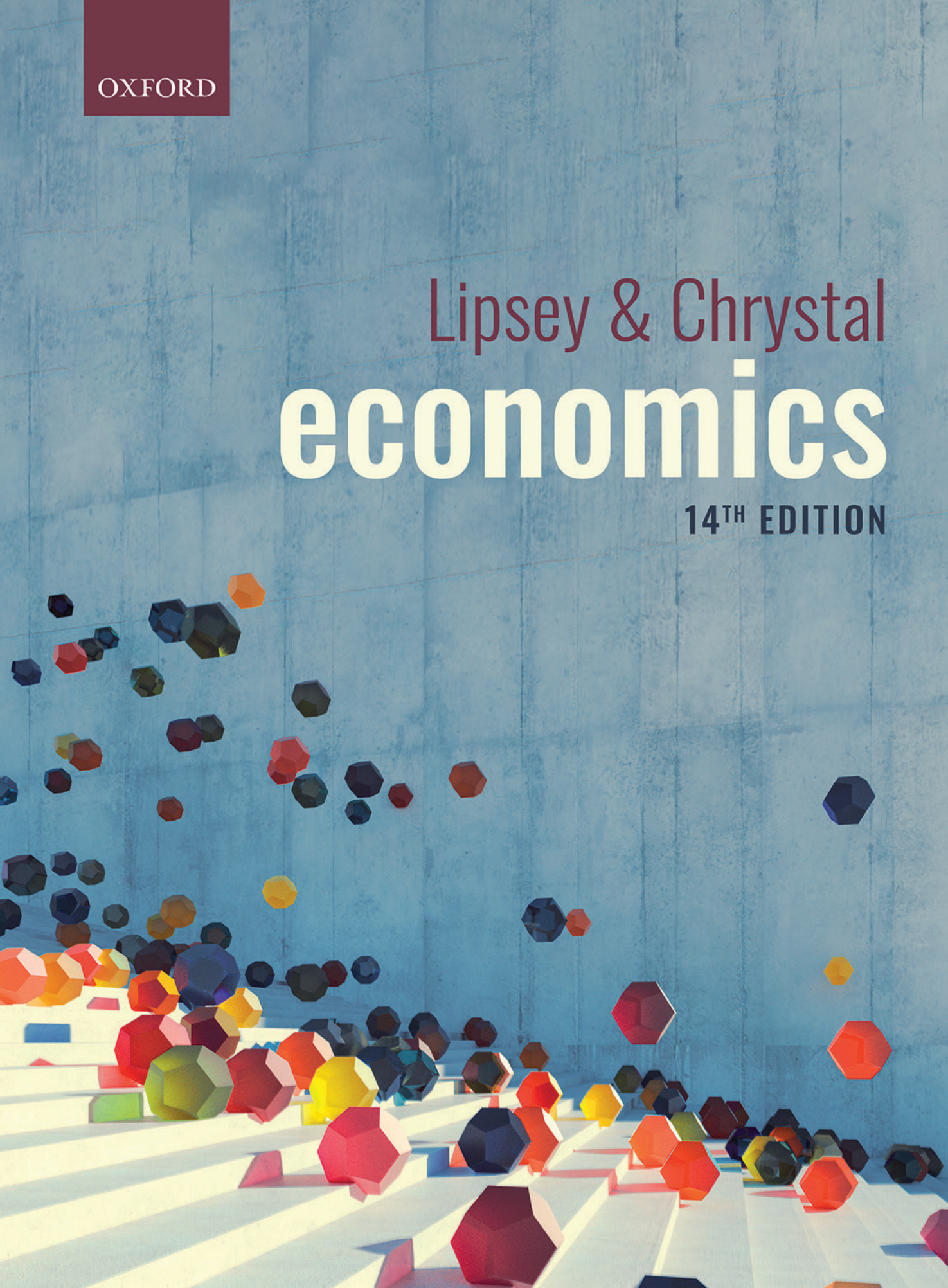
OXFORD

Lipsey & Chrystal

economics

14TH EDITION



ECONOMICS

LIPSEY & CHRYSTAL

ECONOMICS

14TH EDITION

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

Eleventh edition 2007
Twelfth edition 2011
Thirteenth edition 2015

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2019950748

ISBN 978–0–19–879103–4

Printed in Great Britain by
Bell & Bain Ltd., Glasgow

OUTLINE CONTENTS

DETAILED CONTENTS

MICROECONOMICS

PART ONE MARKETS AND CONSUMERS

WHY STUDY ECONOMICS?

Some of you may be excited by the prospect of studying economics, but others among you will answer the question posed in the heading with: 'It was part of my course of study, so I had no choice'. To both the reluctant conscripts and the willing volunteers we offer hope and encouragement.

Economics is highly relevant both to decision-making in most jobs that you are likely to do in life and to understanding many of the major issues facing today's world—free markets versus government intervention, resource exhaustion, pollution and environmental degradation, climate change, artificial intelligence (AI) and the future of work, Brexit, trade wars, government taxes and spending, employment, unemployment and recessions, inflation, income inequality in advanced nations, rapid growth in some of the world's emerging economies, and stagnation among many of the world's poorest nations. Thus, economics is both a preparation for taking decisions in a business or other organization, and training in how to understand many of the 'big issues' of our time.

One of the most important events in the first three-quarters of the twentieth century was the rise of communism. One of the most important events in the last quarter of that century was communism's fall. The century-long battle between free markets and government planning as alternatives for organizing economic activity was settled in almost all countries with a degree of decisiveness that is rare for great social issues. Understanding why market-oriented capitalist economies perform so much better than fully planned or highly government-controlled economies is a core issue in economics. Economic theories are expressly designed to help us understand the successes (and, where they occur, the failures) of free-market economies.

The triumph of market-oriented economies suggests that income- and wealth-creating activities are usually best accomplished through the efforts of private citizens operating in largely unregulated markets. But this is not the end of the story, for at least three fundamental reasons.

First, although market economies certainly work better than fully planned economies, they do not work perfectly. One of today's great social issues is how best to allocate the responsibilities of government, leaving it to do what it can do best and leaving free markets to do what they can do best. Economists ask: 'What are the important roles that governments can play in improving the functioning of a basically market-oriented economy?' During the three decades or so up to 2007 the consensus was moving in favour of more market freedom. However, the global financial crisis of 2007–8 raised questions about the extent to which markets can be left alone. Substantial reforms of banking and financial market regulation have been introduced since the crisis, but there is also an ongoing debate in official and academic circles about the regulatory policies needed to avoid, or at least mitigate, any future crises in the financial system—crises that can have ramifications far beyond the bounds of the financial system, spreading to have adverse effects on the entire economy and the people in it.

Secondly, market economies often produce severe short-term cycles as well as long-term growth. Long-term growth has raised the living standards of the ordinary working person from the horrors and degradations described by Charles Dickens in the nineteenth century to those of the property-owning workers of today, whose living standards are higher than those of 99.9 per cent of all the people of all classes who ever lived on Earth. Yet capitalist growth is uneven growth, because economic activity cycles around its rising trend. In recessions, unemployment is high and living standards typically stand still or even fall temporarily. Although each cycle tends to leave living standards higher

than all previous cycles, the ups and downs of uneven market-driven growth can be upsetting to those affected by it. During the 1990s and 2000s it appeared that the authorities had learned to avoid major cycles in activity, but the recession of 2008–9 made clear that sharp downturns can still occur, that these are costly in terms of lost output and employment, and that governments have only limited power to moderate their effects.

Thirdly, capitalist growth is unequal growth. Although most people agree that it is desirable to create higher incomes and wealth, they also care about how these are distributed among all citizens. Poverty for a minority amid plenty for the majority has always been a problem in wealthy countries—just as poverty for the majority and plenty for the minority has always been a problem for poor countries. A recent phenomenon in some of the richest countries has been rapidly rising real incomes for the few very best off while the average worker sees little or no real income rise. We ask: 'What are the sources of these distributional shifts and can governments do anything to affect the income and wealth distributions that market economies create?'

Economic growth in the world since World War Two was supported by falling barriers to trade and greater international economic integration. Lower transport costs and cheaper communications created opportunities for restructuring of supply chains in a process now called globalization. This process has helped to take billions of people out of poverty in emerging markets, such as India and China, but it has also had some negative effects on real income growth of unskilled workers in the old industrial countries. A backlash against these trends has led to nationalism in trade affairs and protectionism as a political solution in some places. The US trade war with China and the UK's vote to leave the EU are both aspects of this change of direction. How this will develop is far from clear at the time of writing, but economics can help in understanding what is going on and what the policy options are.

Economists have been in the forefront of analysing, explaining, and, where appropriate, offering solutions to all the issues mentioned above—and many more. They have often succeeded in these tasks, mainly because economics has a core of useful theories that explain how markets work and evaluate their performance. Although some economic analysis is extremely abstract, and sometimes even economists wonder about its value, the basic core of economic theory that is the secret of the subject's success can be understood by anyone who is willing to make the effort. This basic theory has an excellent record in illuminating issues in ways that lead to both deeper understanding and useful policy recommendations.

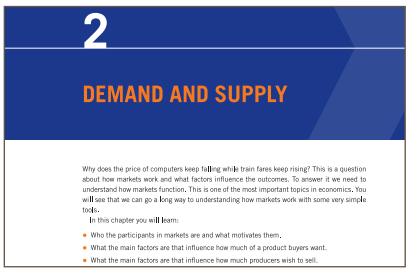
When you start to read this book, you are setting out on the study of a subject that, as the above discussion suggests, is highly relevant to understanding and improving the world in which we live. Approached in the right way, your study will be an adventure. The basic theory must be mastered. Whether or not you find this effort fun, you will find surprisingly early in your studies that theories can be used to understand many practical issues. The world is complex, and fully understanding its economic aspects requires much more economic theory than can be packed into one elementary textbook. But mastery of the subject to the level of this one book will contribute greatly to your understanding of many important issues and many of the policies directed at dealing with them. It will also provide you with a toolbox that will prove useful in the world of work, whatever occupation you eventually enter.

Good luck and good studying!

Richard Lipsey and Alec Chrystal
Vancouver, BC, and Cambridge, England

HOW TO USE THIS BOOK

This book is enriched with learning features designed to help you develop the knowledge and skills you need to study economics. This guided tour shows you how to use your textbook fully and get the most out of your study.


2

DEMAND AND SUPPLY

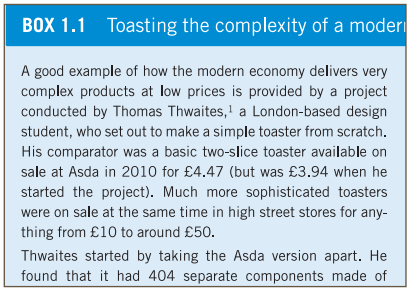
Why does the price of computers keep falling while train fares keep rising? This is a question about how markets work and what factors influence the outcomes. To answer it we need to understand how markets function. This is one of the most important topics in economics. You will see that we can go a long way to understanding how markets work with some very simple tools.

In this chapter you will learn:

- Who the participants in markets are and what motivates them.
- What the main factors are that influence how much of a product buyers want.
- What the main factors are that influence how much producers wish to sell.

Learning objectives

Each chapter opens with a bulleted list of learning objectives outlining the main concepts and ideas you will encounter in this part of the text. These serve as helpful signposts for learning and allow you to clearly track your progress and revision.

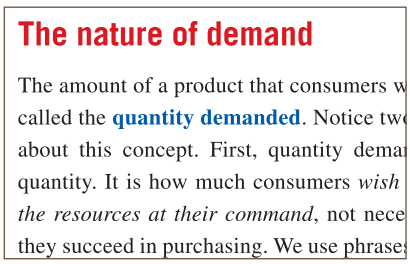

BOX 1.1 Toasting the complexity of a moder

A good example of how the modern economy delivers very complex products at low prices is provided by a project conducted by Thomas Thwaites,[1] a London-based design student, who set out to make a simple toaster from scratch. His comparator was a basic two-slice toaster available on sale at Asda in 2010 for £4.47 (but was £3.94 when he started the project). Much more sophisticated toasters were on sale at the same time in high street stores for anything from £10 to around £50.

Thwaites started by taking the Asda version apart. He found that it had 404 separate components made of

Boxes

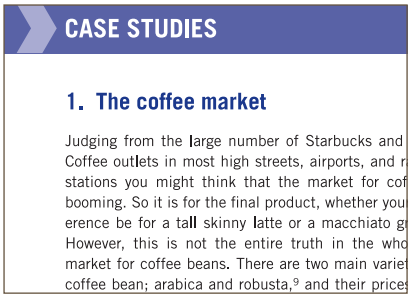
Boxes found throughout the book give you practical illustrations of the theory described in the main body of the text. They include real-life examples and empirical demonstrations as well as more detailed explanations of key points, all to ensure that you fully grasp core economic theory.


The nature of demand

The amount of a product that consumers w
called the **quantity demanded**. Notice tw
about this concept. First, quantity dema
quantity. It is how much consumers *wish*
the resources at their command, not nece
they succeed in purchasing. We use phrase

Key terms

Key terms are printed in bold the first time they appear in the text to alert you to each new concept. These terms are then compiled and defined in the glossary at the end of the book.


CASE STUDIES

1. The coffee market

Judging from the large number of Starbucks and Coffee outlets in most high streets, airports, and r stations you might think that the market for cof booming. So it is for the final product, whether you erence be for a tall skinny latte or a macchiato g However, this is not the entire truth in the whc market for coffee beans. There are two main varie coffee bean; arabica and robusta,[9] and their price

Case studies

All chapters, except Chapter 1, conclude with a minimum of two topical case studies. These cases will help you to contextualize your understanding of core chapter themes and encourage you to apply your learning to real-life situations.


SUMMARY

The complexity of the modern economy

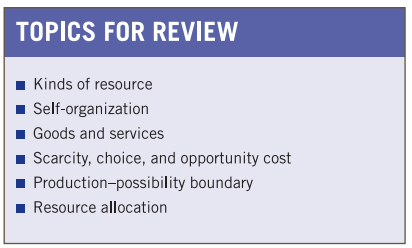
- A market economy is self-organizing in the sense that when individuals act independently to pursue their own self-interest, responding to prices set on open markets, they produce coordinated and relatively efficient economic outcomes.

Resources and scarcity

- Scarcity is a fundamental problem faced by all economies because not enough resources—land, labour, capital, and entrepreneurship—are available t

End-of-chapter summaries

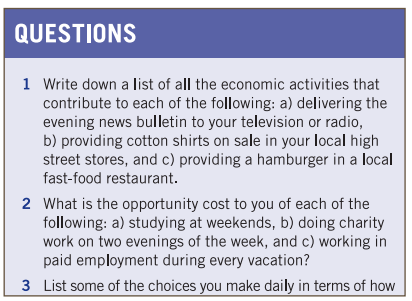
At the end of each chapter, the key points and concepts are summarized to help fix them in your mind. This feature reinforces your understanding of the material you have just covered and are an excellent revision tool.


TOPICS FOR REVIEW

- Kinds of resource
- Self-organization
- Goods and services
- Scarcity, choice, and opportunity cost
- Production–possibility boundary
- Resource allocation

Topics for review

Each chapter ends with a list of topics for review which prompt you to evaluate your understanding of these subjects. They will prove very helpful when you come to revision.


QUESTIONS

1 Write down a list of all the economic activities that contribute to each of the following: a) delivering the evening news bulletin to your television or radio, b) providing cotton shirts on sale in your local high street stores, and c) providing a hamburger in a local fast-food restaurant.

2 What is the opportunity cost to you of each of the following: a) studying at weekends, b) doing charity work on two evenings of the week, and c) working in paid employment during every vacation?

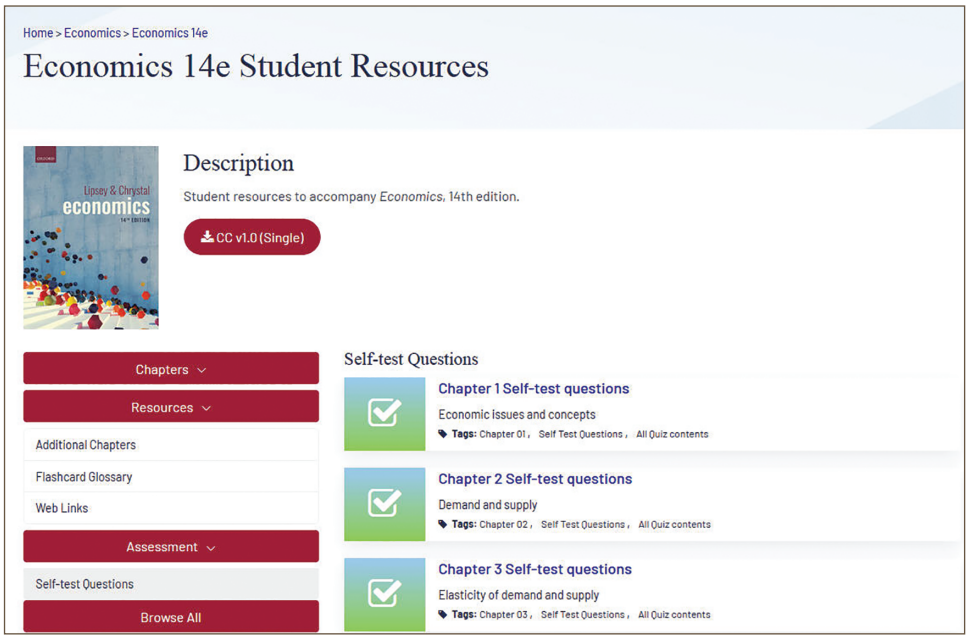
3 List some of the choices you make daily in terms of how

End-of-chapter questions

Problem-solving exercises and essay-type questions at the end of each chapter help you to develop your analytical skills. Use them to test what you have learned so far before moving on to the next chapter, or as a basis for group discussion or further revision.

HOW TO USE THE ONLINE RESOURCES

The Online Resources which accompany this book provides students with ready-to-use learning resources. They are free of charge and are designed to maximize the learning experience. www.oup.com/uk/lipsey14e/



FOR STUDENTS

Self-test questions

A suite of questions for each chapter in the book allows you to test your knowledge of the key themes in the text.

Interactive chapters

Additional course material is provided in an interactive format to assist and support your learning.

Web links

A selection of regularly updated web links, chosen by the authors with comments that explain their value, allows you to conduct further research on those topics that are of particular interest to you.

Flashcard glossary

The key glossary terms from the text are presented in an interactive format to help you remember and revise important concepts.

FOR LECTURERS

PowerPoint slides

A suite of PowerPoint slides is provided for use in lecture presentations. Arranged by chapter theme, the slides may also be used as handouts in class and can be easily adapted to suit your teaching style.

Test bank

A suite of test questions is provided which you can use for assessment purposes.

Instructor's manual

This comprehensive guide for instructors provides you with further discussion of the core themes in each chapter as well as the answers to all the questions set in the textbook.

Figures from the text

You can download all the figures found in the textbook electronically and you can use them to support student learning.

APPROACHES TO STUDYING ECONOMICS

You need to study a book on economics in a different way from how you would study a book on, say, history or English literature. Economic theory has a logical structure that builds on itself from stage to stage. Thus, if you understand some concept or theory only imperfectly, you will run into increasing difficulty when, in subsequent chapters, this concept or theory is taken for granted and built upon. Because of its logical structure, quite long chains of reasoning are encountered; if A then B; if B then C; if C then D; and if D then E. Each step in the argument may seem simple enough, but the cumulative effect of several steps, one on top of the other, may be bewildering on first encounter. Thus, when having followed the argument step by step, you encounter the statement: 'It is now obvious that if A then E', it may not seem at all obvious to you. This is a problem that everyone encounters with chains of reasoning. The only way to deal with it is to follow the argument through several times. Eventually, as the reasoning becomes familiar, it will become obvious that *if A then E*.

Economics has its own technical language or jargon. At first you may feel that you are merely being asked to put complicated names to commonsense ideas. To some extent this is true. It is a necessary step, however, because loose thinking about vaguely formed ideas is a quick route to error in economics. Furthermore, when you begin to put several ideas together to see what follows from them, jargon—a clearly defined term to refer to each idea—becomes a necessary part of your equipment.

A book on economics is to be worked at and understood step by step. It is usually a good procedure to read a chapter quickly in order to see the general run of the argument, and at this stage you might omit the captions to the figures. You then need to re-read the chapter carefully, making sure that the argument is understood step by step. On this reading, you *must* study the captions to all the figures carefully. If you do not understand the captions, you have not understood the economics. You should not be discouraged if, occasionally at this stage, you find yourself spending quite a bit of time on only a few pages.

A pencil and paper are valuable adjuncts to your reading. Difficult arguments should be followed by building up your own diagram while the argument unfolds, rather than relying on the printed diagram, which is complete from the beginning. Numerical examples can be invented to illustrate general propositions.

In short, the technical vocabulary aside, you must seek to *understand* economics, not to memorize it. Theories, principles, and concepts are always turning up in slightly unfamiliar guises. If you have understood your economics, this poses no problem; if you have merely memorized it, it spells disaster.

Email us

Economics is a subject about which one never stops learning. We are grateful to many users—students and teachers—who have taken the trouble to email us pointing out possible errors, making comments, and offering suggestions. We hope that readers will continue to teach us with as many further comments and criticisms as they have in the past. We try to acknowledge every such communication.

OUTLINES FOR SHORT OR MODULAR COURSES

This book provides a comprehensive coverage of basic economics suitable for a full one-year course. We have, however, designed the text to be flexible enough to cover shorter courses. To illustrate, we give our suggestions for the chapter content of several shorter courses.

A short Introduction to Economics course (twenty weeks)

Chapters 1–14 for basic micro, and 15–25 for basic macro.

An Introduction to Microeconomics course (one semester)

Chapters 1–14

An Introduction to Macroeconomics course (one semester)

Chapters 1 and 15–25. (Alternatively, for a course with a growth emphasis include Chapter 26; for a course with an international emphasis, the key chapters are 24 and 27, with addition of our web-based material on developing countries.)

ACKNOWLEDGEMENTS

Finally, we wish to say a word of thanks to some people who have helped make this book possible. In our office, Joanna Lipsey-Ouimet has played an important role is helping us prepare several successive editions of this book. At the OUP, we are indebted to Amber Stone-Galilee and Kathryn Rylance for guiding the project for most of its gestation. Kehinde Badmus and Felicity Boughton have helped enormously in finalizing this edition and getting it into production. This edition has also benefited from many students, especially Cass undergraduates, who have pointed out mistakes or unclear passages in the previous edition. We would also like to thank Lesley Harris for her careful copy editing and Heather Addison for her diligent proof reading. The usual disclaimer of course holds here: for all remaining shortcomings and mistakes, the authors may blame each other, but readers should blame us both.

R.G.L.
K.A.C.

Source Acknowledgements

Grateful acknowledgement is made to all the authors and publishers of copyright material which appears in this book, and in particular to the following for permission to reprint material from the sources indicated:

Bank of England for figures from Robinson, L.: *Rewriting history: understanding revisions to UK GD*P (https://bankunderground.co.uk/2016/06/10/rewriting-history-understanding-revisions-to-uk-gdp/) [last accessed July 2019]; *Asset Purchase Facility Quarterly Report*, Q3, (2018), p. 3 https://www.bankofengland.co.uk/-/media/boe/files/asset-purchase-facility/2018/2018-q3.pdf?la=en&hash=69ACA3A8F372E9F5700DA7FA67F17BC525616E80 [last accessed July 2019]; *Inflation Report* (February 2019) https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2019/february/inflation-report-february-2019 [last accessed October 2019]; *Official Bank Rate History (Data from 1694)* https://www.bankofengland.co.uk/statistics/details/further-details-about-wholesale-baserate-data [last accessed July 2019].

British Petroleum for the figure from *BP Statistical Review of World Energy* (June 2018) p. 20 (https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2018-full-report.pdf) [last accessed July 2019]

Institute for Fiscal Studies for the figure from Joyce, R. and Xu, X. (May 2019) *Inequalities in the twenty-first century: Introducing the IFS Deaton Review*, p. 6

International Monetary Fund for figures from Segoviano, M., Jones, B., Lindner, P., and Blankenheim, J. (January 2015) "Securitization: The Road Ahead." *IMF Staff Discussion Notes* , p. 5 (https://www.imf.org/external/pubs/ft/sdn/2015/sdn1501.pdf) [last accessed October 2019]; International Monetary Fund (October 2018) "Challenges to Steady Growth." *World Economic Outlook*, p. 17

ukpublicspending.co.uk for the figures from www.ukpublicspending.co.uk

Thompson Reuters for the figure from Thompson Reuters (2014) *EuroZone Report*, 2

Every effort has been made to trace and contact copyright holders but this has not been possible in every case. If notified, the publisher will undertake to rectify any errors or omissions at the earliest opportunity.

MICRO ECONOMICS

PART 1

MARKETS AND CONSUMERS

1

ECONOMIC ISSUES AND CONCEPTS

A modern economy produces millions of goods and services to choose from. It provides jobs for most people who want to work. It allows us to travel and communicate easily with anybody around the world. It produces growth in the total output available for citizens to consume. Yet nobody has sat down and planned how all this works. It has evolved as a response to economic forces interacting with individuals and institutions. Our main tasks in this book are to study how such an economy works, why it does not always work well, and what can be done to improve its performance where that is needed.

This book is divided into two main parts. Microeconomics studies how and why basic resources are transformed by producers into all the many goods and services that modern consumers want. Macroeconomics studies the amount of activity in the economy as a whole, and the role of governments in trying to influence that activity. Chapters 2–14 are about microeconomics, while Chapters 15–27 are about macroeconomics.

This introductory chapter provides a broad context for the study of microeconomics, but we also discuss the important links between microeconomics and macroeconomics that have become topical as a result of the financial crisis of 2007–8. It is intended as an introduction for those of you who are immediately going to study microeconomics further, and as providing essential basic micro information for those readers who are going on to study macro now. For you, Chapter 15 provides the main introduction to macroeconomics.

In the present chapter you will learn:

- That a modern market economy uses price signals to solve the complex problems involved in producing all the goods and services that people want.
- How economics explains the choice between competing demands for scarce resources.
- How production, employment, and consumption decisions interact.
- How the market economy usually delivers outcomes desired by consumers.
- How governments sometimes intervene when markets fail to produce the desired results.

THE COMPLEXITY OF THE MODERN ECONOMY

If you want a pint of milk, you go to the shop and buy it. The shop owner is just one part of a complex supply chain that makes this milk available when you want it. When the shop owner needs more milk, she orders it from the distributor, who in turn gets it from the bottling plant, which in its turn gets it from the dairy farmer. The dairy farmer buys cattle feed and electric milking machines and gets power to run all his equipment by putting a plug into a wall socket where the electricity is supplied as he needs it. The milking machines are made from parts manufactured in several different parts of the world, while these in their turn are made from materials mined and refined in a dozen or more different countries.

