision of capital, which makes $k > \bar{k}_2$, is required to break out of the poverty trap.[44] In a developing economy with poverty traps, it is not hard to imagine why there may be no tendency towards convergence.

## 16.9 Endogenous Growth

Increasingly, it has been realized that saving and capital accumulation, population growth, and exogenously driven productivity growth are not the satisfactory ways to explain economic growth. At first glance, they provide an insight into the mechanics of growth but then coherence of explanation must be accompanied by completeness. Take the two emerging giants—India and China. Between 1960 and 2003, China's GDP per capita grew by 5.7 per cent per year on average compared with 2.5 per cent growth in India. India was twice as rich as China in 1969. By 2003, China had become twice as rich as India.[45] More investment has been associated with more rapid growth in China. Since 1978, investment has fluctuated between 35 per cent and 45 per cent of GDP in China, whereas, since the mid-1980s, investment has fluctuated between 20 per cent and 25 per cent of GDP in India. Roughly, China growing a little more than twice as fast as India has been associated with almost twice as much investment as in India.

As regards population growth, there are big differences again between China and India. China adopted a one-child policy in 1980 and imposed financial penalties on families having two or more children. Subsequently, families were allowed to have two children provided the first was a girl. The reduced population enabled the expansion of human capital as fewer children meant better care at home, more and better food, clothing, and medical attention.

India's progress in population control has been much slower than China's, and its population is projected to overtake China's by 2025. Also, female literacy and better health care are much better in China where almost 90 per cent of all females know how to read and write as compared to less than 50 per cent for India. China also sends more than 70 per cent of its youngsters to secondary school compared to 50 per cent in India. Education improves the labour force and is necessary if individuals are to move from low-paid self-employment to better-paid jobs in manufacturing and services.

It is through education that workers acquire skills. Skill acquisition also occurs as workers learn a new technology. The acquisition of skills is considered to be the acquisition of human capital and according to the ENDOGENOUS GROWTH THEORY, the accumulation of human capital is responsible for the growth in the output per capita. The determination of the long-run growth within the model itself, rather than through some exogenous growth variables like technological progress, is the reason behind the name *endogenous growth*.

› Endogenous growth theory does not assume that growth is the result of exogenous improvements to technology but focuses on the determinants of technological progress.

› Human capital refers to the levels of education and skills embodied in the labour force.

› Learning by doing is the process by which individuals and firms learn better ways to produce as a by-product of some other activity undertaken such as investment.

HUMAN CAPITAL is acquired by devoting resources to invest in its acquisition. However, unlike physical capital, human capital acquisition takes place through the active pursuit of learning as well as through the act of production itself. This is what is referred to as LEARNING-BY-DOING. Romer[46] argued that knowledge creation was a side product of investment. A firm that accumulates physical capital at the same time learns how to produce more efficiently. This effect of experience on productivity is called learning-by-doing.

Let $S_i$ be the index of skills or knowledge available to the firm.[47] Then, each firm has a potential production function given by

$$Y_i = AK_i^{\alpha}(S_iL_i)^{1-\alpha} \quad \textbf{(16.23)}$$

Learning by doing occurs through each firm's investment—an increase in a firm's capital stock leads to a parallel increase in its stock of knowledge, $S_i$. This reflects the idea that knowledge and productivity gains come from investment and production. This articulation of how skills are acquired is motivated by the empirical observation of the large positive effects of experience on productivity in many production processes.[48] The Boston Consulting Group[49] has, for instance, been famous for its advocacy of learning-curve strategies. The Boston Consulting Group felt that its clients should increase production in the early stages of the product lifecycle so as to secure learning economies. Firms should use profits from cash cow products—products with a high market share in a stable or declining product market—to fund the increased production of problem-child products—products with a low-market share in a fast-growing market. As problem-child products move down their learning curves, they become cash cows in the next investment cycle.

A key assumption of learning-by-doing models of endogenous growth is that each firm's knowledge is a PUBLIC GOOD that any other firm can access at zero cost. A public good is a good that is non-excludable and non-rival. Once produced, it is difficult to exclude somebody from the consumption of such a good, and, moreover, one person's consumption of such a good does not subtract from another person's consumption (non-rivalness).

› Each firm's knowledge about the production process constitutes a public good. A **public good** is non-excludable and non-rival. A good is excludable when others cannot be excluded from its consumption as the good cannot be provided on a market for a fee. For example, no citizen can be excluded from the production of the good national defence. The good is non-rival if one person's consumption does not subtract from another's. For example the ticket charge for the cinema excludes those who do not pay for a ticket but each individual admitted consumes the services of the cinema without subtracting from the benefit of others.

A classic public good is law and order or national defence, which is equally consumed by all, whereas in the case of a private good, say, food, if one person consumes more of a given supply, another has less available for consumption (rivalness). The endogenous growth theory assumes that knowledge is a *public good*: once discovered, a piece of knowledge escapes or spills out instantly across the whole economy. Of course, firms have incentives to maintain secrecy over their discoveries, and they also file for patents to retain the competitive advantage. Knowledge regarding productivity improvements thus trickle out only gradually. However, for simplicity, we assume that all discoveries are the unintended by-products of investment, and these findings immediately become common knowledge. In a sense, we are presuming that as workers learn to use the new technology in a firm, they acquire skills that can be transferred to another firm. This is because a new way of production can easily be duplicated or copied by other workers in other firms,[50] or because some workers leave one firm and take up a job in another where their acquired skills are widely disseminated. As a result, when firms in the economy invest more or accumulate more CAPITAL PER WORKER, the skills of the workers in each firm go up proportionately as workers improve their performance of specific tasks. This happens as they gain experience and learn the appropriate tolerances for product attributes—the benefits of the learning curve. Each firm benefits from the average level of human capital in the economy; as knowledge spillovers increase, the average level of skills and knowledge also increases. Thus,

› The accumulation of skills in a learning-by-doing approach to growth is a by product of another activity, an increase in the capital stock per worker.

$$\frac{dS_i}{dt} = \frac{d(K/L)}{dt}$$

or by choosing a coefficient of proportionality[51] equal to unity,

$$S_i = \frac{K}{L}$$

Learning-by-doing and knowledge spillovers, thus, allow us to replace $S_i$ in Eq. (16.23) by $K/L$, which makes the production function for firm *i*,

$$Y_i = AK_i^{\alpha}\left(\frac{K}{L}L_i\right)^{1-\alpha}$$

If each firm expands the capital per worker, $K/L$ rises and imparts a spillover benefit that augments the productivity of all firms. Dividing throughout by $L_i$, we can write Eq. (16.24) in per worker terms as,

$$\frac{Y_i}{L_i} = AK_i^{\alpha}\left(\frac{K}{L}\right)^{1-\alpha} L_i^{-\alpha}$$

or,

$$\frac{Y_i}{L_i} = A\left(\frac{K_i}{L_i}\right)^{\alpha}\left(\frac{K}{L}\right)^{1-\alpha}$$

or,

$$Y_i = Ak_i^{\alpha} k^{1-\alpha} \qquad \textbf{(16.24)}$$

Each firm is sufficiently small and, so, disregards its own contribution of increased capital intensity ($K_i/L_i$) to the aggregate capital intensity $K/L$, thereby treating $K/L$ as given.

However, in equilibrium, all firms make the same choices so that $k_i = k$ and $y_i = y$. Then,

$$Y = Ak^{\alpha} k^{1-\alpha}$$

or,

$$y = Ak \qquad \textbf{(16.25)}$$

The average product of capital, then, is as follows:

$$\frac{y}{k} = A \qquad \textbf{(16.26)}$$

› In learning-by-doing (*Ak*) endogenous growth theory, the average product of capital is not subject to diminishing returns and is a constant.

The AVERAGE PRODUCT OF CAPITAL, then, is invariant with $k$ due to the tendency for diminishing returns to capital, which permeated the Solow model, being eliminated here as a result of learning-by-doing and spillover effects. The growth rate of capital stock per capita is given by Eq. (16.21),

or,

$$\text{Growth rate of } k = \frac{sy}{k} - (n+\delta)$$

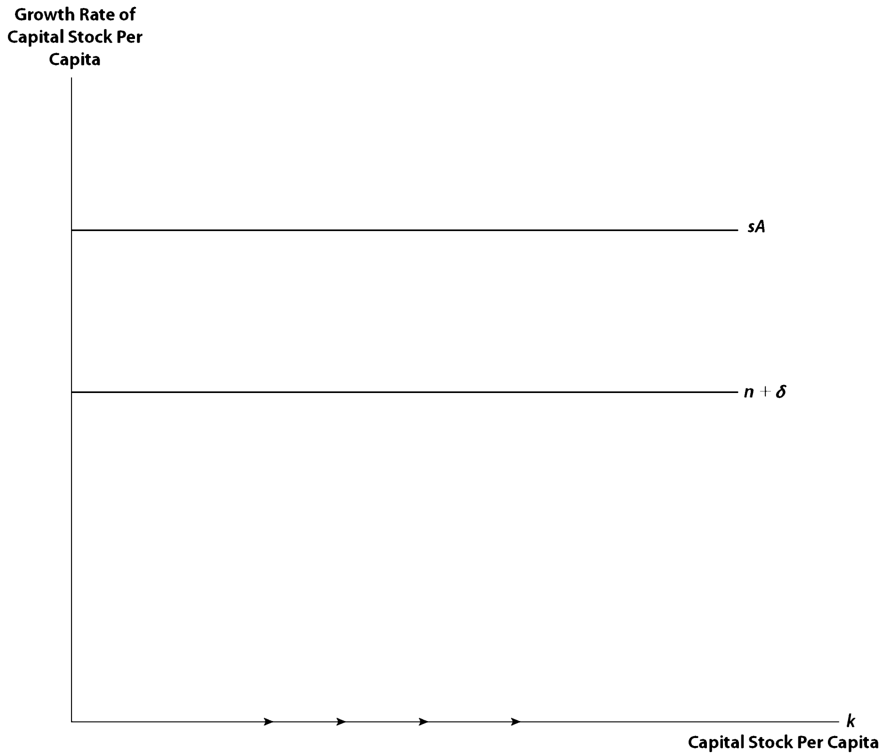
$$= sA - (n+\delta) \qquad \textbf{(16.27)}$$

on substituting from Eq. (16.26). The downward-sloping $s(y/k)$ curve in Figure 16.7 is replaced in Figure 16.12 by the horizontal line $sA$.

The vertical distance between the $sA$ line and line $(n + \delta)$ is the growth rate of capital stock per capita. As the two lines are parallel to one another, the growth rate is constant and independent of $k$. Thus, the capital stock per capita always grows at the steady-state rate $[sA - (n + \delta)]$. Unlike the Solow model, a higher savings rate, $s$, results in a higher rate of long-run per capita growth.

› In endogenous growth theory: (1) a higher savings rate raises the long-run per capita income growth rate permanently, and (2) there is no tendency for convergence.

UNLIKE IN THE SOLOW MODEL, where there is a golden savings rate, here a shift to a higher savings rate at some point of time requires a reduction in consumption per capita at that time, but results in a higher level of per capita consumption in the future. An improvement in technology, $A$, also causes long-run growth to be higher. Finally, unlike the Solow model, the endogenous $Ak$ model does not predict absolute or conditional convergence—the statement that smaller values of $k$ are associated with larger values of the growth of $k$. In fact, for any value of $k$, the derivative of Eq. (16.27) with respect to $k$ is zero.




**› Figure 16.12**
The Endogenous Growth *Ak* Model. **Learning-by-doing and spillover effects result in an average productivity of the capital given by $y/k = A$ that is not subject to diminishing returns. The saving or investment from the average product due to the capital stock, $s(y/k)$, is then constant and does not decline with an increase in the capital stock per capita. The growth rate of the economy is then independent of the level of capital stock per capita and a higher saving rate increases the long-run growth rate of the economy unlike in the Solow model.**

These results follow from relaxing the diminishing returns to capital condition in the Solow model. In an endogenous Romer-type model, the output depends on conventional inputs like capital and labour, but it also depends on an economy's stock of knowledge and skills. The stock of knowledge rises over time as firms accumulate more capital per worker. Each firm has a production function in which the output depends on the firm's private inputs including the firm's stock of private knowledge and skills, and it also depends on the economy's aggregate public stock of skills and knowledge. Thus, the increased capital accumulation per worker in a firm results in an externality where the aggregate stock of knowledge rises and the tendency for diminishing returns is abated.

Growth theory then identifies the ENGINE OF GROWTH as technological change in its exogenous version and the externalities from skill and knowledge acquisition in its endogenous version. This engine of knowledge spillovers is different from most other economic goods. Knowledge is non-rivalrous: my use of a stock of knowledge, such as the use of calculus or the ideas expounded in this book does not bar your simultaneous use of that knowledge. This property of knowledge means that we are free of the scourge of diminishing returns. Of course, producing new knowledge is costly and requires substantial investment that can result in no returns as knowledge generation is risky. But once created, knowledge and skills can be replicated at low cost. It is this externality between the private cost of creating new knowledge and the social cost of replicating it easily that provides the economic fuel for the engine of growth.

› The engine of economic growth is the technological progress that occurs exogenously, or the externalities (by products) of capital accumulation in the form of skills and human capital acquisition.

Growth theory also emphasizes that differences in growth rates across countries or regions are based on *institutional* differences. Knowledge spillovers will be feeble if formal credit markets are scarce, where there is a poor enforcement of property rights, in regions with a high incidence of economic and political instability, and where literacy is low and, thus, impedes skill acquisition. In such countries, it would be difficult for growth to proceed favourably.

There is a huge discrepancy between countries such as Korea and Japan where income levels are high and education is widespread, and countries like Brazil and Mexico where poverty traps prevail, literacy is low, and income inequality is high. A developing country requires to invest in the soft infrastructure of institutions that enable knowledge spillovers and capital accumulation and thus remove actively the obstacles to growth. These include the absence of insurance institutions that discourages risk-taking, the lack of access to credit, the lack of generic skills associated with a standard school education that teach individuals to "learn how to learn", the lack of good health services and disease prevention programmes that increase the productivity of labour, and poorly-defined and enforced property rights that reduce the incentive to invest and even breed corruption. The challenge of modern growth and development is to enhance the institutions that facilitate the engine of growth—capital accumulation and knowledge spillovers—and, thereby, make it run smoothly.

