EDINBURGH
BUSINESS SCHOOL
HERIOT-WATT UNIVERSITY

Economics

Professor Keith G. Lumsden MA, PhD, FRSE

EC-A3-engb 2/2012 (1002)

This course text is part of the learning content for this Edinburgh Business School course.

In addition to this printed course text, you should also have access to the course website in this subject, which will provide you with more learning content, the Profiler software and past examination questions and answers.

The content of this course text is updated from time to time, and all changes are reflected in the version of the text that appears on the accompanying website at http://coursewebsites.ebsglobal.net/.

Most updates are minor, and examination questions will avoid any new or significantly altered material for two years following publication of the relevant material on the website.

You can check the version of the course text via the version release number to be found on the front page of the text, and compare this to the version number of the latest PDF version of the text on the website.

If you are studying this course as part of a tutored programme, you should contact your Centre for further information on any changes.

Full terms and conditions that apply to students on any of the Edinburgh Business School courses are available on the website www.ebsglobal.net, and should have been notified to you either by Edinburgh Business School or by the centre or regional partner through whom you purchased your course. If this is not the case, please contact Edinburgh Business School at the address below:

Edinburgh Business School
Heriot-Watt University
Edinburgh
EH14 4AS
United Kingdom

Tel + 44 (0) 131 451 3090
Fax + 44 (0) 131 451 3002
Email enquiries@ebs.hw.ac.uk
Website www.ebsglobal.net

Economics

Professor **Keith G. Lumsden**, MA, PhD, FRSE is Director of Edinburgh Business School, Heriot-Watt University, Edinburgh.

The concept of a distance-learning MBA was the brainchild of Professor Lumsden while he was at Stanford University, California, where he taught in the Graduate School of Business for 15 years. He pioneered many of the techniques that form the basis of the MBA courses, particularly self-assessment questions and problems, case studies, databases and computer simulations. He wrote the first programmed-learning texts in economics, some of which have been translated into seven languages. He has published 14 books and over 60 articles in professional journals.

Professor Lumsden consults with major companies throughout the world. He is currently engaged in organising and teaching in executive seminars and programmes for many international companies.

Companies that have used the Lumsden materials and methodology include:

- American Express
- Barclays Bank
- British Petroleum
- British Telecom
- Digital Equipment Corporation
- Hewlett-Packard
- IBM
- Marks and Spencer
- Morgan Guaranty
- Rolls-Royce
- Volvo Transport

For comments on and contributions to this course, the author is grateful to Richard Attiyeh, Donald Mackay, Ron McKinnon, Alan Peacock, Charles Ritchie, Alex Scott, Geoff Wyatt and Craig Robinson.

First Published in Great Britain in 1990.

© Keith Lumsden 1990, 2002, 2003, 2008, 2011

The right of Professor Keith G. Lumsden to be identified as Author of this Work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without the prior written permission of the Publishers. This book may not be lent, resold, hired out or otherwise disposed of by way of trade in any form of binding or cover other than that in which it is published, without the prior consent of the Publishers.

Contents

Introduction to the Economics Course

A not unreasonable question you might ask is, 'What use is a knowledge of economics to the practising businessman or businesswoman?' The answer is in two parts: the first is a 'consumption' benefit, the second an 'investment' benefit.

What this course can give you is an understanding of how the market economy functions: what its strengths and weaknesses are; how all the different parts of the economy relate to each other through the market mechanism; and what we as a society can do to modify this market mechanism to serve us better. What this course can also give you is an ability to evaluate the typically biased views of politicians pursuing their own party's policies and criticising the opposition's. From the myriad of statistics emerging from the economy, politicians are past masters at carefully selecting those that support their position and highlighting others with negative connotations as being attributable to mismanagement by the opposition. Almost all politicians also use the tools of economists to predict a rosy future if their policies are enacted and dire results if the opposition prevails. These are some of the principal 'consumption' benefits.

What of the 'investment' benefits? The major benefit that a comprehension of economics can bestow on the practising manager is an understanding of how economists tackle problems and an adoption of the economist's approach. The world in which the economist operates is a very complex world because it is the real world. To understand this world, the economist builds models that attempt to isolate critical factors and strip away peripheral ones. The critical factors remaining must capture the core of the problem being investigated; without this capability, rigorous analysis of the problem is impossible.

Major managerial decisions for the successful company must be treated in a similar fashion. If your company is contemplating setting up a plant abroad and you have to decide on precise location, size of plant, layout of factories, positioning of warehouses, establishment of a research and development centre, a legal department, an international exchange department – to name but a few – you will not find the answers to these problems in an economics textbook. What you will find, however, is a methodology for setting out the major issues to be resolved, and a method and tools for analysing such a complicated problem.

Module 1

Economic Concepts, Issues and Tools

Contents

1.1 Introduction

The pattern of this book is to introduce economic concepts and theories, apply these concepts and theories to real-world problems, and finally to evaluate economic policies based on the theories. Since one issue faced by students in many subjects is not knowing whether they have absorbed their course material, this book has been designed to eradicate this perennial problem. Multiple choice questions at the end of each chapter allow self-testing of how well fundamental concepts and principles have been mastered. Accompanying cases or real problems permit testing of analytical ability in economics and help to develop the higher-order skills of synthesis and evaluation. Those, in turn, can be tested in the essays contained in the Practice Final Examinations which appear at the end of the book.

In all aspects of life, scarcity exists both for individuals and societies. Since all living creatures have wants that they attempt to satisfy but limited means of satisfying these wants, choices have to be made. Economics is the subject that studies how individuals and societies tackle such a problem. For many individuals and societies, the making of choices is not an unpleasant experience. Where to go on holiday, which restaurant to try, which Christmas gifts to choose, how to design the new University library, where to locate new roads and shopping centres are activities that, although normally facing budgetary restrictions, are not matters of life and death. For a large portion of the world's population, however, many choices *are* matters of life and death. Even in more developed countries, scarcity imposes some cruel dilemmas.

In many underdeveloped countries in the world, the problem of scarcity is stark and striking. Insufficient income is generated to keep people alive for very long: children starve and die; medical services are non-existent; housing and sanitation consist of cardboard boxes and open sewers; life is 'nasty, brutish and short'. Attempts to make life better for future generations by building towns and factories can mean fewer resources allocated to the wants of the current generation. How many people would approve, or be willing to take the decision, of sacrificing the lives, say, of 50 per cent of the children in a village so that those who remain – and their children – could have more of their wants satisfied in the future?

In Britain, for example, the number of available kidney machines is less than the number required for all patients suffering from kidney disease. Were more resources allocated to producing additional kidney machines, more lives could be saved. Should more kidney machines be produced? If so, how many, and what should be given up for them in the National Health Service budget? Should expenditure be reduced on geriatric clinics or on ante-natal clinics, or should resources in cancer research be cut back to facilitate the expansion of kidney-machine production? Or should the National Health Service's budget be increased? If so, how large should it be and what should be given up to permit improved healthcare? Would a good trade-off be fewer policemen or firemen, fewer rock concerts or films, less education and recreation facilities, or more work and less leisure?

All of these problems, which necessitate choices, arise from scarcity. The fact that scarcity is relative, that the bulk of the people living in the developed world are incomparably better off than millions of people existing in abject poverty in the poorer countries, does not diminish the need for choice, though the implications of certain choices have more serious consequences.

Different societies in the world have tackled the scarcity–choice problem in different ways. Some countries have relied on a small group of individuals making the major economic decisions: those are the centrally planned economies. Others have utilised the market system in which the forces of demand and supply impersonally resolve the scarcity–choice issue: those are the capitalist, or free-enterprise, economies. Others, primarily societies in the less developed world, have relied on traditional methods of economic organisation in which the positions of individuals in a society determine primarily their jobs and rewards. No country in the world today provides a pure example of one of the above types of economic systems. In countries such as China, Iran, Syria and Vietnam, central authorities have a much greater influence in determining what and how goods and services are produced than the governments of Japan, Germany, the United Kingdom and the USA. In the latter group, however, the governments are more active in tax/transfers policies to redistribute income from the better off to the worse off.

The primary focus of this book is on the functioning of the market or free-enterprise type of economies. The principal reason for this is that about 80 per cent of the world's output is produced by nations that are primarily free-enterprise economies even though they constitute less than 20 per cent of the world's population. The Eastern bloc countries, which under communism adhered to a planning type of economy, moved towards the market system after the fall of the Soviet Union.

The approach is to develop answers to questions such as: How would a completely free-enterprise system work? Would the outcomes for such an economy be desirable for anyone, for a favoured few or for everyone? Are there areas of economic activity where a free-enterprise system will not provide the goods and services people want most? Can government interference in the economy make it function more effectively? Is there reason to expect the government to interfere more than is desirable? Who gains and who loses when the government interferes? In the remainder of this module a number of basic concepts are developed that must be understood to answer these questions.

1.2 Scarcity and Choice

All living creatures have wants, which they attempt to satisfy. The great white shark, for example, wants food for survival. It attempts to satisfy this want by preying on most creatures that swim in the ocean. Some shark hunters, in turn, kill sharks as a means of livelihood, made possible by the wants of others for shark meat, oil, fertilisers and shark's fin soup. Big-game fishermen hunt sharks to satisfy their wants for a dangerous chase or the mounting of a trophy. From this example, it is obvious that wants are not limited to basic subsistence items. Similarly, people's wants for housing, clothing, food, transportation, recreation and art exceed what is needed just to sustain human life. Not many decades ago, large families in what are now high-income countries were raised, as they are today in many poor countries, in houses consisting of a room and kitchen with outside toilet facilities. For most families today in developed countries, houses are spacious and convenient and meet much more than basic shelter requirements. Over time the wants of a society change; fortunately, so does the ability to satisfy them.

Whether human wants are determined by the need to survive (bread and water), the desire for pleasure (caviar and champagne), or gluttony (Roman-style eating orgies) is not a primary concern of an economist. Economists take human wants as given and do not, as economists, pass value judgements about whether people should want the things they do want. (My son believes that hitting a golf ball over 300 metres gives a meaning to life, but *as an economist* I do not question that belief!) The main focus of economics is on how societies can satisfy their wants as fully as possible, given their limited resources for providing the items that satisfy such wants. So, the economist is concerned primarily with how society can best use its available means to achieve its given wants.

It is possible to conceptualise a world in which all of society's wants are fully satisfied, but such a paradise has never been observed. Consider only the materialistic side of such a world. Not only would there have to be a sufficient (infinite?) supply of all the goods and services that each and every person wanted but also sufficient time to consume and enjoy them. The richest person in today's world faces scarcity, for example, in the form of a limited number of years in a lifetime and the unwillingness of owners of unique assets to part with them. Bill Gates has only 24 hours in every day and it is not clear that the Louvre in Paris would be willing to part with the *Mona Lisa* no matter how much he wanted this painting and no matter how much he was willing to pay for it. From a practical viewpoint, it does not matter whether we envisage a world in which human wants are infinite or one in which human wants are finite but greater than resources can satisfy; in both cases we finish up with a world of scarcity. It is only in such a world that economics has any meaning.

Items that satisfy some wants are provided by nature without human effort. Air, rain, beautiful scenery and sunsets can increase human well-being without human effort. They are known as 'free goods'. Other items which increase well-being, however, require considerable coordination of effort on the part of many thousands of individuals. Producing a car or sending a satellite around Mars is not a gift of nature. Thus, with the exception of the free goods provided by nature, the ability of individuals or nations to provide the goods and services that help satisfy wants is limited by the resources under their command. Not all students can earn A grades in every subject because of resource constraints – time and ability. Similarly, life expectancy is low in many parts of the world because of inadequate supplies of food, shelter and medicines.

A 'resource' is defined as anything that helps produce the goods and services people want. The wheat farmer's fields are resources, as are his tractors, fertiliser and seed, all of which are required to produce the wheat that people want. Similarly, the manager of a professional football team requires resources to produce football games that people pay to attend. These resources include a playing field, uniforms, trainers, administrative staff and football players, including the reserves who may only play occasionally during the season. To provide students with a traditional college education requires buildings, professors, computers, textbooks, and a significant input of students' time – another resource. Studying by distance learning renders some of those resources obsolete!

It is the dual existence of insatiable wants and limited resources that produces what economists call the fundamental fact of scarcity. Because of scarcity, all individuals and nations face the same problem, namely the allocation of limited resources among the goods and services that are desired. Students have to decide how to allocate their time among study, drinking, sports, and sleeping activities. Households have to decide how to allocate their incomes among food, clothing, health services, education, transportation, and leisure activities. A nation has to decide how to allocate its labour force, lands, factories, and other resources to produce national defence (jets, nuclear submarines, army, navy and airforce personnel), health services (medical schools, hospitals, kidney machines, AIDS clinics), educational services (pre-school playgroups, university campuses, audio-visual equipment, computers, teachers), transportation services (motorways, aircraft, private cars, railways) and food and drink (wheat, soya beans, steak, beer, whisky).

Economics is the study of how individuals and nations use resources under their command to satisfy their wants as fully as possible or to maximise their welfare (or *utility*) given their resource constraints. Economics, therefore, is concerned with scarcity and choice. Scarcity exists because there are insufficient resources to satisfy all wants fully. The need for choice arises because resources allocated to the satisfaction of one want cannot, by definition, be available to satisfy other wants simultaneously. A child with only $1.00 to spend in a sweet shop has a resource constraint, i.e. $1.00. Given that the child wants many different types of sweets, that child has a choice problem. A dollar spent on jelly babies means no chocolate bar. Similarly, every hour the President of the United States allocates to dealing with foreign policy is one hour less for domestic policy, electioneering, spending with his family, or sleeping.

Some wants can be satisfied without the need for choice. This is possible only when the resources required to produce continuously those goods that satisfy such wants are not scarce. Water and fresh air used to be regarded as free goods. This is no longer true in many parts of the world. These bounties of nature are required at different times, in different locations and in differing amounts. Even if there were enough air and water in the aggregate, resources may be required to obtain the right amounts in the right place at the right time. For example, in many parts of the world the rainfall is sufficient to provide for agricultural wants without human effort. In other parts, such conditions do not hold and water is often stored, piped and sprinkled in fields. In other words, water in such locations is not a free good: the use of the resources required to transport water to the right place at the right time in the right quantities is the price that must be paid for irrigation. Similarly, fresh air in the Alps is a free good; fresh air in Los Angeles often is not. To obtain fresh air in Los Angeles on a smoggy day requires the use of scarce resources to provide for air-conditioning. To condition the air of the entire city of Los Angeles is regarded, currently, as unacceptably expensive.

1.3 Preferences, Resources and Economic Efficiency

To satisfy wants as fully as possible, given limited resources, nations must make choices. No matter how a nation is organised, no matter which political philosophy is followed, the choices all nations face are common. The economic decisions to be made involving choices are:

(a) what goods and services to produce;
(b) how to produce the selected goods and services; and
(c) what share of these goods and services are to be given to each individual or household.

These are the 'what', 'how' and 'for whom' problems. Although common to all nations, they are tackled in a wide variety of ways.

Education provides a striking example involving the above three problems. How much education should be provided in a nation and of what should it consist? How should educational services be produced? Should there be more teachers and smaller classes, a greater use of videoconferencing and computer-assisted learning, or should students study distance-learning courses at home? Who should receive higher education – anyone who wishes, only the most intelligent members of society, or only those who are prepared to pay? In some countries, over 50 per cent of the college-age population participate in the higher-education process. In other nations the proportion of the college-age population going to universities is less than 1 per cent. Some countries do not have universities, and in some areas of the less developed countries no elementary schools exist.