As it is with the milk you drink, so it is with everything else that you buy. When you go to the appropriate shop, what you want is normally in stock. Those who make these products find that all the required components and materials are available when needed—even though these things typically come from many different parts of the world and are made by people who have no direct dealings with each other. The economy is so good at delivering what we want when we want it that we only tend to notice when it goes wrong.

The sales and purchases in which you are involved are only a small part of the amazingly complex set of transactions that take place every day in a modern economy. Shipments arrive daily at our seaports and airports. These include raw materials, such as iron ore and logs, parts, such as computer chips and circuit boards, tools, such as screwdrivers and digging equipment, perishables, such as fresh flowers and fruits, and all kinds of manufactured goods, such as washing machines and TV sets. Rail and road shippers receive and dispatch these goods to thousands of different destinations. Some go directly to consumers. Others are used by domestic firms as inputs into the manufacture of their own products—of which some will be sold in the home market and some exported. Box 1.1 reports on a project that illustrates that the modern economy can deliver complex products at remarkably low prices.

In the past, most of the manufactured goods you bought were made in one country, either at home or somewhere abroad. But today most of the manufactured goods are made

BOX 1.1 Toasting the complexity of a modern economy

A good example of how the modern economy delivers very complex products at low prices is provided by a project conducted by Thomas Thwaites,[1] a London-based design student, who set out to make a simple toaster from scratch. His comparator was a basic two-slice toaster available on sale at Asda in 2010 for £4.47 (but was £3.94 when he started the project). Much more sophisticated toasters were on sale at the same time in high street stores for anything from £10 to around £50.

Thwaites started by taking the Asda version apart. He found that it had 404 separate components made of many different materials. He then set about collecting the raw materials to construct components for a toaster of his own making. For him this meant not just buying finished components like wire and screws, but rather getting the ore necessary for metal parts and converting this into refined metal before moulding it into the needed parts.

It took him nine months of his time and £1187.54 of his own money to construct from scratch a device that would toast a piece of bread. This sum excludes the value of his own time, which would have increased the costs substantially even if he had only paid himself the minimum wage. His product worked far less well and looked much less attractive than the basic Asda model!

The message we take from this is that the modern world economy does a remarkable job of delivering complex products that consumers want at low prices. Key elements of how it does this include a high degree of specialization in production, large volumes of specific outputs that reduce costs per unit, mechanization of routine tasks that save labour time, and globalization of supply chains which means that components can be bought from their cheapest source. All of these ideas and other parts of the story will be developed in the following chapters. Just bear in mind how expensive and difficult it would be to make your own iPad, mobile phone, laptop, or TV, yet the modern economy provides these at prices that many can afford. It is easy to take this for granted but it is actually a remarkable achievement.

[1] Thomas Thwaites, *The Toaster Project*, New York: Princeton Architectural Press, 2011. See also: http://www.metropolismag.com/story/20100317/the-cost-of-convenience and http://www.thetoasterproject.org

in bits and pieces all over the world linked together in what is called a *supply chain*. The parts are then assembled in one country and shipped to consumers throughout the world. So, when the TV set you buy says 'Made in China' what that label really means is 'Assembled in China' from parts made in half a dozen or more different countries.

Most people who want to work have a job. By working they earn incomes, which they then spend on the goods and services that they, and other workers, produce. Some people own businesses that employ workers to assist in making and selling their products. Business owners earn part of their income as profit from their enterprises, though they may also pay themselves some salary as managers if they work in their own firms.

Some activity takes place in the public sector in which the services produced are provided free to consumers, and where wages and salaries are paid for by the government out of general taxation. Public sector workers then spend their incomes on a wide range of goods, just as other types of workers do.

Self-organization

Economics as a subject began when thoughtful observers asked themselves how such a complex set of transactions is organized. Who coordinates the vast array of production, employment, and consumption decisions? Who makes sure that all the activities fit together both nationally and internationally, providing jobs to produce the goods and services that people want and delivering those things to where they are wanted?

The answer is: no one!

The great insight of the early economists was that an economy based on free-market transactions is self-organizing.

By following their own self-interest, doing what seems best and most profitable for themselves, and responding to the incentives of prices set on open markets, people produce a spontaneous social order. In that order, literally thousands of millions of transactions and activities fit together to produce the things that people want within the constraints set by the resources that are available to the nation.

The great Scottish economist and political philosopher Adam Smith, who was the first to develop this insight fully, put it this way:

> *It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages. Nobody but a beggar chuses to depend chiefly upon the benevolence of his fellow-citizens.*
>
> *(Source: Smith 1776,* The Wealth of Nations, *pages 26–7 of 1976 edn.)*

Smith is not saying that benevolence is unimportant—indeed, he praises it in many passages. He is saying, however, that the massive number of economic interactions that characterize any economy cannot all be motivated by benevolence. Although benevolence does motivate some of our actions, such as when we donate to earthquake, famine, and tsunami relief funds, most of our everyday actions are motivated by self-interest. The individual pursuit of self-interest is, therefore, the main behavioural incentive that drives a market economy to behave as it does.

Efficient organization

It is beyond question that a market economy is self-organizing, but it is much debated how efficient the results of that organization are. Another great insight, which was hinted at by Smith and fully developed over the next century and a half, was that this spontaneously generated social order is relatively efficient. Loosely speaking (we will be more precise later), efficiency means that the resources available to a nation are organized to produce the maximum possible total output of the types of goods and services that people wish to consume.

A free-market economy behaves almost as if some power had guided it. This does not literally mean that a supernatural presence runs economic affairs. Instead, it refers to the relatively efficient order that emerges spontaneously out of many independent decisions made by those who make, sell, and buy goods and services. The key to explaining this market behaviour is that these people all respond to the same set of prices, which are determined in markets that reflect the overall conditions of scarcity or plenty. Much of economics is devoted to understanding how this market outcome is generated, and to assessing the efficiency of that outcome.

A planned alternative

A century after Adam Smith, another great economist and political philosopher, Karl Marx, argued that although this market system would produce high total output, it would distribute that output in such a way that, over time, the rich would get richer and the poor poorer. Others took up this line of thinking and argued that when societies became rich enough, they should dispense with the spontaneous social order. They should then replace it by a consciously created system, called a command economy, or communism, in which the government plans all the economic transactions and, in so doing, creates a more equal and just distribution of the total output.

Beginning with the Union of Soviet Socialist Republics (USSR), the governments of many nations established systems in

which conscious government central planning largely replaced the spontaneous order of the free market. For much of the twentieth century two systems—the centrally planned and the market—competed for the favour of undecided governments. Then, within the last two decades of the century, governments of one communist country after another abandoned their central planning apparatus. More and more economic transactions and activities were then left to the market. Seldom has a great social issue been settled with such conclusiveness.

Marx was right about many things, including the importance of technological change in raising living standards over the centuries. Where the Marxists were wrong, however, was in believing that the market could be replaced by central planning as a way of organizing all a nation's economic activities. The task was far too complex and led to inefficiencies and waste, while not satisfying many of the needs of consumers.

In contrast with the failures of command economies, the performance of the free-market price system has generally been impressive. One theme of this book is the success of the market—how the price system works to coordinate with relative efficiency the decentralized decisions made by millions of private consumers and producers. However, this does not mean that doing things better implies doing things perfectly. Another theme of this book is market failure—how and why the unaided market economy can fail to produce socially desirable outcomes.

Box 1.2 discusses some important issues that arise when economists use their analysis to give economic advice to policymakers.

BOX 1.2 Economic advice: positive and normative statements

Economists give advice on a wide variety of topics. If you read a newspaper, watch television news, or listen to commentaries on the radio you will often notice some economists' opinions being reported. Perhaps it is on the prospects for unemployment, inflation, or interest rates, on some new tax, or on the case for privatization or regulation of an industry.

Advice comes in two broad types: normative and positive. A commentator might advise that the government ought to try harder to reduce unemployment or to preserve the environment. This is normative advice. He or she may be using their expert knowledge to come to conclusions about the costs of various unemployment-reducing or environment-saving schemes, but when it is said that the government *ought* to do something, this involves making judgments about the value of the various things that the government could do with its limited resources. Advice that depends on a value judgment is normative—it tells others what they ought to do.

Another type of advice is illustrated by the statement: 'If the government wants to reduce unemployment, then this is an effective way of doing so'. This is positive advice. It does not rely on a judgment about the value of reducing unemployment. Instead the adviser is saying, '*If* this is what you want to do, *then* here are ways of doing it'.

It is difficult to have a rational discussion of issues if positive and normative issues are confused. Much of the success of modern science depends on the ability of scientists to separate their views on *what does*, or *might*, *happen* in the world, from their views on *what they would like to happen*. For example, until the eighteenth century almost everyone believed that the Earth was only a few thousand years old. Evidence then began to accumulate that the Earth was thousands of millions of years old. This evidence was hard for most people to accept since it ran counter to a literal reading of many religious texts. Many did not want to believe the evidence. Nevertheless, scientists, many of whom were religious, continued their research because they refused to allow their feelings about what they wanted to believe affect their search for the truth. Eventually, scientists came to accept that the Earth is about 4.54 billion years old.

Distinguishing what is true from what we would like to be, or what we feel ought to be, true depends to a great extent on being able to distinguish between positive and normative statements.

Normative statements depend on value judgments. They involve issues of opinion, which cannot be settled by recourse to facts. In contrast, **positive statements** do not involve value judgments. They are statements about what is, was, or will be; that is, statements that are about matters of fact.

All six statements listed in Table 1.1 as positive assert things about the nature of the world in which we live. In contrast, the six statements listed as normative require value judgments.

Notice two things about the positive/normative distinction. First, positive statements need not be true. Statement D is almost certainly false. Yet it is positive, not normative. Secondly, the inclusion of a value judgment in a statement does not necessarily make the statement normative.

Table 1.1 Positive and normative statements

Positive		Normative	
A.	Higher interest rates cause people to save more	G.	People should save more
B.	High income tax rates discourage effort	H.	Governments should tax the rich to help the poor
C.	High taxes on cigarettes discourage smoking	I.	Smoking should be discouraged
D.	Road-use charges would increase traffic	J.	The tax system should be used to reduce traffic
E.	People are more worried about inflation than unemployment	K.	Technical change is a bad thing because it puts some people out of work
F.	The burning of fossil fuels is causing global warming	L.	Governments should do more to reduce carbon emissions in order to save the planet from global warming

Statement E is about the preferences that people hold; that is, about their value judgments. We could, however, check to see if people really do worry more about inflation than unemployment. We can observe their answers to survey questions, and we can observe how they vote for parties that give different priority to these objectives. There is no need to introduce a value judgment in order to check the validity of the statement itself.

You can decide for yourself why each of the other statements is either positive or normative. Remember to apply the two tests.

1) Is the statement only about actual or alleged facts? If so, it is positive.

2) Are value judgments necessary to assess the truth of the statement? If so, it is normative.

In short, economists seek to understand how well the market economy works and identify where governments may need to intervene to correct specific aspects of market failure.

Main characteristics of market economies

What then are the main characteristics of market economies that produce this spontaneous self-organization?

- Individuals pursue their own self-interest, buying and selling what seems best for themselves and their families.
- People respond to incentives. Other things being equal, sellers seek high prices while buyers seek low prices.
- Most prices are set in open markets in which suppliers compete to sell to potential buyers.
- People earn their incomes by selling their services to those who wish to use them—their labour services—or by selling things they have produced, or by selling the services of the property that they own.
- All these activities are governed by a legal framework largely created, administered, and enforced by the state.

RESOURCES AND SCARCITY

All the issues discussed so far would not matter much if we lived in an economy of plenty, where there was enough of everything for everybody. But instead we live in a world of scarcity. Most of us want better food, clothing, housing, schooling, holidays, hospital care, and entertainment. But there is not enough to go around. Even the richest economy can only produce a small fraction of the goods and services that people would like to have if everything was free. Hence economics is concerned with the problem of choice under conditions of scarcity. We cannot have everything we want, so we must choose what we will and will not have.

At this stage, it is helpful to clarify a few terms and concepts that you will use in studying economics.

Kinds of resource

An economy's resources can be divided into four main categories:

- All those gifts of nature, such as land, forests, minerals, etc., are commonly called natural resources and called by economists **land**, for short.
- All human resources, mental and physical, both inherited and acquired, which economists call **labour**.
- All those man-made aids to further production, such as tools, machinery, and factories, which are used in the process of making other goods and services rather than being consumed for their own sake. Economists call these **capital**.
- Those who take risks by introducing new products and new ways of making old products. They develop new businesses and forms of employment and are called **entrepreneurs** or **innovators**. The resource they provide is **entrepreneurship**.

Traditionally, these resources have been called **factors of production**, but we shall more frequently refer to them just as different types of input into the production process. Part 3 of this book focuses explicitly on **resource allocation**; that is, why are resources drawn into one activity rather than another—why do some people work in agriculture, others in manufacturing, and yet others in services, for example?

Ownership of resources

Private property is a key institution of a market economy. Individuals own most of the nation's resources. They also own the goods that they produce and the things that they buy. Some assets are owned by the state—roads, schools, public buildings, etc.—but most are, and must be, in private hands. People cannot make contracts to buy and sell what is not theirs. So, without private ownership the market economy cannot function.

Kinds of production

The resources of the economy are used in a production process to make **goods** that are tangible, in that they have a physical existence, such as cars, cans of beans, and shoes, and **services** that are intangible, such as haircuts, TV programmes, car maintenance, and telephone calls. Throughout this book we use the term 'goods' to cover both goods and services, unless we explicitly make a distinction between the two.

A nation's total output of all goods and services over one year is called its gross domestic product or **GDP** for short. The act of making goods and services is called **production**, and the act of using up these goods and services to satisfy wants is called **consumption**. Anyone who makes goods or provides services is called a **producer**, and anyone who consumes them to satisfy his or her wants is called a **consumer**.

Choice and opportunity cost

You might want a mobile phone so that you can call your friends or an iPod so that you can listen to your favourite music. Your parents might have a car to get them to work or to visit your grandparents at the weekend. In general, people value specific goods and services because they help them to satisfy their needs. Goods and services are thus regarded as a means to an end, the satisfaction of wants. Because no economy can produce enough goods and services to satisfy all its citizens' wants, choices must be made.

Most of us have only a specific amount of income that we can spend. If we want to have more of one thing, then we must have less of something else. For example, suppose that a friend of yours is considering whether to go out and have a few drinks with friends. The cost of these extra drinks could be measured as the money cost of so much per pint of beer or glass of juice. A more revealing way of looking at the cost, however, is in terms of what other consumption this person must forgo in order to obtain the drinks. Suppose that he or she decides to give up going to the cinema and use the money instead to buy the drinks. If the price of one drink is, say, one-third of the price of a cinema seat, then the cost of three drinks is one cinema visit; put the other way around, the cost of one cinema visit is three drinks.

Now think of the same problem at the level of a whole society. If the government decides to build more roads and finds the required money by building fewer schools, then the cost of the new roads can be expressed as so many schools per hundred miles of road.

Opportunity cost is a measure of costs expressed as alternatives given up, rather than in terms of money. If some course of action is adopted, there are typically many alternatives that could have been satisfied instead. For example, once the government has decided on its total spending for any given year, this provides it with an aggregate resource constraint and it must then decide how to allocate that spending between various competing parts of the public sector, such as health, education, and the police. So, it could be, for example, that if the government decides that it wants to hire 1,000 extra police it will have to reduce spending on education and 900 fewer teachers can be afforded. The opportunity cost of 1,000 police would be 900 teachers.

Of course, it would not necessarily have to be the education budget that was cut. It could have been the health budget or the defence budget or some other part of the public sector. Indeed, it might be a little bit cut off all the other spending components. The point is that, for any given aggregate resource constraint, if

you want more of one thing you must give up something else. The something else given up is the opportunity cost of what is obtained.

The concept of opportunity cost highlights the choices that must be made by measuring the cost of anything that is chosen in terms of the alternative that could have been chosen instead.

The production–possibility boundary

All the things that governments provide, such as schools, national defence, and roads, are produced in what is called the **public sector**. Everything else, including all the goods and services that consumers buy, is produced in the **private sector**. How should the nation's productive resources be divided between these two sectors? To illustrate this choice, we put all the goods and services that governments provide into one group called 'public sector goods'. The rest are provided by non-government organizations and we call these 'private sector goods'. The balance between public and private provision is determined through the political process by government tax and spending policies. Higher public provision requires higher taxes, and higher taxes reduce private consumption.[2]

The choices that each country must make are illustrated in Figure 1.1. The horizontal axis measures the quantity of public sector goods, while the vertical axis measures the quantity of private sector goods. The curve on the figure shows all those combinations of public and private goods that can be produced if all the nation's resources are fully employed. It is called a **production-possibility boundary**. Points outside the boundary show combinations that cannot be obtained because there are not enough resources to produce them. Points on the boundary are just obtainable; they are the combinations that can just be produced using all the available supplies of resources.

Choice, scarcity, and opportunity cost illustrated

A single country's production-possibility boundary illustrates three concepts that we have already discussed: scarcity, choice, and opportunity cost. Scarcity is shown by the unattainable combinations beyond the boundary. There are some things that we just cannot have. Choice arises because of the need to select one of the attainable points on or inside the boundary. No economy can be at more than one point at one time. Any combination of public sector and private sector goods that is within the blue shaded area is achievable and anything outside is not. Opportunity cost is shown by the negative slope of the boundary. As the economy moves along that boundary, more of one type of good is being obtained at the cost of fewer of the other types.

Increasing opportunity cost

The production-possibility boundary of Figure 1.1 is drawn with a slope that gets steeper as one moves along it from left to right. The increasing slope indicates increasing opportunity cost as more and more private goods must be given up for each additional unit of public goods. Start, for example, at the vertical axis where all production is of private sector goods. A small

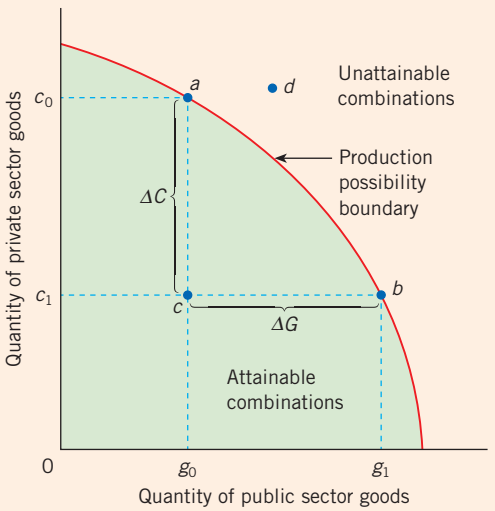


Figure 1.1 A production-possibility boundary

The negatively sloped boundary shows the combinations that are just attainable when all of the society's resources are efficiently employed. In this example the quantity of public sector goods produced is measured along the horizontal axis, and the quantity of private sector goods is along the vertical axis. Any point on the diagram indicates some amount of each kind of good produced. The production-possibility boundary separates the green-shaded attainable combinations, such as *a*, *b*, and *c*, from unattainable combinations, such as *d*. It is negatively sloped because, in a fully employed economy, more of one good can be produced only if resources are freed by producing less of other goods. Moving from point *a* (whose coordinates are c_0 and g_0) to point *b* (whose coordinates are c_1 and g_1) implies producing an additional amount of public sector goods, indicated by ΔG in the figure, at an opportunity cost of a reduction in private sector goods by the amount indicated by ΔC. Points *a* and *b* represent efficient uses of society's resources. Point *c* represents either an inefficient use of resources or a failure to use all the resources that are available.

[2] If the public provision is financed by borrowing rather than by taxes, this merely postpones the need to raise taxes until a later date.

increase in the production of public goods moves the economy along a flat part of the curve, indicating a small reduction in the production of private sector goods. But the loss of private goods gets greater (for each additional unit of public goods) as we move further along the boundary.

The figure can also be used to illustrate another three important economic issues.

Three key issues

What should be produced

How should scarce resources be allocated between the various possible kinds of production? Where to locate on the country's production-possibility boundary is the graphical representation of this key question. Each point on the boundary indicates a specific combination of the possible outputs. Alternative points indicate different allocations of the country's resources, producing different combinations of outputs.

Efficient production

If an economy is located inside its boundary more of everything could be produced. There are two main reasons why an economy may produce inside its production boundary. First, some of its resources may be unemployed. Putting them back to work would raise production of some goods without having to lower the production of anything else. Secondly, although its resources are fully employed, some of them may be inefficiently employed. If they could be used more efficiently, the production of some goods could be increased without having to produce less of anything else. We will have much more to say about inefficient uses of resources in later chapters.

These two possibilities help to reveal the source of opportunity cost.

If all the country's resources are fully employed, and none is employed inefficiently, then more of one good can be produced only by taking resources away from the production of another good.

The lost production of the other good is the opportunity cost of the first.

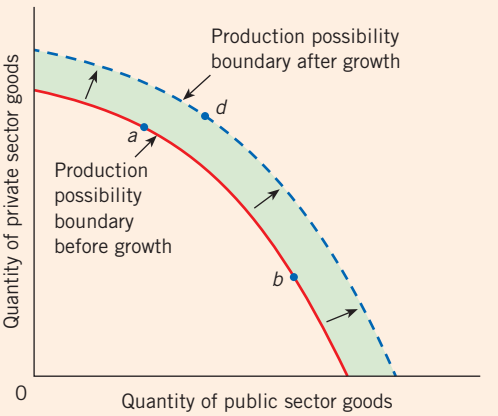


Figure 1.2 The effect of economic growth on the production–possibility boundary

Economic growth shifts the production-possibility boundary outward, allowing more of all commodities to be produced. Before growth in productive capacity points *a* and *b* were on the production-possibility boundary and point *d* was an unattainable combination. After growth point *d* becomes attainable, as do all points within the green band.

Economic growth

There is one other way an economy can get more of everything without having less of anything. If the economy's capacity to produce goods is increasing through time, the production-possibility boundary will be moving outwards over time, as illustrated in Figure 1.2. More of all goods can then be produced. This is what economic growth has accomplished from one decade to the next for the last several hundred years—and sporadically before then, back to the beginning of history. We will study this growth later in the book. In the meantime, we merely note that in the long term growth is driven by technological change. Over the years we learn to make existing products better and more cheaply, and to make many new products that satisfy old needs in new ways and others that create altogether new needs.

Box 1.3 deals with some of the sometimes-confusing terminology that surrounds the concept of the production-possibility boundary.

WHO MAKES THE CHOICES AND HOW?

In the previous section we discussed the production possibilities for a whole country. We now want to consider how the actual outcome is determined among all the possible outcomes. Economic choices must be made, but who makes them and how are they made?

Spending choices

The answer to the question 'Who makes choices?' is that we all do. That is, everybody must decide how to earn their living, how to spend their income, and how to invest their savings.

BOX 1.3 The terminology of production possibilities

We have used the term 'production-possibility boundary'. 'Boundary' emphasizes that the points on the line are maximum points. It is always possible to produce at points inside the line by not employing some factors of production, or by using them inefficiently. Two other terms, 'frontier' and 'curve' are often used instead of 'boundary'.

The words 'production possibility' emphasize the alternative possibilities available to a society. However, the term 'transformation' is often used instead. The idea behind the term 'transformation' is that society can, in effect, 'transform' one product into another by moving resources from the production of one product into the production of the other. Speaking of transforming one product into another involves the idea of opportunity cost. Of course, one good is not literally transformed into another but, by moving resources from producing one type of good to producing another, quantities of the first type of good are sacrificed to gain quantities of the second type.

You can make up six terms by combining the following words:

Production-possibility / or / Transformation *with* Curve / or / Boundary / or / Frontier

All six terms mean the same thing. All six are commonly used.

Firms produce the goods and services that we buy, the owners or managers of those firms must decide what inputs to buy, and all the firms they buy from also have to decide where to source their own inputs.

Hence, the outcome for the economy as a whole is the result of millions of individual decisions made by all the economically active population. To be 'economically active' in this sense you only have to spend even the smallest amount of money on some product, as you are then influencing what is produced. Thus, economic outcomes are driven by the buying and selling decisions of individuals[3] and firms. Later we shall also discuss the role of the government.

Maximizing decisions

The most important characteristic that economists assume about how individuals and firms take their decisions is that everyone tries to do as well as possible for themselves. In the jargon of economics, they are assumed to be *maximizers*. When individuals decide how much of their labour services to sell to producers and how many products to buy from them, they are assumed to make choices designed to maximize their well-being. When managers of firms decide how many inputs to buy from individuals and other firms, and how many of their own products to make and sell to them, they are assumed to be seeking to maximize their profits.

[3] We often refer to the basic decision-making unit for individual people as the 'household'. The term 'consumer' is also widely used, and official data sometimes refers to the 'personal sector'.

Marginal decisions

Individuals and firms make most of their choices *at the margin*. When you enter a shop to buy a newspaper and a pint of milk, you are deciding to buy one more newspaper and one more pint of milk today. You do not have to decide which product to spend your entire income on, and neither do you have to decide what to spend your income on tomorrow, next week or next year. You can, and do, decide to spread your spending around and to make these decisions incrementally (more or less) one at a time. Today you could buy two newspapers and a half-pint of milk. Tomorrow you might buy one newspaper, six eggs, and two pints of milk. These are marginal decisions—decisions to buy a bit more or a bit less. These decisions are made sequentially, not all at once.

Individuals and firms are constantly taking marginal decisions about whether to buy or produce a bit more or a bit less of the things that they consume or make. These decisions are assumed to be taken in order to achieve the most desirable outcome; that is, to maximize some objective.

Production choices

Managers of businesses decide what to produce and how to produce it. Some products, like haircuts, are quite simple to produce. But some other production processes are very complex. A typical car manufacturer assembles a vehicle out of thousands of individual components, many of them made in other countries. It makes some of these components itself. Most are subcontracted to other parts manufacturers, and many of

the major parts manufacturers subcontract some of their work out to smaller suppliers. This kind of production displays two characteristics noted two centuries ago by Adam Smith, and one that is more recent. These are specialization, the division of labour, and globalization.

Specialization

In ancient hunter-gatherer societies, and in modern subsistence economies, most people made most of the things they need for themselves. However, from the time (about 10,000 years ago) that people first engaged in settled agriculture and some began to live in towns, people have specialized in doing specific jobs. Farmers, carpenters, soldiers, and priests had some of the earliest specialized occupations. The allocation of different jobs to different people is known as **specialization of labour**, or the **division of labour**. There are two fundamental reasons why specialization is extraordinarily efficient compared with self-sufficiency.

First, individual abilities differ, and specialization allows each person to do what he or she can do relatively well while leaving everything else to be done by others. This is one of the most fundamental principles in economics. It is called the principle of **comparative advantage**, and a much fuller discussion is found in Chapter 27 where the principle is applied to international trade.