## SUMMARY

- Economic growth is the rate of change of real income (or output) per capita.
- The stylized facts about long-run growth are that (1) output per capita grows at a more or less constant rate, (2) the share of labour and capital in income is constant, and (3) the share of consumption in aggregate income is constant.
- The neoclassical Solow growth theory assumes that (1) the economy is characterized by perfect competition with firms as price takers, and (2) the output is given by a constant return to scale production function such as $Y = AK^{\alpha}L^{1-\alpha}$, where $A$ is a technology parameter.
- Solow used the production function to argue that the change in output is due to the growth of the labour force, the accumulation of capital stock, and technological progress.
- The Solow residual is the estimate of technological progress derived as a residual after accounting for the rate of growth of output due to labour force and capital stock growth.
- Solow demonstrated that the Solow residual or total factor productivity growth contributes the major share to the per capita income growth.
- A steady state is characterized by a point where the amount of the capital stock per capita remains constant over time, $\bar{k} = k_{t+1} = k_t$.
- In a steady state, investment per capita is just sufficient to replace depreciated capital stock and add to the capital stock in pace with population growth so that the capital stock per capita is constant.
- An increase in the investment rate leads to a higher level of the output per capita.
- An increase in the population growth rate reduces the level of the output per capita.
- The golden rule condition for maximum sustainable consumption requires that the slope of the production function equals the amount of capital stock required to keep the capital stock per capita constant. $\partial y/\partial k = n + \delta$.
- The savings rate that sustains the golden rule capital stock per capita is the golden rule savings rate, $sA\bar{k}^{\alpha}_{gold} = (n + \delta)\bar{k}_{gold}$.
- An economy that saves more than the golden rule saving rate is dynamically inefficient as the consumption per capita over time is lower than a feasible alternative that is available to the economy.
- Policy changes in the Solow model have no long-run growth effect and have only level effects on the level of output per capita.
- Economies with lower per capita incomes grow faster than richer economies, and the gap in per capita income between the two economies reduces. This catching up is called convergence.
- Absolute convergence occurs when in economies with the same rates of investment, population growth, and technology levels; the one with a smaller initial per capita income grows temporarily faster than one with an initially higher per capita income.
- Economies converge to their own steady states. Conditional convergence refers to the convergence of economies after we control for, or condition on, differences in steady states.
- Convergence arises due to diminishing returns to capital. The convergence rate is an increasing function of the depreciation rate and an increasing function of the population growth rate.
- In a developing economy, the rate of population growth first increases and then declines as industrialization, urbanization, and improved education occur.
- The diminishing returns to capital is offset as an industrial sector that exploits scale economies, and learning-by-doing supplants the agricultural sector.
- A poverty trap is a tendency for an economy to return to a low-level steady state despite attempts to raise the per capita output.

- An economy can escape a poverty trap with sufficiently large increases in the savings rate, a decline in the population growth rate, or an exogenous increase in the stock of capital.
- The endogenous growth theory does not assume that growth is the result of exogenous improvements to technology but focuses on the determinants of technological progress.
- Human capital refers to the levels of education and skills embodied in the labour force.
- Learning-by-doing is the process by which individuals and firms learn better ways to produce as a by-product of some other activity undertaken such as investment.
- Skills refer to the experience and education of the labour force.
- Each firm's knowledge about the production process constitutes a public good. A public good is non-excludable and non-rivalrous. A good is excludable when others cannot be excluded from its consumption as the good cannot be provided in a market for a fee. The good is non-rivalrous if one person's consumption does not subtract from another's.
- The accumulation of skills in a learning-by-doing approach to growth is a by-product of another activity—an increase in the capital stock per worker.
- In learning-by-doing ($Ak$) endogenous growth theory, the average product of capital is not subject to diminishing returns and is a constant.
- In endogenous growth theory (1) a higher savings rate raises the long-run per capita income growth rate permanently, and (2) there is no tendency for convergence.
- The engine of economic growth is technological progress that occurs exogenously, or the externalities (by products) of capital accumulation in the form of skills and human capital acquisition.

## NOTES

1. Elhanan Helpman, *The Mystery of Economic Growth* (Cambridge, MA: Harvard University Press, 2004).
2. Edward C. Prescott, "Robert M. Solow's Neoclassical Growth Model: An Influential Contribution to Economics," *Scandinavian Journal of Economics* 90, no.1 (1988): 7–12.
3. Robert M. Solow, *Growth Theory: An Exposition* (The Radcliffe Lectures). (Oxford: Clarendon Press, 1970), pp. 8–9.
4. In technical parlance $A$ is said to be Hicks neutral. At a given capital–labour ratio technological progress is Hicks neutral if the marginal rate of technical substitution remains unchanged. It can be shown that Hicks neutrality implies that the output elasticity of capital and labour remains constant.
5. As $P = 1$, in what follows $w/P = w$ denotes the real wage.
6. The first-order condition is $\partial\Pi/\partial L = (1-\alpha)AK^{\alpha}L^{-\alpha} - w = 0$ which gives $w = (1-\alpha)A(K/L)^{\alpha}$ where the right-hand side is the marginal product of labour.
7. See Appendix 16.1, Section A.
8. Robert M. Solow, "Technological Change and the Aggregate Production Function," *Review of Economics and Statistics* (August 1957): 312–320.
9. See Appendix 16.1, Section B.
10. Charles I. Jones, *Introduction to Economic Growth* (New York: W.W. Norton & Co., 1998).
11. When discussing the short-run aspects of employment, full employment was the condition $N^D = N^S$. Here, $N^D = N$ and $N^S = L$ is the total population of the economy.
12. David G. Blanchflower and Andrew J. Oswald, *The Wage Curve* (Cambridge, MA: MIT Press, 1994).
13. Recall that $\varepsilon$ takes values $0 \le \varepsilon \le 1$. If $\beta$ takes the value 0, then, $\varepsilon^{\beta} = 1$. However, $\beta$ taking a larger value, 1 (say), results in $\varepsilon^{\beta} = \varepsilon \le 1$. Generally, that is, as $\beta$ increases, $\varepsilon^{\beta}$ decreases.
14. See Appendix 16.1, Section C for the intervening steps.
15. In intensive form Eq. (16.2) is $Y/L = AK^{\alpha}L^{-\alpha} = Ak^{\alpha}$. Hence, for Eq. (16.11) to be less than Eq. (16.7), we require $[(1-\alpha)Ak^{\alpha}]^{1-\alpha/\alpha+\beta} < 1$. Recall the share of wages is $(1-\alpha)(Y/L) = (1-\alpha)Ak^{\alpha}$. Hence, the term in square brackets is a fraction less than unity.
16. The feasible production function approximates the potential production function when the worker's ability to bargain for higher than competitive wages is weakened. This occurs when $\beta$ takes larger values. As $\beta \to \infty$, we have $(1-\alpha)/(\alpha+\beta) \to 0$, and $\theta = (1-\alpha)^{(1-\alpha)/(\alpha+\beta)} \to 1$. Similarly, $(1+\beta)/(\alpha+\beta) = [(1/\beta)+1]/[(\alpha+\beta)+1] \to 1$. Thus, as $\beta \to \infty$, $y = \theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} \to Ak^{\alpha}$, the equation of the potential production function.
17. Robert M. Solow, "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70 (1956): 65–94.
18. For the case of the feasible production function, see Appendix 16.1, Section D.
19. For the case of the feasible production function, see Appendix 16.1, Section E.
20. As every member of the population is a worker, the labour force participation rate is unity.
21. Equation (16.16) for the case of the feasible production function is, $k_{t+1} = [(1-\delta)/(1+n)]k_t + [s\theta/(1+n)](Ak_t^{\alpha})^{(1+\beta)/(\alpha+\beta)}$, which we obtain by substituting the feasible production function $y = \theta(Ak_t^{\alpha})^{(1+\beta)/(\alpha+\beta)}$ for $Y_t/L_t$ in the previous equation of the text.
22. Equation (16.17) for the case of the feasible production function is $(1+n)(k_{t+1}-k_t) = s\theta(Ak_t^{\alpha})^{(1+\beta)/(\alpha+\beta)} - (n+\delta)k_t$.
23. See Appendix 16.1, Section F, which interprets that $(1+n)(k_{t+1}-k_t)$ is indeed the change in capital stock per capita.
24. The case of the feasible production function is derived in Appendix 16.1, Section G.

25. For the derivation of the feasible production function case, see Appendix 16.1, Section G.
26. From Eq. (16.11), $(\Delta y/y) = \alpha(\Delta k/k)$. Less formally, increases in $y$ are related to increases in $k$ as depicted in Figure 16.2 which implies that $\Delta y/y$ is proportional to $\Delta k/k$.
27. See the derivation to Eq. (16.19).
28. For the case of the feasible production function, see Appendix 16.1, Section I.
29. We derived this expression while exploring the curvature of the production function depicted in Figure 16.2.
30. The golden rule gets its nomenclature from the maxim that "Do not do unto others what you would not like others to do unto you." Translated into the current language, it requires that we do not provide less to future generations than to ourselves.
31. Bosworth and Collins report the growth in total factor productivity for the period 1978–2004 for India to be 1.6 per cent. This implies that $A_t = A_0 + 4.95t$, where, $A_0$ is the technology parameter for 1978. Taking this initial value to be unity, in time $t = 29$, in 2007, $A_t = 145$. For more on this, see, Barry Bosworth and Susan M. Collins, "Accounting for Growth: Comparing China and India," National Bureau of Economic Research, Working Paper no. 12943, Cambridge, MA (2007).
32. That smaller values of $k$ are associated with larger values of the growth of $k$ can be seen technically by differentiating Eq. (16.21) with respect to $k$, which derivative is negative.
33. Robert J. Barro and Xavier Sala-i-Martin, *Economic Growth* (Singapore: McGraw-Hill, 1995).
34. Robert M. Solow, "Applying Growth Theory Across Countries," *The World Bank Economic Review* 15, no. 2 (2001): 283–288.
35. Paul A. Cashin and Ratna Sahay, "Internal Migration Center-State Grants, and Economic Growth in the States of India," *IMF Staff Papers*, 43 (1996): 123–171.
36. M. Govinda Rao and Kunal Sen "Internal Migration, Center–State Grants, and Economic Growth in the States of India: A Comment on Cashin and Sahay," *IMF Staff Papers*, 44 (1997): 283–288.
37. David G. Abler and Jayanta Das, "The Determinants of the Speed of Convergence: The Case of India," *Applied Economics* 30 (1998): 1595–1602.
38. Robert M. Solow, "The Last 50 Years in Growth Theory and the Next 10," *Oxford Review of Economic Policy* 23, no. 1 (2001): 3–14.
39. Tim Dyson, "India's population: The future," in Tim Dyson, Robert H. Cassen, and Leela Visaria, *21st Century India: Population, Environment and Human Development* (New Delhi: Oxford University Press, 2005).
40. The productivity in the education sector for instance as measured by the student–teacher ratio has remained virtually unchanged for decades.
41. William J. Baumol, "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review* 57 (1967): 415–426.
42. There are many ways in which the two curves in Figure 16.11 may intersect. Here we have depicted just one such way.
43. The sudden appearance of a Protestant work ethic would cause people also to save more.
44. The big push approach to development is about the requirement for large doses of savings and investment to overcome existing barriers to growth. See William Arthur Lewis, "Economic Development with Unlimited Supplies of Labour," Manchester School of Economic and Social Studies (Manchester: Blackwell Publishing, 22 May1954): pp. 139–191.
45. For an interesting comparison see Thorvaldur Gylfason (2006), "How do India and China Grow?" *Challenge* 49, no. 1 (January 2006): 74–89.
46. Paul M. Romer, "Increasing Returns and Long Run Growth," *Journal of Political Economy* 94, no. 5 (1986): 1002–1037.
47. Such indices can be constructed by multiplying the years of experience on the job by the years spent in formal education for worker and then adding across workers. Thus experience and education constitute skill.
48. For airframe manufacturing, see Leonard Rapping, "Learning and World War II Production Functions," *Review of Economics and Statistics* 47 (February1965): 81–86.
49. Boston Consulting Group, *Perspectives on Experience*, 1970.
50. This could be due to reverse engineering.
51. The constant of proportionality can be set equal to 1 by an appropriate choice of the units by which we measure our variables.

## TEST YOURSELF

1. What is economic growth and what are the stylized facts about economic growth?
2. Decompose the sources of economic growth and state the contribution of each source to growth.
3. What is the shape of the production function in per capita terms, and why?
4. State the elements of the Solow growth model. How does the capital stock per capita in this model grow over time?
5. What is the steady state of the Solow model? Comment on its stability properties.
6. Why does an increase in the savings rate have only a temporary and no long-run effect on the growth rate? How does it impact the level of output?
7. If there is no technological progress, how do we arrive at the savings rate associated with the maximum steady-state consumption per capita?

8. Why is the Solow model an example of an exogenous growth theory? What is endogenous growth and what are its advantages over exogenous growth as an explanation for the growth of an economy?
9. What is absolute convergence and conditional convergence? Do economies converge? What accounts for this?
10. Why do poverty traps arise, and what can be done to lift an economy out of such traps?