To use a nation's resources as effectively as possible implies producing the goods and services that will satisfy the people's wants as fully as possible, which in turn implies taking account of the people's preferences. Thus to achieve the highest possible level of utility in a society from given resources means assigning priorities to the potential sets of goods and services that could be made available. Families do this continuously. In deciding how to spend the weekly budget, a household will try to choose that set of goods and services (from the many possible sets that could be purchased with that budget) that will provide the highest possible level of utility. The set it hopes to choose is the affordable combination of food, drink, clothing, recreation, etc., compared with all of the sets rejected, that is best suited to the preferences of the household.

If a given amount of goods were allocated randomly among a group of individuals, it might be possible to increase utility levels without increasing the amount of goods, through voluntary exchange. For example, during World War II prisoners in camps sometimes received Red Cross parcels containing tinned milk, jam, butter, tinned meat, chocolate and cigarettes. Each prisoner received an identical package but had his own individual preferences. Prisoners who did not smoke but liked chocolate could raise their utility levels by exchanging cigarettes for chocolate with prisoners who smoked but did not like chocolate. Both groups could become better off through exchange without any increase in the total amount of chocolate and cigarettes. We can see from this example that by exchanging a good with a low level of utility for a good with a high level of utility, an individual can increase his total utility level. The voluntary exchange is possible only if such a condition holds for at least two individuals, so both will benefit. A point will be reached ultimately where all possible exchanges that can increase utility will have been made. Then utility can be increased only if additional goods or resources become available.

This system of exchange is vital to the welfare of most countries in the world today. In the field of international trade, Scotland exports whisky to France, which in turn exports wine to Scotland. The bulk of the world is dependent on the oil-exporting nations to provide fuel for transport, heating and industrial needs. Russia, one of the world's major caviar producers and exporters, buys food from the USA, Canada and Europe. Such exchanges among nations increase people's utilities by accommodating differences in preferences.

It is in the interest of a nation to use as few resources as possible in the production of any one good. The fewer the resources employed in producing a good, the greater will be the resources available for the production of other goods and, consequently, the more fully a nation's wants can be satisfied. The term 'engineering efficiency' (or 'technical efficiency') describes a situation in which a good of stated quality is produced using the fewest possible resources. For building a bridge over a river, there is a minimum amount of steel necessary to ensure that the bridge will not collapse when fully utilised. However, it would be a complete waste of steel to keep adding support structures to a bridge in order for it to bear, say, several 100-ton trucks if no such vehicles will ever use the bridge. It would be a waste – technical inefficiency – since steel has many alternative uses.

The concept of *economic efficiency* is more complex than the concept of engineering efficiency. For economic efficiency to prevail in the allocation of a nation's resources, not only must the least amount of resources be used in the production of each good and service but that set of goods and services that is produced must be the set that satisfies wants as fully as possible. A country could build 100 bridges in a manner that was efficient from an engineering viewpoint but could be using resources poorly from the viewpoint of economic efficiency. This would be the case if the bridges, despite being engineering-efficient, yielded less satisfaction to the society than the other goods that could have been produced with the resources used to produce some or all of the bridges. If building 100 bridges, however, did constitute an economically efficient allocation of resources (think of Venice!) it would mean that not only was each bridge constructed in an engineering-efficient fashion but the alternative goods forgone would have yielded less satisfaction than 100 bridges.

1.4 Marginal Analysis and Opportunity Cost

One possible way to solve problems in which economic efficiency is the goal is to consider the potential welfare from every possible resources allocation and choose that allocation that yields the highest welfare. Consider Table 1.1, which shows the scores a student will achieve in a two-hour economics final examination for different allocations of time between objective questions (maximum score 33 points) and essay questions (maximum score 60 points). The limited resource is two hours of examination time, the goal is score maximisation, and the choices are different allocations of time to objective and essay questions.

Total score is at a maximum when 40 minutes are allocated to objective questions and the remaining 80 minutes allocated to essays. Since no other allocation yields a higher total, this is the most efficient use of the two hours of examination time.

Table 1.1 Examination scores

Objective questions		Essay questions		
Time spent (minutes)	Score	Time spent (minutes)	Score	Total score
120	33	0	0	33
100	32	20	23	55
80	29	40	38	67
60	25	60	48	73
40	20	80	55	75
20	11	100	58	69
0	0	120	60	60

The disadvantages of such an approach are that, first, it may not be feasible always to consider every possible allocation of resources and, second, the calculations could be very time-consuming. As the number of options considered increases, the number of possible permutations approaches infinity. An alternative way to find the optimal allocation of a limited resource is to start with a particular allocation and to determine how total utility or welfare changes as the allocation of resources changes. If total utility increases because of the reallocation, then the change should be made; if there is no change in the allocation of resources that will result in an increase of utility, then the optimal allocation already exists.

In Table 1.1 consider the allocation of time of 80 minutes on objective questions, yielding 29 marks, and the remaining 40 minutes on essay questions, yielding 38 marks. A decrease in time of 20 minutes on objective questions costs the student 4 marks on these questions, but 20 minutes more allocated to essays gains the student 10 marks. The net gain in marks (utility) of 6 (10 − 4) is positive, indicating that the change is worthwhile, i.e. the new allocation of 60 minutes to objective questions and 60 minutes to essay questions is superior to the former 80–40 allocation. But is it optimal? Consider a further reallocation of 20 minutes, i.e. 40 minutes on objective and 80 minutes on essay questions. The net gain is 2 marks (7 − 5). A further change of 20 minutes results in a decrease of 6 marks (−9 + 3). Thus 40 minutes to objective questions and 80 minutes to essays is the optimal time allocations. Making an incremental or small change in that allocation causes benefits and costs (gains and losses) to occur. The gain from the incremental change is called the marginal benefit. The loss from the change is called the marginal cost. Any action or activity in which the marginal benefits exceed the marginal costs will, if undertaken, result in an increase in total utility or well-being.

The technique of considering incremental changes in resource allocation and weighing the marginal benefits and marginal costs is known as *marginal analysis*. Marginal analysis, therefore, is a useful technique for solving problems involving optimisation. Because optimisation problems in economics contain constraints, e.g. income, time, labour or materials, they are called *constrained maximisation problems*. In the analysis of such problems, marginal analysis is used widely by economists, business people and other decision makers, and it can be important in avoiding costly and common mistakes.

In searching for a partner or mate, an individual uses marginal analysis. How? What are the chief characteristics of your ideal partner? He/she should be generous, sensitive, intelligent, healthy, industrious, good-looking, athletic, etc., etc. You start searching for the ideal mate; the longer your search, the greater is the probability you will find a more suitable partner, just as the more often you go fishing the higher the chance you will catch a world-record fish. But searching is not costless; it consumes one of your valuable resources – time. A point will be reached when you will decide that the benefit of additional searching (i.e. an even more suitable partner) is not worth the extra cost (i.e. the time required). Indeed, there will be many countries with a vast supply of potential partners that you will ignore. You will then turn to the 'best' partner you have found to date and, rather than explain the principle of marginal analysis, you will state you have fallen in love.

If you don't like that example, try a motorway. Table 1.2 contains hypothetical data on the benefits and costs of different lengths of a motorway extension.

Table 1.2 Extension of motorway

	1 mile	2 miles	3 miles	4 miles	5 miles
Total benefit ($m)	2	5	9	12	13
Total cost ($m)	1	2.5	4	6	8

One might conclude that the five-mile extension should be built because total benefits ($13 million) exceed total costs ($8 million) and this, therefore, must be an efficient way to allocate resources. The marginal analysis approach shows the error in the logic: for the first mile, total and marginal benefits are equal ($2 million), as are total and marginal costs ($1 million) and so, since marginal benefits exceed marginal costs, the first mile should be built. In adding a second mile, the marginal benefits are $3 million ($5 million − $2 million) and the corresponding costs are $1.5 million ($2.5 million − $1 million). Thus the second mile is worth adding as are, by similar reasoning, the third and fourth miles. However, for the fifth mile the marginal benefit is $1 million ($13 million − $12 million) and the marginal cost is $2 million ($8 million − $6 million). Since the marginal cost exceeds the marginal benefit, the activity – the building of the fifth mile – should not take place despite the fact that the total benefit of five additional miles of motorway exceeds the total cost. Only four miles of additional motorway should be built.

When considering the best use of resources the alternative goods or outputs that could have been produced must be considered, since in allocating a resource to any one activity all alternative uses of that specific resource are sacrificed. Two hours spent in playing tennis are no longer available for watching television, eating or studying. The real cost to society of any activity undertaken is the best alternative forgone. Economists call the best alternative forgone the *opportunity cost* of the activity undertaken. Imagine an economy with adequate natural water supplies that has resources to build an additional school or an additional hospital but not both. It decides to build the school since it concludes that the marginal benefits from a school exceed the marginal benefits from a hospital. The opportunity cost of the school is not the money required to finance the building for the school or the resources themselves; the opportunity cost of the school is the hospital – the best alternative forgone. Assume that the same resources could be used to build a reservoir. The reservoir is a distinctly inferior alternative to the hospital since the country has an ample water supply. Thus the opportunity cost of the school is not the reservoir because the reservoir would provide too low an assessment of the opportunity cost of building the school. Similarly, in the previous example of the allocation of time between objective and essay questions (*see* Table 1.1), the opportunity cost of time allocated to objective questions was time allocated to essay questions, not day-dreaming during the examination, which would have had low marginal benefits in terms of final score.

1.5 Different Economics Systems

There are three methods of solving the scarcity–choice problem: by tradition, by command, and by markets. All societies today utilise each of the methods; what distinguishes different nations is the relative weights or reliance placed on each of the methods.

In traditional economies, the set of goods and services to be produced and the methods of production typically are determined by social custom and habits established over time. The traditional method of resource allocation, common in feudal times, is the principal solution to the economic problem in many underdeveloped societies; it is also prevalent in certain rural and agricultural communities in developing nations. Under this system, most economic activity is concerned with eating, clothing and housing wants. Many of the societies are largely self-sufficient at a relatively low standard of living, with some trade taking place with the 'outside' world. Since there is a heavy reliance on nature for much of the output produced, changes in output are caused chiefly by the vagaries of weather. Too little or too much rain can significantly affect crops and, consequently, the standard of living. Methods of production in agriculture and fishing, for example, are often primitive, with specific tasks being assigned to different members of the community.

How the output is distributed within such societies is often determined by birthrights and other social factors, and changes in the social order occur infrequently and slowly. Examples of traditional organisation are found in the animal kingdom. The tundra wolves, to cite one case, normally lay claim to a certain territory within which they hunt caribou. Only certain wolves hunt and, within the hunting pack, individual wolves perform specific tasks such as herding the prey or killing it. The distribution of the kill is determined by the status of individual wolves. Some have first choice and others are detailed to carry meat to the pregnant females.

While not often recognised as a significant factor in developed nations, the traditional method does still exist, especially in economic organisation within the home. In the last century, economic organisation within the home has undergone significant change. The tradition of men being the primary wage-earner and wives doing the cooking, housekeeping and dressmaking has altered as participation of women in the labour force has approached that of men. Despite dramatic changes, however, the responsibility of weekly shopping and cooking in many homes still lies in the female domain. Many children are assigned certain tasks such as room tidying and setting and clearing the table. The willingness of children to 'volunteer for extra duties' has been noted to increase markedly, albeit briefly, when they pass their driving tests. The male in the home typically takes care of minor repairs and has the responsibility of making sure cars are functioning. In some households the female is in charge of the accounts, in others the male. The pattern varies significantly over households, but whatever the pattern much of the household economy functions along traditional organisational lines.

In the command solution to the scarcity–choice problem, the set of goods and services to be produced is determined by a central ruling body or bodies. Once the set has been determined, it is treated as a target. A plan, often of five years in duration, is drawn up to implement the choice decision. Drawing up such a plan for an economy is an extremely complex problem. Suppose 5000 tractors are assigned a high priority in the plan. The resource requirements for them of machine tools, steel, rubber, paint and labour have to be specified – as they do for all other goods and services in the plan – and matched against available resources. If resources appear insufficient, choices as to what should be omitted or reduced have to be made; if surplus resources exist, the plan has to be revised in the opposite direction. Errors in planning can be extremely costly: if sufficient spare parts for machine tools, for example, are not fully planned for, tractor output could be reduced, food output in turn could be less than planned, which in turn could have repercussions throughout the whole economy. Because of unforeseen elements, a degree of slack, or flexibility, must be built in to every plan. Some allowance is made normally for individuals and managers to adjust the deployment of resources; incentives are often provided to utilise resources more effectively and increase output above the level stipulated in the plan.

The problem of the distribution of goods and services is solved basically by stipulating wage and salary levels for different types of workers and income allowances for people not in the labour force, and also by fixing the prices of goods and services. When the plans 'go wrong', two things happen. First, since there is no compulsion for people to spend all of their incomes, some goods produced and available in the stores may not find buyers, i.e. such goods are not in accordance with consumers' wishes. The prices may be reduced; the goods may be withdrawn. Conversely, consumers may not be able to buy the goods they want because they may not be available in sufficient quantities at the fixed price to satisfy all the consumers' demands. Thus queues are not uncommon in planned economies. When information circulates that a store has a specific good that consumers want, 'first come, first served' becomes the distribution method superimposed upon the price mechanism. *Black markets*, many of which are prohibited in planned economies, form. Penalties can be levied on people dealing in such markets. In a similar vein, foreign travellers to certain planned economies are warned in advance of their trip that foreign-currency dealings with residents are prohibited and subject to fines, confiscation of property, and imprisonment.

In most planned economies significant economic activities, especially in agriculture, are conducted through free markets. Farmers are permitted to sell part of their output in local towns and villages, and are allowed to keep the income earned.

In the market (capitalist/free-enterprise) solution to the scarcity–choice problem, which sets of goods and services are produced is determined by the interaction of consumers who are willing to pay for them and producers who are willing to supply them in response to market prices. Consumers wishing to maximise their welfare bid for the goods that most satisfy their wants. Producers wishing to maximise their welfare (profits) produce the goods and services that they believe consumers most want.

Markets are the places where consumers and producers 'meet' and where prices are determined and exchange of goods and services takes place. Some markets are highly localised – babysitting, window cleaning and everyday shopping are examples. Other markets are international – oil, wheat and gold, for instance.

Competition for resources to produce goods and services determines the prices of these resources and, consequently, the distribution of income among resource owners in capitalist societies. Those people with resources in greatest demand will receive the largest incomes and consequently will be able to buy more goods and services than individuals with fewer resources or those who earn a lower return in resource markets. People with higher incomes will have more 'money votes' and consequently a greater say in the set of goods and services produced. Thus the capitalist system is one in which every money vote, not every person, is equal. It is one that permits concentration of income and wealth. It is one that rewards individuals not according to their needs but according to the value of the contribution of their resources to production. It is one in which individuals have freedom to own resources, to offer them or withhold them from the market, and to choose where such resources will be employed. Those individuals who have no resources that can command a return, or meagre resources that earn for them a low income, have to rely on the charity of upper-income groups or, through government action, a transfer of income to purchase the goods and services they want.

The capitalist system also has to rely on government or some other form of collective action to produce certain goods (such as national defence) that are unlikely to be forthcoming in the correct quantities through individual action. Similarly, collective action is necessary to deal with national or international problems such as pollution. Thus all economies today that fall in the capitalist camp still have significant sectors of government control or planning.

Although there are few, if any, examples of *purely* traditional, planned or market economies in the world today, societies or nations are so classified according to the type of solution predominantly used to allocate resources. A concomitant of how resources are allocated is resource ownership. In market economies, resources are owned by individuals or private groups of individuals, in command economies by the state, and in traditional economies by both the state, or collectively, and by individuals. Again, in the real world the sharp dividing lines are blurred. In all capitalist countries, the national government and state or local governments own resources; in planned economies some private property and resources are owned by individuals.