The second reason is that people's abilities change when they specialize. A person who concentrates on one activity becomes better at it than a jack-of-all-trades. This is called **learning by doing**. Learning by doing is very important in many jobs today where complex tasks are required.

The division of labour

Throughout history, most workers who specialized in making some product made the whole of that product. Over the past several hundred years, many technical advances in methods of production have made it efficient to organize agriculture and manufacturing into large-scale firms organized around a division of labour. There is job specialization within the production process. Few people work alone to make a whole product. Instead, most individuals specialize in making one bit of the final product, which is the outcome of the contributions of the labours of many specialist workers and machines. This is obviously true of most manufactured products such as cars, TV sets, and washing machines, which have many components, and many individuals play different roles in the assembly process. It is also true of most service industries where individuals work in teams to deliver a complex product.

Globalization

Market economies constantly change, largely as a result of the development of new products and new technologies. One important recent development is referred to as **globalization**. Globalized trade is not new—it has been around for hundreds, and in some areas thousands, of years. Since the Industrial Revolution, the usual pattern was manufactured goods being sent from Europe, and later from North America, to the rest of the world, with raw materials and primary products[4] being sent back in return. What is new in the last few decades is the globalization of manufacturing, and in particular the rapid growth of manufacturing in China, India, and some other countries of Southeast Asia.

Two major causes of globalization are the rapid reduction in transportation costs and the revolution in information technology. Firstly, the cost of moving products around the world fell greatly in the second half of the twentieth century owing to air travel, containerization, and the greatly increased size of ships. Secondly, our ability to transmit and to analyse data has been increasing dramatically over the past fifty years, while the costs of doing so have been decreasing equally dramatically. For example, today £1,000 or so buys a powerful computer that fits into a briefcase and has the same computing power as one that in 1970 cost £5 million and filled a large room.

This revolution in information and communication technology (ICT) has made it possible to coordinate economic transactions around the world in ways that were difficult and costly fifty years ago and quite impossible a hundred years ago. This, combined with falling costs of transport, has allowed manufacturing activities to be decentralized. Fifty years ago, if a car was to be assembled in Longbridge, all the parts had to be made nearby, and hence there were many parts producers in the Birmingham area. Today it is possible to make parts anywhere in the world and get them to the place of assembly exactly when they are required (but not Longbridge as assembly there ceased in 2005). As a result, manufacturing, which was formerly concentrated in the advanced industrial countries of Europe and North America, now takes place all over the world. A typical dishwasher, TV set, or car contains components made in literally dozens of different countries. Cars, for example, are made up of tyres, wheel hubs, windows, seats, computers, clocks, media players, speedometers, petrol gauges, locks, doors, handles, pistons, gears, distributors, wiper blades, brakes, brake liners, switches, steering wheels,

[4] In an older usage, products were divided into primary, secondary, and tertiary products. This division is no longer used, but the term 'primary products' remains in use to cover raw materials and agricultural goods.

etc., and we have not even started listing the components of the engine. Many of these are made by different suppliers located in different parts of the world, the general rule being that production of each item takes place where it is cheapest because the costs of transport to the point of assembly are so low as to be almost irrelevant. These bits are part of a supply chain and as a result, although we still know where a product is assembled, it is becoming increasingly difficult to say where it is made.

Also, many *markets* are globalizing. For example, as some fashions have become universal, we can see the same designer jeans, brand labels, and fast food outlets in virtually all big cities. Many *corporations* are globalized, as more and more of them become what the UN calls **transnationals**, but which are also known as **multinationals**. These are massive firms with a physical presence in many countries—McDonald's restaurants are as visible in Paris or Beijing as in London or New York. Many other brands are also virtually universal, such as Coca Cola, Kellogg, Heinz, Nestlé, Guinness, Toyota, Mercedes, BMW, Dell, Starbucks, Apple, Sony, Gucci, and Gap.

The pros and cons of globalization have recently been a subject of great controversy (which we elaborate in several places in later chapters). Notice, however, that globalization is not the only game in town. Globalization has affected financial markets and manufacturing, but an ever-increasing proportion of employment is in service industries, many of which, by the nature of the product, are very local. How far, for example, would you be prepared to travel for an evening meal, a haircut, a concert, or a trip to the theatre? Most people shop in their local town and visit their local doctor and local hospital, though in cases of highly specialized needs some purchase products on the internet from anywhere in the world and others will travel great distances to consult a Harley Street specialist or a US medical expert at the Mayo Clinic. Having said that, it is also important to note that globalization is spreading rapidly in the service industries. When you seek advice on some problem with one of your electronic devices, such as your computer or iPhone, you are more likely to receive it in an Indian or a Philippine accent than an English one. Many firms centralize their book-keeping and billing in one country, often Switzerland. Many US medical clinics send their data records each evening to Ireland where the accounts are kept and returned each morning. Digitized data on such things as fingerprints allow police located anywhere in the world to access information collected anywhere else in the world. Computerized operations allow physicians located in key medical centres to operate on patients located in distant places. Many university degrees and professional qualifications can now be gained through distance learning without ever seeing the bricks and mortar of the granting institutions. And so on for a growing number of service operations.

Markets and money

People who specialize in doing only one thing, whether they are factory workers or computer programmers, must satisfy most of their needs by consuming goods and services produced by other people. In early societies, exchange of products took place by simple mutual agreement between neighbours. In the course of time, however, trading became centred on particular places called *markets*. For example, the French markets or trade fairs of Champagne were well known throughout Europe as early as the eleventh century. Even now, many towns have regular market days. For economists, however, the term 'market' has a much broader meaning. We use the term **market economy** to refer to a society in which people specialize in productive activities and meet most of their material wants through voluntarily agreed exchanges. Most employed people, for example, work for a single employer (at any one time) and buy goods and services in a wide range of outlets (such as shops and restaurants, or by phone and internet, etc.).

Specialization must be accompanied by trade. People who produce only one thing must trade most of it to obtain all the other things they require.

Early trading was by means of **barter**; the trading of goods directly for other goods. But barter is costly in terms of time spent searching out satisfactory exchanges. If a farmer has wheat but wants a hammer, he must find someone who has a hammer and wants wheat. Thus, a successful barter transaction requires what is called a *double coincidence of wants*.

Money eliminates the restrictive system of barter by separating the transactions involved in the exchange of products. If a farmer has wheat and wants a hammer, they do not have to find someone who has a hammer and wants wheat. They merely must find someone who wants wheat. The farmer takes money in exchange. Then they find a person who wishes to trade a hammer and gives up the money for the hammer.

The existence of 'money' greatly expands the possibilities of specialization and trade.

Consumer sovereignty

Many individuals may feel that most power over economic outcomes is in the hands of firms, especially big firms. However, if consumers will not buy a product, it does not pay to produce it. No business can survive for long if it makes things

that nobody wants to buy. If a firm sees an opportunity to satisfy some unsatisfied need, it will develop a product to fill this gap. Even if a need is already satisfied by some product, firms have an incentive to develop products that better satisfy the same need or satisfy it more cheaply. In this, and in many other ways, consumers drive much economic activity, even though the production itself is done by firms. Because these firms are motivated by profits, they respond to (and try to anticipate) consumers' preferences as these are revealed by their purchases in the marketplace.

Box 1.4 provides a discussion of some important methodological issues that arise in economics.

BOX 1.4 Economic theories

Economists seek to understand how the economy works and to make predictions about the consequences of certain changes that are either natural, such as a crop failure, or policy-induced, such as a tax on petrol. To do such things they construct economic theories, many of which you will study throughout this book. First, the things with which the theory is concerned, called its variables, must be carefully defined. Then, assumptions are made about how these variables behave and about other things that may affect them. The implications of these assumptions are then derived. This may be done verbally, geometrically, or mathematically, with the choice between these methods depending on the degree of complexity of the assumptions. These implications are the theory's predictions: if so and so happens, then so and so will follow. For example, if a tax is placed on white bread, its price will rise, and the quantity produced and consumed will fall (as consumers switch to alternative products).

A theory is tested by confronting its predictions with the facts, often called empirical observations. If the facts agree with the theory's predictions, the theory has passed the test and is said to be consistent with the facts. If the theory continues to pass the test, it will be found useful, but it can never be regarded as conclusively proved to be true as there is always the possibility that conflicting facts will be observed in the future. If the facts disagree with the theory's predictions, the theory has failed to pass the test. If the theory fails the test in enough cases, it will be regarded as refuted and will be abandoned.

It is possible to develop theories that have no testable implications. Such theories are empty of empirical content since they are consistent with all possible states of the world. To have empirical content—that is, to be useful in understanding the world in one way or another—a theory must make predictions that are at least capable of being contradicted by evidence.

There is debate among economists as to whether a theory's assumptions should be subject to empirical testing as well as its predictions. In his essay 'The methodology of positive economics', the Nobel Prize winning economist Milton Friedman argued that only the predictions of a theory should be subject to testing, and if a theory passes the test, it does not matter how unrealistic, or otherwise counterfactual, its assumptions are. In his book *The Methodology of Economics: or How Economists Explain*, Mark Blaug, one of the twentieth century's greatest methodologists, criticized the licence that Friedman's 'methodology of positive economics' gave economists to make unrealistic assumptions, provided only that their theories yielded verifiable implications.

Blaug made many arguments against Friedman's position, a few of which follow. Firstly, Friedman failed to note the important requirement that assumptions that are patently counterfactual should be robust in the sense that they can be relaxed without seriously altering the theory's predictions. Secondly, since one of the key purposes of behavioural assumptions is to link a theory`s predictions to observable behaviour, the view of the extreme irrelevance of assumptions implies that it does not make sense to ask which of the assumptions is causing the trouble if a theory that seemed to accord with the facts no longer does so. Thirdly, if a theory that makes empirically correct predictions includes an empirically false assumption, it is interesting to ask why this is so. Perhaps the empirically incorrect assumption is unnecessary and can be stripped away. Maybe there is a second incorrect assumption that cancels out the effect of the first, in which case it surely is worth knowing this. Finally, Friedman's discussion did not recognize the several different senses in which assumptions are used in economics and the different implicit ways in which they should be assessed. Some assumptions merely give the circumstances under which the theory is meant to apply, in which case it is not relevant to criticize the theory because such assumptions are 'unrealistic'. Other assumptions give the behaviour that is driving the results of the theory, in which case we gain by knowing if such assumptions are right or wrong. For example, if some of the predictions of a theory that assumes that all sellers are motivated by altruism rather than self-interest passes some tests, we would still not put much faith in it since we have other evidence that this behavioural assumption is wrong.

GOVERNMENT AND THE MARKET ECONOMY

We have just explained how the interaction of firms and consumers through markets determines what is produced. We now discuss the role of governments in the economy. In order to put this in context we first consider what alternatives there may be to a free-market economy. There are four main types of economic system.

Traditional systems

A **traditional economic system** is one in which behaviour is based primarily on tradition, custom, and habit. Young men follow their fathers' occupations—typically, farming, hunting, fishing, and tool-making. Women do what their mothers did—typically cooking, mending, and fieldwork. There are few changes in the pattern of production from year to year, other than those imposed by the vagaries of nature. The techniques of production also follow traditional patterns, except when the effects of an occasional invention are felt. The concept of private property is often not well defined, and property is frequently held in common, such as common grazing land. Finally, production is allocated between the members of society according to long-established traditions. In short, the answers to the economic questions of what to produce, how to produce, and how to distribute are determined by what has happened in the past. Such a system works best in an unchanging environment. Under static conditions, a system that does not continually require people to make choices can prove effective in meeting economic and social needs.

Traditional systems were common in earlier times. The feudal system under which most people lived in medieval Europe was a largely traditional society. Today only a few small, isolated, self-sufficient communities still retain mainly traditional systems; examples can be found in a few of the most isolated parts of the Canadian Arctic, the Himalayas, the Amazon jungle, and isolated parts of Papua New Guinea.

Command systems

We have already seen that in command systems some central authority determines economic behaviour. It makes most of the necessary decisions on what to produce, how to produce it, and who gets it. Because centralized decision-makers usually lay down elaborate and complex plans for the behaviour that they wish to impose, the terms command economy and **centrally planned economy** are usually used synonymously.

The sheer quantity of data required for the central planning of an entire economy is enormous, and the task of analysing it to produce a fully integrated plan can hardly be exaggerated, even in the age of computers. Moreover, the plan must be a rolling process, continually changing to take account not only of current data but also of future trends in labour supplies, technological developments, and people's tastes for various goods and services. This involves the planners in the notoriously difficult business of forecasting the future.

Five decades or so ago, over one-third of the world's population lived in countries that relied heavily on central planning to deal with the basic economic questions. Today the number of such countries is small. Even in countries where central planning is the official system, as in China, rapidly increasing amounts of market determination are being accepted and encouraged.

Pure market systems

Earlier in this chapter we discussed the basics of a free-market economy. Millions of consumers decide what products to buy and in what quantities. Many firms produce those products and buy the inputs that are needed to make them. Individual decisions collectively determine the economy's allocation of resources between competing uses and the distribution of its output among individual citizens.

In a **pure market economy** all these decisions, without exception, are made by buyers and sellers acting through unhindered markets. The state provides the legal structure and external defence but, beyond that, markets determine all resource allocation and income distribution.

Mixed systems

Fully traditional, fully centrally controlled, and fully free-market economies are useful concepts for studying the basic principles of resource allocation. When we look in detail at any real economy, however, we discover that its economic behaviour is the result of some mixture of central control and market determination, with a certain amount of traditional behaviour as well. The term **mixed economy** refers to an economy in which both free markets and governments have significant effects on the allocation of resources and the distribution of income.

In practice, every economy is a mixed economy in the sense that it combines significant elements of all three systems—traditional, command, and market—in determining economic behaviour.

The proportions of the mixture of free-market determination and government control vary from economy to economy and over time. There is more free-market determination in the UK and the USA than in France and South Korea. There is more free-market determination in the UK today than there was sixty years ago. The mix also varies from sector to sector within any one economy. For example, European agricultural markets have a substantial amount of government control. Under market determination, the average size of a farm would be much larger and agricultural prices much lower than they now are. In contrast, the markets for information and computer technologies are largely free from government intervention. Even the economies closest to free markets have a significant role for government, so it appears that there is no real alternative to a mixed system with major reliance on markets but also with a substantial government presence in many aspects of the economy.

Government in the modern mixed economy

Modern market economies in advanced industrial countries are based primarily on market transactions between people who voluntarily decide whether to transact. Private individuals have the right to buy and sell what they wish, to accept or refuse work that is offered to them, and to move to where they want when they want. But governments create the legal framework that governs transactions.

Key institutions are private property and freedom of contract, both of which must be maintained by active government policies. The government creates laws of ownership and contract, and then provides the courts to enforce these laws. Governments are also responsible for provision of a stable-valued money that is the measuring rod for all prices.

In modern mixed economies governments go well beyond these important basic functions. They intervene in market transactions to correct what are called 'market failures'. These are identifiable situations in which free markets do not work well. For example, natural resources such as fishing grounds and common pastureland tend to be overexploited to the point of destruction under free-market conditions. Some products, called **public goods**, are not provided at all by markets because, once produced, no one can be prevented from using them. Therefore, their use cannot be restricted to those who are willing to pay for them. Defence and law and order are public goods. In other cases, private agents impose costs called **externalities** on others by their economic activities, such as when factories pollute the air and rivers. The public is harmed but has no part in the producers' decisions about what to make and how to make it. These are some of the reasons why free markets sometimes fail to function in desirable ways. They explain why citizens wish governments to intervene and alter the outcome that would result from leaving everything to the market.

There are also some products, like health and education, that could be provided through the market, but which governments have decided should be provided by the state and (in some cases) free of charge. These are not pure public goods, but many countries' governments have decided that at least some minimum provision must be available to all, at least at some basic level, so this cannot be left to the market. These are sometimes referred to as **merit goods**.

The **distribution of income** indicates how the nation's total income is distributed among its citizens. This is largely determined by the price that each type of resource input can command and by how evenly the endowments of these resources are distributed. It could be thought that labour is equally endowed to individuals, because each person has only one body. However, talents are not equally endowed, and people acquire varying skill levels. Ownership of land and other property varies considerably between people.

There are important equity (or fairness) issues that arise from letting free markets determine people's incomes. Some people lose their jobs when firms reorganize to utilize new technologies. Others may keep their jobs, but the market values their services so poorly that they face economic hardship. The old and the chronically ill may suffer if their past circumstances did not allow them to save enough to support themselves. For many reasons of this sort we accept government intervention to redistribute income by taking something from the 'haves' and giving it to the 'have-nots'. Almost everyone accepts that there should be some redistribution of incomes. Care must be taken, however, not to kill the goose that lays the golden egg. Taking too much from the haves risks eliminating their incentive to work hard and produce income, some of which is to be redistributed to the have-nots.

Macro and micro roles of government

So far, we have been discussing the intervention by government in specific markets on a permanent basis. There is another important role for government in the context of managing the economy in order to ensure that aggregate activity is at as high a level as is sustainable, and inflation is avoided. The worldwide recession of 2008–9, which followed from the financial crisis of 2007–8, highlights the importance of the role of the central authorities in **stabilization policy** (see Chapter 15 onwards).

The standard tools of stabilization are monetary and fiscal policies. The former works by influencing the money supply, interest rates, or exchange rates, while the latter works through taxation and government spending.

There has never been a strict separation between macro and micro policies, but in the two decades prior to 2007 there had been a broad consensus that the micro interventions of governments in the market should be kept to a minimum and that macro policy should concentrate on controlling inflation and keeping real activity as close to its sustainable trend level as possible. However, the widespread collapse of financial institutions, many of which had to be bailed out by governments (and some taken into state ownership), has led to an active debate about whether the role of government needs to be extended on a permanent basis. Box 1.5 discusses the underlying issues further.

It is unlikely that governments will wish to keep major financial institutions in state ownership, but it is highly likely that there will be much greater state intervention in the form of tighter regulation of financial market activity.

Direct regulation of private sector firms interferes in the market behaviour of these firms. The reason this is necessary is that large financial institutions, especially banks, have a unique role as guardians of people's savings and as central players in the payments system. If a major bank were to close for business at short notice, it would stop a lot of other economic activity at the same time. This means that there is systemic risk and some of these banks have been considered 'too big to fail'.

Thus, while much activity can be left to the market, the government must intervene to prevent major adverse events, such as a bank collapse, from bringing down many other parts of the economy with it, and so causing a major recession. We discuss the financial crisis in more detail later. However, the point to take from this discussion is that views about the appropriate role of government in the market economy are influenced by both microeconomic and macroeconomic considerations.

BOX 1.5 Government and the market

The 2007–8 world financial crisis and the subsequent recession has reopened old debates about the role of government at both the level of managing the economy as a whole (macroeconomics), and the level of intervention in specific markets (microeconomics), especially financial markets.

At the macroeconomic level, the debate has been characterized as between Keynesians[5] and monetarists, or between Keynesians and classical economists. The Keynesians advocate active use of government spending and taxes (and/or interest rates) to control the level of activity in the economy, while the others see the economy as self-stabilizing and thus oppose active management by governments.

The micro level debate is also between those who see markets working well on their own and those who emphasize the imperfections of markets and recommend various types of correction for 'market failure'. We discuss this further in Chapter 13.

The links between the micro and the macro approaches to these questions are well brought out by the following extract from a report by a UN Commission of Financial Experts chaired by Nobel Prize winner Joseph Stiglitz that was set up to investigate the implications of the 2007–8 financial crisis.

> *Part of the explanation for the current crisis may be found in the underlying economic fundamentals. Another is in the economic theories that motivated the financial and economic policies that produced the crisis. … These same economic doctrines—the belief that economic agents are rational, that governments are inherently less informed and less motivated by sound economic principles and therefore their interventions are likely to distort market allocations, and that markets are efficient and stable, with a strong ability to absorb shocks—also affects macroeconomic policies.*

[5] They are followers of British economist John Maynard Keynes (1883–1946).

›

Box 1.5 *continued*

One of the most important lessons of the Great Depression was that markets are not self-correcting and that government intervention is required at the macroeconomic level to ensure recovery and a return to full employment. In the aftermath of the Great Depression, governments introduced policies that provided automatic stabilizers for aggregate demand and implemented discretionary policy frameworks to reduce economic instability. But, as the Great Depression and earlier panics and crises faded from memory, confidence in the self-stabilizing nature of the market returned.

The fact that the world recovered so quickly from financial crises such as the East Asia crisis of 1997–8 and the global liquidity crisis of August 1998 induced false confidence in the self-correcting nature of market processes. While the recovery was due to public policies, it was credited to market processes. More generally, the historical role for government intervention in recovery and stability was forgotten.

(Source: The Stiglitz Report: Reforming the International Monetary and Financial Systems in the Wake of the Financial Crisis, *New York: New Press, 2010.)*

Another aspect of the relationship between governments and the market occurs in international trade. In the past governments could encourage the local production of an individual commodity by putting a tax, called a tariff, on the importation of that commodity. Today, however, with most manufactured goods being produced in complex international supply chains, it is much harder to encourage domestic production with tariffs because of the interactions among countries. Furthermore, if it is desired to punish some country, say China, for some actual or assumed bad behaviour, restricting one's imports from that country hurts not just that country which is the assembler of goods but all the other countries that produce parts of those goods' supply chain.

The role of governments in the economy is a recurring issue throughout this book. The micro aspects are mainly discussed in Chapters 2–14 and the macro dimensions are covered from Chapter 15 onwards. However, the micro and macro questions can both be stated in the same words: what can be left to the market and what is the appropriate role of government intervention?

CONCLUSION

Economics is about how all the goods and services that we want to buy are produced and how we earn the income to pay for them. It also studies how the production and employment system can go wrong and what can be done to fix it. In particular, economics highlights specific areas in which the economy may fail to produce optimum outcomes and suggests policies that can be followed by governments in order to make improvements to the situation.

SUMMARY

The complexity of the modern economy

- A market economy is self-organizing in the sense that when individuals act independently to pursue their own self-interest, responding to prices set on open markets, they produce coordinated and relatively efficient economic outcomes.

Resources and scarcity

- Scarcity is a fundamental problem faced by all economies because not enough resources—land, labour, capital, and entrepreneurship—are available to produce all the goods and services that people would like to consume. Scarcity makes it necessary to choose among alternative possibilities: what products should be produced and in what quantities.
- The concept of opportunity cost highlights scarcity and choice by measuring the cost of obtaining a unit of one product in terms of the number of units of other products that could have been obtained instead.
- A production-possibility boundary shows all the combinations of goods that can be produced by

an economy whose resources are fully employed. Movement from one point to another on the boundary shows a shift in the amounts of goods being produced, which requires a reallocation of resources.

Who makes the choices and how?

- Economic choices are made by individuals and firms.
- Modern economies are based on the specialization and division of labour, which necessitate the exchange of goods and services.
- Exchange takes place in markets and is facilitated by using money.
- Markets work to coordinate millions of individual, decentralized decisions.
- Supply chains in production have become complex and international.

Government and the market economy

- Three pure types of economy can be distinguished: traditional, command, and free market.
- In practice, all economies are mixed economies in that their economic behaviour responds to mixes of tradition, government command, and price incentives.
- Governments play an important part in modern mixed economies. They create and enforce important background institutions such as private property. They intervene in an attempt to increase economic efficiency by correcting situations where markets do not effectively perform their coordinating functions. They also redistribute income and wealth in the interests of equity.
- The role of the government is also influenced by the need to ensure that the economy as a whole avoids inflation and recessions.

TOPICS FOR REVIEW

- Kinds of resource
- Self-organization
- Goods and services
- Scarcity, choice, and opportunity cost
- Production–possibility boundary
- Resource allocation
- Growth in productive capacity
- Specialization and the division of labour
- Command, traditional, market, and mixed economic systems
- Positive and normative statements

QUESTIONS

1. Write down a list of all the economic activities that contribute to each of the following: a) delivering the evening news bulletin to your television or radio, b) providing cotton shirts on sale in your local high street stores, and c) providing a hamburger in a local fast-food restaurant.
2. What is the opportunity cost to you of each of the following: a) studying at weekends, b) doing charity work on two evenings of the week, and c) working in paid employment during every vacation?
3. List some of the choices you make daily in terms of how you spend your time and how you spend your money.
4. Make a list of all the different goods and services you typically buy in a normal week. In how many different markets does this suggest that a typical individual trades on a regular basis?
5. Explain the concept of opportunity cost and discuss how it relates to the problem of choice between scarce alternatives.
6. Outline the differences between traditional, command, and market economies and explain why the former two have been superseded.
7. In what ways does money facilitate specialization and the division of labour?
8. Why do governments have a role in a market economy? (Revisit this question once you have studied Chapters 13 and 14.)
9. Economics used to be known as the ‘dismal science’ because it pointed out that choices had to be made between scarce alternatives. Assess the prospects of scarcity being eliminated in the foreseeable future.

10 Which of the following statements are positive and which are normative?
a) The health of poor people is worse than the health of rich people.
b) Economic growth in sub-Saharan Africa is affected by the AIDS epidemic.
c) Rich countries should provide medicine more cheaply to Africa.
d) Protectionist policies in rich countries are hurting poor countries and should be abolished.

11 Which one of the following is not a positive statement?
a) If the price of oil rises people will make fewer journeys by car.
b) There would be less pollution if more people used public transport rather than driving to work.
c) If petrol was cheaper there would be more traffic and this would make the roads more congested.
d) People should use public transport more in order to reduce congestion.

12 Which one of the following is not a normative statement?
a) People should be made to pay for non-essential surgery in the NHS in order to divert resources to essential life-saving operations.
b) All treatment in NHS hospitals should be free for those who need it.
c) Government spending on the NHS has exceeded its targets in the budget.
d) Prescriptions should only be free for the over-sixties who cannot afford to pay.

2

DEMAND AND SUPPLY

Why does the price of computers keep falling while train fares keep rising? This is a question about how markets work and what factors influence the outcomes. To answer it we need to understand how markets function. This is one of the most important topics in economics. You will see that we can go a long way to understanding how markets work with some very simple tools.

In this chapter you will learn:

- Who the participants in markets are and what motivates them.
- What the main factors are that influence how much of a product buyers want.
- What the main factors are that influence how much producers wish to sell.
- How buyers and sellers interact to determine the market price.
- Demand and supply forces are present in some form in all markets, but different market structures affect outcomes.

Growers of crops come to a town to sell their produce in farmers' markets. High street stores stock a wide range of goods for individuals to buy. Commodity markets determine worldwide prices for products such as oil, copper, and wheat. Potential buyers or renters of housing units deal with sellers of houses or landlords in housing markets. Workers sell their services to employers in labour markets. In a series of interactions that we described in Chapter 1 as a 'supply chain', some firms make parts of individual products and sell their outputs to other firms who do further work on these parts and then in turn sell them to other firms who assemble all the parts into final goods that they then sell to households. Indeed, a market exists whenever buyers and sellers exchange goods or services—whether raw materials, individual parts, or final commodities—usually for money.