## ONLINE APPLICATION

1. To study economic growth, economists have the benefit of many online data sources. A particularly popular one is the Penn World Tables (PWT). Go to the home page of the Web site of the Centre for International Comparisons at the University of Pennsylvania: (URL: http: //pwt.econ.upenn.edu/).
2. Click on the icon "Penn World Table".
3. Under the Data heading go to sub-section "PWT 6.2 (188 countries, 1950–2004, 2000 as base year)" and click on "data".
4. Select the following: Under countries, select India, Taiwan, Uganda, and the United States. Under variables, select (a) Population, (b) Real GDP per capita (Constant Prices: Chain series), (c) Investment share of RGDPL. For time period, select the years 1990–2003. Take the output of the data, and enter it into a spreadsheet.
5. Calculate the year-on-year population growth rate for each country. Then calculate the average of the population growth rate for the period 1991–2003. The averages you will obtain for India, Taiwan, and Uganda are 1.7 per cent, 0.8 per cent, and 3.1 per cent, respectively. For the United States, it is 1 per cent.
6. Calculate the average of the investment rates for these economies for the period 1991–2003. The averages for India, Taiwan, and Uganda are 11.2 per cent, 19.6 per cent, and 3.3 per cent, respectively. For the United States, the average investment rate is 21.1 per cent.
7. Recall that in the steady state the per capita GDP of an economy is given by the following formula:

$$\bar{y} = A^{1/(1-\alpha)}\left[\frac{s}{n+\delta}\right]^{\alpha/(1-\alpha)}$$

This informed us that countries that have a higher investment rate will tend to be richer and those with a higher population growth rate will tend to be poorer. To test whether the model has an explanatory power in understanding why some countries are richer than others, we define the per capita GDP steady-state value of a country $i$ in terms of the steady-state per capita GDP of the United States as $\hat{y}_i = \bar{y}_i / \bar{y}_{US}$. Relative income per capita in the steady state will, then, be given by the following function:

$$\hat{y}_i = \frac{\bar{y}_i}{\bar{y}_{US}} = \frac{A_i^{1/(1-\alpha)}[s_i/(n+\delta)_i]^{\alpha/(1-\alpha)}}{A_{US}^{1/(1-\alpha)}[s_{US}/(n+\delta)_{US}]^{\alpha/(1-\alpha)}}$$

Assume that $\alpha = 1/3, \delta = 0.06$, and that the relative values of the technology parameter are as follows:

| Country $i$ | India | Taiwan | Uganda |
|---|---|---|---|
| $A_i/A_{US}$ in 1991 | 0.25 | 0.7 | 0.2 |

8. Then, relative GDP per capita in the steady state for an economy may be written as follows:

$$\hat{y}_i = \left(\frac{A_i}{A_{US}}\right)^{1/(1-\alpha)}\left\{\frac{(s_i/s_{US})}{[(n+\delta)_i/(n+\delta)_{US}]}\right\}^{\alpha/(1-\alpha)}$$

Using the data in Step 6, calculate $s_i/s_{US}$, the ratio of the average investment rate of a country with respect to that in the United States for the time period 1991–2003. Using the data in Step 5, calculate $(n + 0.06)_i/(n + 0.06)_{US}$ for the time period 1991–2003.
9. Substitute the calculated values of $s_i/s_{US}$ and $(n + 0.06)_i/(n + 0.06)_{US}$ along with the values in the table above for $A_i/A_{US}$ into the formula for $\hat{y}_i$, and compute the steady-state relative GDP per capita of an economy $i$.
10. Calculate the actual relative value of the GDP per capita of the three economies in the year 2003 from the data you entered on real GDP per capita in Step 4 above. You should obtain the relative values as follows:

$y_{India}/y_{US} = 0.0857, y_{Taiwan}/y_{US} = 0.5702$, and $y_{Uganda}/y_{US} = 0.0319$.

11. Compare the computed steady-state relative GDP per capita $\hat{y}_i = \bar{y}_i/\bar{y}_{US}$ from Step 9 with the actual relative value of GDP per capita $y_i/y_{US}$ computed in Step 10. You should find that the steady-state predicted value in Step 9 overestimates the actual relative value by 1.8 per cent and 0.6 per cent in the case of India and Taiwan, respectively, and underestimates the actual relative value by 2.5 per cent in the case of Uganda.
12. Given the small sample size, the result you obtained in Step 10 indicates that the Solow model is fairly good at explaining the distribution of per capita incomes across countries. Explain how the relative incomes across countries will vary as $A_i/A_{US}$ changes—as technology in an economy $i$ improves—and as the share of capital in income $\alpha$ increases.

# Appendix 16.1 Derivations of Important Equations

Section A

The derivation associated with Eq. (16.3):

Take the logarithm of the production function $Y = AK^{\alpha}L^{1-\alpha}$

$$\log Y = \log A + \alpha \log K + (1-\alpha)\log L$$

Differentiating, $$\frac{dY}{Y} = \frac{dA}{A} + \alpha\frac{dK}{K} + (1-\alpha)\frac{dL}{L}$$

or alternatively, $$\frac{\Delta Y}{Y} = \frac{\Delta A}{A} + \alpha\frac{\Delta K}{K} + (1-\alpha)\frac{\Delta L}{L} \qquad \textbf{(A16.1.1)}$$

Section B

The derivation associated with Eq. (16.5):

Per capita output is $Y/L$. The growth rate of per capita output, then, is the percentage change in $Y/L$,

or, $$\frac{\Delta Y/L}{Y/L} = \frac{\Delta Y}{Y} - \frac{\Delta L}{L}$$

or, the growth rate of per capita output is the growth rate of output less the growth rate of the labour force.

Similarly, $\left(\frac{\Delta K}{K} - \frac{\Delta L}{L}\right) = \frac{\Delta(K/L)}{K/L}$, or the growth rate of the capital–labour ratio.

Section C

The derivation of per capita output for case of a feasible production function associated with Eq. (16.10):

Substituting the equilibrium level of employment $\varepsilon = [(1-\alpha)Ak^{\alpha}]^{1/(\alpha+\beta)}$ into the equation for the production function $Y = AK^{\alpha}(L\varepsilon)^{1-\alpha}$, we obtain,

$$Y = AK^{\alpha}L^{1-\alpha}\left\{\left[(1-\alpha)A\left(\frac{K}{L}\right)^{\alpha}\right]^{1/(\alpha+\beta)}\right\}^{1-\alpha}$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[AA^{(1-\alpha)/(\alpha+\beta)}\right]\left[K^{\alpha}K^{\alpha(1-\alpha)/(\alpha+\beta)}\right]\left[L^{1-\alpha}L^{-\alpha(1-\alpha)/(\alpha+\beta)}\right]$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[A^{(\alpha+\beta+1-\alpha)/(\alpha+\beta)}\right]\left[K^{(\alpha^2+\alpha\beta+\alpha-\alpha^2)/(\alpha+\beta)}\right]\left[L^{(\alpha+\beta-\alpha^2-\alpha\beta-\alpha+\alpha^2)/(\alpha+\beta)}\right]$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[A^{(1+\beta)/(\alpha+\beta)}\right]\left[K^{\alpha(1+\beta)/(\alpha+\beta)}\right]\left[L^{\beta(1-\alpha)/(\alpha+\beta)}\right]$$

Write this expression for the production function in per capita terms:

$$\frac{Y}{L} = (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[A^{(1+\beta)/(\alpha+\beta)}\right]\left[K^{\alpha(1+\beta)/(\alpha+\beta)}\right]\left[L^{\beta(1-\alpha)/(\alpha+\beta)}\right]L^{-1}$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[A^{(1+\beta)/(\alpha+\beta)}\right]\left[K^{\alpha(1+\beta)/(\alpha+\beta)}\right]\left[L^{(\beta-\alpha\beta-\alpha-\beta)/(\alpha+\beta)}\right]$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[A^{(1+\beta)/(\alpha+\beta)}\right]\left[K^{\alpha(1+\beta)/(\alpha+\beta)}\right]\left[L^{-\alpha(1+\beta)/(\alpha+\beta)}\right]$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}\left[A\left(\frac{K}{L}\right)^{\alpha}\right]^{(1+\beta)/(\alpha+\beta)}$$

$$= (1-\alpha)^{(1-\alpha)/(\alpha+\beta)}(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)}$$

or, $$y = \theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} \qquad \textbf{(A16.1.2)}$$

Section D

The derivation of the response of $y$ to change in $k$ for case of a feasible production function $y = \theta(Ak^{\alpha})^{1+\beta/\alpha+\beta}$:

$$\frac{\partial y}{\partial k} = \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}]\, k^{[\alpha(1+\beta)/(\alpha+\beta)]-1}$$

$$= \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}]\, k^{(\alpha+\alpha\beta-\alpha-\beta)/(\alpha+\beta)}$$

$$= \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}]\, k^{-\beta(1-\alpha)/(\alpha+\beta)}$$

Section E

The derivation of the expression for $MP_K$ for the case of a feasible production function $y = \theta(Ak^{\alpha})^{1+\beta/\alpha+\beta}$:

In extensive form, the feasible production function is (see Section C of this appendix),

$$Y = AK^{\alpha}L^{1-\alpha}\left\{\left[(1-\alpha)A\left(\frac{K}{L}\right)^{\alpha}\right]^{1/(\alpha+\beta)}\right\}^{1-\alpha}$$

Hence,

$$MP_K = \frac{\partial Y}{\partial K} = \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}]\{K^{[\alpha(1-\beta)/(\alpha+\beta)]-1}\}[L^{\beta(1-\alpha)/(\alpha+\beta)}]$$

$$= \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}][K^{(\alpha+\alpha\beta-\alpha-\beta)/(\alpha+\beta)}][L^{\beta(1-\alpha)/(\alpha+\beta)}]$$

$$= \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}][K^{-\beta(1-\alpha)/(\alpha+\beta)}][L^{\beta(1-\alpha)/(\alpha+\beta)}]$$

$$= \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}]\left(\frac{K}{L}\right)^{-\beta(1-\alpha)/(\alpha+\beta)}$$

$$= \frac{\alpha(1+\beta)}{\alpha+\beta}[\theta A^{(1+\beta)/(\alpha+\beta)}]\, k^{-\beta(1-\alpha)/(\alpha+\beta)}$$

Section F

The expression for change in capital stock per capita:

The change in the capital stock per capita is given by the following expression:

$$\Delta k = \Delta\left(\frac{K}{L}\right) = \frac{L\Delta K - K\Delta L}{L^2}$$

$$= \frac{K}{L}\left(\frac{\Delta K}{K} - \frac{\Delta L}{L}\right)$$

$$= \frac{K_t}{L_t}\left(\frac{K_{t+1}-K_t}{K_t} - \frac{L_{t+1}-L_t}{L_t}\right)$$

$$= \frac{K_t}{L_t}\left(\frac{K_{t+1}}{K_t} - 1 - \frac{L_{t+1}}{L_t} + 1\right)$$

$$= \frac{K_t}{L_t}\left(\frac{K_{t+1}}{K_t} - \frac{L_{t+1}}{L_t}\right)$$

$$= \frac{K_{t+1}}{L_t} - \frac{K_t}{L_t}\frac{L_{t+1}}{L_t}$$

$$= \frac{K_{t+1}}{L_{t+1}}\frac{L_{t+1}}{L_t} - \frac{K_t}{L_t}\frac{L_{t+1}}{L_t}$$

$$= \frac{L_{t+1}}{L_t}\left(\frac{K_{t+1}}{L_{t+1}} - \frac{K_t}{L_t}\right)$$

$$= (1+n)(k_{t+1} - k_t)$$

Hence, the left-hand side of Eq. (16.17) is the change in the capital stock per capita.

Section G
The derivation associated with Eq. (16.19) for case of the feasible production function:

We have, $$0 = s\theta(A\bar{k}^{\alpha})^{(1+\beta)/(\alpha+\beta)} - (n+\delta)\bar{k}$$

or, $$(n+\delta)\bar{k} = s\theta A^{(1+\beta)/(\alpha+\beta)}\,\bar{k}^{\alpha(1+\beta)/(\alpha+\beta)}$$

or, $$(n+\delta)\,\bar{k}^{1-[\alpha(1+\beta)/(\alpha+\beta)]} = s\theta A^{(1+\beta)/(\alpha+\beta)}$$

or, $$(n+\delta)\,\bar{k}^{(\alpha+\beta-\alpha-\alpha\beta)/(\alpha+\beta)} = s\theta A^{(1+\beta)/(\alpha+\beta)}$$

or, $$\bar{k}^{\beta(1-\alpha)/(\alpha+\beta)} = \frac{s\theta A^{(1+\beta)/(\alpha+\beta)}}{n+\delta}$$

Hence, $$\bar{k} = \left[\frac{s\theta A^{(1+\beta)/(\alpha+\beta)}}{n+\delta}\right]^{(\alpha+\beta)/\beta(1-\alpha)} \quad \textbf{(A16.1.3)}$$

Section H
The derivation for the feasible production function case associated with Eq. (16.20):

We have, $$\bar{y} = \theta(A\bar{k}^{\alpha})^{(1+\beta)/(\alpha+\beta)}$$

Substituting for $\bar{k}^{\alpha}$ from Eq. (16.19) gives us,

$$\bar{y} = \theta A^{(1+\beta)/(\alpha+\beta)}\underbrace{\left\{\left[\frac{s\theta A^{(1+\beta)/(\alpha+\beta)}}{n+\delta}\right]^{[\alpha(\alpha+\beta)]/[\beta(1-\alpha)]}\right\}}_{\bar{k}^{\alpha}}{}^{(1+\beta)/(\alpha+\beta)}$$

$$\bar{y} = \theta A^{(1+\beta)/(\alpha+\beta)}\left\{\left[\frac{s\theta A^{(1+\beta)/(\alpha+\beta)}}{n+\delta}\right]^{[\alpha(\alpha+\beta)]/[\beta(1-\alpha)]}\right\}^{(1+\beta)/(\alpha+\beta)}$$

or, $$\bar{y} = \theta A^{(1+\beta)/(\alpha+\beta)}\left[\frac{s\theta A^{(1+\beta)/(\alpha+\beta)}}{n+\delta}\right]^{[\alpha(1+\beta)/\beta(1-\alpha)]}$$

$$= \theta\theta^{[\alpha(1+\beta)]/[\beta(1-\alpha)]}\,A^{(1+\beta)/(\alpha+\beta)}\,A^{[\alpha(1+\beta)^2]/[\beta(1-\alpha)(\alpha+\beta)]}\left(\frac{s}{n+\delta}\right)^{[\alpha(1+\beta)]/[\beta(1-\alpha)]}$$

$$= \theta^{[\beta-\alpha\beta+\alpha+\alpha\beta]/[\beta(1-\alpha)]}\,A^{[(1+\beta)(\beta-\alpha\beta+\alpha+\alpha\beta)/[\beta(1-\alpha)(\alpha+\beta)]}\left(\frac{s}{n+\delta}\right)^{[\alpha(1+\beta)]/[\beta(1-\alpha)]}$$

$$= \theta^{[\alpha+\beta]/[\beta(1-\alpha)]}\,A^{(1+\beta)/\beta(1-\alpha)}\left(\frac{s}{n+\delta}\right)^{[\alpha(1+\beta)]/[\beta(1-\alpha)]} \quad \textbf{(A16.1.4)}$$

Section I

The derivation for the feasible production function for the maximum consumption sustainable from a given savings rate:

$$\begin{aligned} c &= (1 - s)y \\ &= (1 - s)\theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} \\ &= \theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} - s\theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} \end{aligned}$$

However, in the steady state,

$$(n + \delta)\bar{k} = s\theta(A\bar{k}^{\alpha})^{(1+\beta)/(\alpha+\beta)}$$

Substituting this into the equation for the per capita consumption,

$$c = \theta(Ak^{\alpha})^{(1+\beta)/(\alpha+\beta)} - (n + \delta)k$$

As $k$ is a function of the saving rate, $k = k(s)$, we write,

$$c = \theta[Ak(s)^{\alpha}]^{(1+\beta)/(\alpha+\beta)} - (n + \delta)k(s)$$

Hence, the maximum consumption sustainable from a given saving rate is as follows:

$$\frac{\partial c}{\partial s} = \left[\frac{\alpha(1 + \beta)}{\alpha + \beta}\theta A^{(1+\beta)/(\alpha+\beta)}k^{-\beta(1-\alpha)/(\alpha+\beta)} - (n + \delta)\right]\frac{\partial k}{\partial s} = 0$$

# Bibliography


Alan S. Blinder, *Central Banking in Theory and Practice* (Cambridge, MA: MIT Press, 1998).