Furthermore, as time passes, the classification of a society does not necessarily remain fixed. It changes as the solutions to the problem of what to produce, how to produce it, and how to distribute production alter in response to developments in the overall economic, political and historical context within which the nation exists. For example, in the thirteenth century Britain had a feudal type of economic system operating largely in the traditional mode. The Industrial Revolution of the eighteenth century enabled Britain to become the 'workshop of the world' and to emerge as the first and most advanced of the market economies.

From the end of World War II up to 1978 a larger proportion of allocative decisions were made through the command mode as a result of the Labour Party policies. This trend was reversed, however, as a result of the 1979 and subsequent Conservative Party election victories. Many state-owned resources were privatised and market forces played an ever increasing role in resource allocation. There was no reversal of this role when the 1997 UK general election witnessed the return of the Labour Party. Interestingly, however, the Labour government's attempts to market everyone 'where equal' had efficiency impacts on industries controlled by central authorities, especially the National Health Service and Education. The decline in quality led to many higher-income individuals abandoning state provision of medical service and education for the private sector. The Conservative–Liberal coalition government elected in 2010 has pursued a policy of cutting back on public spending, the outcome of which is a vastly reduced role for government in the allocation of resources.

The last three decades also have witnessed a greater market element operating in the command economies of China, Russia, other Eastern Bloc countries, and many of the emerging nations of the world, this trend culminating in the break-up of the communist trading bloc and the commitment of the old command economies to the market system.

1.6 Production Possibilities Curve

A useful tool for capturing many of the issues in this module is the *production possibilities curve*, which forms a production *frontier*.

Imagine a two-good (guns and butter) economy with a given supply of resources. Assuming all resources are employed in an engineering-efficient fashion, there will be an upper limit to the amounts of guns and butter that can be produced. Suppose Table 1.3 and Figure 1.1 represent the upper limit. If all resources are allocated to butter production, the economy's output will be 210 tons of butter and no guns. Conversely, if all resources are allocated to gun production, the output will be 175 guns and zero tons of butter. Clearly, it will not be in society's interest to produce any output within the production frontier, for example at point z in Figure 1.1, either through adopting inefficient production techniques or through having some resources unemployed. The reason is simple: more output could be produced and consequently wants could be more fully satisfied. Points x or y or any intermediate point on the frontier are preferable to point z because at least as much of each good will be available compared with the 80 guns/144 tons of butter at point z. More generally, for any point *within* the frontier there is at least one point on the frontier that would be preferable.

Table 1.3 Production possibilities

Guns	Butter (tons)
0	210
20	207
40	198
60	186
80	168
100	144
120	119
140	82
160	42
175	0

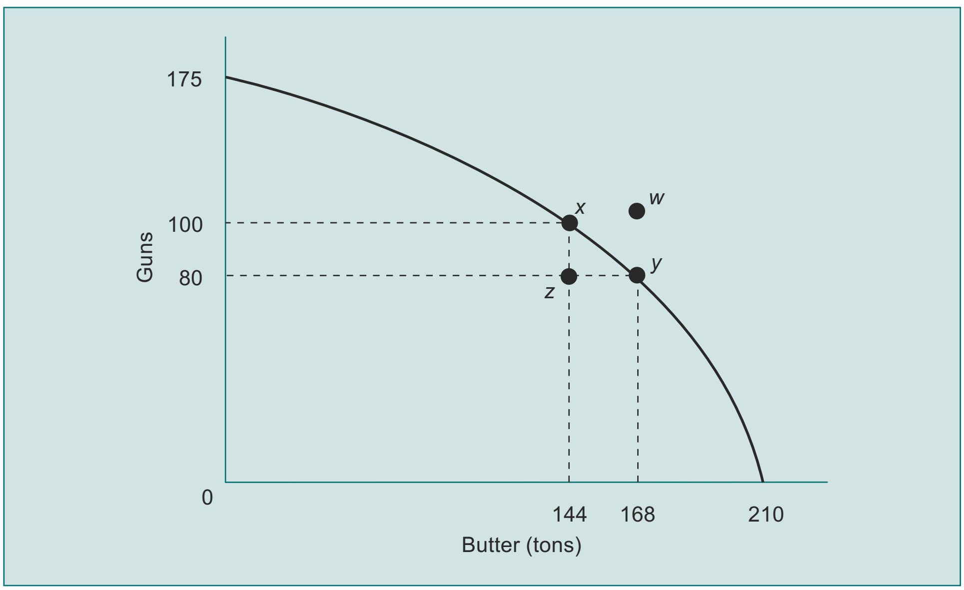


Figure 1.1 Production possibilities curve

Where should society produce on the frontier – at x or at y or at some other point? If the economy starts off at point x, the production of more butter as would occur at point y involves sacrificing some guns. In moving from x to y butter output increases from 144 to 168 tons, a gain of 24 tons, at the expense of 20 guns (100 down to 80). In other words, when the economy is at x the opportunity cost of 24 tons of butter is the best alternative forgone, namely 20 guns. Is such a trade-off worthwhile? The answer depends on whether society values the 24 additional tons of butter more than the 20 guns sacrificed. This is therefore a problem involving marginal analysis: the marginal cost of the move from x to y is 20 guns; the marginal benefit is 24 tons of butter. If the marginal benefit exceeds the marginal cost in society's eyes, the change is worthwhile. Note that the opportunity cost of butter is not constant but varies over the frontier. The more butter there is, the higher is its opportunity cost.

It can be seen from Figure 1.1 that point w, which is preferable to points x or y, is unattainable because insufficient resources or technology exist to make such an output possible. The same figure can be used to illustrate another economic issue, namely *the sacrificing of present consumption for future consumption*.

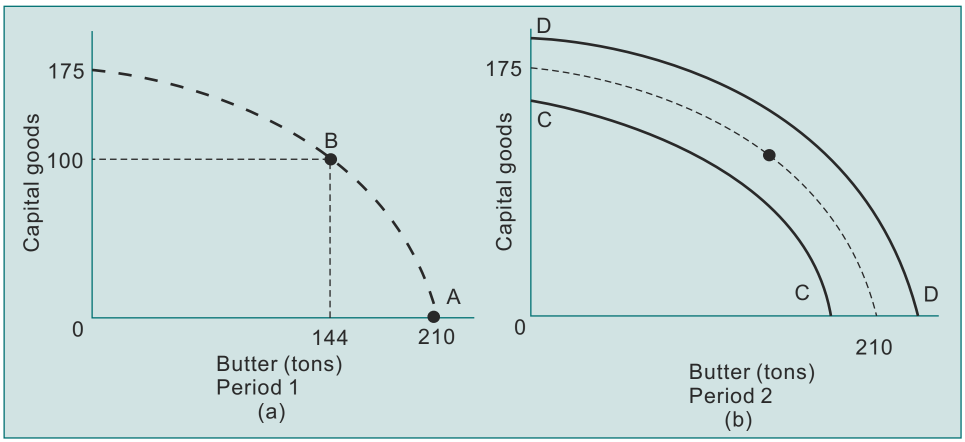


Figure 1.2 Production possibilities over two time periods

In time Period 1 (*see* Figure 1.2a), a new good is introduced – capital goods, consisting of, say, tractors and machine tools. No guns are produced in Period 1. Point A represents no capital goods being produced, only 210 tons of butter. Point B represents 144 tons of butter and 100 capital goods. Is the movement from A to B worthwhile? For the answer, consider Period 2's production possibilities curves (*see* Figure 1.2b). Curve CC represents the situation in which no capital goods were produced in the previous period. Indeed CC represents less output than the curve in Period 1 because some resources depreciated in the production process and were not replenished. Curve DD represents the production possibilities in Period 2 given that point B was chosen in Period 1, i.e. given that 66 tons of butter (210 − 144) were sacrificed in Period 1 to produce 100 machine tools. Whether such a sacrifice is worthwhile will depend upon how society evaluates the loss of 66 tons of butter in Period 1 against the enhanced production possibilities in Period 2.

The measurement of the respective marginal costs and marginal benefits may not be simple. They may involve value judgements, and certain issues are liable to be keenly debated and sometimes even fought over because the burden of the costs that emerge from the choice made may be borne disproportionately by different groups within the society. For example, a family deprived of butter if option B is chosen in Period 1 may see their children die, and that family could argue that such a cost was not worth the benefit that would be enjoyed by others in Period 2.

Economists are no better qualified than doctors, rock stars or students in making such value judgements – these are normative issues. What economists can do is present the options available with their concomitant benefits and costs – these are positive issues. Individuals and society as a whole must make the decisions of their choice.

Learning Summary

This module has taught you that it is the dual existence of insatiable wants and limited resources that produces for all societies the fundamental fact of scarcity, and consequently the need for choice. Resources used to satisfy one want are no longer available to satisfy others. The decisions to be faced by all societies are the 'what', 'how' and 'for whom' decisions.

A useful tool for analysing problems involving maximisation is marginal analysis – a technique that pervades all aspects of economics so that it is important that you master the concept at this stage.

You are now aware that not all societies solve the 'what', 'how' and 'for whom' problems in the same way and you know the differences among them.

One useful tool for analysing output decisions facing an economy at a point in time and also over time is the production possibilities curve. You are now reasonably familiar with this tool – you will meet it again.

Review Questions

Multiple Choice Questions

1.1 Which of the following is correct?

The concept of scarcity means that resources:

A. are not available in sufficient quantities to satisfy any individual's wants.
B. are not available in sufficient quantities to satisfy all wants for them.
C. cannot be increased in quantity to any significant extent.
D. are of primary importance in satisfying the wants of society.

1.2 It has been said that the fundamental fact of scarcity is no longer applicable to the US economy. Which of the following is correct?

This statement is:

A. true, because the US is one of the richest countries in the world.
B. true, because the resources that are scarce in the US can be imported from abroad.
C. false, because not all the wants of all US citizens are fully satisfied.
D. false, because a significant number of families in the US have incomes below what is known as the poverty level.

1.3 Examples of free goods are

I. a free tyre with every four purchased.
II. a crystal drinking glass with every 20 litres of petrol.
III. fresh air in an air-conditioned building in Oxford Street.

Which of the following is correct?

A. II only.
B. II and III only.
C. I, II and III.
D. Not I, not II and not III.

1.4 Which of the following is correct?

The fact that individuals' tastes and preferences differ implies that:

A. no two individuals ever buy the same quantities of any given good in a given time period.
B. free exchange of goods and services can increase society's welfare without an increase in the quantities of goods and services available.
C. an individual's wants can never be fully satisfied until all of a nation's resources are efficiently utilised.
D. a nation's wants can never be fully satisfied.

1.5 A given quantity of goods can be produced in a variety of ways using different amounts of two resources, A and B.

Which of the following statements is correct regarding engineering efficiency?

- Method 1 uses 5 units of A and 10 units of B.
- Method 2 uses 10 units of A and 4 units of B.
- Method 3 uses 6 units of A and 4 units of B.

A. Method 2 is more efficient than method 1.
B. Method 3 is more efficient than method 2.
C. Method 3 is more efficient than method 2 and method 1.
D. Insufficient information exists to determine whether one method is more efficient than any other.

1.6 An economically efficient society that is capable of producing more goods and services in year 2 than in year 1 must

I. have access to superior productive techniques in year 2.
II. have more resources available in year 2.

Which of the following is correct?

A. I only.
B. II only.
C. I or II or both I and II.
D. Neither I nor II.

1.7 A government has completed a cost–benefit study showing that the annual value of the services from an additional 100 miles of motorway would be £4 million and the annual value of the services from an additional airport would be £3 million. To achieve the most efficient use of resources, what should the government do?

A. Construct only 75 miles of motorway (that is three-quarters of 100).
B. Construct only the 100 miles of motorway.
C. Construct both the 100 miles of motorway and the airport.
D. Not necessarily construct either any additional motorway or the airport.

1.8 'The problem facing the government is whether to build a new motorway system or to improve public transport throughout the country during the next three years. Resources for both projects are not available. It must be one or the other.' Which of the following is correct?

The opportunity cost of the new motorway system mentioned in the preceding paragraph is:

A. greater than the economy can afford.
B. an improved public transport system.
C. the money required to pay for it.
D. the resources required to build it.

1.9 The following table shows the extra daily benefit and costs that would be derived from adding successive terminals at an airport in order to increase passenger-handling capacity.

	Additional capacity (passengers)	**Additional benefits (£)**	**Additional costs (£)**
First terminal	1 000	45 000	40 000
Second terminal	1 000	35 000	25 000
Third terminal	1 000	25 000	20 000
Fourth terminal	1 000	15 000	18 000

How large an expansion programme should the airport undertake to maximise net benefit (total benefit minus total cost) and why?

A. Build one terminal since that provides the greatest marginal benefit.
B. Build two terminals since that is the level of expansion where the difference between marginal benefit and marginal cost is greatest.
C. Build three terminals since that is the largest number of new terminals for which marginal benefit exceeds marginal cost.
D. Build four terminals since the programme would involve the largest total benefit and the lowest marginal cost.

1.10 Consider the following statements:

I. For centrally planned economies, since all members of the labour force are allocated jobs, scarcity is not a problem.
II. For capitalist economies, since workers occasionally become unemployed, scarcity is not always a problem.
III. For both centrally planned and market economies, since resources are insufficient to satisfy all wants, scarcity is a problem.

Which of the following is correct?

A. I only.
B. II and III.
C. III only.
D. Not I, not II and not III.

1.11 In the present programme of a National Health Service (medical services provided by the government at zero cost to patients), the total benefits exceed total cost. The government's goal is efficient allocation of resources. Which of the following is correct?

It should undertake an increase in expenditure on medical services only if:

A. the total benefit derived from all medical services would still exceed the total cost.
B. the total benefit derived from all medical services would exceed that of any other good.
C. the benefit from the extra medical services would outweigh the cost of doing with less of other goods.
D. such expenditures improve people's health independently of the costs.

1.12 Consider the following statements:

I. The opportunity cost of a unit of land is the value it would create when put to its best alternative use.

II. The opportunity cost of a good is the value of all other goods that must be forgone in order to produce it.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

Questions 1.13 and 1.14 are based on Figure 1.3.

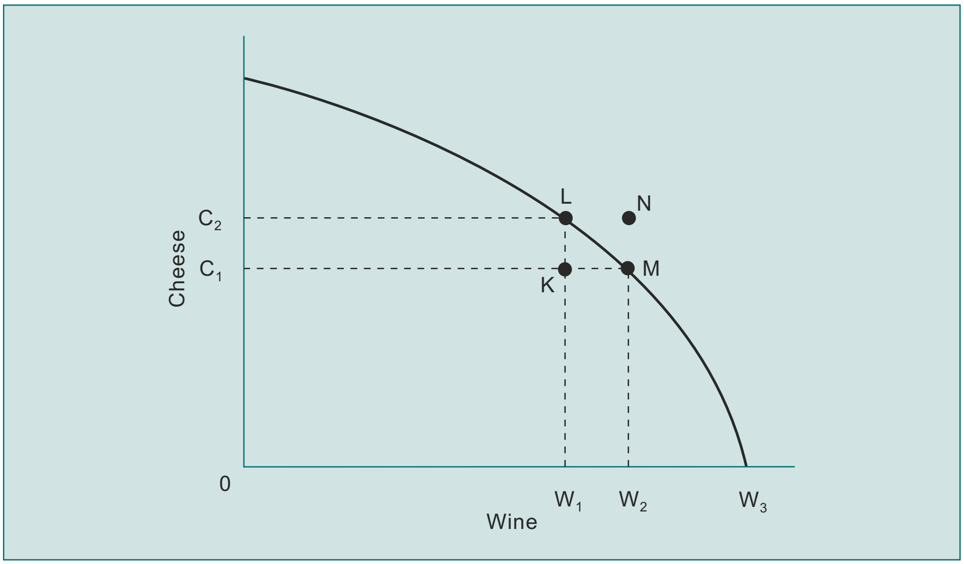


Figure 1.3 Production possibilities curve

1.13 Which of the following is correct?

I. The opportunity cost of cheese when production is taking place at point L in Figure 1.3 is W_1W_2 of wine.

II. At point M in Figure 1.3 a bottle of wine is of equal value to 1lb of cheese.

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

1.14 Of the outputs K, L and N in Figure 1.3, which of the following statements is correct?

A. Only L is attainable.

B. K and L are attainable.

C. Only K is attainable.

D. K, L and N are attainable.

1.15 Which of the following is correct? To satisfy wants as fully as possible, a society must

A. choose the best set of goods and services from all feasible sets.
B. produce various sets of goods and services then choose one set.
C. ensure that free goods are available to all members of society.
D. ensure that luxury goods are produced only when necessities are freely available to all.