Markets do not necessarily happen in one place. There is a separate housing market in each town and city, while there is a single market for some skilled labour (for example, university lecturers in physics) throughout the country. The stock market and the foreign exchange market operate globally using computers and telephones. The book market works partly through bookshops and partly by internet and mail order.

Our aim is to help you understand all these different types of market. We start by focusing on a market in which there are many buyers and many sellers. (Markets with a few, or even just one, of either or both buyers and sellers are considered in later chapters.) Although firms buy many things from other firms such as raw materials and individual parts as well as labour from people, we here focus on those who buy goods and services for their own consumption. What they buy are called *final goods* to distinguish them from raw materials, semi-finished goods, and parts that are typically sold by one firm to another. The buyers that we consider in this chapter can be looked at

BOX 2.1 'Tea prices to soar after droughts'

Tea is a typical example of an agricultural commodity whose price can be understood using demand and supply analysis. This box's title was a headline in a newspaper in 2009. It reflected the fact that there had been adverse weather conditions in India and Kenya and this had led to reduced tea crops. Official estimates were that demand was then running ahead of production (supply) and thus it was expected that wholesale prices of tea would rise. A similar price rise occurred in 2017 as reflected in the following statement from the World Bank:

> *Global tea prices were up slightly in the third quarter of 2017 (q/q) and 15 percent more than a year ago. While last season was affected by weather-induced supply shortfalls, especially in East Africa and Sri Lanka, this season's crop appears to be healthy. Tea prices, which are expected to rise 17 percent in 2017, will likely remain steady in 2018 and 2019.*
>
> *(Source: World Bank,* Commodity Price Review, *October 2017.)*

The 'weather-induced' influences were not necessarily droughts in this case as major floods were the problem in Sri Lanka. However, the main point to notice is that standardized commodities like tea have prices determined in the world market, and the key driver of these prices is the balance between demand and supply. In the case of agricultural commodities like tea, coffee, and cocoa, the demand for these products is relatively steady from year to year, but the supply can be adversely affected by unusual weather conditions, such as droughts, floods, or unseasonal frosts. Occasionally also there can be political disruption in a major producer country which restricts the supply reaching markets and forces up the price.

We expand on the example of tea in the case study at the end of this chapter. Other commodities are discussed later in this chapter and the next.

from the point of view of whole households or individuals. They are often called *demanders* or consumers. The firms that sell these final goods and services are often called *producers*. The product whose markets we choose to analyse for purposes of illustration is a good that has a physical existence, so it must be grown or made by the producing firm. Anyone who makes decisions relevant to our theory is called an **agent**. To make the study of their behaviour more manageable, we deal with just three types of agent: consumers, firms, and government. In this chapter you will encounter consumers as demanders and firms as suppliers. In Chapter 10, when we study the labour market, you will find individuals supplying and firms demanding. Governments play a role in some markets as either producers or demanders. They can also intervene in markets by taxing transactions (with a sales tax, value-added tax, or excise tax) or by imposing regulations that impose maximum or minimum prices. In this chapter and the next few we ignore the potential role of governments, but we look closely at this in Chapters 13 and 14.

Box 2.1 provides an example of a real product, tea, that is typical of the type of good whose markets we will be studying first. We will return to this example as a case study at the end of this chapter, but for now notice how the balance between demand and supply forces is key to any market, and that in markets for agricultural products abnormal weather conditions are often a key short-term influence. Note also that this is about wholesale tea prices and not about the price of a box of PG Tips teabags, though the wholesale and retail markets are connected. How demand and supply interact to determine the market price is what we are about to explain.

DEMAND

Individuals and motives

In formulating our demand theory, the consumers are all assumed to be adult individuals who earn income that they spend purchasing various goods and services.[1]

[1] When real-world data are studied, the spending unit analysed is often not the individual consumer but the household. A household is defined as all the people who live under one roof and who make joint financial decisions. For purposes of developing our analysis of markets, however, we view consumers as individuals.

Most economic theories assume that each individual consumer seeks maximum *satisfaction*, or *well-being*, or *utility*, as the concept is variously called. The consumer is assumed to 'maximize utility' within the limits set by his or her available resources. Utility is hard to measure directly, but we only need to assume that typical consumers know what they like and make spending choices that give them as much personal satisfaction as possible. We discuss utility in more detail in Chapter 4.

The nature of demand

The amount of a product that consumers wish to purchase is called the **quantity demanded**. Notice two important things about this concept. First, quantity demanded is a *desired* quantity. It is how much consumers *wish* to purchase *given the resources at their command*, not necessarily how much they succeed in purchasing. We use phrases such as **quantity actually purchased** or **quantity actually bought and sold** to distinguish actual purchases from quantity demanded. Secondly, note that quantity demanded is a *flow*. We are concerned not with a single isolated purchase, but with a continuous flow of purchases. Therefore, we must express demand as so much per period—for example, 1 million oranges *per day*, or 7 million oranges *per week*, or 365 million oranges *per year*. The important distinction between stocks and flows is discussed in Box 2.2.

The concept of demand as a flow appears to raise difficulties when we deal with the purchases of durable consumer goods (often called consumer durables). It makes obvious sense to talk about a person consuming oranges at the rate of thirty per month, but what can we say of a consumer who buys a new television set every five years? This apparent difficulty disappears if we measure the demand for the *services* provided by the consumer durable. Thus, at the rate of a new set every five years, the television purchaser is using the service (viewing TV programmes) at the rate of 1/60th of a set per month. For most purposes, however, we will be interested in the same question that television manufacturers want to know, which is: how many new TV sets are bought in total per period by consumers as a whole?

The determinants of quantity demanded: the demand function

Five main variables are assumed to influence the quantity of each product that is demanded by each individual consumer:

(1) the price of the product;
(2) the prices of other products;
(3) the consumer's income and wealth;
(4) the consumer's tastes;
(5) various individual-specific and environmental factors.

This list is conveniently summarized in what is called a **demand function**:

$$q_n^{\mathrm{d}} = D(p_n, p_1, \ldots, p_{n-1}, Y, S).$$

BOX 2.2 Stocks and flows

Economics makes extensive use of both stock and flow variables and it is important not to confuse the two. A *flow variable* has a time dimension; it is so much per unit of time. The quantity of free-range eggs purchased in Glasgow is a flow variable. Being told that the number purchased was 2,000 dozen eggs conveys no useful information unless we are also told the period over which these purchases occurred. For example, 2,000 dozen per hour would indicate an active market in eggs, while 2,000 dozen per month would indicate a sluggish market.

A *stock variable* has no time dimension; it is just so much. Thus, the number of eggs in an egg producer's warehouse—for example, 20,000 dozen eggs—is a stock variable. All those eggs are there at one time, and they remain there until something happens to change the stock held by the producer. The stock variable is just a number, not a rate of flow of so much per day or per month.

Economic theories use both flow variables and stock variables, and it takes a little practice to keep them straight. The amount of income earned is a flow—so much per year or per month or per hour. The amount of a consumer's spending is also a flow—so much spent per week or per month. In contrast, the amount of money in your bank account is a stock—just so many pounds sterling. The key test for a variable being a flow is that a time dimension is required to give it meaning. Some other variables are neither stocks nor flows; they are just numbers, (e.g. the price of eggs).

The term q_n^d stands for the quantity that the consumer demands per time period (possibly a week or a month) of some product, which we call 'product n'. The term p_n stands for the price of this product, while $p_1, \ldots, p_{n-1}$ is a shorthand notation for the prices of all other products. The term Y is the consumer's income. The term S stands for a host of factors that will vary from individual to individual, such as age, number of children, place of residence (e.g. big city, small town, country), and other assets (e.g. car owners will demand petrol, while non-car owners will demand train tickets). There are also some environmental factors that will affect demand patterns, such as the state of the weather and the time of year. Although these factors matter in real markets, they are not central to our current analysis. Finally, the precise way in which demand is affected by the variables listed earlier is determined by the tastes of the consumer.

The demand function is just a shorthand way of saying that quantity demanded, which is on the left-hand side, is assumed to depend on the variables that are listed on the right-hand side. The form of the function determines the nature of that dependence.[2]

We will not be able to understand the separate influences of each of the previous variables if we ask what happens when all of them change at once. To avoid this difficulty, we consider the influence of the variables one at a time. To do this, we use a device that is frequently employed in economic theory. We assume that all except one of the variables on the right-hand side of the demand function are held constant. Then we allow this one variable, say p_n, to change and see how the quantity demanded (q_n^d) changes. We are then studying the effect of changes in one influence on quantity demanded, *assuming that all other influences remain unchanged*, or, as economists are fond of putting it, **ceteris paribus** (which means 'other things being equal' or 'holding other things constant').

We can do the same for each of the other variables in turn, and in this way, we can come to understand the effect of each variable. Once this is done, we can combine the separate influences of each variable to discover what will happen when several variables change at the same time—as they often do in practice.

Demand and price

We are interested in developing a theory of how products are priced. To do this, we hold all other influences constant and ask: 'How will the quantity of a product demanded vary as its own price varies?'

A basic economic hypothesis is that the lower the price of a product, the larger the quantity that will be demanded, other things being equal.

This negative relationship between the price of a product and quantity demanded is sometimes referred to as the **law of demand**. Why might this law be true? A major reason is that there is usually more than one product that will satisfy any given desire or need. Hunger may be satisfied by meat or vegetables; a desire for green vegetables may be satisfied by broccoli or spinach. The need to keep warm at night may be satisfied by several woollen blankets, or one electric blanket, or a sheet and a lot of oil burned in the boiler. The desire for a holiday may be satisfied by a trip to the Scottish Highlands or to the Swiss Alps, the need to get there by a plane, a bus, a car, or a train, and so on. Name any general desire or need, and there will usually be several different products that will contribute to its satisfaction.

Now consider what happens when we vary the price of one product, holding all other potential variables constant.

First, suppose the price of the product rises. The product then becomes a more expensive way of satisfying a want. Some consumers will stop buying it altogether, others will buy smaller amounts, still others may continue to buy the same amount, but no rational consumer will buy more of it. (Do not forget the important qualification 'other things being equal'. Of course, they might buy more if there was some offsetting change in something else, such as health, which affects what they would like to buy.) Because many consumers will switch wholly, or partially, to other products to satisfy the same want, less will be bought of the product whose price has risen. For example, as meat becomes more expensive, consumers may switch some of their spending to meat substitutes; they may also forgo meat at some meals and eat less meat at others.

Secondly, let the price of a product fall. This makes the product a cheaper method of satisfying any given want. Consumers will buy more of it and less of other similar products whose prices have not fallen. These other products have become expensive *relative to* the product in question. For example, when a bumper tomato harvest drives prices down, shoppers buy more tomatoes and fewer other salad ingredients, which have now become relatively more expensive than tomatoes.

[2] The 'form of the function' refers to the precise quantitative relation between the variables on the right-hand side of the equation and the variable on the left. For example, $y = f(X)$ is a general function stating that Y is related to X, while $Y = 2 + 3X$ is a form of that function indicating a specific relation between these two variables.

The demand schedule and the demand curve

An individual's demand

A **demand schedule** is one way of showing the relationship between quantity demanded and price. It is a numerical tabulation that shows the quantity that will be demanded at some selected prices.

Table 2.1 shows one consumer's demand schedule for eggs. Alice often eats boiled eggs for breakfast and, as she lives on her own, she often finds omelettes a convenient evening meal. But she does not have a lot of money, so she keeps an eye on the price of eggs when she does her weekly supermarket shopping. The table shows the quantity of eggs that she wishes to buy each month at six selected prices. For example, at a price of £1.50 per dozen Alice demands 3.5 dozen per month. For easy reference each of the price–quantity combinations in the table is given a letter.

Next we plot the data from Table 2.1 in Figure 2.1, with price on the vertical and quantity on the horizontal axis.[3] The smooth curve drawn through these points is called the demand curve, even though it may often be drawn as a straight line. It shows the quantity that Alice would like to buy at every possible price; its *negative slope* indicates that the quantity demanded increases as the price falls.

A single point on the demand curve indicates a single price–quantity relationship. *The whole demand curve shows the complete relationship between quantity demanded and price.* Economists often speak of the conditions of demand in a particular market as 'given' or 'known'. When they do so they are referring not just to the specific quantity that is being demanded at the moment (i.e. not just to a particular point on the demand curve) but to the whole demand curve. The whole demand curve remains in one place as long as all variables other than the price of the product itself remain unchanged.

Table 2.1 Alice's demand schedule for eggs

Reference letter	Price (£ per dozen)	Quantity demanded (dozen per month)
A	0.50	7.0
B	1.00	5.0
C	1.50	3.5
D	2.00	2.5
E	2.50	1.5
F	3.00	1.0

The table shows the quantity of eggs that Alice demands at each selected price, other things being equal. For example, at a price of £1.00 per dozen she demands 5 dozen per month, while at a price of £2.50 per dozen she demands only 1.5 dozen.

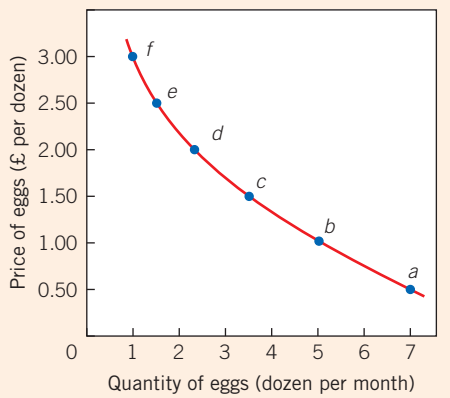


Figure 2.1 Alice's demand curve

This curve relates the price of a commodity to the amount that Alice wishes to purchase. Each point on the figure relates to a row in Table 2.1. For example, when the price is £3.00, 1 dozen are bought per month (point *f*), while when the price is £0.50, 7 dozen are bought (point *a*).

The market demand curve

So far, we have discussed how the quantity of a product demanded by one consumer depends on the product's price, other things being equal. To explain market behaviour, we need to know the total demand of all consumers. To obtain a market demand schedule, we sum the quantities demanded by each consumer at a specific price to obtain the total quantity demanded at that price. We repeat the process for each price to obtain a schedule of total, or market, demand at all possible prices. A graph of this schedule is called the *market demand curve.*

[3] Readers trained in other disciplines often wonder why economists plot demand curves with price on the vertical axis. The normal convention, which puts the independent variable (the variable that does the explaining) on the horizontal axis and the dependent variable (the variable that is explained) on the vertical axis, calls for price to be plotted on the horizontal axis and quantity on the vertical axis. The axis reversal—now enshrined by a century of usage—arose as follows. The analysis of the competitive market that we use today stems from the French economist Leon Walras (1834–1910), in whose theory *quantity* was the dependent variable. However, graphical analysis in economics was popularized by the English economist Alfred Marshall (1842–1924), in whose theory *price* was the dependent variable. Economists continue to use Walras's theory and Marshall's graphical representation, and thus draw the diagram with the independent and dependent variables reversed—to the everlasting confusion of readers trained in other disciplines. In virtually every other type of graph in economics the axes are labelled conventionally, with the dependent variable on the vertical axis.

To avoid unnecessary complication, Figure 2.2 illustrates the summation graphically for only two consumers, Sarah and William. The figure illustrates the proposition that the market demand curve is the horizontal sum of the demand curves of *all* the individuals who buy in the market.

In practice, our knowledge of market demand is usually derived by observing total quantities of sales directly. The derivation of market demand curves by summing individual curves is a theoretical operation. We do it to understand the relation between curves for individual consumers and market curves.

In Table 2.2 we assume that we have data for the market demand for eggs. The schedule tells us the total quantity that will be demanded by all buyers of that product at selected market prices. The data are plotted in Figure 2.3, and the curve drawn through these points is the market demand curve.

Table 2.2 A market demand schedule for eggs

Reference letter	Price (£ per dozen)	Quantity demanded (1,000 dozen per month)
U	0.50	110.0
V	1.00	90.0
W	1.50	77.5
X	2.00	67.5
Y	2.50	62.5
Z	3.00	60.0

The table shows the quantity of eggs that would be demanded by all consumers at selected prices, ceteris paribus. For example, row *W* indicates that, if the price of eggs were £1.50 per dozen, consumers would want to purchase 77,500 dozen per month.

Market demand: a recap

We now summarize what we have learned about demand.

The total quantity demanded depends on the price of the product being sold, on the prices of all other products, on the incomes of the individuals buying in that market, and on their tastes. The market demand curve relates the total quantity demanded to the product's own price, on the assumption that all other prices, total income, tastes, and all other environmental factors are held constant.

Shifts in the demand curve

The demand schedule and the demand curve are constructed on the assumption of *ceteris paribus* (other things held constant). But what if other things change, as surely they must? What, for example, if consumers find themselves with more income? If they spend their extra income, they will buy additional quantities of many products *even though market prices are unchanged*, as shown in Table 2.3. But if consumers increase their purchases of any product whose price has not changed, the new purchases cannot be represented by the original demand curve. The rise in consumer income *shifts* the demand curve to the right, as shown in Figure 2.4. This shift illustrates the operation of an important general rule.

A demand curve shifts to a new position in response to a change in any of the variables that were held constant when the original curve was drawn.

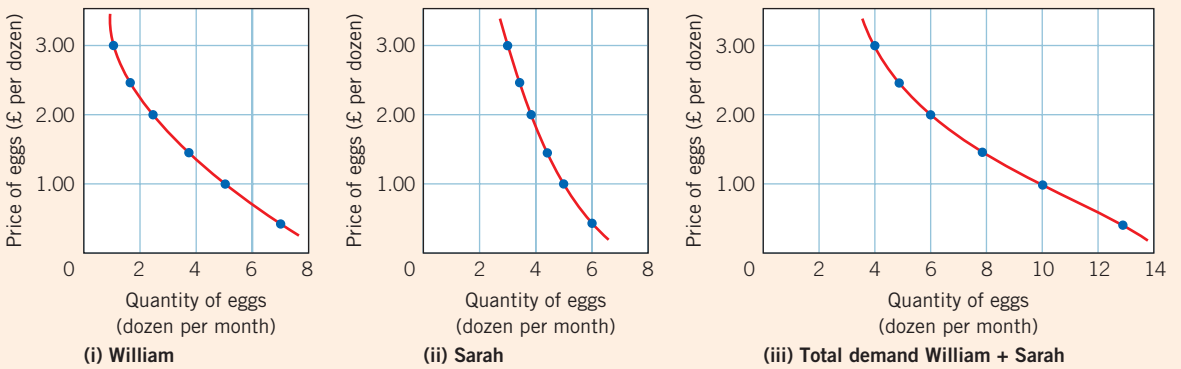


Figure 2.2 The relation between individual and market demand curves

The market demand curve is the horizontal sum of the demand curves of all consumers in the market. The figure illustrates aggregation over two individuals, William and Sarah. For example, at a price of £2.00 per dozen William purchases 2.4 dozen and Sarah purchases 3.6 dozen, and together they purchase 6 dozen.

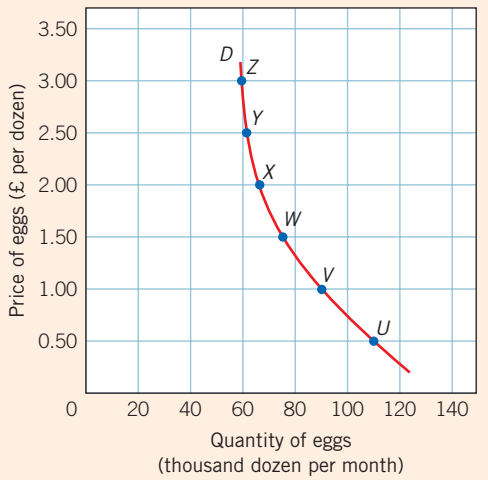


Figure 2.3 A market demand curve for eggs

The negative slope of the curve indicates that quantity demanded increases as price falls. The six points correspond to the six price–quantity combinations shown in Table 2.2. The curve drawn through all the points and labelled *D* is the demand curve.

Any change that increases the quantity of a product consumers wish to buy at each price will shift the demand curve to the right, and any change that decreases the quantity consumers wish to buy at each price will shift the demand curve to the left.

Changes in other prices

We saw that demand curves have negative slopes because the lower a product's price, the cheaper it becomes relative to other products that can satisfy the same needs. Those other products are called **substitutes**. A product becomes cheaper relative to its substitutes if its own price falls. This also happens if the substitute's price rises. For example, eggs can become cheap relative to pizzas either because the price of eggs falls or because the price of pizzas rises. Either change will increase the amount of eggs that consumers are prepared to buy. For example, Alice may eat more omelettes and fewer pizzas whenever she wants a quick meal.

A rise in the price of a product's substitute shifts the demand curve for the product to the right. More will be purchased at each price.

Thus, a rise in the price of pizzas may shift the demand curve for eggs from D_0 to D_1 in Figure 2.4, just as a rise in income did.

Products that tend to be used jointly with each other are called **complements**. Cars and petrol are complements; so are golf clubs and golf balls, bacon and eggs, electric cookers and electricity, an aeroplane trip to Austria and tickets on the ski lifts at St Anton. Since complements tend to be consumed together, a fall in the price of either will increase the demand for both. For example, a fall in the price of cars which causes more people to become car owners will, *ceteris paribus*, increase the demand for petrol.

A fall in the price of one product that is complementary to a second product will shift the second product's demand curve to the right. More will be purchased at each price.

Changes in total income

If consumers receive more income, they can be expected to purchase more of most products even though product prices

Table 2.3 Two alternative market demand schedules for eggs

(1)	Price of eggs (£ per dozen) (2)	Quantity of eggs demanded at original level of personal income (1,000 dozen per month) (3)	Quantity of eggs demanded when personal income rises to new level (1,000 dozen per month) (4)	(5)
U	0.50	110.0	140.0	*U′*
V	1.00	90.0	116.0	*V′*
W	1.50	77.5	100.0	*W′*
X	2.00	67.5	90.0	*X′*
Y	2.50	62.5	81.3	*Y′*
Z	3.00	60.0	78.0	*Z′*

An increase in total consumers' income increases the quantity demanded at each price. When income rises, quantity demanded at a price of £1.50 per dozen rises from 77,500 dozen per month to 100,000 dozen per month. A similar rise occurs at every other price. Thus, the demand schedule relating columns (2) and (3) is replaced by the one relating columns (2) and (4). The graphical representations of these two schedules are labelled D_0 and D_1 in Figure 2.4.

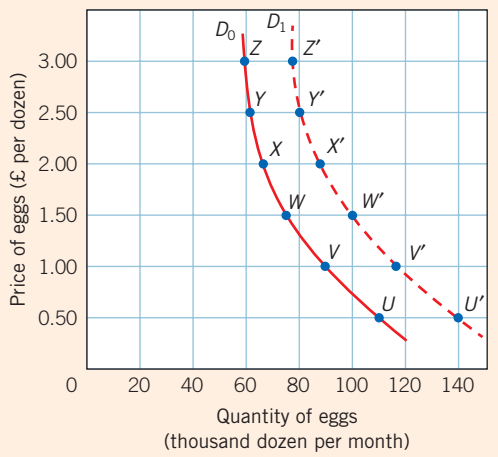


Figure 2.4 Two demand curves for eggs

The rightward shift in the demand curve from D_0 to D_1 indicates an increase in the quantity demanded at each price. The lettered points correspond to those in Table 2.3. When the curve shifts from D_0 to D_1, more is demanded at each price and a higher price is paid for each quantity. For example, at price £1.50 quantity demanded rises from 77,500 dozen (point *W*) to 100,000 dozen (point *W′*), while the quantity of 90,000 dozen, which was formerly bought at a price of £1.00 (point *V*), will be bought at a price of £2.00 after the shift (point *X′*).

remain the same. Such a shift is illustrated in Table 2.3 and Figure 2.4. A product whose demand increases when income increases is called a **normal good**.

A rise in consumers' incomes shifts the demand curve for normal products to the right, indicating that more will be demanded at each possible price.

For a few products, called **inferior goods**, a rise in consumers' income leads them to reduce their purchases (because they can now afford to switch to a more expensive, but superior, substitute).

A rise in income will shift the demand for inferior goods to the left, indicating that less will be demanded at each price.

The distribution of income

If total income and all other determinants of demand are held constant while the distribution of income changes (i.e. some become richer and others become poorer), the demands for normal goods will rise for consumers gaining income and fall for consumers losing income. If both gainers and losers buy a good in similar proportions, these changes will tend to cancel out. However, this will not always be the case.

When the distribution of income changes, demand will rise for those goods favoured by those gaining income and fall for those goods favoured by those losing income.

Individual characteristics

Changes in the characteristics of the individuals who make up the market will cause demand curves to shift. For example, a reduction in the typical number of children per family, as happened in the twentieth century, will reduce the demands for the things used by children at home or in childcare. If the number of retired people increases, there will be a rise in the demand for goods consumed during leisure times.

Environmental factors

Demand for some products is different at different times of year. Some of this is due to weather; for example, demand for electricity is higher in the winter when days are short and the weather is cold,[4] and demand for cold lager and ice cream is higher in the summer during hot weather. Other variations may be due to traditions associated with annual festivals, such as buying presents at Christmas, or timing of school holidays. From the point of view of our theory of demand these are exogenous forces—things that lie outside the theory, affecting demand, sometimes greatly, but not themselves being explained by the theory. Box 2.3 illustrates the major influence that weather can exert on demand.

Changes in tastes

If there is a change of tastes in favour of a product, more will be demanded at each price, causing the demand curve to shift to the right. In contrast, if there is a change of tastes away from a product, less will be demanded at each price, causing the entire demand curve to shift to the left.

Figure 2.5 summarizes our discussion of the causes of shifts in the demand curve. Notice that, since we are generalizing beyond our example of eggs, we have relabelled our axes 'price' and 'quantity', dropping the qualification 'of eggs'. The term *quantity* should be understood to mean quantity per period in whatever units the goods are measured. The term *price* should be understood to mean the price measured in pounds per unit of quantity for the same product.

Movements along demand curves versus shifts

Suppose that you read in today's newspaper that carrot prices have soared because more carrots are being demanded,

[4] In some hot countries, demand for electricity may be even greater in summer than in winter owing to the use of air-conditioners.

BOX 2.3 Weather matters

Other things than price and incomes do influence demand, and sellers must keep abreast of these non-economic causes—which have distinct economic effects.

A newspaper article quotes a person responsible for getting groceries to the stores at the right time and in the right quantities on the effects of a sudden turn in the weather from cold to hot: 'The trigger point is 80 degrees, especially if it is sustained for more than three days. For products such as ice cream, soft drinks, and salads sales can rise by between 70 per cent and 225 per cent on a big change in temperature'. As a result of a sudden upturn in the weather he had recently had to organize another million cases of soft drinks and additional lorry loads of salad and other fast-selling products 'if the shelves [were] not to be bare by lunchtime'.