Alban William Phillips, "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," *Economica* 25(1958): 283–299.

Alberto Giovannini, "Savings and the Interest Rate in LDCs," *World Development* 11, no. 7 (July 1983): 601–607.

Albert Ando and Franco Modigliani, "The 'Life Cycle' Hypothesis of Saving: Aggregate Implications and Tests," *American Economic Review* 53, no. 1 (March 1963): 55–84.

Arthur M. Okun, *The Political Economy of Prosperity* (New York: Norton, 1970).

Assaf Razin and Efraim Sadka, "Fiscal Balance During Inflation, Disinflation and Immigration: Policy lessons," in M. Blejer and T. Ter-Minassian (eds.) *Macroeconomic Dimensions of Public Finance: Essays in Honor of Vito Tanzi* (IMF, 1996).

Assar Lindbeck and Dennis J. Snower, *The Insider-Outsider Theory of Employment and Unemployment* (Cambridge, MA: MIT Press, 1988); "Interactions Between the Efficiency Wage and Insider-Outsider Theories," *Economics Letters* 37 (1991): 193–196.

Avinash K. Dixit and Robert S. Pindyck, *Investment under Uncertainty* (Princeton, NJ: Princeton University Press, 1994); and R. S. Pindyck, "Irreversibility, Uncertainty and Investment," *Journal of Economic Literature* 3 (September 29, 1991): 110–148.

B. Bhaskara Rao and Rup Singh, "Demand for Money in India: 1953–2003," *Applied Economics* 38, no. 11 (2006): 1319–1326.

Bennett McCallum, *Monetary Economics: Theory and Policy* (New York: Macmillan, 1989).

Bruce Greenwald, Joseph E. Stiglitz, and Andrew Weiss, "Information Imperfections in the Capital Market and Macroeconomic Fluctuations," *American Economic Review Proceedings* 74 (May 1984): 194–199.

B. K. Bhoi and S. C. Dhall, "Integration of Financial Markets in India: An Empirical Evaluation," RBI Occasional Papers No. 19 (4), Mumbai 1998, pp. 345–380.

Casey B. Mulligan, "Substitution over Time: Another Look at Life Cycle Labour Supply," *NBER Macroeconomics Annual* 13, 1998.

Carl Shapiro and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74 (1984): 433–444.

Carmen M. Reinhart, "The Mirage of Floating Exchange Rates," *American Economic Review*, May 2002.

Charles Himmelberg, Christopher Mayer, and Todd Sinai, "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions," *Journal of Economic Perspectives* 19, no. 4 (Fall 2005): 67–92.

Charles I. Jones, *Introduction to Economic Growth* (New York: W.W. Norton & Co., 1998).

Charles R. Bean, "Targeting Nominal Income: An Appraisal," *Economic Journal* (December 1983): 806–819.

Charles R. Nelson and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Times Series: Some Evidence and Implications," *Journal of Monetary Economics* 10 (September 1982): 139–162.

David G. Abler and Jayanta Das, "The Determinants of the Speed of Convergence: The Case of India," *Applied Economics* 30 (1998): 1595–1602.

David G. Blanchflower and Andrew J. Oswald, *The Wage Curve* (Cambridge, MA: MIT Press, 1994).

David J. Ravenscraft and F. M. Scherer, *Mergers, Sell-offs and Economic Efficiency* (Washington, DC: The Brookings Institution, 1987).

Donald A. Nichols, "Some Principles of Inflationary Finance," *Journal of Political Economy* 82 (1974): 423–430.

Douglas W. Diamond, "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies* 59 (1984): 393.

Edward C. Prescott, "Theory Ahead of Business Cycle Measurement," Federal Reserve Bank of Minneapolis, *Quarterly Review* (Fall 1986): 9–22.

Edward C. Prescott, "Robert M. Solow's Neoclassical Growth Model: An Influential Contribution to Economics," *Scandinavian Journal of Economics* 90, no. 1 (1988): 7–12.

Edward S. Shaw, *Financial Deepening in Economic Development* (New York: Oxford University Press, 1973).

Elhanan Helpman, *The Mystery of Economic Growth* (Cambridge, MA: Harvard University Press, 2004).

Errol D'Souza, "Prudential Regulation in Indian Banking," in S. Arumugam (ed.) *Indian Capital Markets—Modern Perspectives and Empirical Evidence* (Mumbai: Allied Publishers Ltd., 2001).

Errol D'Souza, "Employment and Human Resource Practices in Public Sector Banks in the 1990s," in Shuji Uchikawa (ed.) *Labour Market and Institution in India–1990s and Beyond* (New Delhi: Manohar Publishers, 2003), pp. 95–120.

Fanny S. Demers, Michel Demers, and Huntley Schaller, "Irreversible Investment and Cost of Adjustment," CEPREMAP Working Paper No. 9416, 1994.

Finn E. Kydland and Edward C. Prescott, "Rules rather then Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 87 (June 1977): 473–492.

Fumio Hayashi, "Why is Japan's Saving Rate So Apparently High?" in Stanley Fischer (ed.) *Macroeconomics Annual* (New York: National Bureau of Economic Research, 1986).

Francesco Giavazzi and Marco Pagano, "Can Fiscal Contractions be Expansionary? Tales of Two Small European Countries," in O. Blanchard and S. Fischer (eds.) *NBER Macroeconomics Annual* (MIT Press, 1990).

Franco Modigliani, "Liquidity Preference and the Theory of Interest and Money," *Econometrica* 12 (1944): 45–88.

Franco Modigliani, "Life Cycle, Individual Thrift and the Wealth of Nations," Nobel Lecture, *American Economic Review* 76, no. 3 (June 1986): 297–313.

Franco Modigliani and Larry Cao Shi, "The Chinese Saving Puzzle and the Life-Cycle Hypothesis," *Journal of Economic Literature* 42, no. 1 (March 2004): 145–170.

Franco Modigliani and Merton H. Miller, "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review* 48 (June 1958): 261–297.

Fumio Hayashi, "Tobin's Marginal q and Average q: A Neoclassical Interpretation," *Econometrica* 50 (January 1982): 213–224.

Geetha Veidyanathan, "Consumption, Liquidity Constraints, and Economic Development," *Journal of Macroeconomics* 15 (Summer 1993): 591–610.

George A. Akerlof, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (August 1970): 488–500.

Gerald P. Dwyer Jr. and Thomas R. Saving, "Government Revenue from Money Creation with Government and Private Money," *Journal of Monetary Economics* 17 (1986): 239–249.

Gerard Caprio Jr., "Banking on Financial Reform? A Case of Sensitive Dependence on Initial Conditions," in Gerard Caprio Jr., Ijak Atiyas, and James A. Hanson (eds.) *Financial Reform—Theory and Experience* (Cambridge, MA: Cambridge University Press, 1996).

GOI, "Report of the Committee on the Financial System" (New Delhi: Government of India), 1991.

Gregg A. Farrell, James A. Brickley, and Jeffry M. Netter, "The market for corporate control: the empirical evidence since 1980," *Journal of Economic Perspectives* 2 (1988): 49–82.

Guillermo A. Calvo and Carmen M. Reinhart, "Fear of Floating," *Quarterly Journal of Economics* 117 (2002): 379–408.

Harvey Leibenstein, "The Theory of Underemployment in Backward Economies," *Journal of Political Economy* 65 (April 1957): 91–103.

Irving Fisher, "The Debt Deflation Theory of Great Depressions," *Econometrica* (October 1933).

J. Marcus Fleming, "Domestic Financial Policies Under Fixed and Floating Exchange Rates," IMF Staff Papers, no. 9 (1962): 369–377.

Jacob A. Frenkel, Assaf Razin, and Chi-Wa Yuen, *Fiscal Policies and Growth in the World Economy* (MIT, MA: MIT Press, 1996).

James M. Poterba and Lawrence H. Summers, "Dividend Taxes, Corporate Investment, and 'Q'," *Journal of Public Economics* 22 (1983): 135–167.

James Tobin, "The Interest Elasticity of Transactions Demand for Cash," *Review of Economics & Statistics* 38 (August 1956): 241–247.

James Tobin, "*Notes on Optimal Monetary Growth*," *Journal of Political Economy* 76, no. 4, Pt. 2 (July-August 1968): 833–859.

James Tobin, "Asset Accumulation and Economic Activity," Yrjo Jahnsson Lectures, Basil Blackwell, Oxford (1980).

Jeffrey A. Sachs, "The Debt Overhang of Developing Countries," in Guillermo A. Calvo, Ronald Findlay, Pentti Kouri, and Jorge Braga de Macedo (eds.) *Debt, Stabilization, and Development* (Oxford: Basil Blackwell, 1989).

Jeffrey Sachs and Felipe Larrain, "Why Dollarisation is More Straitjacket than Salvation," *Foreign Policy* 116 (Fall1999): 80–92.

John B. Taylor, "Staggered Wage Setting in Macro Model," *American Economic Review* 69, no. 2 (1979): 108–113.

John F. Muth, "Rational Expectations and the Theory of Price Movements," *Econometrica* (July 1961).

John Kareken and Robert Solow, "Lags in Monetary Policy," in *Stabilisation Policies*, Commission on Money and Credit (Prentice Hall, 1963).

John M. Keynes, "The General Theory of Employment, Interest and Money," *The Collected Writing of John Maynard Keynes* (London: Macmillan, 1992).

John M. Keynes, *The General Theory of Employment, Interest, and Money* (New York: Harcourt Brace Jovanovich, 1964).

John R. Hicks, "Mr. Keynes and the Classics: A Suggested Interpretation," *Econometrica* (April 1937): 147–159.

Joseph E. Stiglitz and Andrew Weiss, "Credit Rationing in Markets with Incomplete Information," American Economic Review (1981): 393–410.

Joseph G. Altonji, "Intertemporal Substitution in Labour Supply: Evidence from Micro Data," *Journal of Political Economy*, Part 2, 94 (June 1986): S176–S215.

Ken Froot and Richard Thaler, "Anomalies: Foreign Exchange," *Journal of Economic Perspectives* 4, no. *3* (1990): 179–192.

Kenneth J. Arrow and Gerard Debreu, "Existence of Equilibrium for a Competitive Economy," *Econometrica* 22 (1954): 265–290.

Kenneth Rogoff, "The Optimal Degree of Commitment to an Intermediate Monetary Target," *Quarterly Journal of Economics* 100, no. 4 (1985): 1169–1189.

Kenneth S. Rogoff, Aasim M. Husain, Ashoka Mody, Robin Brookes, and Nienke Oomes, "Evolution and Performance of Exchange Rate Regimes," IMF Working Paper No. WP/2003/243, 2003.

Lawrence Ball, "Interpersonal Substitution and Constraints on Labour Supply: Evidence from Panel Data," *Economic Inquiry* 28 (October 1990): 706–724.

Lawrence F. Katz, *Efficiency Wage Theories: A Partial Ecvaluation*, in NBER Macroeconomics Annual (Cambridge, MA: MIT Press, 1986), Vol. 1, pp. 235–276.

Lawrence H. Summers, B. P. Bosworth, J. Tobin, and P. M. White, "Taxation and Corporate Investment: A Q Theory Approach," Brooking Papers on Economic Activity, no. 1 (1981): 67–140.

Lawrence Katz and Alan Krueger, "The High-Pressure U.S. Labour Market of the 1990s," *Brookings Papers on Economic Activity* 1 (1999): 1–88.

Larry Selden, "A New Approach to the Joint Consumption-Portfolio Problem," *Journal of Money, Credit, and Banking* 12, no. 3 (August 1980): 429–447.

Leonard Rapping, "Learning and World War II Production Functions," *Review of Economics and Statistics* 47 (February1965): 81–86.

Luis Serven, "Does Public Capital Crowd out Private Capital? Evidence from India," The World Bank, Policy Research Working Paper No. 1613, 1996.

Luis Serven, "Macroeconomic Uncertainty and Private Investment in Developing Countries—An Empirical Investigation," Policy Research Working Paper, The World Bank, Report No. 2035 (December 1998).

Luis Serven and Andres Solimano, "Private Investment and Macroeconomic Adjustment: A Survey," *World Bank Research Observer* 7, no. 1 (1992): 95–114.

Martin Eichenbaum and Charles L. Evans, "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates," *Quarterly Journal of Economics* 110, no. 4 (1995): 975–1009.

Martin Feldstein and Charles Horioka, "Domestic Saving and International Capital Flows," *The Economic Journal* 90, no. 358 (June 1980): 314–329.

Martin J. Bailey, "The Welfare Cost of Inflationary Finance," *Journal of Political Economy* 64 (April 1956): 93–110.

Martin Rama, "Empirical Investment Equations in Developing Countries," in L. Serven and A. Solimano (eds.) *Striving for Growth After Adjustment* (Washington, DC: The World Bank, 1993).

Matthew B. Canzoneri, "Monetary Policy Games and the Role of Private Information," *American Economic Review* 75 (December 1985): 1056–1070.

Maurice Obstfeld and Kenneth Rogoff, "Mirage of fixed Exchange Rates," *Journal of Economic Perspectives* 9 (1995): 73–95.

Maxwell J. Fry, *Emancipating the Banking System and Developing Markets for Government Debt* (London and New York: Routledge, 1997).

Michael Bruno, *Crisis, Stabilization and Economic Reform: Therapy by Consensus* (Oxford: Clarendon Press, 1993).

Michael J. Athey and Wende D. Reeser, "Asymmetric Information, Industrial Policy, and Corporate Investment in India," *Oxford Bulletin of Economics and Statistics* 62, no. 2 (2000): 267–292.

Michael Rothschild and Joseph E. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90 (1976): 629–649.

M. J. Manohar Rao, "The Demand for Money in India," in K. L. Krishna (ed.) *Econometric Applications in India,* Oxford India Themes in Economics (New Delhi: Oxford University Press, 1997).

M. J. Manohar Rao and Raj Nallari, *Macroeconomic Stabilization and Adjustment* (New Delhi:Oxford University Press, 2001).

M. N. Murty and A. Soumya, "Macroeconomic Effects of Public Investment in Infrastructure in India," IGIDR, Working Paper No. 2006–2003, Mumbai, 2006.

Milton Friedman, "A Monetary and Fiscal Framework for Economic Stability," *Essays in Positive Economics* (Chicago: Chicago University Press, 1953).

Milton Friedman, *A Theory of the Consumption Function* (Princeton, NJ: Princeton University Press, 1957).

Milton Friedman, "Employment, Growth, and Price Levels," 86th Congress, First Session, (25–28 May 1958), p. 609.