1.16 Two individuals who both had baskets containing oranges and apples were observed exchanging voluntarily three apples for two oranges. Which of the following are correct?

A. There were more apples than oranges before the exchange in one individual's basket.
B. Both individuals benefited from the exchange.
C. One individual benefited more than the other from the exchange.
D. One individual prefers oranges to apples and the other prefers apples to oranges.

1.17 Which of the following is correct? An economy is operating in an economically efficient manner when

I. goods and services are being produced using the least amount of resources.
II. no individual's utility could be increased.
III. society's wants are being satisfied as fully as possible.

A. I and II only.
B. I and III only.
C. II and III only.
D. I, II and III.

1.18 Which of the following statements is correct?

I. What, how and for whom goods and services are produced are problems common to traditional, command and market economies.
II. In traditional, command and market economies all land is owned privately but other resources may be privately and/or collectively owned.
III. In traditional and command economies each individual receives the same amount of goods and services, whereas in market economies individuals receive an amount of goods and services equivalent to the resources each owns.

A. I only.
B. I and II only.
C. I and III only.
D. I, II and III.

Case Study 1.1: Marginal Analysis and Noise Pollution

This case is an application of the concept of marginal analysis. Before you tackle it you should understand:

a. total benefit;
b. total cost;
c. marginal benefit;
d. marginal cost; and
e. the rule for efficient resources allocation.

The case shows that the application of the basic economic concept of marginal analysis can demonstrate the error in an apparently common-sense conclusion.

How to Deal with Noise Pollution

There are certain sounds or noises that give some people pleasure – the disco in full swing, or the grand opera chorus. There are other noises, equally loud, that cause displeasure – motorcycles without silencers, and jet aircraft. The traffic on many motorways passing through built-up areas causes noise levels that result in residents' complaints. Often action is taken to reduce such noise to tolerable levels.

How can marginal analysis help a society make sensible decisions in reducing noise pollution? Noise can be measured on a decibel (dB(A)) scale. For noise levels below 68 on the scale, people seldom complain, whereas permanent ear damage can result at levels above 90 decibels if an individual is exposed to such levels over a period of time. (You might think about this if you are a rock music fan, and you might also note the earmuffs worn by ground personnel at airports.)

Suppose that the residents of a housing estate decide to investigate reducing the noise level from the motorway passing near their housing estate. The average noise level is 85 on the dB(A) scale. The options shown in the table below are available to the residents to reduce motorway noise. They may choose one or more.

Options		Marginal change in noise level (db(A))	Cost ($)*
1 (a)	Build concrete wall 2 metres high	−10	160 000
1 (b)	Build concrete wall 3 metres high	−12	190 000
1 (c)	Build concrete wall 4 metres high	−13	230 000
2	Reduce speed limit by 10 mph and enforce new limit through police control	−3	90 000
3	Resurface Road	−2	70 000
4	Rebuild the road 4 metres below its existing level	−10	2 500 000

*Total cost divided by estimated number of years of useful life.

The residents decide that they are prepared to pay, collectively, $31 000 annually for each dB(A) that noise is reduced until a level of 68 is reached, and nothing thereafter for further noise reduction.

The local council, conducting its own investigation, proposes that a 4-metre wall be built (Option 1(c)) but that no other option be considered. Its reasoning is that the annual total cost of the 4-metre wall would be $230 000 and the annual total benefit $403 000, and it points out that no other option yields such a large total benefit.

1 In light of all this information, assess the following questions.

1. What is the difference in monetary terms between the total benefit and the total cost of each option?
2. Which options should be excluded on the basis that total cost is higher than total benefits?
3. For which options does marginal benefit exceed marginal cost?
4. Which options should be chosen to reduce noise pollution in an economically efficient manner?
5. What is the mistake the Council made in arriving at its decision?
6. Does the economic solution to the problem mean that the residents will suffer no noise pollution?

Module 2

An Overview of Economics

Contents

Because of the broad nature of this module's contents, it has no review questions or case studies.

2.1 Introduction

Economics can be conveniently subdivided into two main branches of study: microeconomics and macroeconomics. Because it is a complex and difficult subject, we shall begin by making simplifying assumptions about the world in which we live and gradually relax those assumptions as our knowledge of fundamental concepts and ideas grows. The purpose of this module is to provide an overview of economics.

'Micro' is derived from the Greek work *mikros*, meaning small, and 'macro' from *makros*, meaning large. Microeconomics is concerned with the behaviour of particular items: e.g. the price of personal computers relative to the price of horror comics, the salaries of lawyers relative to the wages of farm labourers, the employment of resources in the car industry compared with the shipping industry, the behaviour of households in spending their incomes, the behaviour of business firms in hiring resources to produce different goods and services, and the role of government in producing certain goods and services. Macroeconomics is concerned with the big picture, with aggregates in the economy: e.g. the level of a nation's output, the national unemployment rate, the inflation rate. In addition, macroeconomics is concerned with government policy making, the use of taxes and transfers, the level of government expenditure, and control of the money supply.

International economics also has micro and macro components. The micro portion is concerned with exports and imports of goods and services, tariffs, quotas, and exchange rates. The macro portion is concerned with the *balance of trade* (i.e. all exports and imports summed), the *balance of payments* (i.e. the balance of trade plus capital flows), and policy making to affect exchange rates.

The basic theory of microeconomics is the theory of demand and supply. The interaction of demand and supply determines the price of a commodity and the quantity bought and sold. Resource allocation occurs in response to price signals. The basic theory of macroeconomics is the theory of the circular flow of national income. How well off a nation can be depends upon the nation's potential to produce goods and services, which in turn depends upon the quality and quantity of its stock of natural and man-made resources and the quality and quantity of its labour force. How well off a nation is depends upon how fully it utilises its potential.

International trade allows a nation to exchange the goods and services it produces for goods and services produced by other nations. International flows of goods, services and resources allow world output to be larger than it otherwise would be.

2.2 Strengths and Weaknesses of the Capitalist System

Every nation faces the same problem: using its resources to produce those goods and services that will maximise its well-being or utility. Every nation has a capital stock. It consists of nature's endowment (arable land, forests, rivers, mineral deposits) and man-made resources, our inheritance from past generations (roads, buildings, factories, machine tools). Every nation also has a labour force – that proportion of the population that is able and willing to work at going wage rates and working conditions (*see* Figure 2.1).

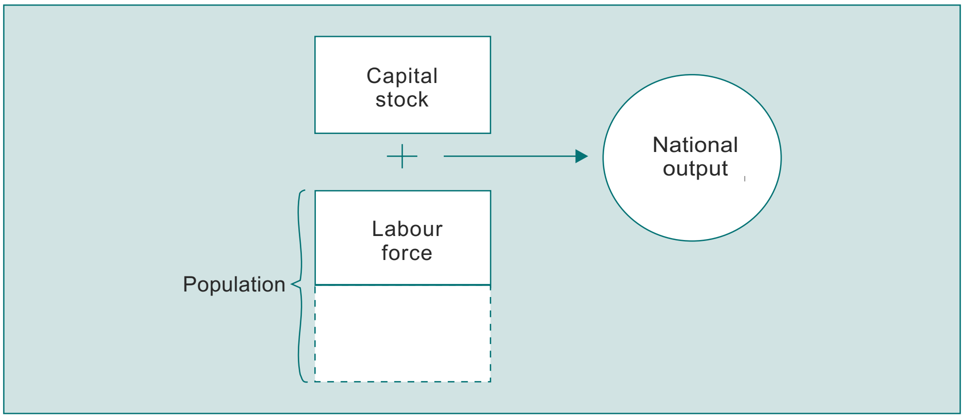


Figure 2.1 Resources and output

Part of that labour force consists of entrepreneurs and managers whose job it is to combine capital and labour to produce the goods and services the people want most. How large that flow of goods and services will be depends upon the quality and quantity of the capital stock and the labour force and how well these are deployed. In market economies it is the price mechanism that provides the signals that indicate how these resources should be deployed to produce the goods and services that will maximise society's utility given the resource constraint. How does the system operate?

To understand the underlying principles let us assume we have something that does not exist anywhere in the world, namely a pure market economy in which all resources are owned by households (consumers) and all production is carried out by firms – also, of course, owned by households. There is no government sector, no taxes and transfers, and no international trade. In Figure 2.2 some consumers form entities called firms. Those firms hire resources owned by the consumers – land and labour – and build factories and have other firms build machine tools. These firms then use those resources to produce goods and services, which flow back to the consumers.

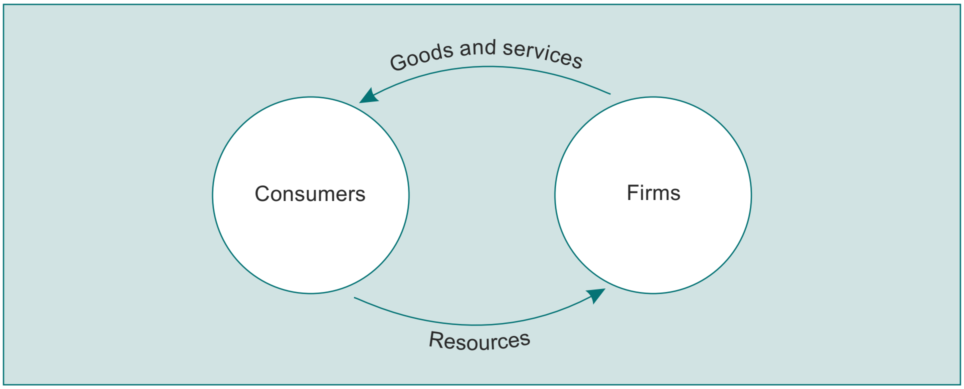


Figure 2.2 A simple economy

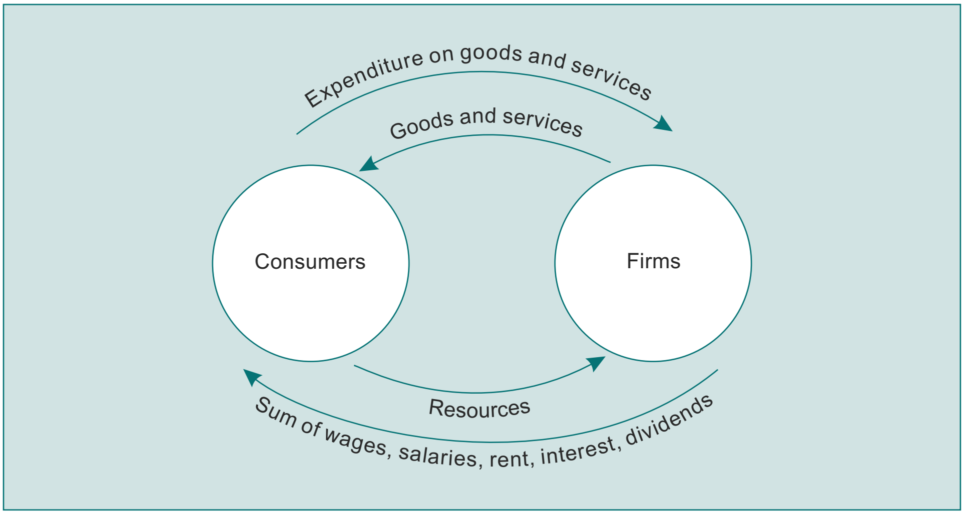


Figure 2.3 A simple economy with money

To make the situation more realistic and workable, we introduce money in Figure 2.3. The money could be gold or IOUs or dollar bills – it does not really matter so long as people are prepared to accept it as a medium of exchange. The firms still hire resources from the consumers but now pay them for the use of these resources. They pay wages and salaries for the use of labour, rent for the use of land, and interest and dividends for the use of capital. The consumers in turn pay for the goods and services produced by the firms. They pay for these goods and services with the money they have received from hiring out their resources.

Let's imagine that you decide to start a firm from the savings you have accumulated by hard work. You might also borrow money from neighbours with a promise to pay interest to start your firm, or you may make your neighbours part-owners of the firm with a promise to pay dividends from future profits. You decide to manufacture wheelbarrows with square wheels. You build a factory and buy machine tools and raw materials with your savings, hire labour and start producing your wheelbarrows. After a week or so you start to worry; you are producing wheelbarrows, they are piling up in inventory but no one wants to buy them – why? Because they are useless: they cannot be pushed, and when full cannot be pulled. Your chief engineer at the first managers' meeting suggests that the product could be significantly improved if round wheels were substituted for square wheels. You concur. Now will you sell your wheelbarrows? That will depend upon the quality of your wheelbarrows and the price of your wheelbarrows compared with the quality and price of wheelbarrows produced by other firms, and also on the demand for wheelbarrows in the market. Thus the first painful lesson of the market economy is driven home: if you produce goods and services that no one wants, you will go out of business because there will be a limit to your ability to keep spending money on resources producing goods that generate no revenues. The second painful lesson is that even if there is a demand for the product you are producing, your product had better be at least as good as the same product produced by your competitors and priced no higher than theirs.

Let us change the scenario. Let us imagine that you invent the first wheelbarrow and your prototype has a round wheel. Orders come flooding in faster than you can cope with, and consumers anxious to obtain this marvellous machine keep offering higher prices than you initially advertised. Each week you sell to the highest bidders and your profit per wheelbarrow – the difference between cost of production and selling price – keeps increasing. What will happen is not very surprising. Other resource owners, observing your success, move into the wheelbarrow business. They too build factories, hire resources and start producing wheelbarrows, hoping to share in the large profits in this industry. As more and more wheelbarrows are produced, consumers discover that they do not have to bid as high as previously to obtain a wheelbarrow. Thus the price of wheelbarrows starts to fall. When will it stop falling? It will stop falling when resources stop moving into the industry. When will they do that? They will do that when the returns that could be earned in other industries are higher. These returns include payments to the entrepreneurs who are attempting to ensure the resources are deployed optimally. Then the wheelbarrow industry will settle down with the 'correct' number of firms satisfying the demand for wheelbarrows. There will be a '*going price*' for a standard wheelbarrow, and each producing company will have to be as efficient as its competitors just to ensure it stays in business.

What is happening in the wheelbarrow industry in our example is happening in all other industries simultaneously. Resources are constantly moving from industry to industry as consumer tastes and preferences change, as technology changes, and as new industries emerge and old industries die out. Resource owners are motivated to seek the highest returns from their resources since this strategy, if successful, will yield the highest income. Likewise entrepreneurs or businesses are motivated to hire resources at the lowest prices, combine them as efficiently as possible to produce the goods and services people want, and sell them at the highest prices possible since this will yield the maximum return on their efforts. Consumers with the incomes earned from hiring out their resources (for many, their labour services only) will seek out the lowest prices for the goods and services they want. The less they have to pay for any given good means the more they have to spend on all other goods and services and consequently the higher the real standard of living they will enjoy.