Here are some of the ways in which the article said the weather affects UK demand.

- Drink sales respond immediately to temperature change.
- After two days of good weather we might think about buying a bike, but it must be nice for more than a week before we start buying suntan lotion.
- Curiously, hot weather increases the sales of plain coleslaw much more than coleslaw with pineapple.
- Soft drink sales depend not only on heat but also on humidity.
- Rain and mild temperatures suit insurers best—we drive less and so have fewer and less serious prangs.
- Builders like storms—they interrupt work but generate huge business volumes repairing and replacing roofs.
- Every one degree colder adds 4 per cent to gas demand and increases electricity demand by about 5,000 megawatts—enough to supply the whole of Sheffield.

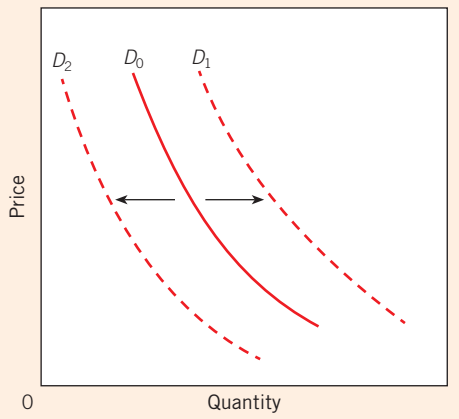


Figure 2.5 Shifts in the demand curve

A shift in the demand curve from D_0 to D_1 indicates an increase in demand; a shift from D_0 to D_2 indicates a decrease in demand. An increase in demand can be caused by a rise in the price of a substitute, a fall in the price of a complement, a rise in income, a redistribution of income towards groups who favour the commodity, or a change in tastes that favours the commodity. A decrease in demand can be caused by a fall in the price of a substitute, a rise in the price of a complement, a fall in income, a redistribution of income away from groups who favour the commodity, or a change in tastes that disfavours the commodity.

perhaps following a report that carrot consumption gives protection against some disease. Then tomorrow you read that the rising price of carrots is greatly reducing the typical consumer's demand for carrots as shoppers switch to potatoes, courgettes, and peas. The two statements appear to contradict each other. The first associates a rising price with a rising demand; the second associates a rising price with a declining demand. Can both statements be true? The answer is that they can be, because they refer to different things. The first refers to a *shift* in the demand curve; the second refers to a movement *along a* demand curve in response to a change in price.

Consider first the statement that the increase in the price of carrots has been caused by an increased demand for carrots. This statement refers to a shift in the demand curve for carrots. In this case, the demand curve must have shifted to the right, indicating more carrots demanded at each price. As we will see later in this chapter, this shift will increase the price of carrots.

Now consider the statement that fewer carrots are being bought because carrots have become more expensive. This refers to a movement along a given demand curve and reflects a change between two specific quantities being bought, one before the price rose and one afterwards.

So, what lay behind the two stories might have been something like the following.

1. A rise in the perceived health-giving properties shifts the demand curve for carrots to the right as more and more are demanded at each price. This in turn is raising the price of carrots (for reasons we shall soon study in detail). This was the first newspaper story.

2. The rising price of carrots is causing each individual consumer to cut back on his or her purchase of carrots. This causes a movement upward to the left along any particular demand curve for carrots. This was the second newspaper story.

To prevent the type of confusion caused by our two newspaper stories, economists have developed a specific vocabulary to distinguish shifts of curves from movements along curves. **Demand** refers to one *whole* demand curve. **Change in demand** refers to a *shift* in the whole curve; that is, a change in the amount that will be bought at *every* price.

An increase in demand means that the whole demand curve has shifted to the right; a decrease in demand means that the whole demand curve has shifted to the left.

Any one point on a demand curve represents a specific amount being bought at a specified price. Therefore, it represents a particular quantity demanded. A movement along a demand curve is referred to as a **change in the quantity demanded**.

A movement down a demand curve is called an increase (or a rise) in the quantity demanded; a movement up the demand curve is called a decrease (or a fall) in the quantity demanded.

To illustrate this terminology, look again at Table 2.3. First, at the original level of income, a decrease in price from £2.00 to £1.50 increases *the quantity demanded* from 67,500 to 77,500 dozen a month. Secondly, the increase in average consumer income *increases demand* from what is shown in column (3) to what is shown in column (4). The same contrast is shown in Figure 2.4, where a fall in price from £2.00 to £1.50 increases the quantity demanded from the quantity shown by point X to the quantity shown by point W. An increase in total consumers' income increases demand from curve D_0 to curve D_1.

SUPPLY

We now look at the supply side of the market for final goods. The suppliers are **firms**, which are in business to make the goods and services that consumers want to buy. Some firms make these goods wholly themselves but more commonly they purchase many parts from various suppliers and assemble the final product to sell to consumers.

Firms' motives

Economic theory gives firms several attributes.

First, each firm is assumed to make consistent decisions, as though it was run by a single individual decision-maker. This allows the firm to be treated as the agent on the production or supply side of product markets, just as the consumer is treated as the individual unit of behaviour on the consumption or demand side.

Secondly, firms hire workers and invest capital and entrepreneurial talent in order to produce goods and services that consumers wish to buy. As we are focusing here on markets for consumer goods and services, we are leaving for later consideration markets where firms sell to other firms, where they sell to governments, and where individuals sell their labour to firms.

Thirdly, firms are assumed to make their decisions with a single goal in mind: to make as much profit as possible. This goal of *profit maximization* is analogous to the consumer's goal of utility maximization.

The nature of supply

The amount of a product that firms are able and willing to offer for sale is called the **quantity supplied**. Supply is a desired flow: how much firms are willing to sell per period, not how much they actually sell.

Here, we make a start on analysis of supply, establishing only what is necessary for a theory of price. In later chapters we study the behaviour of individual firms, and then aggregate individual behaviour to obtain the behaviour of market supply. For our present purposes, however, it is enough to go directly to market supply, the aggregate behaviour of all the firms in a particular market.

The determinants of quantity supplied: the supply function

Three major determinants of the quantity supplied in a market for final goods are:

(1) the price of the product;

(2) the prices of inputs to production, including the parts that they purchase from other firms;

(3) the state of technology.

This list can be summarized in a **supply function**:

$$q_n^s = S(p_n, F_1, \ldots, F_m),$$

where q_n^s is the quantity supplied of product n, p_n is the price of that product, $F_1, \ldots, F_m$ is shorthand for the prices of all inputs into production, and the state of technology determines the form of the function S. (Recall, once again, that the form of the function refers to the precise quantitative relation between the variables on the right-hand side of the equation and the one on the left.)

Supply and price

For a simple theory of price, we need to know how quantity supplied varies with a product's own price, all other things being held constant. Therefore, we are only concerned with the *ceteris paribus* relationship $q_n^s = S(p_n)$; that is, between the quantity firms wish to supply and the price of the product itself. We will have much to say in later chapters about this relationship. For the moment, it is sufficient to state the hypothesis that, holding other things constant, *the quantity of any product that firms will produce and offer for sale is positively related to the product's own price, rising when the price rises and falling when the price falls.*

In Chapter 5 we link this hypothesis to the profit-maximizing behaviour of firms. In the meantime, all we need to note is that the basic reason behind this relationship is the way in which costs are assumed to behave as output changes. The cost of increasing output by another unit is assumed to be higher, the higher is the existing rate of output. So, for example, if the firm is already producing 100 units per week, the cost of increasing output to 101 units per week might be £1, while if 200 units were already being produced, the cost of increasing output to 201 units might be £2. Clearly, the firm will not find it profitable to increase output if it cannot at least cover the additional costs that are incurred. As the price of the product rises, the firm can cover the rising costs of more and more additional units of output. As a result, higher and higher prices are needed to induce firms to make successive increases in output. The result is a positive association between market price and the firm's output. This assumed relationship is discussed in much more detail in in Chapter 5.

The **supply schedule** given in Table 2.4 is analogous to the demand schedule in Table 2.2. It records the quantity that all producers wish to produce and sell at several alternative prices, rather than the quantity consumers wish to buy.

The six points corresponding to the six price–quantity combinations shown in Table 2.4 are plotted in Figure 2.6. The curve drawn through the six points is a **supply curve** for eggs. It shows the quantity produced and offered for sale at each price.[5]

[5] Since we are not considering individual firms in this chapter, all supply curves are market curves showing the aggregate behaviour of the firms in the market. Where that is obvious from the context, the adjective 'market' is usually omitted.

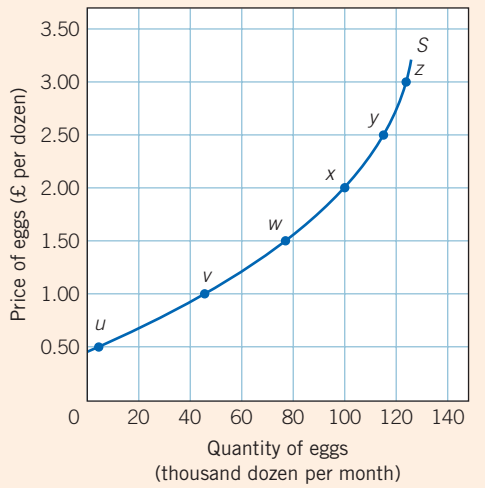


Figure 2.6 A supply curve for eggs

This supply curve relates quantity of eggs supplied to the price of eggs; its positive slope indicates that quantity supplied increases as price increases. The six points correspond to the price–quantity combinations shown in Table 2.4. The curve drawn through these points, labelled *S*, is the supply curve.

The supply curve in Figure 2.6 has a positive slope. This is a graphical expression of the following assumption.

The market price and the quantity supplied are positively related to each other.

Shifts in the supply curve

A shift in the supply curve means that, at each price, a different quantity is supplied. An increase in the supply at each price is illustrated in Table 2.5 and plotted in Figure 2.7. This change

Table 2.4 A market supply schedule for eggs

Reference letter	Price (£ per dozen)	Quantity supplied (1,000 dozen per month)
U	0.50	5.0
V	1.00	46.0
W	1.50	77.5
X	2.00	100.0
Y	2.50	115.0
Z	3.00	122.5

The table shows the quantities that producers wish to sell at various prices, ceteris paribus. For example, row *y* indicates that if the price were £2.50 per dozen, producers would wish to sell 115,000 dozen eggs per month.

Table 2.5 Two alternative market supply schedules for eggs

(1)	Price of eggs (£ per dozen) (2)	Original quantity supplied (1,000 dozen per month) (3)	New quantity supplied (1,000 dozen per month) (4)	(5)
U	0.50	5.0	28.0	*u′*
V	1.00	46.0	76.0	*v′*
W	1.50	77.5	102.0	*w′*
X	2.00	100.0	120.0	*x′*
Y	2.50	115.0	132.0	*y′*
Z	3.00	122.5	140.0	*z′*

An increase in supply means that a larger quantity is supplied at each price. For example, the quantity that is supplied at £2.50 per dozen rises from 115,000 dozen to 132,000 dozen per month. A similar rise occurs at every price. Thus, the supply schedule relating columns (2) and (3) is replaced by the one relating columns (2) and (4).

appears as a rightward shift in the supply curve. A decrease in the supply at each price causes a leftward shift.

There is an important general rule for supply curve shifts similar to the one stated earlier for demand curves.

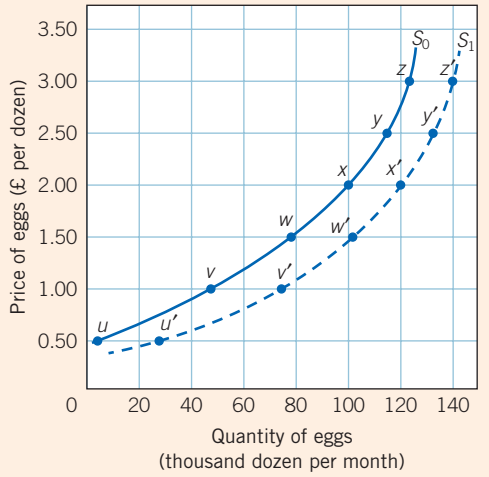


Figure 2.7 Two supply curves for eggs

The rightward shift in the supply curve from S_0 to S_1 indicates an increase in the quantity supplied at each price. For example, at the price of £1.00 the quantity supplied rises from 46,000 dozen to 76,000 dozen per month.

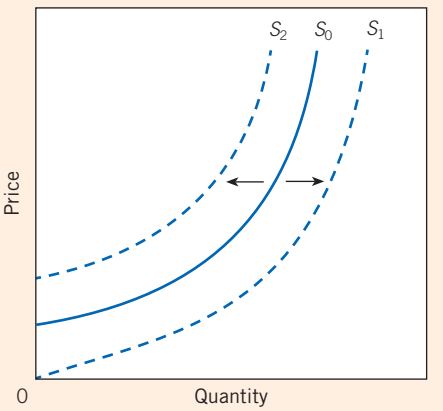


Figure 2.8 Shifts in the supply curve

A shift in the supply curve from S_0 to S_1 indicates an increase in supply; a shift from S_0 to S_2 indicates a decrease in supply. An increase in supply can be caused by improvements in technology or decreases in the costs of inputs that are important in producing the commodity. A decrease in supply can be caused by increases in the costs of inputs that are important in producing the commodity or by changes in technology that increase the costs of production (although such changes are rare).

When there is a change in any of the variables (other than the product's own price) that affects the amount of a product that firms are willing to produce and sell, the whole supply curve for that product will shift.

The major possible causes of such shifts are summarized in the caption to Figure 2.8 and are considered briefly in the next two sections.

Prices of inputs

All things that a firm uses to produce its outputs—such as, in the case of an egg producer, chicken feed, labour, and egg-sorting machines—are called the firm's *inputs*. Other things being equal, the higher the price of any input used to make a product, the less will be the profit from making that product. Thus, the higher the price of any input used by a firm, the lower will be the amount that the firm will produce and offer for sale at any given price of the product.

A rise in the price of any input shifts the supply curve to the left, indicating that less will be supplied at any given price; a fall in the price of inputs shifts the supply curve to the right.

Technology

At any time, what is produced and how it is produced depend on the technologies in use. Over time, knowledge and

production technologies change, and so do the quantities of individual products that can be supplied.

A technological change that decreases costs will increase the profits earned at any given price of the product. Since increased profitability leads to increased production, this change shifts the supply curve to the right, indicating an increased willingness to produce the product and offer it for sale at each possible price.

Movements along supply curves versus shifts

As with demand, it is essential to distinguish between a movement along the supply curve (caused by a change in the product's own price) and a shift of the whole curve (caused by a change in something other than the product's own price). We adopt the same terminology as with demand: **quantity supplied** refers to a specific quantity actually supplied at a particular price of the product, and **supply** refers to the whole relationship between price and quantity supplied. Thus, when we speak of an *increase* or a *decrease in supply*, we are referring to shifts in the supply curve such as the ones illustrated in Figures 2.7 and 2.8. When we speak of a *change in the quantity supplied*, we mean a movement from one point on the supply curve to another point on the same curve.

THE DETERMINATION OF PRICE

So far, we have considered demand and supply separately. We now outline how demand and supply interact to determine price in competitive markets.

The concept of a market

A **market** can be defined as an area over which buyers and sellers negotiate the exchange of some product or related group of products. It must be possible, therefore, for buyers and sellers to communicate with each other and to make meaningful transactions over the whole market. Some markets are local, such as the farmers' market in a county town, while others cover the entire world, such as the market for petroleum or wheat. In recent times, the technology of computers has greatly increased the number of world markets. For example, many commodities are advertised for sale to a worldwide audience on eBay.

Individual markets differ in the degree of competition among the various buyers and sellers. In the next few chapters we will confine ourselves to markets in which the number of buyers and sellers is sufficiently large that no single one of them has any appreciable influence on price. This is a very rough definition of what economists call *perfectly competitive markets*. Starting in Chapter 7, we will consider the behaviour of markets that do not meet this competitive requirement.

The graphical analysis of a market

Table 2.6 brings together the demand and supply schedules from Tables 2.2 and 2.4. Figure 2.9 shows both the demand and the supply curves on a single graph; the six points on the demand curve are labelled with uppercase letters, while the six points on the supply curve are labelled with lowercase letters, with each letter referring to a common price on both curves.

Quantity supplied and quantity demanded at various prices

Consider first the point at which the two curves in Figure 2.9 intersect. Both the figure and Table 2.6 show that when the

Table 2.6 Demand and supply schedules for eggs and equilibrium price

Price per dozen (£)	Quantity demanded (1,000 dozen per month)	Quantity supplied (1,000 dozen per month)	Excess demand (quantity demanded minus quantity supplied) (1,000 dozen per month)
0.50	110.0	5.0	105.0
1.00	90.0	46.0	44.0
1.50	77.5	77.5	0.0
2.00	67.5	100.0	–32.5
2.50	62.5	115.0	–52.5
3.00	60.0	122.5	–62.5

Equilibrium occurs where quantity demanded equals quantity supplied so that there is neither excess demand nor excess supply. These schedules are repeated from Tables 2.2 and 2.4. The equilibrium price is £1.50. For lower prices there is excess demand; for higher prices there is excess supply, which is shown as negative excess demand.

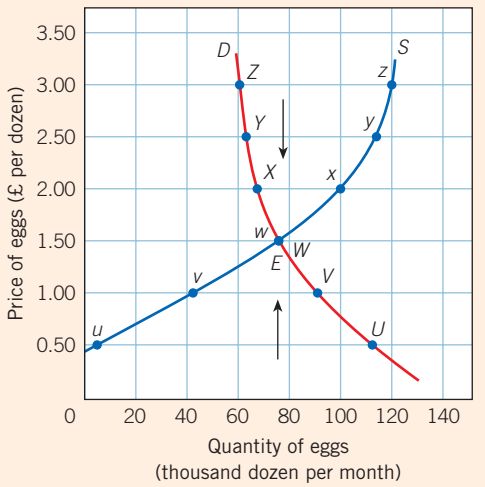


Figure 2.9 Determination of the equilibrium price of eggs

The equilibrium price corresponds to the intersection of the demand and supply curves. Point *E* indicates the equilibrium. At a price of £1.50 per dozen quantity demanded (point *W*) equals quantity supplied (point *w*). At prices above equilibrium there is excess supply and downward pressure on price. At prices below equilibrium there is excess demand and upward pressure on price. The pressures on price are represented by the vertical arrows.

market price is £1.50, the quantity demanded is 77,500 dozen, and the quantity supplied is the same. At that price consumers wish to buy the same amount as producers wish to sell. Provided that the demand curve is negatively sloped, and the supply curve positively sloped throughout their entire ranges, there will be no other price at which the quantity demanded equals the quantity supplied.

Now consider prices below £1.50. At these prices consumers' desired purchases exceed producers' desired sales. It is easily seen that at all prices below £1.50 the quantity demanded exceeds the quantity supplied. Furthermore, the lower the price, the larger is the excess of the one over the other. The amount by which the quantity demanded exceeds the quantity supplied is called the **excess demand**, which is defined as quantity demanded *minus* quantity supplied ($q^d - q^s$). This is shown in the last column of Table 2.6.

Finally, consider prices higher than £1.50. At these prices, consumers wish to buy less than producers wish to sell. Thus, quantity supplied exceeds quantity demanded. It is easily seen, and again you should check a few examples, that for any price above £1.50 quantity supplied exceeds quantity demanded. Furthermore, the higher the price, the larger is the excess of the one over the other. In this case there is negative excess demand ($q^d - q^s < 0$). This is also shown in the last column of Table 2.6.

Negative excess demand is usually referred to as **excess supply**, which measures the amount by which supply exceeds demand ($q^s - q^d$).

Changes in price when quantity demanded does not equal quantity supplied

Whenever there is excess demand, consumers are unable to buy all they wish to buy; whenever there is excess supply, firms are unable to sell all they wish to sell. In both cases some agents will not be able to do what they would like to do. How will they react?

There is a key driving force in markets, which is called the **law of price adjustment** (not to be confused with the *law of demand* that we introduced earlier, which says that demand curves have a negative slope). This law predicts what will happen to the price in a competitive market when there is either excess demand or excess supply.

When supply exceeds demand, the market price will fall. When demand exceeds supply, the market price will rise.

Excess supply means that producers cannot sell all that they wish to sell at the current price. They may then begin to offer to sell at lower prices, such as through clearance sales or discounts. If purchasers observe the glut of unsold output, they may begin to offer lower prices. For either or both reasons, the price in the market will fall.

If, at the current price, consumers are unable to buy as much as they would like to buy, they may offer higher prices in an effort to get more of the available supply for themselves. Suppliers are unable to produce a greater quantity of the product in the short run, but they can ask higher prices for the quantities that they are producing and will make more profit if they do so. For either or both reasons, prices will rise.

This law of price adjustment makes considerable sense and conforms to common experiences of how markets work—shortages of any product tend to lead to price rises, while gluts tend to lead to price falls. Most importantly, it implies that prices will move towards the level at which demand and supply will be equal.

This is a necessary condition for the market to exhibit *stability*. Whenever the current price is not the one that equates demand and supply, the law of price adjustment ensures that the price will move towards the market-clearing price rather than away from it. Thus, it is not enough that there exists a price for which demand is equal to supply. Stability of the market also requires some mechanism to return the price to the market-clearing level whenever it is away from that point. The

combination of a negatively sloped demand curve and a positively sloped supply curve with the law of price adjustment will guarantee a stable market, if any market in this product exists (that is, the demand and supply curves intersect at some positive price and quantity).

The equilibrium price

In our hypothetical example, for any price of eggs above £1.50 the price will fall, while for any price below £1.50 the price will rise. At a price of £1.50 there is neither excess demand associated with a shortage, nor excess supply associated with a glut; the quantity supplied is equal to the quantity demanded. Once supply and demand are equal, there is no tendency for the price to change because suppliers are just able to sell all that they want, and demanders are just able to buy all that they want. Nobody has any incentive to change the price.

The price of £1.50, where the supply and demand curves intersect, is the price towards which the actual market price will tend. It is called the **equilibrium price**: the price at which quantity demanded equals quantity supplied. The amount that is bought and sold at the equilibrium price is called the **equilibrium quantity**. The term 'equilibrium' means a state of balance; it occurs when desired purchases equal desired sales and there are no forces tending to make anything change. Box 2.4 discusses the implications of inflation for our interpretation of market price.

When quantity demanded equals quantity supplied, we say that the market is in **equilibrium**. When quantity demanded does not equal quantity supplied, we say that the market is in **disequilibrium**.

Summary

We have now developed one of the most famous and powerful theories in all of economics, and it is worth summarizing what we have done.

Assumptions concerning a competitive market

- The law of demand: demand curves have negative slopes throughout their entire range.
- The theory of supply: supply curves have positive slopes throughout their entire range.

BOX 2.4 Prices in periods of inflation

Up to now we have developed the theory of the prices of individual products under the assumption that all other prices remain constant. Does this mean that the theory is inapplicable during an inflationary period when almost all prices are rising? Fortunately, the answer is no.

We have mentioned several times that what matters for demand and supply is the price of the product in question relative to the prices of other products. The price of the product expressed in money terms is called its money price; the price of a product expressed in relation to other prices is called its relative price.

In an inflationary world changes in a product's relative price can be measured by changes in the product's own price relative to changes in the average of all other prices, which is called the *general price level*. If, during a period when the general price level rose by 40 per cent, the price of oranges rose by 60 per cent, the price of oranges rose relative to the general price level. Oranges became *relatively* expensive. However, if the price of oranges had risen by only 30 per cent when the general price level had risen by 40 per cent, their relative price would have fallen. Although the money price of oranges rose, oranges became *relatively* cheap.

In Lewis Carroll's famous story *Through the Looking-Glass*, Alice finds a country where everyone must run in order to stay still. So it is with inflation. A product's price must rise as fast as the general level of prices just to keep its relative price constant.

It has been convenient in this chapter to analyse a change in a price in the context of a constant price level. However, the analysis is easily extended to an inflationary period. Any force that raises the price of one product when other prices remain constant will, given general inflation, raise the price of that product relative to the average of all other prices. Consider the example of a change in tastes in favour of eggs that would raise their price by 20 per cent when other prices were constant. However, if the general price level goes up by 10 per cent, then the price of eggs will rise by 32 per cent.[6] In each case the price of eggs rises 20 per cent *relative to the average of all prices*.

In price theory, whenever we talk of a change in the price of one product, we mean a change *relative* to the general price level.

[6] Let the price level be 100 in the first case and 110 in the second. Let the price of eggs be 120 in the first case and x in the second. To preserve the same relative price, we need x such that $120/100 = x/110$, which makes $x = 132$.

• The law of price adjustment: prices rise when demand exceeds supply and fall if supply exceeds demand. They remain unchanged when demand and supply are equal.

Implications

• There is no more than one price at which quantity demanded equals quantity supplied: equilibrium is unique.

• Only at the equilibrium price will the market price remain constant.

• The market is stable in the sense that forces exist to move the price towards its market-clearing level.

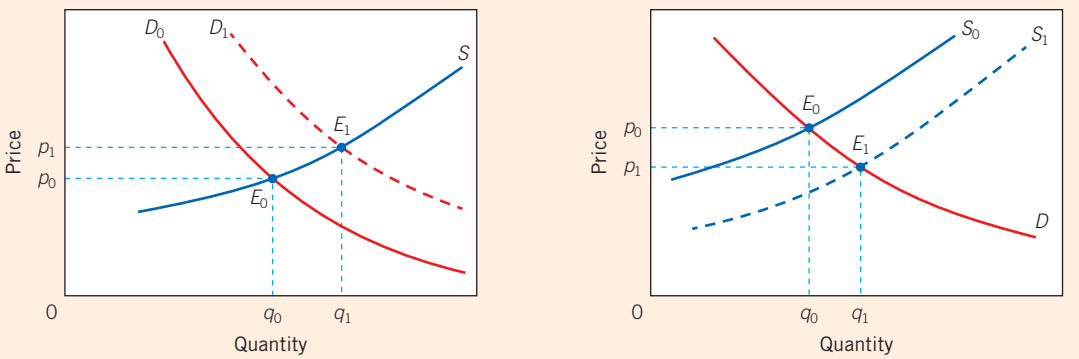
Collectively these forces are sometimes known as the *laws of demand and supply*.

The predictions of demand and supply analysis

Earlier in this chapter, we studied shifts in demand and supply curves. Recall that a rightward shift in the relevant curve means that more is demanded or supplied *at each market price*, while a leftward shift means that less is demanded or supplied *at each market price*. How does a shift in either curve affect price and quantity?

The answers to this question are the predictions of our supply and demand theory. We wish to see what happens when an initial position of equilibrium is upset by some shift in either the demand or the supply curve, and a new equilibrium position is then established. This will enable us to derive predictions about what will happen in any market when something changes, and this is why we study economics—so that we can anticipate what will happen in specific markets when some change happens.