Milton Friedman, "The Role of Monetary Policy," *American Economic Review* 58(1) (1968): 1–17; Edmund S. Phelps, "Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time," *Economica* 34 (1967): 254–281.

Morris Goldstein and Mohsin S. Khan, "Income and Substitute Effects in Foreign Trade," in Ronald W. Jones and Peter B. Kenen (eds.) *Handbook of International Economics* 2 (Elsevier, 1985).

M. Govinda Rao and Kunal Sen "Internal Migration, Center–State Grants, and Economic Growth in the States of India: A Comment on Cashin and Sahay," *IMF Staff Papers* 44 (1997): 283–288.

N. Gregory Mankiw, "Real Business Cycles: A New Keynesian Perspective," *Journal of Economic Perspectives* (Summer 1989): 79–90.

Nadeem U. Haque and Peter Montiel, "Consumption in Developing Countries: Test for Liquidity Constraints and Finite Horizons," *Review of Economics and Statistics* 71 (August 1989): 408–415.

Pami Dua and B. L. Pandit, "Interest rate determination in India: domestic and external factors," *Journal of Policy Modeling* 24 (2002): 853–875.

Parthasarathi Shome, "India: Tax Policy for the Ninth Five Year Plan," Report of the Working Group on Tax Policy of the Steering Group on Financial Resources, New Delhi, 1997; and the Kelkar Committee Report.

Patrick Bolton and Xavier Freixas, "A Dilution Cost Approach to Financial Intermediation and Securities Markets," Working Paper, Universitat Pompeu Fabra, 1998.

Paul A. Cashin and Ratna Sahay, "Internal migration centerstate grants, and economic growth in the sates of India," *IMF Staff Papers* 43 (1996): 123–171.

Paul Einzig, *The History of Foreign Exchange* (London and Basingtoke: Macmillan St. Martin's Press, 1990).

Paul Krugman, "Purchasing Power Parity and Exchange Rates: Another Look at the Evidence," *Journal of International Economics* 8 (August 1978): 397–407.

Paul Krugman and Lance Taylor, "Contractionary Effects of Devaluation," *Journal of International Economies* (8 August 1978): 445–456.

Paul M. Romer, "Increasing Returns and Long Run Growth," *Journal of Political Economy* 94, no. 5 (1986): 1002–1037.

Peter B. Kenen, "The International Financial Architecture: What's New? What's Missing?" *Institute for International Economics* (Washington, DC: IIE, 2001).

Prema C. Athukorala and Kunal Sen, *Saving, Investment and Growth in India* (New Delhi: Oxford University Press, 2002).

Reid W. Click, "Seigniorage in a Cross-Section of Countries," *Journal of Money, Credit and Banking* 30 (May 1998), 154–171.

Reserve Bank of India, "Report of the Committee to Review the Working of the Monetary System," Bombay, 1985.

Ricardo Hausmann, "Should There be Five Currencies or One Hundred and Five?" *Foreign Policy* 116 (Fall 1999): 65–79.

Richard A. Brealey and Stewart C. Myers, *Principles of Corporate Finance* (New York: McGraw Hill, 1991).

Robert A. Mundell, "The Appropriate Use of Monetary and Fiscal Policy Under Fixed Exchange Rates," IMF Staff Papers, no. 9 (1962): 70–77.

Robert A. Mundell, "Notes on the Development of the International Macroeconomic Model," in A. Arnon and W. Young (eds.) *The Open Economy Macromodel: Past, Present and Future* (Boston, MA: Kluwer Academic Publishers, 2002), pp. 1–16.

Robert Barro, "Democracy and Growth," *Journal of Economic Growth* 1 (January 1996).

Robert J. Barro, "Are Government Bonds Net Wealth?" *Journal of Political Economy* 82 (1974): 1095–1117.

Robert J. Barro, *Getting It Right-Markets and Choices in a Free Society* (Cambridge, MA: MIT Press, 1996), pp. 58.

Robert J. Barro and David B. Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy* 91 (August 1983): 589–610.

Robert J. Barro and Xavier Sala-i-Martin, *Economic Growth* (Singapore: McGraw-Hill, 1995).

Robert E. Hall, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy* 96 (1978): 971–987.

Robert Eisner and Robert Strotz, "Determinants of Business Investment," Commission on Money and Credit, Impacts of Monetary Policy (Englewood Cliffs, NJ: Prentice-Hall, 1963), pp. 60–138.

Robert Feenstra, "Functional Equivalence Between Liquidity Costs and the Utility of Money," *Journal of Monetary Economics* 17 (1986): 271–291.

Robert Hubbard, "Capital Market Imperfections and Investment," *Journal of Economic Literature* 36 (1998): 193–225.

Robert Lane, *The Market Experience* (New York: Cambridge University Press, 1991).

Robert M. Solow, "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70 (1956): 65–94.

Robert M. Solow, "Technological Change and the Aggregate Production Function," *Review of Economics and Statistics* (August 1957): 312–320.

Robert M. Solow, *Growth Theory: An Exposition* (The Radcliffe Lectures) (Oxford: Clarendon Press, 1970), pp. 8–9.

Robert M. Solow, "Applying Growth Theory Across Countries," *The World Bank Economic Review* 15, 2 (2001): 283–288.

Robert M. Solow, "The Last 50 Years in Growth Theory and the Next 10," *Oxford Review of Economic Policy* 23, 1 (2001): 3–14.

Robert P. Flood and Peter M. Garber, "Collapsing Exchange Rate Regimes: Some Linear Examples," *Journal of International Economics* 17 (August 1984): 1–13.

Ronald I. McKinnon, *Money and Capital in Economic Development* (Washington, DC, Brookings Institution, 1973).

Ross Levine, "Financial Development and Economic Growth: Views and Agenda," *Journal of Economic Literature* 35 no. 2 (1997): 688–726.

Rudiger Dornbusch, "Fewer Monies, Better Monies," NBER Working Paper No. 8324 (Cambridge, MA: NBER, 2001).

Rudiger Dornbusch, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy* 84 (1976): 1161–1176.

Rudiger Dornbush, "Policies to Move from Stabilization to Growth," Proceedings of the World Bank Annual Conference on Development Economics, 1990.

Scott Freeman, "Transactions Costs and the Optimal Quantity of Money," *Journal of Political Economy* 93 (February 1985), 146–157.

Shankar Acharya, "Macroeconomic Management in the Nineties," *Economic and Political Weekly* (20 April 2002): 1515–1538.

Sidney S. Alexander, "Devaluation Versus Import Restriction as an Instrument for Improving the Foreign Trade Balance," International Monetary Fund Staff Papers No. 1, 1951, pp. 379–392; "Effects of a Devaluation on a Trade Balance," International Monetary Fund Staff Papers No. 2, 1952, pp. 263–278.

S. S. Tarapore, "Financial Sector Reform: Retrospect and Prospect," RBI Bulletin (May 1996): 299–306.

Simon Kuznets, *National Income: A Summary of Findings* (New York: National Bureau of Economic Research, 1946).
Stanley Fischer, "Exchange Rate Regimes: Is the Bipolar View Correct?" *Journal of Economic Perspectives* 15, no. 2 (Spring 2001): 3–24.
Stanley Fischer, Ratna Sahay and Carlos A. Végh, "Modern Hyper- and High Inflations," *Journal of Economic Literature* (September 2002): 837–880.
Stephen P. Magee, "Currency Contracts, Pass Through and Devaluation," Brookings Papers on Economic Activity 1, 1973, 303–325.
Stephen S. Everhart and Mauricz A. Sumlinski, "Trends in Private Investment in Developing Countries–Statistics for 1970–2000," International Finance Corporation, Discussion Paper No. 44 (Washington, DC: IFC, 2001).
Stewart. C. Myers, "The Capital Structure Puzzle," *Journal of Finance* 39, no. 3 (July 1984): 575–592.
Stewart C. Myers and Nicholal S. Majluf, "Corporate Financing Decisions when Firms have Investment Information that Investors Do Not," *Journal of Financial Economics* (June 13, 1984): 187–220.
Svend Laursen and Llyod A. Metzler, "Flexible Exchange Rates and the Theory of Employment," *Review of Economics and Statistics* 32 (November 1950): 281–299.
Swarnjit M Arora, Mohsen Bahmani-Oskooee, and Gour Goswami, "Bilateral J-Curve between India and Her Trading partners," *Applied Economics* 35 (2003): 1037–1041.
Tarlok Singh, "Testing J-Curve Hypothesis and Analyzing the Effect of Exchange Rate Volatility on the Balance of trade in India," *Empirical Economics* 29 (2004): 227–245.
Thomas J. Sargent, "The Ends of Four Big Inflations," in Robert E. Hall (ed.) *Rational Expectations and Inflation* (New York: Harper & Row, 1986).
Thomas J. Sargent and Neil Wallace, "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics* 2 (1976): 169–183.
Thomas R. Saving, "Monetary Policy Targets and Indicators," *Journal of Political Economy* 75, no. 4 (1967): 446–456.
Thorvaldur Gylfason (2006), "How do India and China Grow?" *Challenge* 49, no. 1 (January 2006): 74–89.
Tim Callen and Paul Cashin, "Assessing India's External Position," in T. Callen, P. Reynolds, and C. Towe (eds.) *India at the Crossroads-Sustaining Growth and Reducing Poverty* (Washington, DC: IMF, 2001).
Tim Dyson, "India's population: The future," in Tim Dyson, Robert H. Cassen, and Leela Visaria, *21st Century India: Population, Environment and Human Development* (New Delhi: Oxford University Press, 2005).
Toni M. Whited, "Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data," *Journal of Finance* 47, no. 4 (September 1992), pp. 1425–1460.
Vijay Joshi and Ian Malcolm David Little, *India's Economic Reforms* (New Delhi: Oxford University Press, 1996).
Willem Buiter and Urjit Patel, "Excessive Budget Deficits, a Government-Abused Financial System, and Fiscal Rules," in Suman Bery, Barry Bosworth, and Arvind Panagariya (eds.) *India Policy Forum*, No. 2 (New Delhi: Sage Publications, 2006).
William Arthur Lewis, "Economic Development with Unlimited Supplies of Labour," Manchester School of Economic and Social Studies (Manchester: Blackwell Publishing, 22 May1954): pp. 139–191.
William J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics* 66, no. 4 (November 1952): 545–556.
William J. Baumol, "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review* 57 (1967): 415–426.
William Poole, "Optimal Choice of Monetary Policy Instrument in a Simple Stochastic Macro Model," *Quarterly Journal of Economics* 84, no. 2 (May 1970): 197–216.

# Glossary

## A

Absorption approach: An approach to trade balance determination that emphasizes changes in domestic income and expenditure. (139)

Acceleration hypothesis: The thesis that a government that attempts to sustain the unemployment rate below the natural rate will cause accelerating inflation. (225)

Adaptive expectations: A situation where Individuals revise their price expectations in every period to incorporate a fraction of the errors in expectation ($p-p^e$) from the previous period. (214)

Adjustment costs: The costs of integrating new capital goods into the production process. The larger the investment, the larger will be the cost of adjustment. (96)

Aggregate demand (*AD*) curve: The inverse relationship between the real quantity demanded of final goods and services and the general price level. (268)

Aggregate income: The sum of all incomes accruing to the citizens of a country within a given period of time–a year or a quarter. (4)

Aggregate supply (*AS*) curve: The positive relationship between the real quantity of final goods and the services produced in an economy and the general price level. (271)

Appreciation: Occurs when the price of the foreign currency in terms of the domestic currency decreases leading to a decrease in the exchange rate. (36)

Arbitrage: The proposition that the expected rates of returns on two assets must be equal. (108)

Automatic stabilizers: Expenditures that automatically increase, or taxes that automatically decrease, when the aggregate income and output decline. (327)

Average propensity to consume (APC): The ratio of consumption to disposable income. (59)

Average propensity to save (APS): The ratio of savings to disposable income. (60)

## B

Balance of (merchandise) trade: The value of exports less the value of imports of goods. (19)

Balance of payments: It summarizes a country's transactions—monetary payments for goods and services, and acquisition of financial assets—with the rest of the world during an accounting period. (34)

Bond: A bond is a legal promise to repay a debt. The principal amount or the amount originally lent is paid at some fixed date in the future called the maturation date. A bondholder receives regular interest, or coupon payments, until the bond's maturation date. (161)

Borrowed reserves: The reserves that central banks supply directly to commercial banks in the form of advances of credit. (53)

## C

Capital account: A summary of the country's financial asset transactions with the rest of the world. (35)

Capital Controls: Legal restrictions on the ability of citizens to hold and exchange assets denominated in the currencies of other nations. (307)

Capital good: A long-lived good that is used in the production of other goods and services. (8)

Capital theory: Informs us about the optimal stock of capital to be used as an input. (95)

Classical *AS* curve: Occurs when the *AS* curve is vertical, and the quantity of output produced depends on real factors with no effect of nominal variables on output. (209)

Consumer price index (CPI): The cost of a standard basket of goods and services consumed by a typical domestic consumer. (23)

Consumption augmentation: The outcome of resorting to the international capital markets to finance investment whose productivity enables a higher present as well as future consumption. (355)

Consumption function: A positive function of the aggregate disposable income and is negatively related to the nominal interest rate. (238)

Consumption smoothing theory: A preference maximization theory where individuals are forward looking. (63)

Consumption smoothing: Occurs when the consumption in each period is set equal to the permanent income. (70)

Consumption tilting: The case where the consumer's impatience for the present versus future consumption is not equal to the interest rate at which income is discounted. (69)

Convergence: The catch-up phenomenon whereby poor countries would tend to grow faster than rich countries and close the gap between the two groups of countries. (451)

Credibility: The degree to which individuals and markets believe that a policy announcement will be implemented even when there are short-term economic costs that may be imposed. (114)

Crowding out: Occurs when the stimulative effects of fiscal policy are offset by a reduction in the interest-sensitive consumption and investment due to a rise in the interest rate. (251)

Currency board: The requirement that the exchange rate is fixed and the monetary authority commits to buying or selling the domestic currency without any limit at this pre-announced exchange rate. (149)

Current account balance: The statement of a country's payments to and from the rest of the world for goods and non-factor services transacted in, for factor incomes from abroad, and for transfers. (20)

Current account surplus: An accumulation of financial assets by residents due to lending abroad. (348)

Cyclical unemployment: The unemployment associated with periods of low levels of economic activity. (215)

## D

Deficit overall balance: The net decline in the country's stock of international reserves over the year. (35)

Dependency rate: The ratio of dependents (the young and the old population) to the working population. (80)

Depreciation: Depreciation occurs when the price of the foreign currency in terms of the domestic currency rises leading to an increase in the exchange rate. (36)