We shall explore in detail in later modules precisely how prices are determined in goods markets and resource markets. For the purposes of this overview of economics, it is sufficient to realise the crucial role that prices play in the process of the allocation of resources. Prices tell consumers how far their budgets can stretch and consequently how well off they are. Prices tell business firms how much revenue they will receive from different levels of sales. But prices of resources are also business firms' costs and consequently, for a given level of production, inform businesses what their total costs will be. Putting expected revenues and expected costs together determines expected profits and indicates whether such a business venture is potentially worthwhile.

All of these prices mentioned, whether they be of goods and services or resources, are determined in a market economy by the forces of competition – consumers competing amongst themselves to buy and businesses competing amongst themselves to sell. Those prices ultimately determine how resources are allocated in a market economy to produce the goods and services consumers want most. In a perfect world the basket of goods and services produced is the best basket; it contains what people have voted for with their dollar votes; the goods and services in the basket are all produced using the least amount of resources possible; the prices of these goods and services reflect the value to society of the scarce resources used in their production. Unfortunately, the world is not perfect and, even if it were, the price mechanism would still not produce the ideal basket of goods and services. The first question is why? The second question is can we 'interfere' with the price mechanism in a market economy and come closer to producing this optimal basket of goods and services?

Before we tackle those questions, it is worthwhile observing some facts about the world in which we live. Reliance on market forces in the world today is growing as many socialist government economies are attempting to switch from a group of planners deciding what goods and services to produce to the market mechanism making such decisions. Even within market economies, socialist parties are rapidly embracing the merit results of the market system. The reason is simple: the market economy may not be perfect but, given the apparent basic nature of mankind, namely a 'selfish' motivation to improve his lot in life, the market economy currently appears to offer the best chance of converting such a desire into reality.

In the immediately post-World War II period (1950s), however, the market economy appeared to have more opponents than supporters. Socialist parties were voted into power so as to correct the evils and inequities caused by the free-market 'jungle', and many resource-allocation decisions formerly undertaken by the market became the province of governments or central planners. What were, and still are, the criticisms of the market economy?

2.2.1 Criticisms of the Market Economy

Criticisms of the market economy can be summarised under five headings, as follows.

(a) **Wrong goods and services.** 'You don't have to be an economist to realise the capitalist system produces the wrong bundle of goods and services. There are plenty of flick knives, horror comics and pornographic magazines available in the economy. There is a shortage of hospitals, schools and clean streets.'

(b) **The fallacy of consumer sovereignty.** 'Economists of a free-market economy may tell you that consumers voting with their dollar bills ultimately determine how resources are allocated in order to produce the most desired goods and services, but this is a myth. It is the power of advertising agencies, hired by companies, that persuades consumers to buy goods and services they do not really want.'

(c) **The pollution problem.** 'If the capitalist system is so great, how come the free-market economies are the world's biggest polluters – noise, air and water!'

(d) **Poverty amongst plenty.** 'Within the USA, if the market economy is one of mankind's greatest inventions, how is it possible for such an economy to laud the merits of the price system when 35 million people in its population live in poverty?'

(e) **Inflation and unemployment.** 'Economists used to tell us that we had a choice: the economy could run at full employment but we would suffer from inflation, or we could have a zero inflation rate but that would necessitate some unemployment. What we have witnessed in the past is the worst of both worlds – inflation plus unemployment, i.e. stagflation. Will stagflation re-emerge?'

These are the major attacks. Are they supported empirically?

(a) Flick knives, horror comics and pornographic literature are not difficult to find; they are available in most cities of the capitalist world. Is there a shortage of hospitals, good schools and clean streets? The answer is 'yes'. In many countries there are queues for medical operations, university and college places are limited, and many streets are dirty and ill-paved.

(b) Is 2 per cent of total national output spent on advertising excessive? That is the level of advertising for several capitalist economies, the bulk of which is spent on newspapers, magazines, direct mail and at the point of sale, rather than on TV and radio. Some advertising is obviously useful: it provides information about the characteristics of goods, including price, and also informs us about new products – for example, buying 'worthless' breakfast cereal to keep the children happy because each package contains a plastic green turtle, or being aware of the insignificant differences between pure (generic) aspirin and the many branded aspirins compared with the price differences. In contrast to these instances, companies have gone bankrupt despite massive advertising campaigns to persuade consumers to buy their products.

(c) Pollution is undoubtedly a current major concern in the world's capitalist economies. It is not, however, a pressing concern to the Brazilian farmer engaged in cutting down a couple of acres of rain forest so that he can grow sufficient crops to keep his family alive. The world's major polluters are the developed economies. Los Angeles, in the USA, has a population of c.15 million people and an accompanying population of 8 million cars and trucks – the major cause of pollution.

    China's industrialisation is fuelled primarily by coal. This practice causes major pollution problems in industrialised cities, and China is criticised by mature capitalist nations that no longer rely on coal as the major energy source but that relied heavily on coal in their early industrial growth period.

    Furthermore, many beaches throughout the world are health hazards, fish cannot breed in many of our rivers, and acid rain is blamed for the destruction of forests. We also have noise pollution in factories, at airports, and in towns and cities. We are warned continually by 'green parties' and scientists of the greenhouse effect. Pollution exists.

(d) While the poverty levels in capitalist countries are 'insignificant' compared with poverty levels in countries such as Bangladesh and the Sudan, where mass starvation is the normal state, there is no doubt that many people in capitalist economies live below normally accepted definitions of the poverty line. Many of these people are homeless, ill-clad, ill-fed and in poor health, unable to find jobs and dependent upon the state or charities for a meagre income.

(e) In the mid-1960s, governments and economists congratulated themselves on the success of their policies because many economies achieved full employment accompanied by very little inflation. However, in the 1970s and 1980s we observed and experienced in many capitalist countries stagflation, that is, high unemployment and simultaneously high inflation. The 1990s and early 2000s saw healthy economic growth, large increases in government expenditure and rapid rises in house prices. The financial crisis of 2008 brought economic growth to an abrupt halt. Banks failed, and many had to be rescued by governments; bankruptcies were rife, and the housing bubble collapsed.

By becoming involved in the wars in Iraq and Afghanistan, the USA and Britain, and to a lesser extent other United Nations members, allocated significant resources to the war effort. This additional expenditure, when added to the growing government expenditures in many other areas, meant burgeoning budget deficits. Unrest in the Middle East and the overthrow of autocratic rulers left some nations without recognised governments or rulers. The resulting chaos interrupted oil supplies, which led to rapidly rising oil prices and inflation in a world of economic depression.

By 2010 in many capitalist countries, despite rising unemployment, rising prices and growing pessimism about the economic future, governments elected the growing budget deficit as public enemy number one. As you will learn in greater detail in the macroeconomic portion of the course, to cure a budget deficit (i.e. where government expenditure exceeds government income) expenditure has to be reduced and/or income has to be increased. Lower government expenditure and higher taxes, however, will not help the unemployment problem – they will make it worse. In addition the policy shift by capitalist countries was occurring when exogenous shocks (outside world events) were not helping the world situation. Tsunamis, the nuclear accident in Japan, the worst drought in 50 years, earthquakes, hurricanes, landslides and volcanoes caused many deaths but also led to severe shortages of foodstuffs in both the short and the longer run.

Some of the problems above are not man-made, but there is little doubt that empirical evidence supports some criticisms of the market economy. What can be done about those weaknesses? People on the political 'left' maintain the answer is relatively simple: have the government solve the problems where the price mechanism fails by:

(a) banning 'bad' goods and services and providing 'good' goods and services, which the market mechanism does not;
(b) establishing advertising standards to provide useful information to households and to prohibit misleading and persuasive adverts;
(c) setting and enforcing pollution standards;
(d) establishing poverty levels and providing goods and services to the poor so that every household enjoys a reasonable living standard; and
(e) controlling wages and prices in times of inflation, and becoming an employer of the last resort, i.e. creating jobs through government expenditure when unemployment threatens.

People on the political 'right' also maintain that the solution is simple: problems exist in the world's market economies because there is too much government in these economies. If greater reliance were placed on market forces, the bulk of the problems or apparent weaknesses would disappear.

Who is correct? The answer is that there are some areas of economic activity where reliance on the market mechanism will not provide an efficient allocation of resources and consequently will prevent a nation enjoying the optimal basket of goods and services. On the other hand, significant inefficiency and incentive problems arise when government controls areas of economic activity. Before you decide where you lie on the political spectrum of the 'correct amount of government interference' in the market economy, let us consider the areas of market failure or partial failure.

2.2.2 Public Goods

Before we discuss public goods, let us consider private goods. A *private good* is one that you buy and consume and for which the act of consumption affects no one else. For example, consider a good bottle of malt whisky retailing for $50. In making the decision to buy and consume whisky, you will apply marginal analysis, i.e. is the extra utility you receive from consuming the malt greater than the utility you would receive from allocating that $50 to other goods and services? If the answer is 'yes', you buy the bottle. In attempting to make such a purchase, all you have to do is find someone who has a bottle of good malt whisky but who would prefer to exchange it for $50. Such people are not hard to find – you find them in liquor stores. Both you and the liquor store owner benefit from the exchange of one bottle for $50; if this were not the case, no exchange would take place.

A *public good* is different. If you buy a public good, other people can consume and enjoy it without paying for it. If someone else buys it, you can consume it without paying for it. This causes a problem as far as the efficient allocation of economic resources is concerned.

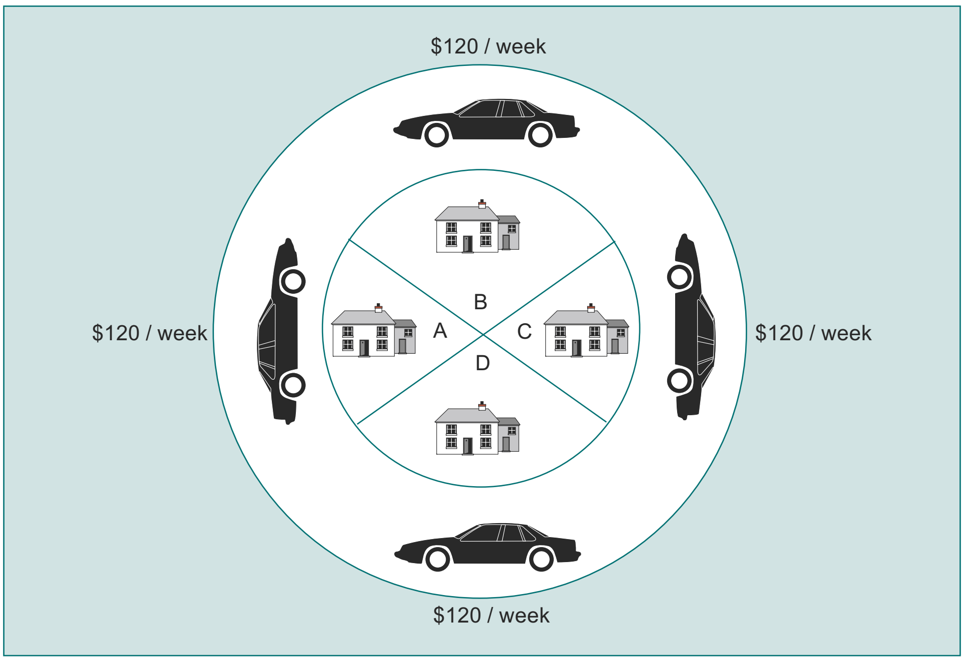


Figure 2.4 An exclusive housing estate

Consider the housing estate depicted in Figure 2.4. A, B, C and D live in an exclusive housing community. Independently, each hears that a series of robberies has occurred in neighbouring estates. Again, independently, each considers whether or not it is worthwhile to hire a security firm to patrol the area. Each uses marginal analysis in weighing up the extra cost of such protection, $120 per week, against the benefits of not being robbed. Each concludes that protection is a good buy and suddenly there are four protection vehicles patrolling the street. The robber also uses marginal analysis to decide that the ABCD community is not a fruitful area of robbing activity, given the amount of protection.

A invites the neighbours over for a glass of wine. He suggests that B, C and D get rid of their protection, since one patrolling vehicle is adequate, and that each share the $120 cost, i.e. $30 each. All agree. Drinking white wine becomes a weekly event. However, one Sunday A pours each a glass of the excellent white wine and asks if anyone would object if he had a glass of red wine. Then A announces he is not only giving up white wine but he is also giving up protection against robbery. B, C and D, however, agree to carry on, each now paying $40. The patrol vehicle continues to circle the street, thus protecting A, the non-contributor, as well as B, C and D, who are paying for the services. We have a public good in the community. A few weeks later, B drops out and C and D now bear the cost, each paying $60. The following week C drops out and poor old D, who wishes to remain protected, carries the whole $120 cost. A, B and C are 'free riders'. Note that D is no worse off than he was originally, i.e. paying $120 for protection; however, A, B and C are able to consume the protection service without having to pay for it. Note also that had A, B, C and D been ignorant of each other's actions, there would have been too many resources in the protection business in this community – three unnecessary security vehicles – and an inefficient use of resources.

Thus when public goods are involved, two issues have to be faced: how much of a given public good should be bought and who should pay for it? The price mechanism will not solve this problem. Is it possible to have too many resources committed, e.g. four security vehicles, but also possible to have too few as each person waits for his neighbour to buy the good? Do we consume many public goods? The answer is 'yes'. National defence, police and fire services, the judicial system and lighthouses are all examples of public goods. To solve the scarcity–choice problem in this context requires some type of collective action, and government, be it local, state or national, is often the collective agent that makes the decision.

Different political parties offer different bundles of public goods in their election manifestos and different tax structures to pay for them. For example a political party might stress the need for a large national defence commitment, including the most sophisticated weapons systems, and private health services, and a tax structure in which the highest marginal tax rate (i.e. the amount of income taxes payable on the last dollar earned) is 30 per cent. An opposition party might argue for minimal national defence, a national health service and a very progressive tax structure with high marginal tax rates, e.g. 60 per cent for individuals earning more than $100 000 per annum. Hopefully, we are showing in this simple but realistic example how societies and their political systems can modify the market mechanism to affect both the allocation of resources and the distribution of disposable (after tax/transfer) income.

2.2.3 Externalities

Externalities occur when the actions of an individual or firm confer benefits or costs on individuals or firms not directly involved in those actions. I smoke cigarettes. Why? Because I enjoy smoking. I have noted the opinion of King James I of the United Kingdom, who was also, of course, King James VI of Scotland, on smoking. He said it was 'loathsome, hateful and dangerous'. I have read also the medical reports, as will many of the people who will smoke about three trillion cigarettes this year. I have read also the medical reports on sunbathing, butter, barbecued steaks and drinking alcoholic beverages, and have implicitly conducted marginal analysis; and I still smoke. However, my smoking in public imposes costs on non-smokers. They don't like it and/or it can affect their health.

I have just finished a meal in a restaurant and I decide to have a cigarette with my coffee. All the neighbouring tables are filled with non-smokers. Should I smoke? Should I be allowed to smoke? Allowed by whom? The restaurant owner or the law? Suppose on the market there exists a $20 widget that, if placed on the end of a cigarette, prohibits the emission of noxious fumes. Who should buy the widget in the restaurant, the smoker or the non-smokers? The smoker, the non-smokers will argue. But why should the smoker buy clean air for the non-smokers? If the non-smokers want clean air – or caviar – in the restaurant, shouldn't they be prepared to pay for it?