To discover the effects of the demand and supply shifts that we wish to study, we use the method known as **comparative statics**. We start from a position of equilibrium and then introduce the change to be studied. The new equilibrium position is determined and compared with the original one. The differences between the two positions of equilibrium must result from the change that was introduced, for everything else has been held constant.



(i) The effect of shifts in the demand curve

(ii) The effect of shifts in the supply curve

Figure 2.10 The predictions of demand and supply theory

The predicted effects on equilibrium price and quantity of shifts in either demand or supply are as follows:

An increase in demand. In part (i) assume that the original demand and supply curves are D_0 and S, which intersect to produce equilibrium at E_0, with a price of p_0 and a quantity of q_0. An increase in demand shifts the demand curve to D_1, taking the new equilibrium to E_1. Price rises to p_1 and quantity rises to q_1.

A decrease in demand. In part (i) assume that the original demand and supply curves are D_1 and S, which intersect to produce equilibrium at E_1, with a price of p_1 and a quantity of q_1. A decrease in demand shifts the demand curve to D_0, taking the new equilibrium to E_0. Price falls to p_0 and quantity falls to q_0.

An increase in supply. In part (ii) assume that the original demand and supply curves are D and S_0, which intersect to produce an equilibrium at E_0, with a price of p_0 and a quantity of q_0. An increase in supply shifts the supply curve to S_1, taking the new equilibrium to E_1. Price falls to p_1 and quantity rises to q_1.

A decrease in supply. In part (ii) assume that the original demand and supply curves are D and S_1, which intersect to produce an equilibrium at E_1, with a price of p_1 and a quantity of q_1. A decrease in supply shifts the supply curve to S_0, taking the new equilibrium to E_0. Price rises to p_0 and quantity falls to q_0.

The four main predictions of demand and supply are derived in Figure 2.10. The analysis of that figure generalizes our specific discussion about eggs. Because it is intended to apply to any product whose price is determined by the forces of demand and supply and not the decisions of specific individual agents, the horizontal axis is simply labelled 'quantity' and the vertical axis 'price'.

The predictions of supply and demand theory are as follows:

1. A rise in the demand for a product (a rightward shift of the demand curve) causes an increase in both the equilibrium price and the equilibrium quantity bought and sold.

2. A fall in the demand for a product (a leftward shift of the demand curve) causes a decrease in both the equilibrium price and the equilibrium quantity bought and sold.

3. A rise in the supply of a product (a rightward shift of the supply curve) causes a decrease in the equilibrium price and an increase in the equilibrium quantity bought and sold.

4. A fall in the supply of a product (a leftward shift of the supply curve) causes an increase in the equilibrium price and a decrease in the equilibrium quantity bought and sold.

In Figures 2.5 and 2.8 we summarized the many events that cause demand and supply curves to shift. Using the four predictions derived in Figure 2.10, we can understand the link between these events and changes in market prices and quantities. To take one example, a rise in the price of butter will lead to an increase in both the price of margarine and the quantity bought. This is because a rise in the price of one product causes a rightward shift in the demand curves for its substitutes, and Prediction 1 tells us that such a shift causes price and quantity to increase.

The theory of the determination of price by demand and supply is beautiful in its simplicity and yet powerful in its range of applications. Box 2.5 gives some simple applications of demand and supply to real market events.

BOX 2.5 Demand and supply: what really happens

Here are examples of newspaper headlines or extracts that illustrate how demand and supply shifts are relevant to explaining what is happening to price and/or output of a specific product. As an exercise, you should draw a demand and supply curve and then shift the relevant curve in each case to see how the theory explains the outcome.

- OPEC countries once again fail to agree on output quotas. Output soars and prices plummet.
- Oil prices fall as world GDP growth declines sharply.
- Drought in Australia forces up the price of wool.
- The price of cashew kernels has fallen nearly 6 per cent in 10 months as Vietnam has begun to challenge India and Brazil, the world's two largest exporters.
- How deep is the art market's recession? In today's unforgiving economic climate, the sales of contemporary, impressionist, and modern works of art took hits at this week's auctions. Sales totalled just under £60 million compared with £500 million just one year ago. Many paintings on offer went unsold, and those that did sell went for well under their predicted price.
- Coffee prices at the London Commodity Exchange staged another spectacular rise, putting them above their level at the start of the year. The president of the Association of Coffee Producing Countries said that the supply shortages that are underpinning prices would last quite some time.
- Increased demand for macadamia nuts causes price to rise above competing nuts. A major producer now plans to double the size of its orchards during the next five years.
- World steel prices fell as China's increased production capacity came on stream just as world demand fell owing to the financial crisis.
- The price of oats surged 40 per cent in a week as torrential rains in Canada left fields unplanted, raising fears of lower supply of key agricultural commodities.

BUT ARE MARKETS REALLY LIKE THAT?

As we built up our analysis of markets in this chapter, you might have been saying to yourself: 'This is all very well in theory but the markets I know about are not like that'. There is one important respect in which you would be correct. But we now want to persuade you that the doubts you might have about this analysis are not a real problem. Demand and supply

analysis is applicable to many more markets than you might think if you just looked at those prices that react continually to changes in either demand or supply.

Administered prices and auction prices

One important worry that you may have when reading the analysis of markets is that most of the markets in which consumers operate, including you and us, do not work in the way we describe. For example, if you wish to attend your local cinema for a peak-time showing of a much-hyped film you may find queues outside the cinema to buy tickets. There may be more people who want to go at that time than the number of seats available for sale. According to our theory, this means that there is an excess demand and so the price of cinema seats should rise until enough people are discouraged and all those left in the queue are just able to obtain a seat.

Of course, this does not happen. What happens is that the cinema continues to sell tickets at its existing prices until it has sold all the seats and it then puts up a notice saying: 'sold out' or 'house full'.

Similarly, if you go into your local supermarket or department store, you will find the prices of all the goods clearly labelled. You can buy as much of each product as you like at the price set by the store, but the price does not change according to how many people are buying the product on that day. If some product is very popular the store will run out and the shelf will be empty, but the store does not adjust the price to ensure that there is just enough supply to meet the demand.

To go further we must understand that market institutions vary with product and participants. Let us think about different ways in which prices are set and then we shall try to give some reasons for these differences.

The prices that most obviously fit our theory are referred to as **flexible prices** or **auction prices**, as they adjust on a continuous basis to equate demand and supply. Prices in the foreign exchange market, the markets for crude oil, minerals, and grains are flexible as they can change minute by minute while the market is open. Prices that are set by the supplier, who then just waits to see how much of the product sells at that price, are known as **administered prices** or **fixed prices**. Most consumer goods and services are sold at administered prices.

Demand and supply theory is also applicable to many administered prices

Although prices in most retail outlets are set by the retailer, this does not mean that these prices do not adjust to market forces over time. On any specific day we find that all products have a price ticket on them. However, this price may be different from day to day or week to week. The price that the farmer gets from the wholesaler is much more flexible from day to day than the price that the retailer charges consumers. If, for example, bad weather leads to a poor potato crop, then the price that supermarkets must pay to their wholesalers for potatoes will go up and this will be reflected in the prices they mark on potatoes in their stores. Thus, these prices do reflect the interaction of demand and supply in the wider marketplace for potatoes. Although they do not change in the supermarket from hour to hour to reflect local variations in demand and supply, they do change over time to reflect the underlying conditions of the overall production of and demand for the goods in question. For example, fresh strawberries sell at very different prices in mid-winter than in mid-summer. In the summer they are locally grown, but in winter they will have been flown in from the other side of the world.

Even within a supermarket that sets prices on all its produce there will be times when they mark down prices in order to get rid of stock. This may be as it approaches its sell-by date, or if they have new lines arriving the next day. Department stores often have sales at lower prices in order to get rid of stock that has not sold and to make room for new products. When the costs of producing such durable consumer goods as TV sets and refrigerators rise or fall, their prices follow, even if there is a substantial interval between the two sets of changes. However, in all these cases, the supplier is still setting a price and then (in effect) saying: 'Take it or leave it at this price'.

Why are most retail prices administered?

If administered prices do eventually respond to reflect demand and supply conditions, why are they set at a fixed price in the first place? The answer is that this is a more efficient way to organize a retail marketplace. Auction markets work well where all the potential buyers can be assembled in one place (the auction room) or are connected by communication equipment (telephones or computers), so that they can simultaneously bid for the product. The price is set so that the highest bidder gets the goods, and all the goods available are sold.

Imagine the chaos, however, if all the people who shop in your local supermarket must turn up at the same time and make bids for their weekly shopping basket. This is clearly not feasible. Imagine also what the checkout queues would be like if every shopper had to negotiate the price of each item in their shopping trolley as they check it out. Again, this would be a very time-consuming way of shopping. There have been some attempts to organize some retail markets by collecting bids through the internet, such as via eBay. However, it seems

unlikely that this form of shopping is going to replace the supermarket and the department store any time soon.

Mixed pricing

Some markets do have a mixture of administered prices and a degree of price negotiation. This is efficient because these are usually markets for items that are large and ones that you do not buy very often. Cars, for example, have 'list prices' but there is usually some leeway for negotiation about the price of the car itself, about the extras it includes, or about the trade-in price of your old car. When new models are introduced, stocks of the old models may be sold off at lower than the original list price. Houses also are typically listed at an 'asking price' but there is some negotiation around this price, and if several people are chasing the same house there may be what amounts to an auction where the house goes to the highest bidder. Indeed, in Scotland a sealed-bid auction system is the norm. Bidders must enter a written bid on the same day without knowing what others have bid, and the highest bidder gets the property.

Many manufactured goods are put on sale in shops at a fixed price. But these prices are influenced by supply factors such as costs of production, and the rent and wages paid by the retailer. Demand influences will certainly affect the price. A very popular item that the retailer is finding hard to get may have its ticket price raised, while unsold items will be marked down for clearance at some stage. Clothing and other fashion items also typically stay on sale for a period at a fixed price but are then sold off in clearance sales to make way for new fashions or new styles. Indeed, while most clothes retailers have 'sales' around twice a year, many also have permanent racks of discounted items within their store. A similar example is wine shops that regularly have 'bin end' sales of unwanted stock to make space for their new stock.

So far, the discussion has concerned administered prices of retailers who sell to consumers. Although demand and cost conditions will also exert considerable influence on the producers of the goods that end up in shops and also the many services that exist in a modern economy, we need to know more than is contained in the theory of competitive markets that we have just developed to understand the pricing and output decisions of most of these producers. These decisions will be the topic of Chapters 6, 7, and 8.

Summary

The theory of price determination by demand and supply is useful in understanding the working of many different types of market, but some care is needed. The theory works exactly as described in markets where prices are set impersonally by market forces and adjust more or less continually in response to changes in demand and supply. Agricultural commodities at the wholesale level, raw materials, such as iron ore and crude petroleum, and many other similar products fall into this group, as do markets for foreign exchange and company shares. Although most retailers sell at administered prices, these change as the prices that they pay their suppliers change. Thus, the price of foodstuffs will change from week to week, or even day to day, as the wholesale prices that the supermarkets must pay their suppliers change. So variations in demand and supply do explain variations in prices—sometimes immediately and sometimes with lags that depend on such things as who is setting prices and how often it is efficient for these to be changed, and sometimes we need to know more than is contained in the theory of competitive markets to understand pricing and output decisions of those who actually make the goods and services that are eventually sold to consumers.

Relationships between different markets

Although each of the individual markets referred to previously is distinct, all are interrelated, and we need to see why.

The separation of individual markets

Markets are separated from each other in three main ways: by the product sold, by natural economic barriers, and by barriers created by governments. One example of each type of separation is given in the following.

1. The market for men's shirts is different from the market for refrigerators because different products are sold in each.
2. The market for cement in the UK is distinct from the market for cement in the western USA. The costs of transporting cement are so high that UK purchasers would not buy American cement even if its market price in the western USA was much lower than its market price in the UK.
3. The market for textiles used to be separated between many countries because government-imposed trade restrictions severely limited the amount that firms in one country could sell to consumers in another or the prices at which they could sell.

Because markets are distinct, we can use demand and supply analysis to study the behaviour of markets one at a time.

The interlinking of individual markets

Although all markets are to some extent separated, most are also interrelated. Consider again the three causes of market separation: different products, spatial separation, and government intervention. Firstly, the markets for different kinds of product are interrelated because all products compete for consumers'

income. Thus, if consumers spend more in one market, they will have less to spend in other markets. Secondly, the geographical separation of markets for similar products depends on transport costs. Products whose transport costs are high relative to their production costs tend to be produced and sold in geographically distinct markets. Products whose transport costs are low relative to their production costs tend to be sold in what amounts to one world market. But whatever the transport costs, there will be some price differential at which it will pay someone to buy in the low-priced market and ship to the high-priced one. Thus, there is always some potential link between geographically distinct markets, even when shipping costs are high. Thirdly, markets are often separated by policy-induced barriers, such as tariffs (which are taxes paid when goods come into a country from abroad). Although high tariffs tend to separate markets, they do not do so completely. If price differences become large enough, it will pay buyers in the high-price market to import from the low-price market and producers in the low-price market to export to the high-price one, even though they must pay the tariff as a result.

Because markets are interrelated, we must treat them as a single interrelated system for many purposes. *General equilibrium analysis* studies markets as a single interrelated system in which individual demands and supplies depend on all prices, and what happens in any one market will affect many other markets—and in principle could affect all other markets.

CASE STUDIES

1. A storm in a teacup

In Box 2.1 we introduced the wholesale tea market. The central theme of this story was that adverse weather conditions often affected crops and the resulting supply shortages led to higher prices. Figure 2.11 shows the wholesale price of tea from 1971 to 2019. There are price spikes on four occasions before the 2000s which are due to specific supply shortages due to adverse weather or political problems in supplying countries. The

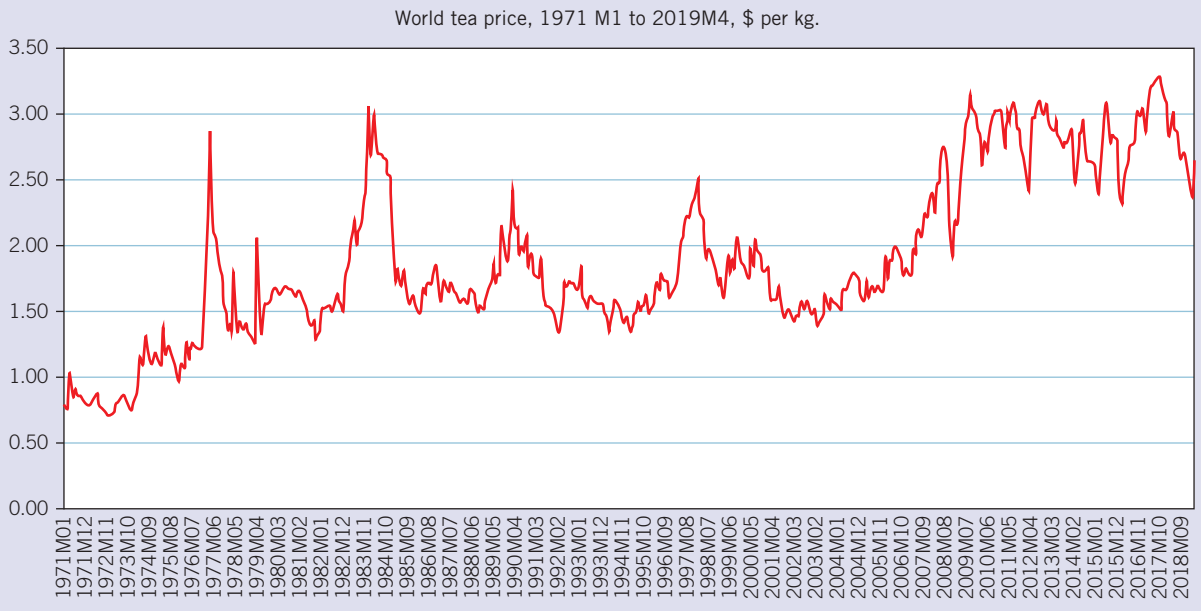


Figure 2.11 Wholesale price of tea, 1971M1–2019M4

Source: World Bank Commodity Price Database as of May 2019, the price is the average of three major markets.

rising price of tea in 2008 can be explained by the excess demand reported previously.

Almost all commodity prices fell in the final quarter of 2008 owing to the recession in most major countries, associated with a collapse in industrial production, world trade, and consumer demand. Against this background the fall in tea prices was relatively modest; for example, the price of crude oil fell from over $140 per barrel to around $40 per barrel at the same time as the tea price fell from $2.76 per kg to $1.93.

The drought conditions in many producing countries (mentioned in Box 2.1) were the main driver of sharply rising tea prices in 2009. Production appeared to be recovering in 2010 and the price dropped back into its earlier range. However, the respite was temporary, and it became clear that the drought conditions of 2008–9 were far from temporary. Industry analysts started to talk about the implications of longer-term climate change for tea-growing areas.

The best tea grows on cool hillsides at around 1,000 feet above sea level. However, climate change has meant that the weather has become too hot and dry in some traditionally successful production areas, such as Assam. Moving location cannot be done quickly as tea plants take between four and twelve years to seed, while plants grown from cuttings do not give crops for at least three years. This means that it takes some time to move production to cooler higher ground. The combination of lower production in old plantations and slowness to expand in new areas helps explain why supply remained restricted between 2010 and the end of 2012.

What explains the price reductions in 2013? There seem to be two main elements to this. The first is that greater supply finally came on stream in response to the high tea prices of 2010–12. The second is that economic disruption in Egypt and the civil war in Syria led to a sharp fall in demand in two important tea-drinking nations. The sharp price rise in 2015 is related to political disruption in Kenya as well as further temporary weather events. A sharp recovery in production in 2016 was then followed in 2017 by the weather-related disruption mentioned in Box 2.1. Prices fell in 2018 as production recovered.

The central message conveyed by this example is that by studying the determinants of demand and supply in a specific market we can explain the main movements of the prices in that market.

A common feature of the markets for many agricultural products is that they normally have stable demand conditions, as demand for food is not very sensitive to either income or price changes (though major wars and recessions can have an impact), while supply can be vulnerable to extreme weather conditions and can thus involve sudden adverse supply shifts. In the case of tea, climate change has been the recent problem. We shall see in the next chapter another example, coffee, where supply is vulnerable to abnormally harsh winters in the main producing countries.

In contrast to markets for agricultural products, markets for metals tend to be characterized by stable supply conditions and demand that varies with the business cycle. The following case study provides supporting evidence for this statement.

2. Keeping the lid on the tin

Figure 2.12 shows the price of tin from January 1960 to April 2019. The inflation of the 1970s and the oil price rises of 1973 and 1979 were associated with many commodity prices rising in dollar terms and tin was no exception. After the oil price and many other commodity prices fell in the mid-1980s the price of tin was flat for a decade or so, though there was a slight rise in the first half of the 1990s as the world came out of recession, and there was a dip between 2000 and 2003 during the post-millennium slowdown. These price variations would look significant if drawn on a larger scale, but they are dwarfed by the subsequent price swings.

Tin is mined in most of the continents of the world even though it is relatively rare as an element in the Earth's crust. Production facilities take time to build, but once in place they can produce at steady levels until the area is exhausted. As most mining is underground it is not subject to adverse weather conditions, and as many countries have tin mines the market is not normally much affected by production problems in any one country. These factors explain the stable price of tin over the decade from 1990, when there was a balanced growth of both demand and supply.

What happened next? Tin is an input into many manufactured products. It is used in the production of bronze, pewter, and die-casting alloys, and in modern engineering to make tungsten more machinable. The largest uses for tin are to produce solders and for tin plating (providing a coating for many iron and steel products). The demand for tin soared in 2006–8 as demand soared for the products in which tin is an input. These include aircraft, ships, trains, white goods (washing machines, dishwashers, tumble driers, refrigerators), cars, commercial buildings, and housing. The demand growth was worldwide, but the biggest single contributor over this period was China whose output was growing at around 10 per cent a year in real terms.

Speculative purchases by investment institutions probably also helped boost the price of tin during the first half

Figure 2.12 The World Price of Tin, 1960M1–2019M4

Source: World Bank Commodity Price Database.

of 2008. These institutions did not want tin for their own use, but at the time they thought that prices were likely to continue rising. They could buy a tin contract in the present and sell it later at a higher price (if they were right and prices did rise) without ever taking delivery of the tin itself. Of course, if prices fell they would lose money, so this was a risky thing to do. However, some institutions specialize in taking such bets and being quick to sell if the price moves against them.

In any event, the boom in tin prices burst in the late summer of 2008, following the global financial crisis and at around the same time as many other commodity and asset prices collapsed. This collapse was linked to the worldwide recession that set in during the latter half of 2008. This price fall came from the realization that world demand was falling, and so production would fall and demand for industrial commodities would also fall.

From the spring of 2009 the world economy recovered strongly and with it so did the demand for tin. Added to the previous uses of tin this time around was the surge in demand for tablet computers that swept the world. The tin price was driven even higher than in 2008. Indeed, the price went so high in 2009 that many producers looked actively into reopening old mines and existing producers tried to expand their production capacity. Accordingly, supply increased somewhat, and the price fell below its peak.

Thus the 2006–9 tin price boom and bust can be explained very simply by a strong rightward shift in the demand curve combined with a positively sloped supply curve (due to the rise in the cost of producing tin as the rate of extraction rises) followed by a sharp leftward shift in the demand curve. The 2009–11 boom can also be attributed to a sharp rightward shift in the demand curve, and the subsequent moderation of price can be thought of as the result of a modest rightward shift of the supply curve as new production capacity came on stream.

Growth of demand and supply were balanced in 2012–14 and the price stayed within a narrow range, but in 2015/16 new production from Myanmar came on stream as this country started to reintegrate with the

world economy after several decades of isolation, and the increased supply depressed the market price. But by August 2016 the following assessment appeared in the financial press:

> *A peak in Myanmar's production could help push the global tin market into a deficit in a few years, as production has also declined in other producing countries such as Peru and the Democratic Republic of Congo. In Indonesia, a rule banning the export of unprocessed materials has also hurt exports .... an analyst said he expects the market to move into a deficit of 5,000 tonnes next year. That could leave China short of supply. Myanmar production is likely to level off because the mainly open pit mining .... has now moved underground, where the tin content of the ore is lower. But an analyst who has visited the mines, warned that it was unclear how much tin there was, as large parts of the area remained unexplored.*
>
> *(Source:* Financial Times*, 29 August 2016 online version: https://www.ft.com/content/808c277a-6b53-11e6-a0b1-d87a9fea034f)*

From late 2016 the tin price settled down and traded in a narrow range from then until mid-2019 when our data end.

In summary, the movements in demand for tin are partly explained by the boom and bust in world activity, or what we will call after Chapter 15 the world *business cycle*. Metal prices tend to be pro-cyclical (they rise in a boom and fall in a slump) as demand for the metal is high when industrial production is high and vice versa. The price of tin clearly fits this pattern. Supply is not weather related but there can be swings in supply due to political disruption or to the emergence of a new producer. The interaction between world demand and supply drives the market price.

CONCLUSION

In many of the markets in which we are interested, the analysis of how demand and supply interact to determine the market-clearing price is an essential tool. More detailed analyses that we are going to do between now and Chapter 14 are designed to build a fuller understanding of the forces affecting different types of market and market structure, and the motives and behaviour of market participants. But we have already gone a long way towards an understanding of how markets work.

SUMMARY

- The decision-taking units in economic theory are called agents. They are (a) individuals, for demand in markets for final goods and services and for supply in labour markets, (b) firms, for supply in goods and services markets and demand in labour and capital markets, and (c) governments, for supply and demand for some goods and services and for regulation and control of the private sector. Given the resources at their command, everyone is assumed to maximize his or her satisfaction, and each firm is assumed to maximize its profit.

Demand

- In the market for final goods and services an individual consumer's demand curve shows the relation between the price of a final product and the quantity of that product the consumer wishes to purchase per period. It is drawn from the assumption that all other prices, income, and tastes remain constant. Its negative slope indicates that the lower the price of the product, the more the consumer wishes to purchase.
- The market demand curve is the horizontal sum of the demand curves of all the individual consumers. The demand curve for a normal good shifts to the right when the price of a substitute rises, when the price of a complement falls, when total income rises, when the distribution of income changes in favour of those with large demands for the product, and when tastes change in favour of the product. It shifts to the left with the opposite changes.
- A movement along a demand curve indicates a change in quantity demanded in response to a change in the product's own price; a shift in a demand curve indicates a change in the quantity demanded at each price in response to a change in one of the conditions held constant along a demand curve.

Supply

- The supply curve for a product shows the relationship between its price and the quantity that producers wish to produce and offer for sale per period. It is drawn on the assumption that all other forces that

influence quantity supplied remain constant, and its usual positive slope indicates that the higher the price, the more producers wish to sell. A supply curve shifts in response to changes in the prices of the inputs used by producers, and to changes in technology. The shift represents a change in the amount supplied at each price. A movement along a supply curve indicates that a different quantity is being supplied in response to a change in the product's own price.

The determination of price

- At the equilibrium price the quantity demanded equals the quantity supplied. Graphically, equilibrium occurs where the demand and supply curves intersect. At any price below equilibrium there will be excess demand and price will tend to rise; at any price above equilibrium there will be excess supply and price will tend to fall.
- A rise in demand raises both equilibrium price and quantity; a fall in demand lowers both. A rise in supply raises equilibrium quantity but lowers equilibrium price; a fall in supply lowers equilibrium quantity but raises equilibrium price.

But are markets really like that?

- Most retail markets do not have prices adjusting continuously to changes in demand and supply. Prices are set at a specific level and shoppers can buy as much as they want at this price. Prices are administered by the seller.
- Even administered prices do adjust to demand and supply forces, but it would not be efficient for these prices to be set either by auction or by negotiation, or to change every minute. Price changes do happen in response to persistent changes in the conditions affecting both demand and supply, but they do not do so instantaneously.

TOPICS FOR REVIEW

- Quantity demanded and the demand function
- The demand schedule and the demand curve for an individual and for the market
- The law of demand
- Shifts in the demand curve and movements along the curve
- Substitutes and complements
- Quantity supplied and the supply function
- The supply schedule and the supply curve
- Shifts in the supply curve and movements along the curve
- Excess demand and excess supply
- Equilibrium and disequilibrium prices
- The law of price adjustment
- Auction prices and administered prices

QUESTIONS

1. What is the equilibrium market price and quantity for each of the following pairs of demand and supply curves:
   i. Demand: $p = £100 - 2q$; supply: $p = £0 + 3q$
   ii. Demand: $p = £100 - 2q$; supply: $q = 30$
   iii. Demand: $p = £100$; supply: $p = £20 + 5q$
2. Use demand and supply curves to analyse what is happening in each of the following situations.
   - The price of coffee has risen because of a frost in Brazil reducing the coffee crop.
   - A fall in air fares from the UK has raised demand for hotel rooms on the Spanish coast.
   - Further falls in chip prices have led to a reduction in the price of laptop computers.
   - An exceptionally cold winter in North America has led to a higher price of oil.
   - A disease in British beef has necessitated the slaughter of large numbers of cattle.
3. List the 'markets' in which you regularly buy goods or services. How are the price and quantity determined during your transaction? Do these prices change on a day to day basis or only infrequently? If prices do not adjust continually, what happens when there is an excess demand or supply?
4. Outline the main determinants of quantity demanded and quantity supplied and explain how these interact to determine the market price.
5. Explain the main differences between administered prices and auction prices and discuss which markets are most suitable for these two different mechanisms.