Depth: The maximum amount or limit that a dealer stands ready to buy or sell at posted prices. (148)

Direct lending: Also called discount window lending, it is the lending of reserves by the central bank to commercial banks. (40)

Discount rate: The interest rate that the central bank charges commercial banks to borrow reserves. (54)

Discretionary monetary policy: Results from the deliberate and conscious choices by the central bank on money supply (or the interest rate) in order to influence output and prices in the economy. (255)

Diversification: The process of reducing risk by holding a number of financial securities in a portfolio. (378)

Dollarization: Adoption of a foreign currency as its legal tender by a country. (149)

## E

Economic growth: The rate of change of the real income (or output) per capita. (435)

Efficiency wages: The wages that maximize firm profits when workers may work or shirk. (295)

Engine of economic growth: The technological progress that occurs exogenously of capital accumulation in the form of skills and human capital acquisition. (459)

Excess reserves: The cash assets held in the vault of banks as a contingency against a need for liquidity due to unpredictable withdrawals and payments. (52)

**Exchange rate**: The value of a given currency relative to another. (35)

**Exogenous variable**: A variable that is taken as given and is not explained within an economic model. (238)

## F

**Final demand**: The goods and services consumed by the ultimate user and equals final purchases by consumers and producers and the changes in stocks. (10)

**Financial products**: The contracts that allow investors to delegate the management of their assets to others. (373)

**Financial repression**: The prevalence of credit controls, interest rate controls, entry barriers into banking, operational restrictions on banks, the predominance of state-owned banks, and restrictions on international financial transactions. (380)

**Fiscal deficit**: The government expenditure that is not related to the repayment of the debt, $G$, plus repayment of the debt, $rB$, less government revenues, $T$. (329)

**Fisher effect**: Describes the full adjustment of the nominal interest rate to a change in the inflation rate. (165)

**Frictional unemployment**: The short-term unemployment associated with matching workers with jobs. (215)

**Full employment**: A state where cyclical unemployment is zero. (216)

## G

**GDP deflator**: The nominal GDP divided by the real GDP. (22)

**Golden Rule**: The condition for maximum sustainable consumption that does not provide less consumption to future than to current generations. (449)

**Government saving**: It is equal to net government revenues less government spending. (32)

**Gross domestic product (GDP)**: The market value of the final goods and services produced in a country during a given time period. (6)

**Gross national product (GNP)**: Equals GDP plus net factor incomes from abroad. (12)

**Growth rate**: The percentage change in aggregate income or output over a unit of time, usually a year. (21)

## H

**Hedge**: A future transaction agreeing to purchase or sell a given quantity of an asset at a future date at a price agreed today in order to offset the existing risk. (377)

**Human capital**: Refers to the levels of education and skills embodied in the labour force. (456)

## I

**Idiosyncratic risk**: The risk that is specific to a financial instrument or project that cannot be diversified away and comprises default risk, legal risk, and operational risk. (378)

**Impossible trinity**: Refers to the dilemma that a country cannot simultaneously have control over a fixed exchange rate and a monetary policy, and maintain an open-capital account. (152)

**Income effect**: The change in the quantity demanded of a good or an asset that results from the effect of a change in the price on the purchasing power of the individual's budget, holding all other factors constant. (77)

**Inside lags**: The delay between the time a policy change is needed and the time the policy change is implemented. (277)

**Insiders**: Incumbent employees who have market power in the determination of wages. (299)

**Intermediate target**: A macroeconomic variable that the central bank can influence with some predictability and is related to its ultimate objectives. (401)

**Intertemporal analysis**: The analysis about the trade-offs when the present choices affect the alternatives available in the future. (63)

**Intertemporal choices**: The choices that take into consideration the consequences for the current time period as well as for future periods. (3)

**Investment equals savings (*IS*)**: It is given by the downward-sloping combination of the nominal interest rate and aggregate income, which depicts equilibrium in the goods market. (306)

**Investment expenditure**: The expenditure resulting in the increase in the stock of capital goods is investment expenditure. (8)

**Involuntary unemployment**: Refers to a situation where the workers are willing to work at the going wage and yet firms do not employ them. (272)

**Isoquant**: The set of input combinations that yields a given level of output. (93)

## J

**J-curve**: The phenomenon where a depreciation of the currency causes the trade balance to initially worsen before it eventually improves. (137)

## L

**Law of One Price**: States that if the international arbitrage is unhindered (the market is efficient) then identical assets (with the same risk and liquidity characteristics) must earn the same rate of return. The Law applied to trade in goods is known as Purchasing Power Parity. (143)

**Learning by doing**: The process by which individuals and firms learn better ways to produce as a by-product of some other activity undertaken such as investment. (456)

**Liquidity**: The ease and speed with which agents can convert assets into purchasing power. (378)

***LM***: The upward-sloping curve in nominal interest rate – aggregate income space depicting equilibrium between the demand for money or liquidity (L) and money supply (M). (245)

## M

**M1**: The sum of currency with the public and demand deposits—the most liquid assets that can be used in transactions. Also called narrow money. (44)

**M2**: $M1$ plus time deposits, with a maturity of less than a year. Also called broad money. (44)

**M3**: $M1$ plus all time deposits regardless of their maturity. (44)

**Marginal product**: The change in output that occurs in response to a unit change in the application of an input. (92)

**Marginal propensity to consume (MPC)**: The amount by which consumption rises when disposable income rises by INR 1. (59)

**Marginal rate of technical substitution**: The rate at which an input can be substituted or exchanged for another without altering the level of output. (93)

**Monetary base**: The money issued by the central bank; also called high-powered money. (39)

**Monetary union**: Occurs when a group of countries share a common currency and jointly manage the currency through a common central bank. (150)

**Money illusion**: A situation when individuals formulate decisions on the basis of changes in nominal values rather than on nominal values adjusted for the price level (real values). (210)

**Money multiplier**: The increase in the money supply as a result of an increase by one rupee in the central bank base money. (45)

**Money stock**: The market value of a sum of liquid assets that are used to make transactions in the economy. (39)

**Money supply**: The money multiplier times the reserve money. (42)

**Moral hazard**: The hidden action arising after a contract is endorsed. (374)

## N

**National income**: The sum of the net domestic product valued at factor cost plus the net factor incomes from abroad. (13)

**National intertemporal budget constraint**: It states that the present value of the aggregate consumption expenditure equals the present value of the national production net of investment. (350)

**National saving**: The sum of private and public savings. (32)

**Natural rate of unemployment**: The sum of the seasonal, frictional, and structural unemployment. Analytically it is the unemployment rate at which price expectations are realized. (215)

**Net domestic product:** GDP less depreciation. (11)

**Net investment:** It is the gross investment less depreciation. (9)

**Nominal GDP:** The sum of the quantities of goods produced times their current price. (22)

**Nominal incomes:** Incomes measured in terms of current rupee values. (21)

**Nominal interest rate:** The real rate of interest plus the rate of inflation. (165)

**Non-accelerating inflation rate of unemployment (NAIRU):** The rate of unemployment at which the current rate of inflation will continue, neither accelerating nor decelerating. This is the point at which the Phillips curve cuts the X-axis. (225)

**Non-borrowed reserves:** The reserves the central bank supplies to commercial banks through open-market operations. (53)

## O

**Okun's Law:** States that each percentage point increase of cyclical unemployment is associated with a three percentage point increase in the output gap. (216)

**Open-market operation:** The purchase or sale of government bonds by the central bank for the purpose of increasing or decreasing the money supply. (40)

**Output gap:** The difference between the potential output and the actual output as a fraction of potential output. (216)

**Outside lag:** The time required for a policy change to affect the economy. (277)

**Overall balance:** The sum of the current account balance and the capital account balance. (35)

## P

**Pareto efficiency:** A condition in which no change in outcome that will make some members of society (such as the unemployed) better off without making some other members of society (such as the employed) worse off is possible. (286)

**Pass-through effect:** Arises when a change in the exchange rate results in a change in the domestic prices of imported goods. (138)

**Per capita income:** The aggregate income divided by the population of the country. (4)

**Permanent income:** The constant stream of income that equals the present discounted value of the fluctuating stream of current and future incomes. (70)

**Phillips curve:** The mirror image of the *AS* curve that posits a negative trade-off between the price level and the unemployment in its static version and between inflation and the unemployment rate in its dynamic version. (220)

**Policy ineffectiveness, proposition:** States that the government cannot systematically exploit the Phillips curve to guide the economy to a point on it. (227)

**Potential output:** The maximum sustainable amount of output that an economy can produce. (216)

**Poverty trap:** The tendency of an economy to return to a low-level steady state despite attempts to raise the per capita output. (455)

**Private saving:** The value of the private sector's disposable income minus its consumption. (32)

**Production function:** Describes how inputs like capital and labour are transformed into output. (92)

**Prudential regulation:** Refers to the set of laws or rules designed to minimize the risks to the financial system. (379)

**Purchasing power parity:** States that if arbitrage across national borders is unhindered, the price of a good or service in one nation should be the same as the exchange rate adjusted price of the same good in another nation. (145)

## R

**Real exchange rate:** The price of the average foreign good or service relative to the price of the average domestic good or service, when the prices are expressed in terms of a common currency. (131)

**Real GDP:** The aggregate income measured in the prices of a base year. It measures aggregate income in constant rupees or adjusted for inflation. (22)

**Real incomes:** The incomes measured in physical terms, that is, in terms of the quantities of goods and services. (21)

**Real rate of interest:** The nominal interest rate minus the rate of inflation. (236)

**Rental cost of capital:** Also referred to as the user cost of capital; The cost of using capital over a given period of time (a year) and is the sum of the real interest rate and depreciation rate times the unit price of the capital good. (91)

**Rental:** The price of a unit of housing times the rate of interest plus depreciation. (108)

**Required reserves:** The result of the central bank regulation requiring banks to hold specified fractions of their deposits—either as vault cash or as funds on deposit at the central bank. (40)

**Reserve or base money:** The sum of the domestic credit advanced by the central bank and its foreign exchange assets. (307)

**Resource constraint:** Depicts the locus of all possible input combinations that can be purchased for a given level of resources available to spend on purchasing these inputs. (92)

**Ricardian equivalence:** The proposition that in terms of the impact on the aggregate income it is irrelevant whether increased government expenditure is financed by taxes or bonds. (257)

**Risk pooling:** The process by financial intermediaries to protect themselves from excessive risk through sharing risks of independent investment projects. (378)

## S

**Seasonal unemployment:** Varies predictably with the seasons. (214)

**Seigniorage:** The use of money creation as a revenue device by the government. (175)

**Shock:** A shock is an unexpected disturbance to the economy that affects the demand or supply side of the economy. (71)

**Slope of an isoquant:** The marginal rate of technical substitution which, in turn, is the ratio of the marginal products of labour to the marginal product of capital. (94)

**Solow residual:** The estimate of the technological progress derived as a residual after accounting for the rate of growth of output due to labour force and capital stock growth. (437)

**Spot exchange rates:** The exchange rate when two parties agree to an exchange of financial instruments and execute the deal immediately. (142)

**Stationary state:** A state where the growth rate of a variable is constant, or, where there is no tendency for the variable to change with time. (223)

**Statutory liquidity ratio:** Specifies the required holding of government securities as a fraction of the assets of the commercial bank. (255)

**Steady state:** The state when various quantities grow at constant rates. (177)

**Steady-state debt/GDP ratio:** That value of the debt/GDP ratio that satisfies the government budget constraint and is independent of time. (331)

**Sterilization:** The central bank policy of altering the domestic credit in an equal and opposite direction to the variation in foreign exchange reserves so that the monetary base remains unchanged. (43)

**Structural unemployment:** A long-term unemployment that does not vary with the pace of economic activity. (215)

**Substitution effect:** The change in the quantity demanded of goods or assets that results from a change in the price making it more or less expensive and thereby more or less attractive relative to other goods or assets, holding constant the effect of the price change on the purchasing power of the individual's budget constraint. (77)

**Sunk cost:** An investment in an asset with no alternative use, also referred to as a specific asset. (112)

**Surplus overall balance:** The net increase in the country's stock of international reserves over the year. (35)

Systemic risk: The default risk associated with general economy-wide or macroeconomic conditions that affect all borrowers and lenders. (385)

## T

Terms of trade: The relative price that is given by the price of exports (P) divided by the price of imports (EP*). (140)

Tobin's marginal Q: The ratio of the marginal value of the capital stock (computed by adding the market value of the additional equity and the debt raised by the firm to finance the addition to capital stock) to the rental cost of capital. (98)

Total investment expenditure: Sum of business investment, residential investment, and inventory investment. (107)

Treasury bill: A government bond with a maturity period not exceeding one year. (255)

Two-period budget constraint: The present value of the income in the two periods less the discounted value of interest lost on account of holding money balances rather than bonds. (167)

## V

Value of waiting: The expected present value of the irreversible mistake when undertaking physical investment that would be revealed in the future should the returns fall short of the costs. (113)

## W

Walras' Law: States that the sum of the excess demand for money, bonds, and current output must equal zero. (236)

Wealth of the private sector: Consists of its net claims on agents outside that sector—the stock of money and the stock of government bonds. (163)

Wholesale price index (WPI): A price index of goods and services produced domestically. (23)

# Index


## A

Absolute convergence, 451
Absorption approach, 139–140
Active crawling peg, exchange rate parity in, 150
*AD* curve. *See* Aggregate demand curve
Adaptive expectations hypothesis, 214
Adjustable peg, exchange rate parity in, 150
Adjustment costs, 96
  level of investment to maximize profits, 98
  Tobin's marginal Q, 98–99
Adjustment lags, 138
Affective systems, 67
Aggregate demand curve, 195, 236, 267–270, 290–291, 327
  policy lags, 276–279
  shifts, 268–270
Aggregate domestic income, 11, 13
Aggregate economy, 79
Aggregate expenditure, components, 15–19
Aggregate income,
  definition, 4
  omissions in estimation of, 5–6
  and welfare, 4–5
Aggregate investment, 33, 238–239
Aggregate output, 8–9
Aggregate savings, 33, 238–239
Aggregate supply, mistaken expectations of, 274–276
  with/without money illusion, 207
Aggregate supply curve, 207, 271
Aggregation, across individuals, 79–80
Agricultural output, 14
Amount of reserves, 54
Animal spirits, 242
Anticipated future shock, 73–74
APC. *See* Average propensity to consume
APS. *See* Average propensity to save
Appreciation, 36
Arbitrage, 108, 148
*AS* curve. *See* Aggregate supply curve
Asian Crisis of 1997, 152, 319, 332
Asset market condition, 102
Asset prices, impact of change in exchange rates, 142–145
Asymmetric information, 101
Audit costs, 375
Automatic stabilizers, 327
Average propensity to consume, 59–60
Average propensity to save, 60