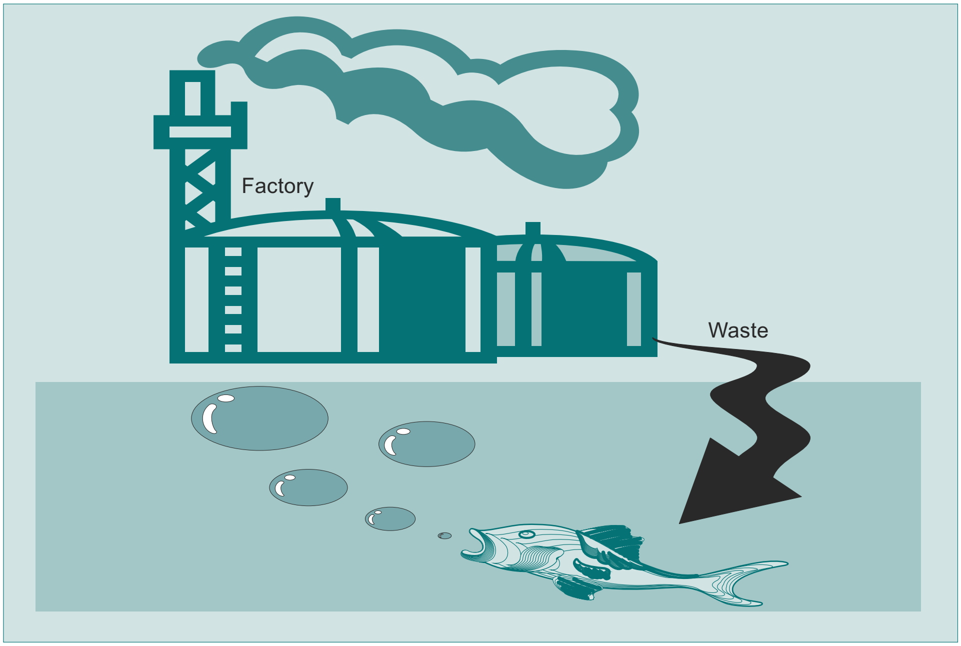


Figure 2.5 A riverside detergent factory

Further, consider Figure 2.5, which shows a detergent factory situated on the banks of a river. The river has only two economic functions in this society: dumping garbage (in this case, the detergent factory's waste) and salmon fishing. Also, in this society all detergent factories are located on river banks (to get rid of their waste) and the industry is highly competitive. If any detergent factory were to locate inland and have to process its waste, its costs would be higher than all its competitors and it would be forced out of business.

Back to the river. The salmon fishers approach the owner of the detergent factory and request that she stop polluting the river because the pollution is killing fish. The owner responds that processing waste would cost her an extra $100 per week and if the fishermen compensate the factory by $101 per week, it will stop polluting the river. The fishermen take their case to the government. Consider two possible outcomes.

First the detergent company, and all other detergent companies, are prohibited from polluting the river by the government. What will happen? The cost of producing detergent will increase, the selling price of detergent will increase, the real incomes of detergent users will fall, but we shall have a non-polluted river. However, in the second scenario the government legislates that to sell salmon caught in the river, the river must be pollution-free and that making sure the rivers are pure is the responsibility of the fishermen. The fishermen clean up the river. Now the cost of catching salmon will rise, the price of salmon will rise, and the real income of salmon eaters will decline, but again we shall have a clean river. Remember we – voting households – are the government. How are you going to vote on who should clean up the river? Being a rational utility-maximising individual, you will give an answer that will depend upon whether you are a detergent user or a salmon eater!

There is yet another possibility, which is that the detergent factory owner buys the salmon-fishing rights to the river. Then the problem completely disappears. What will the outcome be and why? The factory owner, a profit maximiser, will calculate, using marginal analysis, whether cleaning up the river and selling salmon will add more to her costs than it will to her revenues. If it will, the river will remain polluted. If, on the other hand, stopping dumping waste in the river (cost of $100 per week, remember) is less than the revenue she will receive from catching and selling salmon, then she will stop her factory polluting the river. Thus when the factory owner owns the fishing rights, polluting the river is a cost she bears and consequently is a cost she must take into account when attempting to maximise profit. When the factory owner does not own the fishing rights, polluting the river is a cost borne by the fishermen and society but not by the factory owner producing only detergent, so that she ignores such a cost and consequently the price of detergent does not reflect the real cost to society of producing detergent, i.e. society is implicitly subsidising detergent users.

The *externality* in the river pollution case is the extra cost society bears in the production of detergent, i.e. the polluted river. It is a cost not reflected in the price of detergent when the factory owner owns the fishing rights, i.e. the salmon given up becomes a cost of producing detergent and will be reflected in its price.

How much pollution should there be in an economy? Not only must ‘society’, i.e. all households taken together, decide on the optimal levels of air, water and noise pollution but it must also decide who in society should pay for achieving these levels. Resources used in stopping pollution mean fewer resources for all other goods and services.

2.2.4 Economies of Scale

A third cause of market failure involves *economies of scale*. Economies of scale arise when, as a firm’s level of output expands, the unit cost of producing output falls.

Consider Figure 2.6. The average cost curve is derived by dividing total cost of production by the amount of output. Which is preferable, to have one firm producing $0q_2$ at a cost per unit of $0c_1$ or two firms each producing $0q_1$ at a cost per unit of $0c_2$? If we assume $0q_2 = 2 \times 0q_1$ and $0c_2 = 2 \times 0c_1$, we shall have the same quantity of output in both cases but the two-firm case will require twice the amount of resources.

Thus the answer appears clear: have only one firm produce the output because the resources saved can be allocated to the production of other desirable goods and services.

Suppose, however, that one firm could supply the whole market. This firm would be called a *monopolist* (sole supplier) and would not have to compete with other firms in selling its output. It would have monopoly power and could charge a price higher than the price that a competitive situation would produce. Thus the price of monopolistically produced goods will not reflect the cost to society of the scarce resources involved in their production. The monopolist will earn above-normal profit but, because of economies of scale, resource owners will realise that to enter the market and compete on favourable cost terms with the established monopolist they will have to rival the monopolist in terms of level of output; in other words if they enter as a small firm producing an output level of $0q_1$, for example, they will be at a severe cost disadvantage. Given a limited market for the good in question, there may be 'room' for only one firm. Thus we have a problem: from the viewpoint of minimising the amount of resources used in the production of the good in question, we only want one firm; however, this firm will sell the good at a price in excess of the cost of resources used in its production.

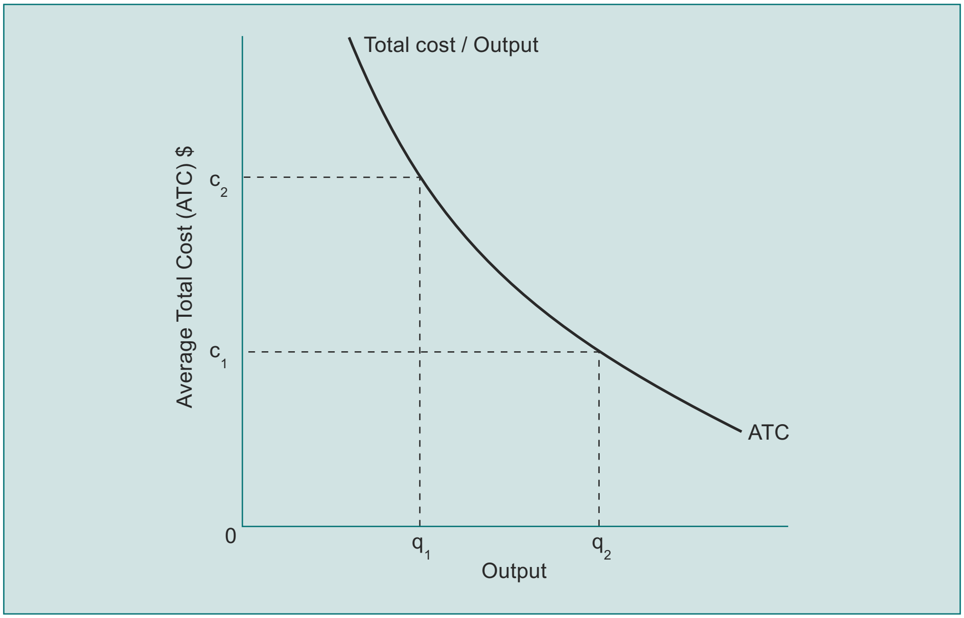


Figure 2.6 Economies of scale

If we want to have only one firm and also the 'right' price, then society will have to regulate the monopoly because the price mechanism will not perform this task. Public utilities normally experience economies of scale. Most cities only have one water supply, one sewerage system, one gas supply, one electricity supply and one telephone company. To regulate such monopolists to achieve an allocation of resources that is efficient in overall terms is not a simple task. Suppose the government legislates that a monopolist can exist and be the sole supplier in a market but that the price will be set by the government. What price will the government set? It would be reasonable – and desirable, you might argue – to set a price that covered all costs and yielded a normal return to resource owners, i.e. a normal profit to keep the necessary resources in that business.

At this stage it is important to realise that economies of scale are a form of market failure requiring a 'collective' solution. The problem that immediately presents itself is that, with a guaranteed return, there is no incentive to minimise costs.

2.2.5 Income Distribution

The fourth weakness of the price mechanism is a different animal altogether, involving the distribution of income. As seen earlier in this module in Figure 2.3, income in the form of wages, salaries, rent, interest and dividends flow from firms to resource owners. Thus, how large an individual's or household's income can be will be determined by the value of resources owned. Households whose resources are in short supply and that are eagerly sought after by firms will have 'large' incomes. At the other end of the spectrum, households with no marketable resources will have zero income. The demand for resources is a derived demand, i.e. derived from the goods and services produced from these resources. For most households in capitalist or market economies, their principal resource is their labour. A proportion of the population, however, owns stocks and shares, or equity, in companies, firms or financial institutions, i.e. they are part-owners of the capital stock of these companies and earn interest and dividends on such resources. Others own land and property, which in turn earn rent.

In market economies, returns to resource owners are determined by the forces of demand and supply, just as the prices of goods and services are determined by demand and supply forces. The income of star football, basketball and baseball players, golfers, or tennis players may appear 'high' but in a market economy these are 'equilibrium wages' or returns, determined in competitive markets. You may feel that in tennis the Williams sisters and Pete Sampras earn enormous incomes. Relative to most of us they do, but that shouldn't persuade you that your children would be well advised to pursue a professional tennis career. If they do not make the top 20 in the world, they might be better advised to take up a career in accounting – life may not be as glamorous or exciting but the expected lifetime earnings are undoubtedly higher.

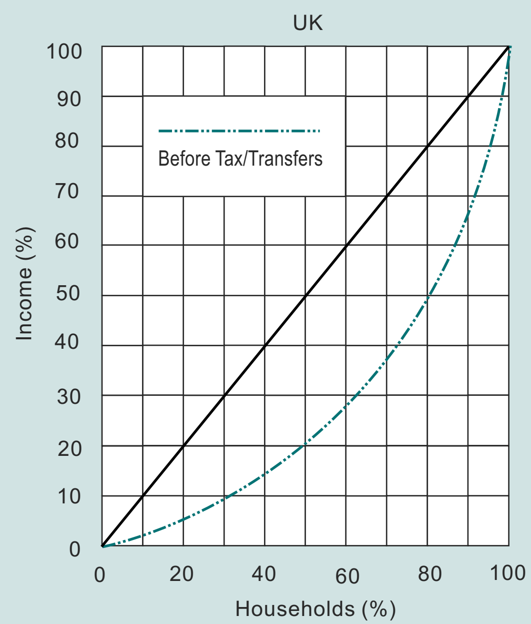


Figure 2.7 Lorenz curve: Income distribution in the UK (pre-tax and pre-transfers)

One way of measuring income distribution in a nation is shown in Figure 2.7. The graph showing income distribution is known as a *Lorenz curve*, named after its inventor Max Otto Lorenz who devised the graphical presentation at the turn of the twentieth century. The curve shows the cumulative percentage of incomes of any given cumulative percentage of households. Households are numbered on the horizontal axis from the lowest to the highest incomes. Thus the lowest 20 per cent of UK households receive 5 per cent of total income accruing to all UK households; the lowest 50 per cent receive 20 per cent of the national income; and the lowest 90 per cent receive 65 per cent – i.e. the highest 10 per cent receive 35 per cent of national income. If every household received an identical income, the resulting Lorenz curve would be represented by the solid 45° line.

How should income be distributed? Or, to put the question another way, what proportion of the goods and services produced should flow to what proportion of the population? The curve of income distribution shown in Figure 2.7 is typical of that which is determined by free markets. Is this just? Is this equitable? Questions of justice or equity are value judgements. Your opinion is as valid as a professional economist's opinion, or as a clergyman's, or as a film star's. There is no 'right' or 'wrong' income distribution in economics. How income should be distributed is not an unimportant question but it is a question not subjectable to economic analysis. Economists can evaluate the marginal benefits and marginal costs of adopting various strategies or policies to change the income distribution, but they have no special expertise, only personal opinions, in stating what the 'ideal' income distribution is.

The Lorenz curves in Figure 2.8 show the UK income distribution before and after taxes and transfers. Is the after-tax/transfer distribution preferable to the market-determined distribution? Should there be even higher taxes and larger transfers in order to move the curve towards the 45° line? Again these are value judgement questions to which there are no correct answers. Why, therefore, are we stating that a weakness of the market economy is the income distribution that results from market forces? The answer can be found in our political process. Every major political party in market economies today has, as part of its political platform, a commitment to alter the market-determined income distribution. To the extent that the vast majority of the population supports one of these parties and by definition or default approves of its platform, the majority of households, therefore, finds the distribution of income emerging from market forces to be unacceptable, and, through the voting process, implicitly votes to alter that distribution in favour of lower income groups and at the expense of higher income groups. How to alter the market income distribution to a more desirable distribution is the subject of much debate and government expenditure, and one to which we shall return. It is sufficient to note at this point that to bring about a significant change in the market-determined distribution of income requires collective action.

2.2.6 End Comment

The four preceding weaknesses of the market mechanism will be discussed in greater detail in later modules, as will the very significant strengths of the price mechanism, in the microeconomics portion of this course. We shall also study the way societies attempt to compensate for such weaknesses by collective or government action.

The remaining area of the market economy that was included in our original list of criticisms concerned stagflation, the dual existence of inflation and unemployment. Taken separately, inflation and unemployment can each be thought of as undesirable attributes of a market economy, as can imbalanced government budgets and imbalances in the balance of international payments. Such topics are macroeconomic in nature, to which we now give a brief overview.

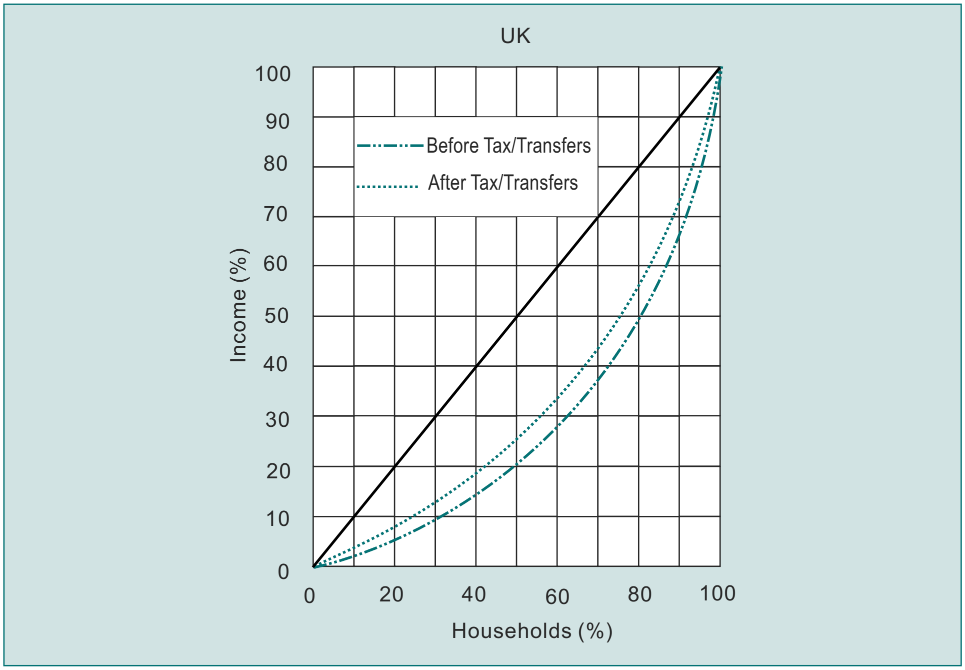


Figure 2.8 Lorenz curve: Income and distribution in the UK (pre- and post-tax and transfers)

2.3 Macroeconomics – an Overview

In Figure 2.9 we consider a generic macroeconomy.