6 Outline the conditions that are required to achieve equality between demand and supply at a market-clearing price.

7 Explain the effect on market price and quantity in the market for mobile phones of each of the following: consumer incomes rise; technical improvements reduce production costs; the price of fixed-line calls falls sharply.

8 Suppose that the world price of cocoa rose by 20 per cent last year and the quantity supplied (and demanded) fell by 5 per cent. Which one of the following could explain this?

a) There was an increase in supply from a new producer country.
b) Incomes in the main consuming countries rose strongly.
c) The price of coffee, which is a substitute, fell.
d) There was a reduction in supply of cocoa on world markets due to civil unrest in the Ivory Coast (one of the main producers of cocoa).
e) There was a reduction in supply of coffee on world markets due to an early frost in Brazil.

9 Suppose that the price of takeaway pizzas has risen sharply owing to a rise in the world price of wheat. What will happen to the demand curve for takeaway Chinese meals, assuming that these are substitutes for pizzas. Choose one answer.

a) It will become flatter.
b) It will become steeper.
c) It will shift to the left.
d) It will shift to the right.
e) The demand curve will not move but the supply curve will shift to the left.

10 Suppose that the world price of coffee fell by 10 per cent last year and the quantity purchased fell by 5 per cent. Which one of the following could explain this?

a) There was a rise in supply from a producer country, but no change in demand conditions.
b) Demand for coffee rose owing to higher incomes in consumer countries.
c) The price of tea rose, and tea is a substitute for coffee.
d) There was a reduction in supply of coffee owing to bad weather in the main producing countries and price elasticity of demand is –0.5.
e) There was a reduction in demand owing to a recession in the main consumer countries.

3

ELASTICITY OF DEMAND AND SUPPLY

In this chapter we develop some very important concepts that will help you better understand markets. The previous chapter concentrated on the market for final goods, but the analysis in this chapter applies to all markets irrespective of what is bought and sold and independent of the market structure within which this happens. The key points you will learn are:

- That the sensitivity of quantity demanded to a change in price is measured by the elasticity of demand, and what determines the size of elasticity.
- How elasticity is measured at a point or over a range.
- How income elasticity is measured and how it varies with different types of goods and services.
- How elasticity of supply is measured and what it tells us about conditions of production.
- Some of the difficulties that arise in trying to estimate elasticities from sales data.

The demand and supply analysis of Chapter 2 helps us to understand the direction in which price and quantity would change in response to shifts in demand or supply. In most real-world situations economists and business analysts are not going to get away with saying things like: 'If we raise our price sales will fall' or 'If incomes rise this year our demand will increase'. The question they need to answer is: 'By how much?' Fortunately, tools exist to help them, and you, to answer this question in the many different circumstances in which it may be asked. These tools measure the responses of the quantity demanded and the quantity supplied to changes in the variables that determine them, particularly prices and incomes.

To illustrate why we want to have these measures, consider the effects of a new government tax intended to reduce emissions of greenhouse gases by taxing each litre of petrol that is refined. This will shift the supply curve of petrol to the left because less will be offered at each price, or equivalently any specific quantity will only be offered at a higher price. If the demand for petrol is as shown in part (i) of Figure 3.1, the effect of the government's policy will be to increase petrol prices slightly, while greatly reducing the quantity refined and consumed. However, if the demand is as shown in Part (ii) of Figure 3.1, the effect of the policy will be to increase petrol prices greatly, but to reduce petrol production and consumption by only a small amount. If the purpose of the tax is to reduce the amount that is produced and consumed, then the policy will be a great success when the demand curve is similar to the one shown in part (i), but a failure when the demand curve is similar to that shown in Part (ii). However,

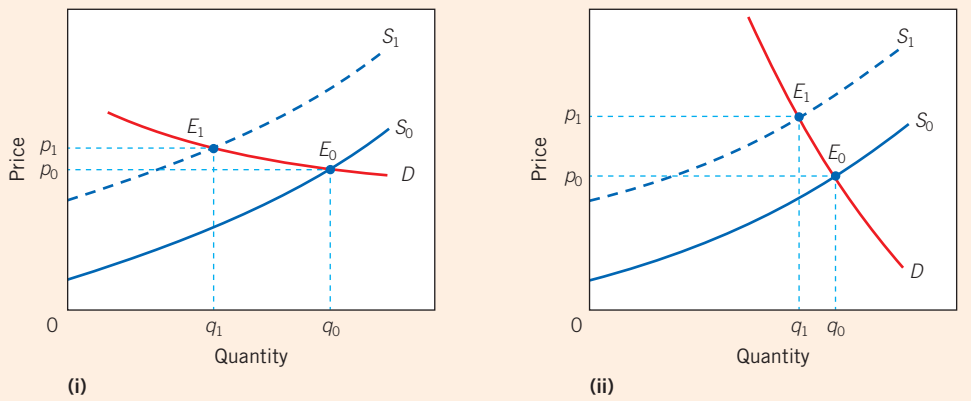


Figure 3.1 The effect of the shape of the demand curve

The flatter the demand curve, ceteris paribus, the less the change in price and the greater the change in quantity. Both parts of the figure are drawn on the same scale. Both show the same initial equilibrium price p_0 and quantity q_0, the same shift of the supply curve from S_0 to S_1, and a new equilibrium at p_1 and q_1. In part (i) the effect of the shift in supply is a slight rise in the price and a large fall in quantity. In part (ii) the effect of the identical shift in the supply curve is a large rise in the price and a relatively small fall in quantity.

if the main purpose of the tax is to achieve a large increase in tax revenue, the policy will be a failure when demand is as shown in part (i) but a great success when demand is as shown in part (ii).

This example shows that it is often not enough to know whether quantity rises or falls in response to some change. It is important to know by how much. To measure this, we use the concept of *elasticity*.

Box 3.1 contains a real report of events in the world market for coffee. We return to this product, and to another one, sugar, later in the chapter, but we introduce it at this stage to alert readers to the fact that understanding the concept of elasticity will help a great deal in understanding developments in real-world markets. We will only be able to spell out the relevance of elasticity to such markets once we have explained the concept in more detail.

BOX 3.1 Real markets that elasticity will help understand

The following is an extract from a report by the World Bank relating to events in the world wholesale market for coffee. We will return to this example later in the chapter to explain how the concept of elasticity is important in understanding the coffee market, and many other similar markets.

The World Bank's Beverage Price Index ticked up marginally in the third quarter of 2017 (q/q), but is almost 13 percent lower than a year earlier. Higher Robusta coffee, cocoa, and tea prices were balanced by receding Arabica coffee prices. Following a sharp, but short-lived, tumble below $3.00/ kg in July, Arabica prices recovered. On a quarterly average basis, however, prices changed little. Robusta prices edged up less than 2 percent in the third quarter. Forecasts for 2017-18 suggest that the Arabica market will be in surplus if an anticipated Brazilian bumper crop materializes. The Robusta market, however, may face tightness due to a shortfall in Vietnam, the world's largest Robusta supplier. Arabica prices are on track to experience a 6 percent slide in 2017, while Robusta prices are projected to jump 15 percent. Prices of both coffees are expected to be steady in 2018.

(*Source:* Commodity Markets Review, *World Bank, October 2017.*)

DEMAND ELASTICITY

In the first part of this chapter we deal with quantity demanded and start by considering its response to changes in a product's own price.

Price elasticity of demand

In Figure 3.1, we were able to compare the responsiveness of quantity demanded along the two demand curves because they were drawn on the same scale. But you should not try to compare two curves without making sure that the scales are the same. Also, you must not leap to conclusions about responsiveness of quantity demanded based on the apparent steepness of a single curve. The hazards of so doing are illustrated in Figure 3.2. Both parts of the figure plot the same demand curve, but the choice of scale on the 'quantity' and 'price' axes serves to make one curve look steep and the other flat.

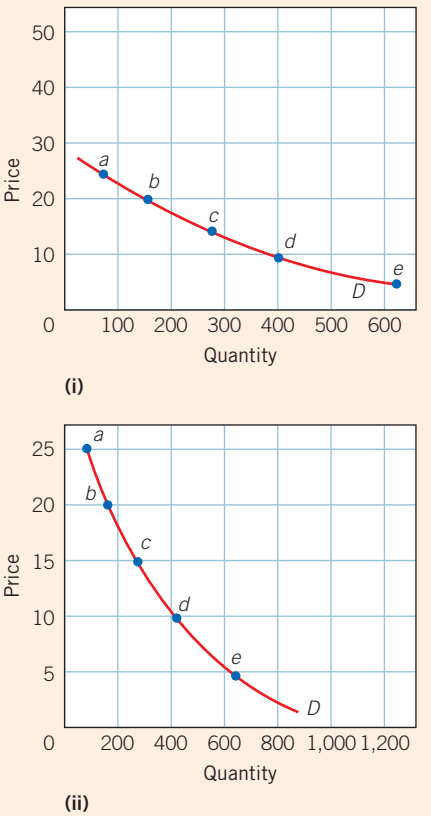


Figure 3.2 One demand curve drawn on two different scales

Suitable choice of scale can make any demand curve appear steep or flat. Parts (i) and (ii) plot the same demand curve. Because the same distance on the quantity axes stands for twice as much in part (ii) as in part (i), and the same distance on the price axes stands for half as much, the curve is steeper when plotted in graph (ii) than when plotted in graph (i).

Measuring the responsiveness of demand to price

We wish to get a measure of responsiveness that is independent of the units in which we measure our quantities and prices as well as the way we draw our graphs; to do this we deal in percentage changes. A given percentage change in the amount of petrol purchased will be the same whether we measure it in gallons or litres. Similarly, although we cannot easily compare the absolute changes in kilos of carrots and barrels of oil, we can compare their two percentage changes.

These considerations lead us to the concept of the **price elasticity of demand**, which is defined as the percentage change in quantity demanded *divided by* the percentage change in price that brought it about.[1] This elasticity is usually symbolized by the lower-case Greek letter eta (η):

$$\eta = \frac{\text{percentage change in quantity demanded}}{\text{percentage change in price}} \qquad (3.1)$$

Many different elasticities are used in economics. To distinguish η from the others, the full term 'price elasticity of demand' can be used. Since η is by far the most commonly used elasticity, economists often drop the adjective 'price' and refer to it merely as *elasticity of demand*, or sometimes just *elasticity*. However, when more than one kind of elasticity could be involved, η should be given its full title.

The sign of the measure

Because of the negative slope of the demand curve, price and quantity will always change in opposite directions. One change will be positive and the other negative, making the measured elasticity of demand negative. This would pose no problem except for two unfortunate habits of economists. First, either by carelessness or design, the minus sign is often dropped, and elasticity is reported as a positive number. Secondly, it is almost

[1] Elasticity is an example of what mathematicians call a *pure number*, which is a number whose value is independent of the units in which it is calculated. Slope, $\Delta p/\Delta q$, is not a pure number. For example, if price is measured in pence, $\Delta p/\Delta q$ will be 100 times as large as $\Delta p/\Delta q$ along the same demand curve where price is measured in pounds sterling.

universal practice when comparing two elasticities to compare their absolute, not their algebraic, values.[2] For example, if product X has an elasticity of −2 while product Y has an elasticity of −10, economists will say that Y has a greater elasticity than X (despite the fact that −10 is *less than* −2). As long as it is understood that absolute and not algebraic values are being compared, this usage is acceptable. After all, the demand curve with the larger absolute elasticity *is* the one where quantity demanded is more responsive to price changes. For example, an elasticity of −10 indicates greater response of quantity to price than does an elasticity of −2.

This need not cause confusion so long as you remember the following.

Demand elasticity is measured by a ratio: the percentage change in quantity demanded divided by the percentage change in price that brought it about. For normal, negatively sloped demand curves, elasticity is negative, but the relative size of two elasticities is usually assessed by comparing their absolute values.

Table 3.1 shows the calculation of two demand elasticities, one that is quite large and one that is smaller. The larger elasticity indicates that quantity demanded is highly responsive to a change in price. The smaller elasticity indicates that the quantity demanded is relatively unresponsive to a change in price.

Interpreting price elasticity

The value of price elasticity of demand ranges from zero to minus infinity. In this section, however, we concentrate on absolute values, and so ask by how much the absolute value *exceeds zero*.

Elasticity is zero if quantity demanded is unchanged when price changes, namely when quantity demanded does not respond to a price change. A demand curve of zero elasticity is shown as curve D_1 in Figure 3.3. It is said to be *perfectly* or *completely* inelastic.

If there is some positive response of quantity demanded to a change in price, the absolute value of elasticity will exceed zero. The greater the response, the larger the elasticity. Whenever this value is less than one, however, the percentage change in quantity is less than the percentage change in price and demand is said to be **inelastic**.

When elasticity is equal to one, the two percentage changes are equal. This case, which is called **unit elasticity**, is the boundary between elastic and inelastic demands. A demand curve having unit elasticity over its whole range is shown as D_3 in Figure 3.3.

When the percentage change in quantity demanded exceeds the percentage change in price, the elasticity of demand is greater than one and demand is said to be **elastic**. When elasticity is infinitely large, there exists some small price reduction that will raise quantity demanded from zero to infinity. Above the critical price, consumers will buy nothing. At the critical price, they will buy all that they can obtain (an infinite amount, if it were available). The graph of a demand curve with infinite price elasticity is shown as D_2 in Figure 3.3. Such a demand curve is said to be *perfectly* or *completely elastic*. (This unlikely looking case will turn out to be important later when we study

Table 3.1 Calculation of two demand elasticities

	Original	New	% change	Elasticity
Good A				
Quantity	100	95	−5	$\frac{-5}{10} = -0.5$
Price	£1	£1.10	10	
Good B				
Quantity	200	140	−30	$\frac{-30}{20} = -1.5$
Price	£5	£6	20	

Elasticity is calculated by dividing the percentage change in quantity by the percentage change in price. With good A, a rise in price of 10p on £1, or 10 per cent, causes a fall in quantity of 5 units from 100, or 5 per cent. Dividing the 5 per cent reduction in quantity by the 10 per cent increase in price gives an elasticity of –0.5. With good B, a 30 per cent fall in quantity is caused by a 20 per cent rise in price, making elasticity –1.5.

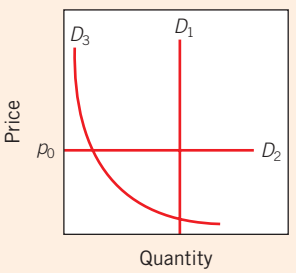


Figure 3.3 Three constant-elasticity demand curves

Each curve has a constant elasticity. D_1 has *zero elasticity*: the quantity demanded does not change at all when price changes. D_2 has *infinite elasticity at the price* p_0: a small price increase from p_0 decreases quantity demanded from an indefinitely large amount to zero. D_3 has *unit elasticity*: a given percentage increase in price brings an equal percentage decrease in quantity demanded at all points on the curve; it is a rectangular hyperbola for which price *times* quantity is a constant.

[2] The absolute value is the magnitude without the sign. Thus, for example, –3 is smaller in algebraic value than 2 but larger in absolute value.

the demand for the output of a single firm with many competitors all producing an identical product.)

Box 3.2 summarizes the discussion of this and subsequent sections. The terminology in the table is important, and it is worth becoming familiar with it at some stage, but you may want to come back to it once you have read the rest of the chapter.

Elasticity and total spending

How does consumers' total spending on a specific product react when the price of the product is changed? The total spending of the product's buyers is equal to the money received by the product's sellers plus any taxes that the government levies on the product. For simplicity, we ignore any taxes, so that sellers' receipts are equal to buyers' spending.

A simple example can be used to show that buyers' total spending and sellers' receipts may rise or fall in response to a decrease in price. Suppose 100 units of a product are being sold for £1 each. The price is then cut to £0.90. If the quantity sold rises to 110, the total spent falls from £100 to £99. But if quantity sold rises to 120, total spending rises from £100 to £108.

The change in total spending brought about by a change in price is directly related to the elasticity of demand. If elasticity is less than unity (so demand is inelastic), the percentage change in price will exceed the percentage change in quantity. The price change will then be the more influential of the two changes, so that total spending will change in the same direction as the price changes, rising as price rises and falling as price falls. However, if elasticity exceeds unity (demand is elastic), the percentage change in quantity will exceed the percentage change in price. The quantity change will then be the more influential change, so that the total amount spent will change in the same direction as quantity changes (that is, in the opposite direction to the change in price).

1. When elasticity of demand exceeds unity (demand is elastic), a fall in price increases total spending on the good and a rise in price reduces it.

2. When elasticity is less than unity (demand is inelastic), a fall in price reduces total spending on the good and a rise in price increases it.

BOX 3.2 The terminology of elasticity

TERM	SYMBOL	NUMERICAL MEASURE OF ELASTICITY	VERBAL DESCRIPTION
Price elasticity of demand (supply)	η (ε_S)		
Perfectly or completely inelastic		Zero	Quantity demanded (supplied) does not change as price changes
Inelastic		Greater than zero, less than one	Quantity demanded (supplied) changes by a smaller percentage than does price
Unit elasticity		One	Quantity demanded (supplied) changes by the same percentage as does price
Elastic		Greater than one, but less than infinity	Quantity demanded (supplied) changes by a larger percentage than does price
Perfectly, completely, or infinitely elastic		Infinity	Purchasers (sellers) are prepared to buy (sell) all they can at some price and none at a higher (lower) price
Income elasticity of demand	η_y		
Inferior good		Negative	Quantity demanded decreases as income increases
Normal good		Positive	Quantity demanded increases as income increases:
Income-inelastic		Less than one	less than in proportion to income increase
Income-elastic		Greater than one	more than in proportion to income increase
Cross-elasticity of demand	η_{xy}		
Substitute		Positive	Quantity demanded of some good and the price of a substitute are positively related
Complement		Negative	Quantity demanded of some good and the price of a complement are negatively related

3. When elasticity of demand is unity, a rise or a fall in price leaves total spending on the good unaffected.[3]

You can check points 1 and 2 for yourself using the example in Table 3.1. Calculate what happens to total spending on the product when price changes in each case. In the case of good *A*, whose demand is inelastic, you will see that a rise in price raises total spending (and thus also the revenue of sellers). In contrast, the rise in the price of good *B*, whose demand is elastic, lowers total spending (and sellers' revenue).

Some complications

We now need to look a little more closely at the elasticity measure. Let us first write out in symbols the definition that we have been using, which is percentage change in quantity divided by percentage change in price:

$$\eta = \frac{(\Delta q/q) \times 100}{(\Delta p/p) \times 100}.$$

We can cancel out the 100s and multiply the numerator and denominator by $p/\Delta p$ to get

$$\eta = (\Delta q/q) \times (p/\Delta p).$$

Since it does not matter in which order we do our multiplication (i.e. $q \times \Delta p = \Delta p \times q$), we can reverse the order of the two terms in the denominator and write

$$\eta = (\Delta q/\Delta p) \times (p/q). \qquad (3.2)$$

We have now split elasticity into two parts: $\Delta q/\Delta p$, the ratio of the *change* in quantity to the change in price, which is related to the slope of the demand curve, and p/q, which is the ratio of the *level* of the price to the quantity at which we make our measurement.

Figure 3.4 shows a straight-line demand curve. If we wish to measure the elasticity at a point, we take our p and q at that point and consider a price change, taking us to another point, and measure our Δp and Δq between those two points. The slope of the straight line joining the two points is $\Delta p/\Delta q$. However, the term in eqn (3.2) is $\Delta q/\Delta p$, which is the reciprocal of $\Delta p/\Delta q$. (This involves just turning the ratio upside down. For example, the reciprocal of 2/3 is 3/2, and the reciprocal of a whole number is its inverse: the reciprocal of 4 is ¼.) Thus, the first term in the elasticity formula (3.2) is the reciprocal of the slope of the straight line joining the two price–quantity positions under consideration. The second term is the ratio of price to quantity at the point where elasticity is measured.

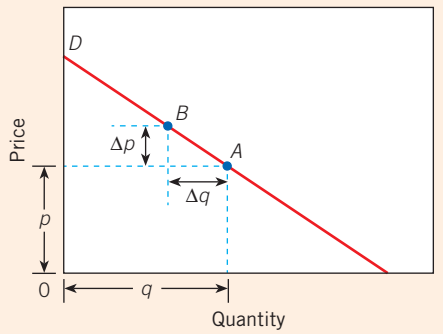


Figure 3.4 Elasticity on a linear demand curve

Elasticity depends on the slope of the demand curve and the point at which the measurement is made. Starting at point *A* and moving to point *B*, the ratio $\Delta p/\Delta q$ is the slope of the line, while its reciprocal $\Delta q/\Delta p$ is the first term in the percentage definition of elasticity. The second term is p/q, which is the ratio of the coordinates of point *A*. Since the slope $\Delta p/\Delta q$ is constant, the elasticity along the curve varies with the ratio p/q, which is zero where the curve intersects the quantity axis and 'infinity' where it intersects the price axis.

Now we can use the expression in eqn (3.2) to discover several things about our elasticity measure.

First, the elasticity of a negatively sloped straight-line demand curve varies from infinity at the price axis to zero at the quantity axis. A straight line has a constant slope, so that the ratio $\Delta p/\Delta q$ is the same anywhere on the line. Therefore, its reciprocal $\Delta q/\Delta p$ must also be constant. We can now infer the changes in η by inspecting changes in the ratio p/q as we move along the demand curve. At the price axis $q = 0$ and p/q is undefined. However, if we let q approach zero, without ever quite reaching it, we see that the ratio p/q becomes very large. Thus, elasticity increases without limit as q approaches zero. Loosely, we say that elasticity is infinity when q is zero. Now move the point at which elasticity is being measured down the demand curve. As this happens, p falls and q rises steadily; thus, the ratio p/q is falling steadily, so that η is also falling. At the q axis the price is zero, so the ratio p/q is zero. Thus, elasticity is zero.

Secondly, with a straight-line demand curve the elasticity measured from any point (p, q), according to eqn (3.2), is independent of the direction and magnitude of the change in price and quantity. This follows immediately from the fact that the slope of a straight line is a constant. If we start from some point

[3] Algebraically, total spending is price *times* quantity. If, for example, the equilibrium price and quantity are p_1 and q_1, then total spending is p_1q_1. On a demand curve diagram price per unit is given by a vertical distance and quantity by a horizontal distance. It follows that, on such a diagram, total spending is given by the *area* of a rectangle, the length of whose sides represent price and quantity. Total revenue (receipts) to the supplier and total spending by consumers are identical in these examples.

(p, q) and then change price, the ratio $\Delta q/\Delta p$ will be the same whatever the direction or the size of the change in p.

Our third point takes us back to the beginning of this chapter, where we warned against judging elasticity from the apparent shape of a demand curve. We often want to compare elasticities of two different demand curves, but we have just seen that the elasticity of a straight-line demand curve varies as we move along it. So how can we compare two numbers both of which are ranging from zero to infinity for every straight-line demand curve? Fortunately, if two demand curves intersect, their elasticity can be compared *at the point of intersection* merely by comparing the slopes of the two curves. The steeper curve is the less elastic. Figure 3.5 shows two intersecting curves and proves that the steeper curve is less elastic than the flatter curve when elasticity is measured at the point where the two curves intersect. The intuitive reason is that at the point of intersection p and q are common to both curves, so all that differs in the elasticity formula is their relative slopes. This is a valuable result which we will use many times in later chapters.

Measured at the point of intersection of two demand curves, the steeper curve has the lower elasticity.

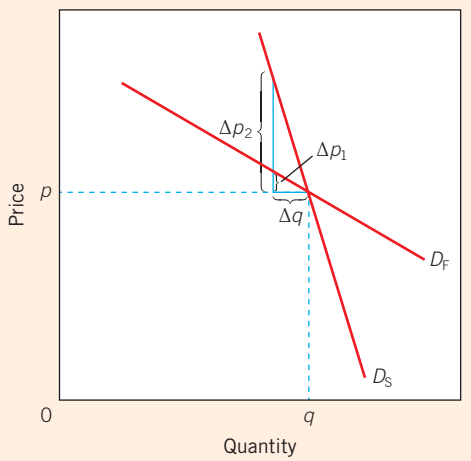


Figure 3.5 Two intersecting demand curves

At the point of intersection of two demand curves, the steeper curve has the lower elasticity. At the point of intersection p and q are common to both curves, and hence the ratio p/q is the same. Therefore, elasticity varies only with $\Delta q/\Delta p$. The absolute value of the slope of the steeper curve, $\Delta p_2/\Delta q$, is larger than the absolute value of the slope $\Delta p_1/\Delta q$ of the flatter curve. Thus, the absolute value of the ratio $\Delta q/\Delta p_2$ on the steeper curve is smaller than the ratio $\Delta q/\Delta p_1$ on the flatter curve, so that elasticity is lower.

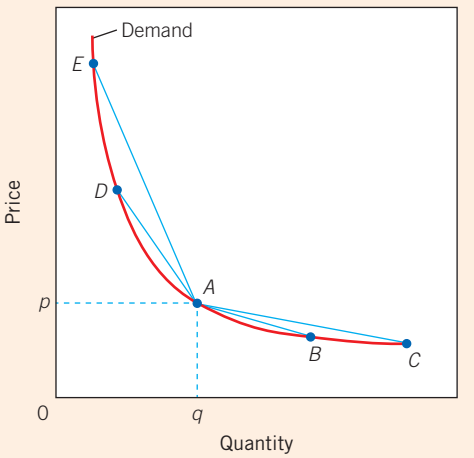


Figure 3.6 Elasticity on a non-linear demand curve

Elasticity measured from one point on a non-linear demand curve and using the percentage formula varies with the direction and magnitude of the change being considered. Elasticity is to be measured from point A, so the ratio p/q is given. The ratio $\Delta p/\Delta q$ is the slope of the line joining point A to the point reached on the curve after the price has changed. The smallest ratio occurs when the change is to point C and the highest ratio when it is to point E. Since the term $\Delta q/\Delta p$ in the elasticity formula is the reciprocal of this slope, measured elasticity is largest when the change is to point C and smallest when it is to point E.