## B

Balanced Budget and Emergency Deficit Control Act, 339
Balance of (merchandise) trade, 19
Balance of payments accounts, 34–38
  crisis, 194–197
Balance sheet
  of the Central Bank, 39, 42, 306
  commercial banks, 40–41
  consolidated, of a banking system, 41–42
  consolidation, 41
  of households and firms, 162
  and money supply, 39–45
  of private sector, 41, 163
Banks
  aspect of human capital issues in, 394–395
  asset holding decisions and liberalization policy, 392
  balance sheets and money supply, 39–45
  capital adequacy, 393
  Central Bank, 39–55
  Commercial banks, 39–55
  and intermediate targets, 402
  and macroeconomic uncertainty and instability, 392
  policy of monetary targeting, 404–411
  response to financial reforms, 390–393
  solvency, 393
Barter economy, 161
Base year, 21
Basle banking norms, 383
Black economy. *See* underground economy
Bond prices, 161–162, 172
Bonds, 161, 375
  expected capital gain on, 173
  proportion of wealth in, 172
  riskiness of the portfolio, 173
Borrowing constraints
  and consumption, 88–89
  and joint families, 75
Budget constraints, 165–171, 204–205, 282
  of a government, 329–333
  two period, 186
Business cycle, 278
  Keynesian view of, 279–286

## C

Call option, financial, 113
Capital account, 35
  convertibility, 311
Capital account balance, 359
Capital gain, 102
Capital good, 8
Capital markets, 354–355
Capital mobility, 308
Capital theory, 95
Cash-holding ratio, 46
Cash reserve ratio, 383
Central bank, relevance in modern times, 43
CEO overconfidence, 100
Chakravarty Committee, 391, 403
China
  financial aggregate in, 402
  fiscal rules, 338
  growth accounting, 437
  savings rate in, 80
Cobb–Douglas production function, 94, 435–436
Cognitive systems, 67
Collateral financing, 105–107
Competitiveness, of a country's goods and services, 133
Conditional convergence, 452
Consumer price index, 23–24
Consumption, 70
  and borrowing constraints, 88–89
  coefficient of variation in, 62
  current account position when all output is consumed, 350
  current and lagged, 62
  and disposable income, 61–62
  effects of shocks on, 71–75
  Hall's result on, 77
  and housing prices, 81
  individual's decision-making, effects of, 79–80
  interest rate changes, effects of, 77–79
  intertemporal budget constraint facing a consumer, 63–64
  Keynes' theory of, 59–63
  preference maximizing choice and real money balances, 184–185
  present value of life-cycle, 65
  and savings portfolios, 81–85
  smoothing, 63–71, 69
  smoothing under stochastic income, 76
  in a stochastic environment, 75–77
Consumption augmenting, 355
Consumption smoothing, 355
Consumption smoothing theory, 63
Consumption tilting, 69–70, 355
Convergence, of economies, 451–453
Convertibility Law, governing Argentina's currency board, 149
Cost of investment, adjustment, 96
Covered interest parity, 143
CPI. *See* Consumer price index
Crawling peg, exchange rate parity in, 150
Credibility, defined, 114
Credit aggregate, 402
Credit facility, 40
Crowding out, 249–252, 315
CRR. *See* Cash reserve ratio
Currency board, 149
Currency-money supply behavioural ratio, 47
Current account, 34
Current account balance, 20
  borrowing/lending, determinants of, 354–356
  changes in the ratio of foreign assets to GNP, 366–367
  determination of, 352–358
  expansionary fiscal contractions, effect, 364
  expenditure-output determination, 361–363
  fiscal adjustment, 363–365
  global imbalances, 352
  and intertemporal accounting, 347–352
  *IS–LM–BP* model, 358
  Japan, 365
  in a medium-run model, 358–365
  permanent, 367
  and shocks, 357–358
  short-run Mundell–Fleming type model, 358
  sustainability in the context of, 365–369
  United States, 365
Current account deficit, 348
Current account surplus, 348
Current balance rules, 339
Cyclical unemployment, 214–219

## D

Data lag, 277
Debt contracts, 375
Debt/GDP ratio, 330–331
Default risk, 378
Delta (δ), higher values of, 68
Demand shocks, 280
Dependency rate, 80
Deposit-money supply behavioural ratio, 47
Depreciation, 9, 36
Depreciation rate (δ), 99, 117n
Depth, of country's financial sector, 148
Devaluation
  and adjustment lags, 138
  contractionary, 141
  in developing countries, 139–141
  evidence of, in India, 141
  imported input cost channel, impact, 141
  income-distribution effect, 141
Developing countries
  devaluation effect, 139–141
  and economic growth, 453–456
  investments in, 111, 114–116
  working capital finance and effect of fiscal policy, 273
Development banking, 402
Dirty float, 38
Discount rate, 54
Discount window, 53
Discretionary fiscal policy, 328–329
Discretionary monetary policy, 255
Diversification, 378
Dollarization, 149–150
Domestic expenditure, 17
Domestic financial instrument, 142–143
Dynamic or time inconsistency, 416

## E

Economic growth
  aggregate output, 438–441
  capital stock growth and, 448
  comparative statics, 445–448
  convergence of economies and, 451–453
  determinants of long-run, 441–443
  endogenous, 456–460
  engine of, 459

Economic growth (*Continued*)
golden rule of, 449–451
long-run, 435
neoclassical growth equation, 442
and population growth, 447, 454
and poverty, 453–456
Romer-type model of, 456–459
Solow diagram, 443–445
sources, 435–438
Economic regulation, 379
Efficiency wages, 294
Egypt's underground economy, 7
Elasticity of demand, measure of, 134
Emotions and intertemporal decisions, 67
Employment, equilibrium, 440
Endogenous growth theory, 456–460
Equilibrium exchange rate, 144
Equilibrium in the product market, 238
Equity
cost of finance in the form of, 103
value of the firm's, 102
Euler equation for consumption, 185
Euler's theorem for homogeneous production functions, 127
The Euro, 151
Euroization, 149
*Ex ante* real interest rate, 237
Excess reserves, 52
Exchange rate, 35
determination, 144
equalization of the domestic and foreign rates of return, 143
policy, 341
regimes, 148–153
Exchange rate crisis in Argentina, 319
Exogenous variable, 238
Expansionary fiscal policy, 315, 317
Expenditure-limiting rules, 339
Expenditures
Government, 16
by the private sector, 16
Expenditure switch, 139
Exports, impact of change in exchange rates, 132–134
External finance, in investment decisions, 101–102

## F

Fear of floating, 153
Final demand, 10
Final purchases, 10
Financial capital, 91
Financial markets
adverse selection, 373–374
asymmetric-information situation in, 373
banks, 375–377
capital markets, 376–377
contracts, 375
debt markets, 375–376
financial products, 373
functions of, 377–379
hidden information, 373–374
in India, 379
intermediaries, 379
moral hazard, 374
reforms, 379
transactions in, 373
Financial products, 373
Financial repression, 380
and major reforms in India, 379
policy dimensions, 381
Financial sector reforms
bank response to, 390–393
financial markets, 373–379
and human capital, 394–395
implications of supply of loans by banks, 385–389
in India, 379–384
Financing hierarchy, of the firm, 103
collateral, 105–107
debt financing, 104–105
equity financing, 101–103
liquidity constraints and mergers, 105–107
of mature firms, 105
Firm's factor accumulation decision, 123–27
Firm-specific capital, 112
Fiscal deficits, 329
sustainability of, 255
Fiscal Responsibility Act, 332
Fiscal Responsibility and Budget Management Act, 338–340
Fiscal restructuring, 340–344
Fiscal rules
for adjustment, 338–340
in Brazil, 338
Fiscal surpluses, 332
Fisher effect, 165
Fixed exchange-rate system, 131, 148
fiscal policy under, 316–317
monetary policy under, 315–316
Fixed-price bond, 161
Flexible exchange-rate system, 131
fiscal policy under, 314–315
monetary policy under, 312–314
Flow of goods, 8
Foreign-exchange market, 35, 313
equilibrium of, 144–145
Foreign financial instrument, 142
Foreign savings, 33
Forward exchange rates, 143
FRA. *See* Fiscal Responsibility Act, 332
FRBMA. *See* Fiscal Responsibility and Budget Management Act
Frictional unemployment, 215
Friedman, Milton, 70
Friedman mistaken expectations framework, 274–276
Full employment, 216

## G

GDP, 7, 11
and components, 12
emerging economies, 19
factor cost by industry, 14
growth rate of, 21
interest rate and monetary targeting, implications, 411–416, 429–431
nominal, 403, 411–416
of the public sector, 16
versus GNP, 12
GDP deflator, 22, 28–29
inflation rates based on, 25
GNP. *See* Gross national product
Government budget constraint and debt dynamics
debt/GDP ratio, 330–331
fiscal deficit, 329
fiscal restructuring, 340–344
fiscal rules for adjustment, 338–340
intertemporal budget constraint, 329
primary deficit, 329, 331–333
primary surpluses, 333
steady-state debt/GDP ratio, 331
sustainable primary deficit/GDP ratio, 337–338
Government expenditures, 16
Government savings, 32
Great Depression, The, 280–281
Gramm-Rudman-Hollings legislation. *See* Balanced Budget and Emergency Deficit
Control Act
Gross national product, 12, 348
Growth rate, 21

## H

Harassment, 299
Hedge, 377
Hicks, John, 245
Housing
demand and supply of, 108–109
difference between renting and owning, 118n
investment expenditure in, 109–110
Human capital, 456
Hyperinflation, 179

## I

Idiosyncratic risks, 378, 385
Imperfect capital mobility, 308
Implementation lag, 277
Imported goods
impact of devaluation, 140
real value of, in terms of domestic goods, 134
Impossible trinity, 152
Income effect, 77, 140, 206
Income growth, in market-oriented economy, 80
Incomes policy, 343
India
consumption function for, 61
current account balance, 20
debts and deficits, 333–337
devaluation, 141
as emerging economy, 19
monetary policy in, 403
financial markets in, 379
financial reforms, 379–384
foreign trade, 20
growth accounting, 437
income-tax laws, 340–341
inflation tax in, 338
investments, 115–116
net factor payments abroad, 13
seigniorage position, 181
sterilization process, 320
unemployment, 215, 217–218
Inflation, 401
discretion and surprise (unanticipated), 420–422
expectations, 226–227
measure of, 23–26
Information asymmetry problems, in financial markets, 104
Inside lags, 277
Insider–outsider theories, to employment, 299–301
and macroeconomic policy, 300
Insiders, 299
Inter-bank interest rate, 52–53
Interest payments and profits, 10
Interest rates, 99, 403
charged on loans, 387
and consumption, 77–79
liberalization, on government securities, 390
market determination of, 390
pass-through effects of, 277
targeting, 404–411
Intermediate good, 7
Intermediate target, defined, 401
Interporal constraints and preferences, 63–67
Intertemporal accounts
budget constraint in two-period model, 349–352
current, determination of, 352–358
of a nation, 347–349
in terms of the trade balance, 351–352
Intertemporal analysis, 62
Intertemporal budget constraint, 65, 338
Intertemporal choices, 2
Intertemporal indifference curve, slope of an, 65, 66
Inventory demand for money, 198–201
Inventory investment, 8, 110
Investment, 8
and adjustment costs, 95–100
crowding out effects, 252
in developing countries, 111, 114–116
expenditure in the economy, sources, 33–34, 110
financial structure and, 100–107
implications for a $Q$ model of, 117n
irreversibility and, 111–114
marginal benefit of investing an additional unit, 113
overconfidence and, 100
rate of increase and economic growth, 446
residential and inventory, 107–111
and savings, 34
spending of the firm, 104
traditional theory of, 112
Involuntary unemployment, 272
Irreversibility of investment, theory, 111
*IS* curve, 268
aggregate investment and savings, 238–239
of an open economy, 306
case of expansionary monetary policy, 313–314

condition for equilibrium in the product market for a closed economy, 238–240
consumption function, 238
government expenditures, 238, 241
impact of fiscal and monetary policy, 247–254
increase in the government expenditure at a given price level, 269
in India, 255–256
investment expenditure, 238
in Japan, 263–265
Keynesians and Monetarists views of shifts, 274
parametric changes of, 240–242
Ricardian equivalence, 256–259
shocks, 404–406
slope of, 242–243
tax increase and, 241

*IS-LM* model, 234–263
*IS-LM-BP* model, 358
*IS-LM* in India, 255
*IS-LM* in Japan, 263
Isoquant, 93
Isoquant map, 94

## J

Japanese economy, 254
J-curve, definition, 137
J-curve effect, 136–138, 141

## K

Keynesian aggregate supply curve, 211–212
AD-AS equilibrium, 271–274
Keynesian *AS* curve, 271

## L

Labour market
aggregate supply with/without money, 207–210, 230–233
bargained real-wage curve, 439
classical supply curve, 209
distribution of wages, 210
effect of price increases, 208–209
effect of wage increases, 207
equilibrium, 208
intertemporal supply decisions, in real business cycle, 283–284
optimal supply, 206
Phillips Curve, dynamic version, 221–227
Phillips Curve, static version, 219–221
profit maximization and, 203
short-run demand curve, 203–204
unemployment, 210–219
utility function and labour supply, 204–207

Lags, in monetary policy, 276–279
Laspeyre price index, 23
Law of One Price, 143
Legal risk, 378
Lemons problem, 101
Lender of last resort, 40
Liberalization policy, 392–393
Life-cycle income
and consumption, 79
and savings, 80

Lifetime budget constraint, 261n
Liquidity, 378
*LM* curve, 268
case of expansionary monetary policy, 313–314
definition, 243
demand for real money balances, 244–245
impact of fiscal and monetary policy, 247–254
impact of varying the money supply, 245
shift, 270
in India, 255–256
in Japan, 263–265
Keynesians and Monetarists views of shifts, 274
Ricardian equivalence, 256–259
shocks, 407–409
slope of, 245–247

Loans and deposits, supply and demand
asset portfolio of the bank, 388–389
risk and supply of credit, 385–388

## M

*M*1, 44
*M*2, 44
*M*3, 44
Maastricht Treaty, 151, 339
Managed exchange rate regime, 38
Managed float, 150, 152
Marginal product, 92
of capital, 97

Marginal propensity to consume, 59–60
Marginal rate of substitution, 65
Marginal rate of technical substitution, 93
Marginal rate of time preference, 66
Market Stabilization Scheme, 320
Market value of a firm, 100
Marshall–Lerner condition, 134–137, 158–159
Mature firm financing, 105
Miller, Merton H., 79, 117n
Modigliani, Franco, 79, 117n
Monetary base, 39–40, 46
approaches to increase, 54–55
central bank policy and, 54