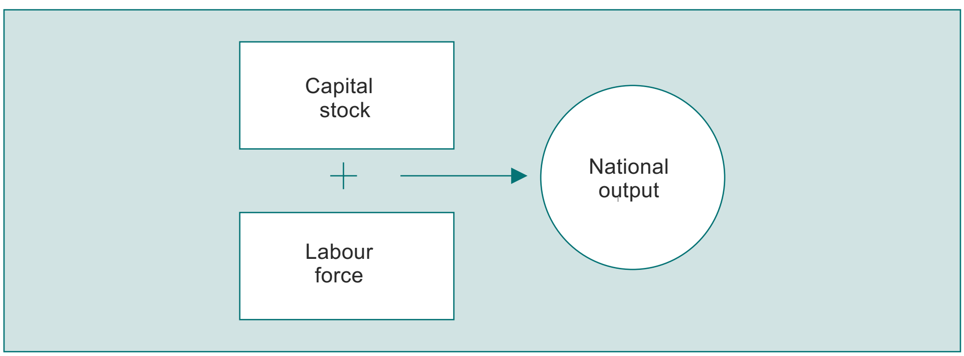


Figure 2.9 Resources and national output

How large the flow of national output can be will be derived by the quality and quantity of the capital stock and the quantity and quality of the labour force. It is obvious that if part of the capital stock is lying idle and part of the labour force is unemployed, national output will be less than its potential. As a result, households will enjoy fewer goods and services than the economy could produce. Such a situation is shown in Figure 2.10.

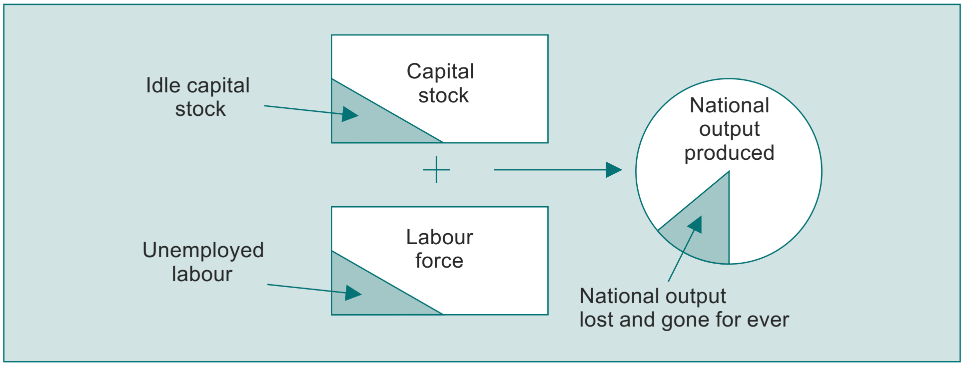


Figure 2.10 Unemployed resources

Why would any nation ever allow such a situation to exist? There are many reasons, as we shall find out. Let us return to our fully employed economy. Where does all the output go? Figure 2.11 shows the breakdown.

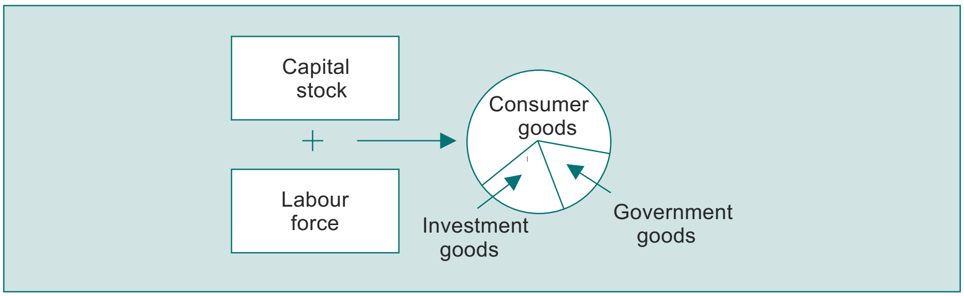


Figure 2.11 Distribution of national output

How large is each sector? That depends on the country. Table 2.1 gives the allocation of national output (Y) among consumption expenditure (C), investment expenditure (I), government expenditure (G) and net exports (exports − imports) for the seven major capitalist economies (the G7 economies) and the 'Big Four' emerging economies in the world (the BRIC countries, consisting of Brazil, Russia, India and China). These data are averages for the last two decades. While the net export sector (exports − imports) is relatively small in each case, this does not imply exports and imports are unimportant. Indeed, the exports or imports sector as a percentage of national output is large for some economies.

Table 2.1 Percentage allocation of national output for G7 and BRIC economies, 2000–2010

Country	C/Y	G/Y	I/Y	(Exports − Imports) /Y
Brazil	59.0	24.0	16.8	0.2
Canada	60.0	17.5	21.0	1.5
China	35.0	37.5	20.0	7.5
France	57.0	25.0	20.0	–2.0
Germany	57.0	15.5	21.0	6.5
India	60.0	21.0	24.5	–5.5
Italy	59.0	21.5	20.0	–0.5
Japan	57.0	11.0	31.0	1.0
Russia	48.0	24.0	19.0	9.0
UK	64.0	21.0	17.0	–2.0
USA	70.0	20.0	15.0	–5.0

Figure 2.12 contains data for a selection of countries showing the importance of the international sector as measured by the ratio of exports to national output. These data show the growing importance of international trades. In the USA and Germany, for example, ratios have doubled approximately. The emergence of China as a key player in world trade in the late 1990s can also be seen.

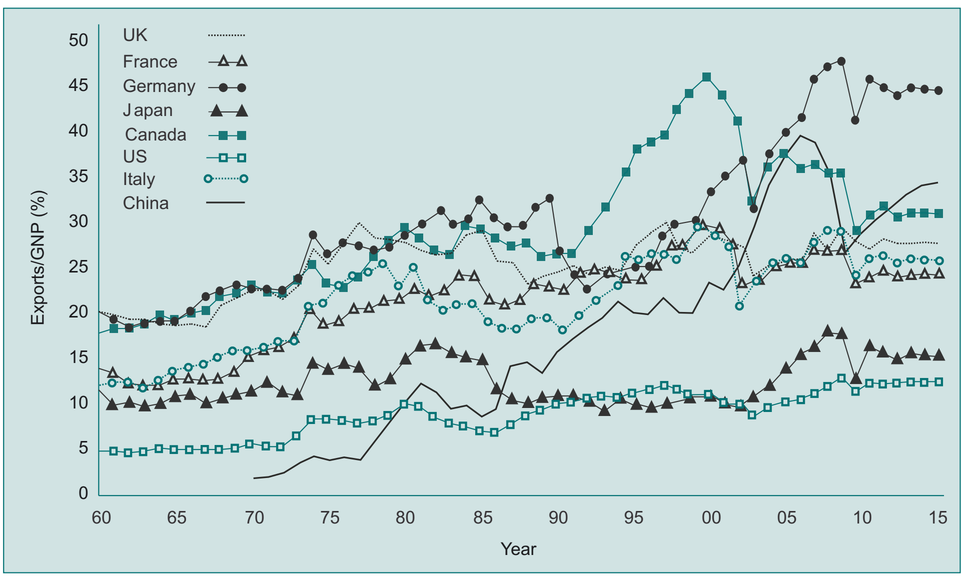


Figure 2.12 Exports as a percentage of GNP for the G7 economies and China, 1960–2015

To understand basic macroeconomics, we shall ignore – temporarily – the international sector. The reason for the large consumption sector is obvious. Consumption is the basic rationale for economic activity, i.e. using scarce resources to produce the goods and services that we enjoy. Why produce government goods? They are part of the goods and services we also enjoy – national defence, fire and police protection, are public goods. Thus they can be considered consumption goods also. But why produce investment goods? The reason is that we wish to enjoy a higher standard of living in the future. If we did not produce any investment goods, our capital stock would decline, buildings and factories would disintegrate, and machine tools would wear out. National output, remember, is a function of the quality and quantity of the capital stock and the labour force. Thus if we wish to enjoy increased living standards in the future, we must sacrifice some consumption today and allocate a proportion of our resources to increase the quality and quantity of the capital stock and/or the labour force.

In addition, as technological change takes place, it does not occur in a vacuum but is embodied in new capital equipment. For example, we live currently in the computer age, a quite remarkable technological innovation. Consider the Internet. We cannot take an old Olivetti typewriter and 'add on' an Internet package. The marvellous economies and increases in output realisable from developments in the Internet world are achievable only by building computers that can exploit the technological breakthrough. Similarly you can't attach a jet engine to the Wright brothers' aeroplane or attach a petrol engine to a horse's rear end to revive the horse and carriage trade. Technological change has contributed more to improved living standards in the past two centuries than has the increase in the 'size' of the labour force or the 'size' of the capital stock. To realise the benefits of new inventions requires incorporation of technological change into new capital goods. How much current consumption should be sacrificed for future consumption, therefore, is a decision every society must face. For less developed economies with low living standards, this can be a very painful choice. It can be a choice of life or death today for some members of society, for the benefit of future generations.

Let us return to our economy with unemployed resources, ignoring the desirability of economic growth temporarily. Figure 2.13 is a repeat of Figure 2.10. How can we, or governments representing us, get the economy to full employment and avoid wasting resources? The tools available come under two classifications: *fiscal policy* and *monetary policy*. Fiscal policy consists of changing government expenditure and/or tax rates. Monetary policy consists of changing the money supply or interest rates. Let us consider each in turn via some examples.

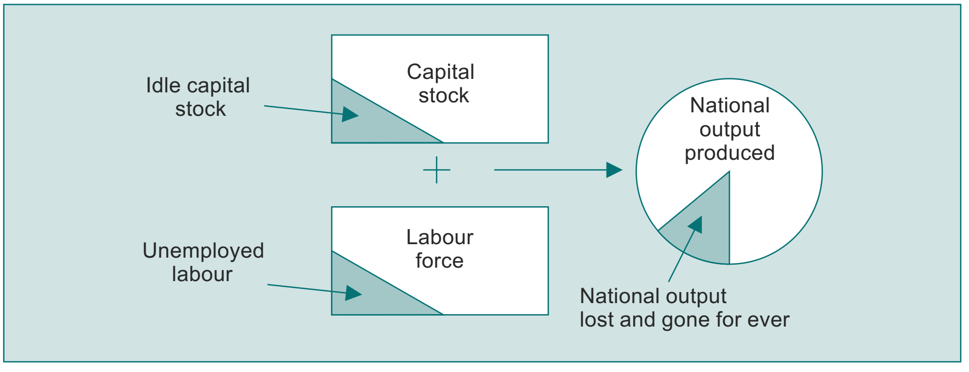


Figure 2.13 Unemployed resources

To cure the unemployment problem (people plus part of the capital stock), the government announces a large road-building programme and invites tenders for 1000 miles of motorway. A contract is awarded. Construction companies hire unemployed labour and order new road-building equipment, and the road is built. The people involved in building the road spend part of the income earned on a large variety of goods and services. This in turn stimulates further economic activity. Firms have to produce more goods and services to meet the increased demand, which involves hiring more people and perhaps buying more capital goods, which in turn means hiring more people and buying even more capital goods. Thus a multiplier process produces many more goods and services and creates many more jobs than created by the initial expenditure on the road system. Alternatively, the government elects to cut income tax rates. Households now have a higher disposable income (i.e. income − taxes). They spend more, i.e. buy more goods and services, which in turn creates more revenue for the firms supplying these goods and services and therefore more income for the labour force and owners of these firms, who also will spend more of their enhanced incomes. In other words, the multiplier process is in operation again.

Suppose the decision is made to ignore fiscal policy and concentrate on monetary policy. The money supply is increased, the price of money (the interest rate) falls, and this in turn can affect both firms and households. Take firms first. The cost of borrowing money falls. Investment projects that were not profitable when the expected rate of return was 15 per cent per annum and the cost of borrowing was 16 per cent become profitable if the rate of interest falls to 14 per cent. Thus investment expenditure increases. The firms that produce these capital goods experience increased orders, hire more people, and again the multiplier process goes into operation. Now let's consider households. Lower interest rates mean increased demand for houses, washing machines, cars – anything affected by the cost of borrowing. Increased demand leads to more goods and services being produced and ultimately more jobs, and consequently a reduced rate of unemployment of both the labour force and the capital stock.

This all sounds quite simple. All we have to do is measure the unemployment rate of the labour force and the capital and estimate by how much we have to stimulate the economy through fiscal or monetary policy and – bang – right there and then we achieve full employment, making the best use of our resources. You will not be surprised to learn that the real world is not quite so simple. What are the problems? There are many and they are complex.

First, we are not shooting at a fixed target. The potential of the economy changes over time; it is not static but very dynamic. Technological change occurs continuously; new industries expand; old industries decline. People emigrate; people immigrate. People join the labour force; people retire from the labour force. Firms decide to expand their capacity and build new factories and new office blocks; others contract. Foreign nations decide to invest in certain countries; some countries impose barriers to keep out foreign investment.

Second, it is difficult to predict households' and firms' behaviours. Tastes and preferences change; what was in fashion last year is not in fashion this year. Firms make investment decisions based on what they expect households to buy in the future. Some get it right; some get it wrong. Predicting the level and direction of investment expenditure is fraught with hazards; investment expenditure is the most volatile component of national income.

Third, we have no control over the outside world on which we are dependent for the supply of goods and services. Suppose we are an oil-importing economy in which oil and its derivatives are essential to many of our industries. The price of oil doubles; in the immediate future we have no alternative fuels. The concomitant effects can be severe. Households buy the same amount of oil and related products despite the twofold price increase. Given their budget constraints, the demands for other goods and services plummet, workers are laid off and the economy is in a recession.

Fourth, discovering exactly – or even approximately – how close the economy is to producing at its potential involves time lags; taking 'corrective' action involves time lags; benefiting from the implementation of the 'correct' fiscal and monetary policies also involves time lags before full utilisation of resources is achieved – and remember the making of fiscal and monetary policy is done in economies in which potential national output is increasing.

Fifth, no modern capitalist system lives in isolation; what is happening in other economies can have a significant effect. A recession in the rest of the world, for instance in a market for our exports, affects domestic employment of both labour and capital goods. Uncertainty about the state of our economy can cause capital flights affecting the balance of payments, the exchange rate, the demands for our exports with a 12- to 18-month lag, employment opportunities and the utilisation of our capital stock.

These are the concepts and issues that we shall analyse and discuss in the macroeconomic section of the course. Lest you believe such problems are not too complex, imagine how you deal with such complexities if the starting position of your economy is one of high wage and price inflation, a large government budget deficit, high unemployment, a large balance of trade deficit and the stated objectives are a full-employment economy, zero inflation, a balanced budget, exports equal to imports and a stable exchange rate. That's economics!

Module 3

Demand

Contents

3.1 Introduction

In this module we shall start by studying how a household allocates its income to make itself as well off as possible. For most households, income is a scarce resource. It can be supplemented by using savings and/or by borrowing against future income, but, temporarily ignoring such supplements, we shall assume a household has a limited amount of income in some given time period. Thus a household has a budget constraint. What the household attempts to do is to allocate this limited income on goods and services that create the greatest household satisfaction, or, using economic jargon, to *maximise utility subject to a budget constraint*.