The fourth point is that when eqn (3.2) is applied to a non-linear demand curve, the elasticity measured at any one point varies with the direction and magnitude of the change in price and quantity. Figure 3.6 shows a non-linear demand curve with elasticity being measured at one point. The figure makes it apparent that the ratio $\Delta q/\Delta p$, and hence the elasticity, will vary according to the size and the direction of the price change. This result is very inconvenient. It happens because the ratio $\Delta q/\Delta p$ gives the average reaction of q to a change in p over a section of the demand curve, and, depending on the range that we take, the average reaction will be different.

A more precise measure

The measure defined in eqn (3.2) gives the elasticity over some range, or *arc*, of the demand curve. This measure is sometimes used in empirical work where elasticity is measured between two observed price–quantity situations. In theoretical work, however, it is normal to use a concept that gives a unique measure of the elasticity at each specific point on the demand curve. Instead of using the changes in price (Δp) and quantity (Δq) over some range of the curve, this elasticity measure

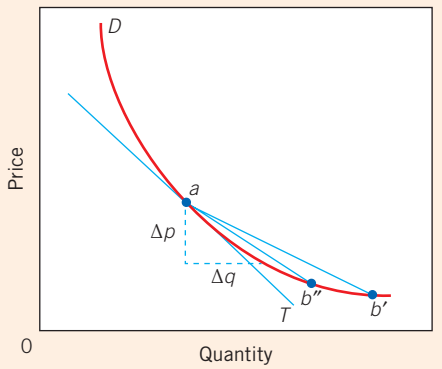


Figure 3.7 Elasticity by the exact method

When elasticity is related to the slope of the tangent to the demand curve at some point, there is a unique measured value of elasticity at that point. In this method the ratio $\Delta q/\Delta p$ is taken as the reciprocal of the slope of the line that is tangent to point *a*. Thus, there is only one measured elasticity at point *a*. It is p/q multiplied by $\Delta q/\Delta p$ measured along the tangent *T*. There is no averaging of changes in *p* and *q* in this measure because only one point on the curve is used.

uses the concept of how quantity is *tending* to change as price changes at each specific point on the curve.

If we wish to measure the elasticity in this way, we need to know the reaction of quantity to a change in price at each point on the curve, not over some range on the curve. We use the symbol dq/dp to refer to this concept and define it as *the reciprocal of the slope of the straight line (i.e. $\Delta q/\Delta p$) that is tangent to the demand curve at the point in question*. Figure 3.7 illustrates the use of this measure to calculate the elasticity of demand at the point *a*. It is the ratio p/q (as it has been in all previous measures) now multiplied by the ratio $\Delta q/\Delta p$ measured along the straight line that is tangent to the curve at *a*.[4] This definition can now be written as

$$\eta = (dq/dp) \times (p/\Delta q). \qquad (3.3)$$

This elasticity measure is the one normally used in economic theory. Elasticity measured by the percentage formula $(\Delta q/\Delta p)(p/q)$ can be regarded as an approximation to this expression. It is obvious from inspecting Figure 3.7 that the elasticity measured from $(\Delta q/\Delta p)(p/q)$ will come closer and closer to that measured by $(dq/dp)(p/q)$, the smaller the price change used to calculate the value of $\Delta q/\Delta p$. Thus, if we consider the percentage definition of elasticity as an approximation to the precise definition, the approximation improves as the size of Δp diminishes. Box 3.3 further investigates some of the properties of the percentage definition and shows a practical way of avoiding some of its undesirable aspects.

[4] Although the expression dq/dp, as we have defined it, is the differential calculus concept of the derivative of quantity with respect to price at the point (p, q), you can understand the concept without knowing calculus.

What determines elasticity of demand?

The main determinant of elasticity is the availability of substitutes. Some products, such as margarine, cabbage, Coca Cola, and the Ford Fiesta, have quite close substitutes—butter, other green vegetables, Pepsi, and the Vauxhall Corsa. When the price of any one of these products changes, *the prices of the substitutes remaining constant*, consumers will to some extent substitute one product for another. When the price falls, consumers buy more of the product and less of its substitutes. When the price rises, consumers buy less of the product and more of its substitutes. More broadly defined products, such as all foods, all clothing, cigarettes, and petrol, have few if any satisfactory substitutes. A rise in their price can be expected to cause a smaller fall in quantity demanded than would be the case if close substitutes were available.

A product with close substitutes tends to have an elastic demand; one with no close substitutes tends to have an inelastic demand.

Closeness of substitutes—and thus measured elasticity—depends both on how the product is defined and on the time period under consideration. This is explored in the following sections. One common misconception about demand elasticity is discussed in Box 3.4.

Definition of the product

There is no substitute for food; it is a necessity of life. Thus, for food taken as a whole demand is inelastic over a large price range. However, it does not follow that any one food—say Weetabix or Heinz tomato soup—is a necessity in the same sense. Each of these has close substitutes, such as Kellogg's Cornflakes and Campbell's tomato soup. Individual food products can have elastic demands, and they frequently do.

Durable goods provide a similar example. Durables as a whole have less elastic demands than individual durable goods. For example, after a rise in the price of TV sets, some consumers might replace their personal computer or their hi-fi system instead of buying a new TV. Thus, although their purchases of TV sets fall, their total purchases of durables fall by much less.

BOX 3.3 Measuring elasticity over a range

We have seen that the percentage formula gives different answers for the elasticity at any point on a non-linear demand curve depending on the size and the direction of the change being considered. Many textbooks just give the percentage elasticity formula, without warning the reader about this property. Inquisitive students usually discover this property with a shock the first time they try to calculate some elasticities from numerical data.

One common way in which students discover the problem is when they try to calculate the elasticity on a unit-elasticity curve. Using the percentage formula, the answer never comes out to be one. For example, the demand curve

$$p = £100/q \qquad \text{(i)}$$

is a unit-elastic curve because expenditure pq remains constant at £100 whatever the price. But if you substitute any two prices into the demand curve (i) and calculate the elasticity according to the percentage formula, you will never get an answer of 1, whatever two prices you take. For example, the equation tells us that if price rises from £2 to £3, quantity falls from 50 to 33.3. If we take the original price as £2, we have a price change of 50 per cent and a quantity change of –33.3 per cent, making an elasticity of –0.667. If we take the original price as £3, the elasticity comes out to be –1.5.

This is unsatisfactory. The problem can be avoided when measuring elasticity between two separate points on the curve by taking p and q as the average values between the two points. This measure has two convenient properties. First, it is independent of the direction of the change and, secondly, it gives a value of unity for any point on a demand curve whose true value is unity.

In the previous example the average p is £2.50 and the average q is 41.667. This makes the percentage change in price 40 per cent ((1/2.5) × 100) and the percentage change in quantity also 40 per cent ((16.667/41.667) × 100). So, elasticity is correctly measured as 1. Whatever two prices you put into eqn (i), you will always get a value of unity for the elasticity, if you use the average of the two prices and of the two quantities when calculating the elasticity. Readers who enjoy playing with algebra can have fun proving this proposition.*

The best approximation to the correct measure when elasticity is measured between two separate points on a demand curve is obtained by defining p and q as the average of the prices and quantities at the two points on the curve.

This is the best way to measure elasticities given readings from any two points on a curve when that is all that is known. As we have seen in the text, for theoretical purposes the way out of the problem is to measure the ratio $\Delta q/\Delta p$ as the slope of the tangent to one point on the curve rather than between two points on the curve. To do this we need to know a portion of the demand curve around the point in question.

In practice economists do not usually estimate elasticity based on only one observation. It is more common to report an elasticity measure that is valued at the mean of two p and q data points. This is analogous to the averaging we suggest here.

* What you need to prove is that

$$\frac{q_2-q_1}{p_2-p_1}\cdot\frac{(p_1+p_2)/2}{(q_1+q_2)/2} = 1.$$

Because many specific manufactured goods have close substitutes, they tend to have price-elastic demands. A Marks & Spencer own-brand raincoat could be expected to be price elastic, but all clothing taken together will be inelastic. This is because it is easy to substitute a Marks & Spencer raincoat with a John Lewis or Selfridges raincoat, but you have to wear something in the winter and you cannot avoid wearing clothing altogether when its price in general rises relative to, say, the price of food.

Any one of a group of related products will tend to have an elastic demand, even though the demand for the group as a whole may be inelastic.

Long-run and short-run elasticity of demand

Because it takes time to adjust fully to some price changes, a demand that is inelastic in the short run may prove elastic when enough time has passed. For example, before the first Organization of Petroleum Exporting Countries (OPEC) oil price shocks of the early 1970s, the demand for petrol was thought to be highly inelastic because of the absence of satisfactory substitutes. But the large price increases over the 1970s led to the development of smaller, more fuel-efficient cars and to less driving. The elasticity of demand for petrol was measured as –0.6 soon after the price rose. However, when the first five years of quantity adjustment had been allowed for the elasticity had become –1.2.

BOX 3.4 Elasticity and income

It is often argued that the demand for a product will be more inelastic the smaller the proportion of income spent on it. The argument runs as follows. When only a small proportion of income is spent on some product, consumers will hardly notice a price rise. Hence, they will not react strongly to price changes one way or the other.

The most commonly quoted example of this alleged phenomenon is salt. However, salt is a poor example for the argument being advanced. Although it does take up a very small part of consumers' total expenditure, it also has few close substitutes. Consider another product, say one type of mint. These mints no doubt account for only a small portion of the total expenditure of mint suckers, but there are many close substitutes—other types of mints and other sucking sweets. The makers of Polo mints, for example, know that if they raise Polo prices greatly, mint suckers will switch to other brands of mint and to other types of sucking sweets. Thus, they face an elastic demand for their product.

Similar considerations apply to any one brand of matches. If the makers of Swan Vesta matches raise their prices significantly, people will switch to other brands of matches rather than pay the higher price.

What this discussion shows is that *goods with close substitutes will tend to have elastic demands whether they account for a large or a small part of consumers' incomes.*

However, there is another aspect of the influence of income. To see this, consider any good that has an inelastic demand. A rise in its price causes more to be spent on it. If consumers spend more on that product, they must spend less on all others taken as a group. But the higher the proportion of income spent on the product, the less likely they are to spend more on it when its price rises. After all, if a consumer spends all his or her income on potatoes, demand must have unit elasticity. As price rises, purchases must then fall in proportion since the consumer has only a given income to spend. Thus, *for a good to have a highly inelastic demand it must have few good substitutes, and it must not take up too large a proportion of consumers' total expenditure.*

For many products the response of quantity demanded to a given price change, and thus the measured price elasticity of demand, will tend to be greater the longer the time span considered.

The different quantity responses can be shown by different demand curves. Every demand curve shows the response of consumer demand to a change in price. For products such as cornflakes and ties, the full response occurs quickly and there is little reason to worry about longer-term effects. Therefore, a single demand curve will suffice for these products. Other products are typically used in connection with highly durable appliances or machines. A change in price of, say, electricity or petrol may not have its major effect until the stock of appliances and machines using these products has been adjusted. This adjustment may take a long time, making it useful to identify two kinds of demand curve for such products. A *short-run demand curve* shows the response of quantity demanded to a change in price, *given* the existing quantities of the durable goods that use the product, and *given* existing supplies of substitute products. A different short-run demand curve will exist for each such structure of durable goods and substitute products. The *long-run demand curve* shows the response of quantity demanded to a change in price after enough time has passed to allow all adjustments to be made.

The relation between long-run and short-run demand curves is shown in Figure 3.8. Assume, for example, that there is a large rise in the price of electricity. The initial response will be along the short-run demand curve. There will be some fall in quantity demanded, but the percentage drop will be less than the percentage rise in price, making short-run demand inelastic. Over time, however, many people will replace their existing electric cookers with gas cookers as they wear out. New homes will be equipped with gas rather than electric appliances more often than they would have been before the price rise. After further time, some factories will switch to relatively cheaper sources of power. When all these types of long-run adaptation have been made, the demand for electricity will have fallen a great deal. Indeed, over this longer period, the percentage reduction in quantity demanded may exceed the percentage increase in price. If so, the long-run demand for electricity will be elastic.

The long-run demand curve for a product that is used in conjunction with durable products will tend to be substantially more elastic than any of the short-run demand curves.

It should not be concluded that long-run elasticities always exceed their short-run value. Estimates of demand for automobiles themselves suggest that the long-run price elasticity of demand may be as low as −0.2, while the short-run

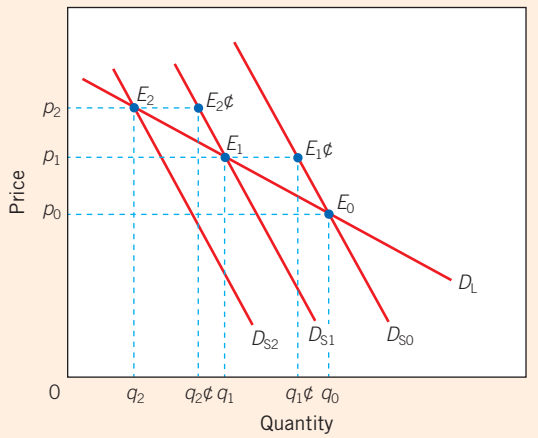


Figure 3.8 Short-run and long-run demand curves

This long-run demand curve is more elastic than the short-run curves. D_L is the long-run demand curve showing the quantity that will be bought after consumers become fully adjusted to each given price. Through each point on D_L there is a short-run demand curve. It shows the quantities that will be bought at each price when consumers are fully adjusted to the price at which that short-run curve intersects the long-run curve. So, at every other point on the short-run curve consumers are not fully adjusted to the price they face, possibly because they have an inappropriate stock of durable goods. When consumers are fully adjusted to price p_0, they are at point E_0 consuming q_0. Short-run variations in price then move them along the short-run demand curve D_{S0}. Similarly, when they are fully adjusted to price p_1, they are at E_1 and short-run price variations move them along D_{S1}. The line D_{S2} shows short-run variations in demand when consumers are fully adjusted to price p_2.

elasticity is of the order of −1.2 to −1.4.[5] This evidence comes from the United States where having a car is essential, especially for those living in rural areas, so demand is not very sensitive to price in the long run. However, it is usually easy to postpone buying a new car for some time, so higher prices do reduce spending on cars more than in proportion to the rise in price over the first several months after a price rise. The demand for the cars of one specific maker is even more price elastic in the short run. Chevrolet, for example, was estimated to face a short-run price elasticity of −4 if it were to raise its own price while other producers held their prices constant.

[5] See http://www.mackinac.org/article.aspx?ID=1247.

Durable goods may have a higher price elasticity of demand in the short run than in the long run, as it is possible to postpone the purchase of a replacement for some time.

These insights will prove valuable in several of the chapters that follow.

Other demand elasticities

So far, we have discussed *price elasticity of demand*, the response of the quantity demanded to a change in the product's own price. The concept of demand elasticity can be broadened to measure the response to changes in *any* of the variables that influence demand. How much, for example, do changes in income and the prices of other products affect quantity demanded?

Income elasticity

Economic growth has raised the real income of the average citizen of Europe and North America quite dramatically over the past two centuries. At low levels of income, most money is spent on such basics as food, clothing, and shelter. As income rises an increasing proportion of expenditure tends to fall on manufactured goods, particularly such durables as cars, TV sets, and refrigerators. At yet higher levels of income more and more of any additional income goes to services such as foreign travel, entertainment, and education.

The responsiveness of demand for a product to changes in income is termed **income elasticity of demand**, and is defined as

$$\eta_y = \frac{\text{percentage change in quantity demanded}}{\text{percentage change in income}}$$

For most products, increases in income lead to increases in quantity demanded, and therefore income elasticity is positive. If the resulting percentage change in quantity demanded is larger than the percentage increase in income, η_y will exceed unity. The product's demand is then said to be **income-elastic**. If the percentage change in quantity demanded is smaller than the percentage change in income, η_y will be less than unity. The product's demand is then said to be **income-inelastic**. In the boundary case, the percentage changes in income and quantity demanded are equal, making η_y unity. The product is said to have a *unit income elasticity of demand*.

While virtually all observed price elasticities are negative, income elasticities are observed to be both positive and negative.

We have already encountered the link between income changes and demand in Chapter 2. We argued that a change in income would shift the demand curve for a product. If a rise in income causes more of it to be demanded (other things being equal) so that the demand curve shifts rightward, as it does with most products, we call it a normal good. If a rise in income causes less of it to be demanded, which means a leftward shift in the product's demand curve, we call it an inferior good. So normal goods are those that have positive income elasticities, while inferior goods are those that have negative income elasticities. Finally, the boundary case between normal and inferior goods occurs when a rise in income leaves quantity demanded unchanged, so that income elasticity is zero.

The important terminology of income elasticity is summarized in Figure 3.9 and Box 3.2. It is worth spending a bit of time familiarizing yourself with this terminology. Figure 3.9 illustrates all possible reactions by showing a product whose income elasticity goes from zero to positive to negative. No specific good is likely to show a pattern exactly like this. Most goods will have a positive income elasticity at all levels of income—people demand more as they get richer. (It should be obvious that no good can have a negative income elasticity at *all* levels of income.) A graph that directly relates quantity demanded to income, such as Figure 3.9, is called an *Engel curve* after Ernst Engel (1821–1896), the German economist who used this device to display the relationship between household income and spending on necessities.

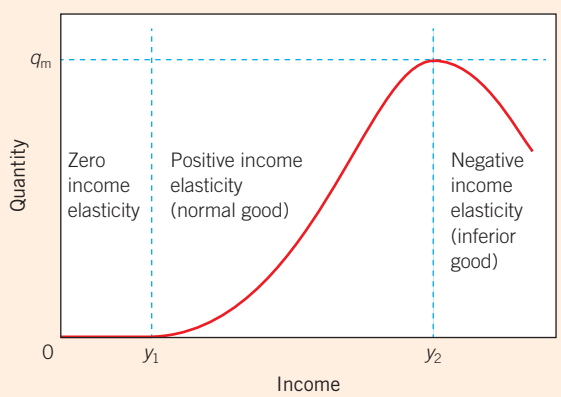


Figure 3.9 The relation between quantity demanded and income

Normal goods have positive income elasticities; inferior goods have negative elasticities. Nothing is demanded at income less than y_1, so for incomes below y_1 income elasticity is zero. Between incomes of y_1 and y_2, quantity demanded rises as income rises, making income elasticity positive. As income rises above y_2, quantity demanded falls from its peak at q_m, making income elasticity negative.

Income elasticities may be larger in the long run than in the short run, just as many price elasticities are. This is because people may take time to adjust spending patterns as their resources change over time. You may, for example, get a pay rise this year, but not buy a bigger car until it is clear that your income has reached a sustainable higher level, or your current car is sufficiently depreciated. This explains why one study found that the income elasticity of fuel demand is between 1.1 and 1.3 in the long run, and between 0.35 and 0.55 in the short run.

Cross-elasticity

The responsiveness of quantity demanded for one product to changes in the prices of other products is often of considerable interest. Producers of, say, beans and other meat substitutes find the demands for their products rising when cattle shortages force the price of beef up. Producers of large cars find their sales falling when the price of petrol rises dramatically.

The responsiveness of demand for one product to changes in the price of another product is called **cross-elasticity of demand**. It is defined as

$$\eta_{xy} = \frac{\text{percentage change in quantity demanded of product } x}{\text{percentage change in price of product } y}$$

Cross-elasticity can vary from minus infinity to plus infinity. Complementary goods have negative cross-elasticities and substitute goods have positive cross-elasticities.

Mobile phones and the calls that can be made on them, for example, are complements. A fall in the price of calls causes an increase in the demand for both handsets and calls. Thus, changes in the price of calls and in the quantity of handsets demanded will have opposite signs—price of calls goes down and demand for handsets goes up. In contrast, mobile calls and fixed-line calls are substitutes: a fall in the price of mobile calls increases the quantity of mobile calls made but reduces the quantity demanded of fixed-line calls. Therefore, changes in the price of mobile calls and in the quantity of fixed-line calls demanded will have the same sign. The terminology of cross-elasticity is also summarized in Box 3.2.

Box 3.5 provides some examples of the importance of taking elasticity into account when making many practical decisions.

BOX 3.5 Elasticity matters

Elasticities may seem rather boring concepts to learn about and to calculate. But they are powerful tools. Practical people who scorn theory ignore elasticity considerations at their peril. Here are a few cautionary tales.

- Not long ago the treasurer of a professional association introduced a motion at the annual meeting 'to increase membership fees by 10 per cent to increase our revenues by 10 per cent'. You could have told them that they were optimistically assuming that the elasticity of demand for membership was zero. If the elasticity differed at all from zero, revenues would rise by less than 10 per cent. If the elasticity proved to be greater than unity, they would suffer a loss of revenue as a result of the rise in fees.
- A local bus company raised its prices by 10 per cent to cover increased costs. It was pleased to see its revenues rise by 5 per cent. The next year it confidently raised its prices again and was surprised and dismayed to find that its revenues fell by 2 per cent. The manager was reported in the local press as saying, 'It is hard to do business when our customers are so erratic'. You could have told him that there was nothing erratic about the customers' behaviour. The manager was unreasonably assuming that the elasticity of demand for bus rides was constant over the whole relevant range. All that happened was that the second increase took fares into the range where the market demand curve was elastic.
- Many countries have recently experienced water shortages, and some have decided to install water meters and pricing by usage in order to reduce demand from households. However, the evidence later emerged that domestic water consumption has a very low price elasticity of demand. For internal use this figure could be as low as –0.04. If this number is correct a 10 per cent rise in water prices would only lower domestic usage by less than half of one per cent. Water providers may have slightly better luck in lowering external usage (for the garden and car washing) as the price elasticity for this is significantly higher at –0.3. This is still inelastic, but a 10 per cent rise in price would achieve a 3 per cent cut in usage. Thus, pricing would not be an effective way of reducing demand in the case of water, but it may generate enough revenue to justify investing in new supply capacity.

SUPPLY ELASTICITY

We have seen that elasticity of demand measures the response of quantity demanded to changes in any of the variables that affect it. Similarly, elasticity of supply measures the response of quantity supplied to changes in any of the variables that influence it. Because we wish to focus on the product's own price as a variable influencing its supply, we will be mainly concerned with *price elasticity of supply*. The usual practice is to drop the adjective 'price', referring to 'elasticity of supply' or 'supply elasticity' whenever there is no ambiguity in this usage.

Supply elasticities are important in economics. Our treatment is brief for two reasons: first, much of what has been said about demand elasticity carries over to the case of supply elasticity and does not need repeating; secondly, we will have more to say about the determinants of supply elasticity later in this book when we have looked more closely at the production decisions of firms.

A definition

The **price elasticity of supply** is defined as the percentage change in quantity supplied divided by the percentage change in price that brought it about. Letting the lower-case Greek letter epsilon, ε, stand for this measure, its formula is

$$\varepsilon_s = \frac{\text{percentage change in quantity supplied}}{\text{percentage change in price}}$$

Supply elasticity is a measure of the degree of responsiveness of quantity supplied to changes in the product's own price.

Since supply curves normally have positive slopes, supply elasticity is normally positive.

Interpreting supply elasticity

Figure 3.10 illustrates three cases of supply elasticity. The case of zero elasticity is one in which the quantity supplied does not

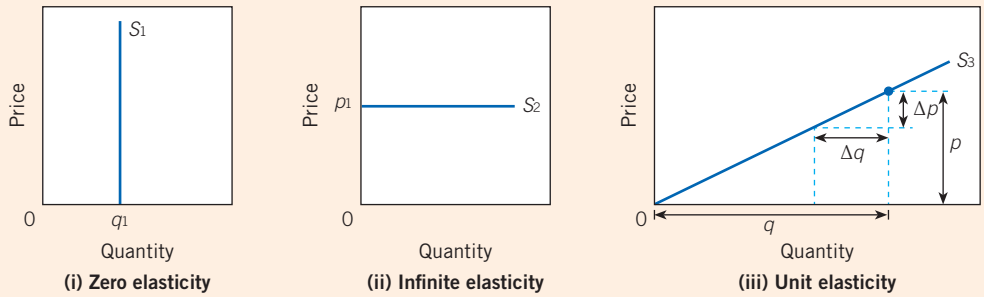


Figure 3.10 Three constant-elasticity supply curves

All three curves have constant elasticity. Curve S_1 has a *zero elasticity*, since the same quantity, q_1, is supplied whatever the price. Curve S_2 has an *infinite elasticity at price* p_1; nothing at all will be supplied at any price below p_1, while an indefinitely large quantity will be supplied at the price of p_1. Curve S_3, as well as all other straight lines through the origin, has a *unit elasticity*, indicating that the percentage change in quantity equals the percentage change in price between any two points on the curve.

change as price changes. This would be the case, for example, if suppliers persisted in producing a given quantity and dumping it on the market for whatever it would bring. Infinite elasticity occurs at some price if nothing is supplied at lower prices, but an indefinitely large amount will be supplied at that price. Any straight-line supply curve drawn through the origin, such as the one shown in part (iii) of Figure 3.10, has an elasticity of unity. The reason is that, for any positively sloped straight line, the ratio of p/q at any point on the line is equal to the ratio $\Delta p/\Delta q$ that defines the slope of the line. Thus, in the expression $(\Delta q/\Delta p)(p/q)$ the two ratios are equal, so the product is unity.

The case of unit supply elasticity illustrates that the warning given earlier for demand applies equally to supply. Do not confuse geometric steepness of supply curves with elasticity. Since *any* straight-line supply curve that passes through the origin has an elasticity of unity, it follows that there is no simple correspondence between geometrical steepness and supply elasticity. The reason is that varying steepness (when the scales on both axes are unchanged) reflects varying *absolute* changes, while elasticity depends on *percentage* changes. The terminology of supply elasticity is summarized in Box 3.2.

What determines elasticity of supply?

What determines the response of producers to a change in the price of the product that they supply? First, the size of the response depends in part on how easily producers can shift from the production of other products to the one whose price has risen. If agricultural land and labour can be readily shifted from one crop to another, the supply of any one crop will be more elastic than otherwise. Here also, as with demand, length of time for response is critical. It may be difficult to change quantities supplied in response to a price increase in a matter of weeks or months, but easy to do so over a period of years. An obvious example concerns the planting cycle of crops. Also, new oilfields can be discovered, wells drilled, and pipelines built over a period of years, but not in a few months. Thus, the elasticity of supply of oil is much greater over five years than over one year, and greater over one year than over one month. Secondly, elasticity is strongly influenced by how costs respond to output changes. This issue will be looked at in more detail in later chapters.

MEASUREMENT OF DEMAND AND SUPPLY

Much of what economists do to earn a living uses measurements of demand and supply elasticities. Will a fare increase help to ease the deficit of London Underground or the Panama Canal? The answer requires knowledge of price elasticity of demand. The United Nations Food and Agriculture Organization (FAO) and producers' co-operatives use income elasticities of demand to predict future changes in demand for food. Over the past decade, many industries have estimated their products'