Monetary expansion, in a fixed exchange rate regime, 316
Monetary policy
Barro–Gordon model rule, 422–423
in case of demand shocks, 404–409
in case of supply shocks, 409–411
and credibility, 424–426
in a deterministic environment, 404
discretion *versus* rules, 416–426, 432–433
goals and intermediate targets, 401–403
implication on nominal GDP, 411–416, 429–431
implications, 277–278
rules, 424

Monetary union, 150
Money
allocation to balances, 168
cost of holding on, 167
crowding out effects, 252
demand for, in the Indian economy, 174–175
demand for, 171, 174
determinants of demand for, 168–169
effect of rise in interest rates, 170, 174
financial assets and budget constraints, 165–171
in the form of bonds, 161–162
functions of, 161
inventory demand for, 198–201
marginal utility of, 189
net wealth of private sector, 162–164
nominal and real exchange rates, 164–165
parametric changes on demand, 169–171
portfolio share of, 171
relationship between inflation and demand, 177
and stock market, 169
as a store value, 171–175
theories on stock of (base) money, 175–181
trade-off between holding and consumption, 167
utility function, 188–189

Money multiplier, 45–51
formula for, 57
value of, 48

Money stock, 39
sources of, 43

Money supply, 39, 42, 307
applications, 45
and balance sheets of the central bank and the commercial banks, 39–45
implications for, 48
increase from open-market purchase, 51
and the *LM* curve, 245
model, 49
money multiplier, 45–51
open-market operations and reserves, 51–55

MPC. *See* Marginal propensity to consume
MRS. *See* Marginal rate of substitution
MRTS. *See* Marginal rate of technical substitution
MRTP. *See* Marginal rate of time preference
MSS. *See* Market Stabilization Scheme
Mundell–Fleming approach
to an open economy, 347
current balance, determination of, 307
fiscal contraction, impact of, 347, 364
under a flexible exchange rate, 347
to income determination, 358
model of open-economy, 312
short-run, current account balance, 358, 364

Myers–Majluf model, of corporate finance, 101

## N

Narasimham Committee recommendations, 382, 391
National income, 8, 13, 85n
in India, 14
in the United States, 14–15

National saving, 32
Natural rate of employment and output, 213
Natural rate of unemployment, 215
NDP. See Net domestic product
NDP at factor cost, 11
NDP at market prices, 11
Neoclassical Solow growth theory, 436
Net domestic product, 11
Net exports, 17, 133
Net factor incomes from abroad, 12
Net investment, 9
Net non-monetary liabilities, 40
Net wealth, of the private sector, 162–164
Nicaraguan hyperinflation, 1986-1991, 179
Nigeria's underground economy, 7
NMLCB. *See* Net non-monetary liabilities
Nominal exchange rate, 131–133, 164–165
Nominal expenditure, 22
Nominal income, 21
Nominal interest rate, 236–237
Nominal wealth, of the private sector, 166
Non-labour income, 205
Non-performing asset, 384
NPA. *See* Non-performing asset
NX. *See* Net exports

## O

Okun's Law, 216
Open economy model
balance of payments of, 307–312
capital flow, 317–321
capital mobility, 307–312
currency appreciation/depreciation effect, 309
fiscal policy under fixed exchange rates, 316–317
fiscal policy under flexible exchange rates, 314–315
macroeconomic accounting for, 305–307
monetary policy under fixed exchange rates, 315–316
monetary policy under flexible exchange rates, 312–314
Mundell–Fleming model, 305–306, 312

Open-market operations, 40, 51–55, 255
Open-market purchase by the central bank, effect, 50
Operational risk, 378
Optimal capital stock, 94, 120
Optimal consumption, 68
Optimum capital stock, 95
Original sin, 149
Output, measuring of
aggregate expenditure, 15–21
aggregate income categories, 11–15
aggregate output, 8–9
feasible, 437
GDP, 7
goods/services included, 6–7
per capita as a function of capital stock per capita, 442
relation with income, 9–11

Output gap, 216
Output sold, 9
Outright operation, 51
Outside lag, 277
Overall balance account, 35
Overall-balance rules, 339
Owner-occupied housing, market for, 108

## P


Parallel economy. *See* underground economy
Pareto efficiency, 286
Passive crawling peg, exchange rate parity in, 150
Pass through effects, 138
Pecking order, in corporate finance, 101
Per capita income, defined, 4
Perfect capital mobility, 308
Permanent income, 67–71
Permanent income, 70
Perpetuities, 161
  price of, 162
Personal disposable income, 59
Phillips Curve, 419
  dynamic version, 221–227
  long-run and short-run, 223–225
  static version, 219–221
Physical capital, 91
Post-reform intercept dummy, 115
Potential or full employment output, 216
Poverty traps, 455
PRD. *See* Post-reform intercept dummy
Preferences, 65
Premium, associated with new equity issues, 101, 102
Price elasticity of demand for imported goods, 134–135
Price level
  and *AD*, 271–274
  determination of, 275
Price of capital goods, 98
Primary deficit, 329
Primary surpluses, 333
Private demand shocks, 280
Private disposable income, 165
Private savings, 32
Private sector expenditure, 16
Privatization, 341–342
Production function, 92
  capital, cost of, 91
  derivative equations, 464–467
  feasible, 439
  firm's factor accumulation decision, 123–128
  for given marginal product of capital, 97
  isoquant for production output, 93–94
  labour, cost of, 91–92
  marginal productivity, 92–93
  marginal rate of technical substitution, 93
  production function, 92
  in profit (II) maximization, 94
  resource constraint, 92
  slope of isoquant, 94
Production Processes, 8
Productivity shocks, 281
Profit (II) maximization, 91–94, 203
Profit maximization and the optimal capital stock, 91
Profit-maximizing firms, 203
Prudential regulation, 379
public debt, in New Zealand, 332
Public debt rules, 339
Public-sector savings, 33
Purchasing power parity, 145
  absolute version, 146
  Big Mac index, 147
  relative version, 146
  weakness of, 147


## Q


Quasi-fiscal deficits, 341


## R


Rational expectations theory, 226–227
Real business cycles, 279
Real business cycle theorists, 281
Real exchange rate, 131, 164–165
  and demand for imported goods, 135–136
  the J-curve, 137–38
  and net exports, 132–134
Real expenditure, 22
Real income, calculation, 21–23
Real interest rate, 236–237
Real wages, , 284–285
Recognition lag, 277
Rental cost of capital, 91–92
Rental rate, on housing, 108
Repurchase agreement, 51
Required reserve ratio, 41, 52
Required reserves, 40
Reserve money, 40, 47, 306–307
  components, 44
Reserve requirements, 39
Reserves, 51–55
  bank, 52
  borrowed, 53
  demand and supply of, 52–53
  discount window, 53
  non-borrowed, 53
  total supply of, 53
Residential investment and inventory, 107
Residential investment spending, 108
Resource constraint, 92
Revenues, 10
Ricardian proposition, 256–259
Risk pooling, 378
Risky asset, 82–83
  increased savings allocated, 84


## S


Savings, 166
  defined, 71
  effect of interest rate increase on, 78
  golden rule for, 450–451
  impact of shocks, 71–75
  implication of life-cycle theory, 75
  and portfolio choice, 81–85
  precautionary motive, 75
  rate, 60
  rate in China, 80
  rate in Taiwan, 81
  sources and uses, 31–34
Schumer–Graham bill, 318
Seasonal unemployment, 214
Seigniorage, 175
  India's position, 181
  means of obtaining, 179–181
  optimal level, 176–179
  and real money demand, 187–193
Self-employed individuals, in India, 14
Shadow economy. *See* underground economy
Shirking of workers, 294–296
Shock, 280–282, 347
  anticipated, 73
  defined, 71
  and monetary targeting, 404–411
  permanent negative, 74–75
  temporary current negative, 72
  temporary current positive, 72
Short-run profits, 102
Slopes
  of an indifference curve, 65–69, 77, 85, 185, 205–206
  of an isoquant, 93, 94
  of an open-economy *IS* curve, 306
  of the *BP* curve, 308–309, 322n
  of the budget line, 77, 168, 188, 205, 283
  of the income constraint, 169
  intertemporal budget constraint, 64–65
  of the *IS* schedule, 242–243
  of *LM* and *BP* Curves, 324–325
  of the *LM* schedule, 245–247
  marginal productivity of labour, 127
  of money supply and demand determination, 47–48, 52–53
  of Phillips Curve, 223–224
  relation between consumption and disposable income, 61
  of the resource constraint, 92–93, 122
SLR eligible securities, 391
Smooth consumption, 62
Social welfare function, for policymakers, 417–418
Soft pegs, 150
Solow growth model, 442, 452
  stability in, 444
  steady state in, 445
Solow residual, 437
Speculative attack, 194
Spot exchange rate, 36, 142
Staggered wage contracts, 297–298
Standard choice theory, 67
Statutory liquidity ratio, 255, 327
Steady-state debt/GDP ratio, 331
Sterilization process, 43, 316, 320
Sticky wages, 296
Stochastic environment expectations, 75–77
Structural unemployment, 215
Substitution effect, 77, 206
  of a rise in the interest rate, 78
Sunk cost, 112
Supply shocks, 280–281
Sustainability, 365–369
Sustainable primary deficit/GDP ratio, 337–338
Systemic risk, 385


## T


Taylor's monetary policy rule, 424
Technology shocks, 282
Terms of trade, 140
Time preference, 67–71
Tobin's marginal *Q*, 98–99, 103, 113
  and rate of interest, 129
Total investment expenditure, 107
Trade balance
  approaches to, 139–141
  and asset influence, 142–145
  exchange rate regimes, 148–153
  and permanent current account, 368
  purchasing power parity, 144–148
  and real exchange rates, 131–138, 157–159
Trade deficit, 351, 367
Trade-offs between alternatives, 63
Transfers, by government, 32
Treasury bill or T-bill, 255, 327
net factor payments abroad, 13
Two-period budget constraint, 167


## U


Uncovered interest parity, 143
Underground economy, 7
  in Mexico, 7
  in Thailand, 7
Unemployment, 210–219
  and wages in India, 217
  efficiency wage approach, 293–296
  Friedman's views on, 211–214
  and inflation trade-offs, 421
  insider–outsider approach, 299–301
  in long-run and staggered wage setting, 296–298
  quasi-equilibrium rate of, 297
  in shirking model, 293–296
  trade-off with inflation, 219–227
Unit capital rental, 91
Unofficial economy. *See* underground economy
UPS. *See* Usual principal status
Usual principal status, 217
Utility function, 68
  of preferences, 65
Utility of consumption, 68


## V


Value added, 7, 10
Value of marginal product of capital, 98
Value of the firm, 102
Value of waiting, 113
Volatility
  of exchange rate, 148–149
  of inflation, 115
  of output growth, 115


## W


Wage adjustment, 298
Wage curve, 439
Wages, efficiency of, 298
Walras' Law, 235–236
Wholesale price index, 23–24
WPI. *See* Wholesale price index

## Dear Student:

We hope that this textbook has been a useful resource and learning tool for you. We would appreciate hearing about your experiences with this textbook and its supplements. We will credit all contributions and, in addition, the three best contributions will be offered gifts worth Rs. 1000. The feedback you provide will help us continue to develop better textbooks. We thank you for giving us your valuable time, suggestions, and comments.

1. Several features are present in the text to make the material more accessible and interesting. Using other texts as the basis of comparison, please rate the following features on a 5-point scale (1—least useful and 5—most useful):

| | | | | | |
|---|---|---|---|---|---|
| Critical Questions | 1 | 2 | 3 | 4 | 5 |
| Margin Notes and Definitions | 1 | 2 | 3 | 4 | 5 |
| Examples | 1 | 2 | 3 | 4 | 5 |
| MacroFocus Boxes | 1 | 2 | 3 | 4 | 5 |
| Tables/Figures/Charts | 1 | 2 | 3 | 4 | 5 |

Any other comment: ..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................

2. The following features are present at the end of each chapter. Please rate each of them on a 5-point scale (1—least useful and 5—most useful).

| | | | | | |
|---|---|---|---|---|---|
| Summary | 1 | 2 | 3 | 4 | 5 |
| Test Yourself | 1 | 2 | 3 | 4 | 5 |
| Online Applications | 1 | 2 | 3 | 4 | 5 |

Any other comment: ..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................

3. The following supplement resources are present with this book. Please rate each of them on a 5-point scale (1—least useful and 5—most useful).

| | | | | | |
|---|---|---|---|---|---|
| Companion Web site | 1 | 2 | 3 | 4 | 5 |
| PowerPoint Slides | 1 | 2 | 3 | 4 | 5 |

Any other comment: ..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................

4. How would you describe the writing style and readability in these chapters? ....................................................
..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................

5. Is the text easier or more difficult to read and understand than your current text(s)? In what way? .................
..........................................................................................................................................
..........................................................................................................................................
..........................................................................................................................................

*Continued on back* ➲

6. What is your opinion about the figures/ tables/ boxes in the chapters? How can they be improved? ..................
....................................................................................................................................................................
....................................................................................................................................................................
....................................................................................................................................................................

7. Do you feel that the examples and exercises helped you understand the concepts? Is the number of these sufficient? Please give instances of some specific examples that you feel should be added. .....................................
....................................................................................................................................................................
....................................................................................................................................................................
....................................................................................................................................................................

8. Overall, what did you like most about this book? And what did you like the least? How can this book be improved? ........................................................................................................................................................
....................................................................................................................................................................
....................................................................................................................................................................
....................................................................................................................................................................

9. Please tell us about the textbooks that are currently used in your course.

Book No. 1 (Title) ..........................................................................................................................................

Author ..........................................................................................................................................................

Publisher .....................................................................................................................................................

Book No. 2 (Title) ..........................................................................................................................................

Author ..........................................................................................................................................................

Publisher .....................................................................................................................................................

Please fill in the form below using CAPITAL LETTERS.

INSTITUTE: ..................................................................................

INSTRUCTOR'S NAME: ...........................................................

YOUR NAME (OPTIONAL): ....................................................

E-MAIL: ..........................................................................................

DATE: ..............................................................................................

May we have permission to quote your comments in promotions for *Macroeconomics*?

☐ Yes ☐ No

*Please send us your feedback, via e-mail or post, to:*

PEARSON EDUCATION
A division of Dorling Kindersley (India) Pvt Ltd
Head office: 482 FIE, Patparganj, Delhi 110 092, India

E-mail: hedfeedback@pearsoned.co.in