There are many ways to allocate the limited income. The household has many wants that it wishes to satisfy. Companies or firms have produced, or can produce, many goods and services that they would like the household to buy. But the household cannot buy them all because it has a budget constraint. An allocation of $200 to the purchase of food in a week is no longer available for car repairs or beauty treatment.

Subject to legal considerations, most people in market economies are free to allocate their income as they wish. It is this consumer freedom that is the cornerstone of a market economy. As we shall see, it is consumer spending patterns that indicate to companies which goods and services households consider most desirable, i.e. it is the 'voting' with dollar bills on the part of households that encourages companies to produce the goods and services people want most. Companies respond to consumers' preferences as expressed through their spending patterns by hiring the necessary resources to produce the desired goods and services. But they do not do so because of love for their fellow man; they do so because they can make a profit on such transactions. These are the driving forces of a capitalist or market or free-enterprise economy: consumers trying to maximise utility from the way they spend their incomes, and companies or firms trying to maximise profit from hiring resources to produce the goods and services consumers want most.

Is our assumption of consumer rationality – that households know their own best interests – valid? Some people argue 'no' and further argue that there should be limits on the goods and services for which consumers can freely bid through a market mechanism. What do you think? Should consumers be allowed to buy alcohol and tobacco? What about flick-knives, horror comics and pornographic literature? What about soft drugs, hard drugs, sporting rifles and sub-machine-guns?

In a similar vein, should consumers be forced to buy certain goods and services? What about seat belts and crash helmets if they drive cars or motorcycles? What about paying taxes for national defence and police services, and what about paying taxes to build schools and have teachers if those paying have no children? What about national insurance contributions? What about car insurance, medical insurance, property insurance?

The answers to some of these questions are value judgements; but for others, as we shall see, a market economy – with each household pursuing its own self-interest – will not produce the optimal basket of goods and services through the market mechanism. No matter what your personal opinion about the desirability of alcohol, tobacco and drugs, for example, people in market economies do allocate part of their income to these goods and, in response to consumer spending, manufacturers do produce such goods.

We shall start by assuming that consumers know their own best interests and study how consumers signal with dollar votes to indicate their preferences to manufacturers. Manufacturers are concerned with people's preferences only to the extent that these preferences are backed by dollar votes. (You will receive little sympathy from a Mercedes-Benz dealer by informing him how much you would like to own the latest Mercedes sports car unless you back up this want with the required number of dollar bills.)

In the ensuing discussion we shall make no distinction between the household and the individual, even though many households consist of more than one individual, each with differing tastes and preferences. In other words, we shall assume the household acts as an individual consumer and knows its tastes and preferences. Many teenagers might well dispute this point and argue that insufficient weight is given to their wants in allocating the household budget – for example, too many carrots and not enough Coca-Cola, and undoubtedly insufficient resources allocated to the family car (or cars) when the driving test is passed, despite an increased willingness on their part to put out the rubbish and perform a multitude of household tasks.

3.2 The Theory of Consumer Choice

The rational consumer attempts to maximise utility or satisfaction subject to a budget constraint. Thus the household will change its expenditure pattern, i.e. buy a different combination of goods and services in some given time period, only if the change yields a higher level of utility. This trial-and-error process will continue until no change can yield a higher level of satisfaction. Only then will the household be in equilibrium. Of course, tastes and preferences change over time, albeit slowly, and new goods and services are constantly appearing on the market, but the principle of attempting to maximise utility remains. When will the consumer be in equilibrium? This state will be reached when the utility received from the last dollar spent on any good or service just equals the utility received from the last dollar spent on any other good or service. Then, and only then, will there be no gain in total utility from a reallocation of expenditure.

As an individual consumes more of a good in a given time period, total utility increases, but at a decreasing rate. The extra utility derived from consuming an additional unit of a good is known as *marginal utility*. Marginal utility decreases as the consumption of a good increases. This is known as the *Law of Diminishing Marginal Utility*. A tennis player at the end of a five-set match will obtain more utility from the first mouthful of water than from the second, more from the second than from the third and so on. At some stage, say by the sixth mouthful, his thirst will have gone and he will obtain no more utility from additional water – and at that point he stops drinking.

The existence of diminishing marginal utility means that an individual can increase total utility by purchasing a combination of goods and services rather than by allocating the budget to one good only. Total utility can be increased by reallocating expenditure from goods with lower marginal utility to goods with higher marginal utility. Suppose you were dining out on a $40 budget; it is highly unlikely that you would choose 20 helpings of soup at $2 per helping, no matter how delicious the soup. The reason is simple: the law of diminishing marginal utility is at work. You are much more likely to select a variety of items from the menu and wine list in order to maximise total utility given your tastes and preferences and your $40 budget constraint. It is the desire to maximise satisfaction that stops you telling the waiter to bring you any $40 worth of food in any order.

The rule for efficient consumer behaviour, i.e. for choosing the combination of goods that would yield the highest level of utility from a given budget is that the budget is allocated so that the last dollar spent on each good yields the same utility. What if one good (good A) is twice as expensive as another good (good B)? In that case the last unit of good A consumed will have to yield twice the level of utility of the last unit of good B consumed. Thus to be in equilibrium, i.e. maximising total utility, the following condition must hold:

$$\frac{\text{Marginal Utility of Good A}}{\text{Price of Good A}} = \frac{\text{Marginal Utility of Good B}}{\text{Price of Good B}} = \ldots = \frac{\text{Marginal Utility of Good Z}}{\text{Price of Good Z}}$$

That is:

$$\frac{\mathrm{MU_A}}{\mathrm{P_A}} = \frac{\mathrm{MU_B}}{\mathrm{P_B}} = \cdots = \frac{\mathrm{MU_Z}}{\mathrm{P_Z}}$$

Suppose an individual allocates his income on two goods only, goods A and B, such that:

$$\frac{\mathrm{MU_A}}{\mathrm{P_A}} > \frac{\mathrm{MU_B}}{\mathrm{P_B}}$$

He could increase total utility by buying more of A and less of B – why? Given the Law of Diminishing Marginal Utility, as more of A is bought, the marginal utility of A decreases. Simultaneously, as less of B is bought, its marginal utility increases. Thus by a reallocation of the budget to more A and less B, $\mathrm{MU_A}$ declines and $\mathrm{MU_B}$ increases. There will be one combination of A and B purchases at which:

$$\frac{\mathrm{MU_A}}{\mathrm{P_A}} = \frac{\mathrm{MU_B}}{\mathrm{P_B}}$$

With this combination, the last dollar spent on A and the last dollar spent on B provide the same utility, and so total utility is maximised. Any other combination yields a lower level of total utility.

Suppose you decide to visit the latest wine bar in town with a budget of $10. This wine bar serves only local white wine at $1 per glass and smoked salmon sandwiches at $0.50 each. You like both wine and smoked salmon. How should you allocate your $10 on wine and sandwiches? To maximise utility you should consume wine and sandwiches in such amounts that the last dollar spent on each yields the same utility, i.e.

$$\frac{\text{MU wine}}{\$1} = \frac{\text{MU sandwiches}}{\$0.50}$$

The utility of the last glass of wine must therefore be twice the utility of the last smoked salmon sandwich.

Further, suppose that the owner of the wine bar, attempting to attract your custom in the future, gives you a $5 voucher that you must spend that night on wine and sandwiches. With this additional budget you can now purchase more of each, both wine and sandwiches. However, the marginal utility of both wine and sandwiches will decline as you consume more – remember the Law of Diminishing Marginal Utility – although total utility will rise. In your new equilibrium position it must still be true that:

$$\frac{\text{MU wine}}{\$1} = \frac{\text{MU sandwiches}}{\$0.50}$$

although this of course assumes you can still make rational choices despite the amount of wine consumed!

Assume now that you are in equilibrium in your wine and smoked salmon world on your $10 per night budget, and thus:

$$\frac{MU_w}{P_w} = \frac{MU_s}{P_s}$$

To attract more custom, the owner of the wine bar reduces the price of a glass of wine, i.e. P_W goes down. As far as you are concerned this will cause disequilibrium, with

$$\frac{MU_w}{P_w} > \frac{MU_s}{P_s}$$

To restore equilibrium, i.e. to make

$$\frac{MU_w}{P_w} = \frac{MU_s}{P_s}$$

again, you must reallocate your budget between wine and sandwiches – buying more wine and fewer sandwiches than previously. As you consume more wine, the marginal utility of wine will decrease, and as you consume fewer sandwiches the marginal utility of sandwiches will increase, and when equality is restored you will be back in equilibrium but better off than before because the reduction in the price of one of your staple foods is similar, but not identical, to the owner giving you the $5 voucher. You will have experienced an increase in real income, and your total utility will have increased in both cases, so that, even though your money income has not increased, you can buy more with it.

What these examples show is that both income (budget) and prices affect the quantities of goods you buy. These are not the only factors, however, that affect your purchases or demands for different goods and services.

3.3 Individual (Household) Demand

The quantity of a particular good that an individual buys in any given time period depends on a large number of factors. For example, the number of vacations or holidays purchased by a household in a year may depend upon:

(a) the price of a week in Greece versus two weeks in Spain;
(b) the prices of other goods that could be bought instead of a vacation, for example a second-hand car;

(c) household income;
(d) size and age distribution of the family;
(e) state of health;
(f) expected weather conditions; and
(g) the price of airline tickets.

Similarly, the amount of steak a household buys in a week will depend upon a large number of factors. The price of steak will be important, but so also will be the price of alternatives, e.g. chicken or lamb chops. Household income, if low, may dictate no steak at all is purchased. If you like barbecuing, the weather may be a factor, as may be the number of guests you have invited to dinner.

To isolate and measure the influence of any one factor on potential purchases of a particular good, it is necessary to hold constant all other factors affecting the purchase of that good. Suppose a university were considering an increase in tuition fees and wanted to investigate the impact this would have on student applications. Further, suppose that, simultaneously, educational grants and family income were increasing. This would make accurate calculation of the impact of increased fees difficult. Similarly if an electronics company's sales of personal computers (PCs) are lagging and it is considering an increase in its advertising budget, measuring the effectiveness of such expenditure will have to be disentangled from all factors affecting PCs, e.g. economic recession, competing products, after-sales service levels, expectations of more advanced machines, foreign exchange rates, interest rates and the investment plans of customers.

In studying demand, we shall select one variable initially and assume all other factors affecting demand for the good in question are held constant. We shall then investigate how changes in the selected variable affect demand for this good. Economists use the Latin expression *ceteris paribus* (with other things remaining equal or unchanged, sometimes abbreviated *cet. par.*) when investigating such a relationship. We shall encounter this expression throughout the study of economics.

Because of the importance of prices in providing information to consumers and producers of goods and services, and because it has an effect on their behaviour, we shall concentrate initially on the relationship between the price and the quantity of a good that an individual would be willing to buy in a given time period. This relationship is known as *individual demand* for a good.

Consider Table 3.1 and Figure 3.1.

Table 3.1 Individual demand for salmon in a week

Price ($)	Quantity (kilos)
10	0
8	1
6	2
4	3
2	4

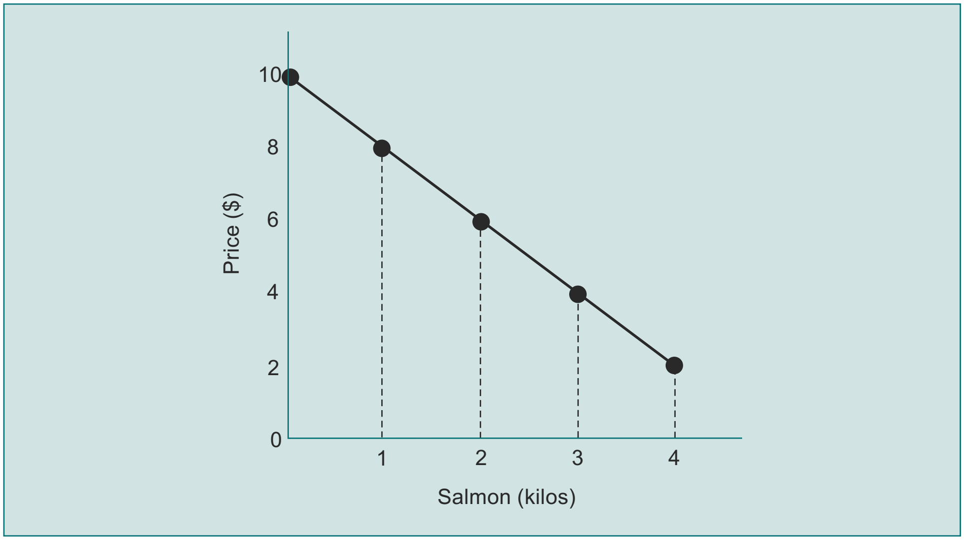


Figure 3.1 Individual demand for salmon in a week

The observations in Table 3.1 have been plotted in Figure 3.1 and joined by a smooth curve. Points lying on the curve between any two given points are approximations of the quantities of salmon the individual would buy if prices lay between any two of the noted prices in Table 3.1. For example, 2.5 kilos would be purchased if the price were $5 per kilo.

Let us be quite clear what Figure 3.1 is telling us. First, we are considering a time period of one week. Were we to consider the individual's *monthly* demand curve for salmon, i.e. over approximately four weeks, we would expect larger quantities of salmon to be purchased at each price less than $10 per kilo, and the curve would shift to the right. Second, we are holding constant all factors other than price, which might affect salmon purchasers. For example, we are assuming that the price of halibut, a potential substitute for salmon, does not change. Were a glut of halibut to appear on the fish market, causing its price to halve, the individual might buy no salmon in a given week and switch to the cheaper halibut. Third, the demand curve tells us that the higher the price, the smaller is the quantity that would be bought. Specifically, if the price of salmon were $8 per kilo the individual would buy 1 kilo in the week, whereas if the price were $4 per kilo the individual would buy 3 kilos.

Thus, the demand curve does not describe an event, i.e. it does not tell us how much salmon an individual actually bought; it is a *hypothetical relationship* concerning what our individual would be prepared to buy at each price, in a given time period, *cet. par.*

Our individual is now in a fish shop, mindful of the household budget constraint, looking at potential purchases. The salmon looks fresh and is priced at $6 per kilo. Just as he is about to buy 2 kilos, the fishmonger replaces the $6 per kilo sign with a $4 per kilo sign, explaining that it is Saturday and he does not want to be left with unsold salmon over the weekend. Our purchaser is delighted and elects to buy 3 kilos now instead of 2. Has the individual's demand curve changed? No. All he has done is move to a new position on his demand curve. If everything determining the position of the demand curve remains unaltered, i.e. if all the *cetera* remain *pares*, then neither the position nor shape of the demand curve changes.

The typical demand curve is negatively inclined to the price axis: at higher prices lower quantities would be purchased; at lower prices higher quantities would be purchased. There are two reasons for this negative slope. First, at a higher price the individual is unable to purchase as many units of goods normally purchased because of the budget or income constraint. Second, at higher prices the individual substitutes relatively cheaper goods for the good in question. The decrease in quantity of a good (good X) that would be purchased at a higher price, is therefore, a combination of two effects: an income/budget effect and a substitution effect.

The income and substitution effects can be isolated. A higher price of good X will, *cet. par.*, lower an individual's total utility because it decreases the quantities of goods he (or she) is able to purchase, i.e. he suffers a loss in real income. An equivalent loss in total utility could be caused by money income (budget) being reduced by a certain amount and prices of all goods, including X, remaining unchanged. Such a reduction in money income would lower the purchase of good X, i.e. the income effect leads to a reduction in the quantity of X purchased.

The difference between the lower quantity of X purchased because of a price increase in X and the decrease in quantity due to the income effect is the substitution effect. Alternatively, the substitution effect is the decrease in the quantity of good X purchased that results from the change in the price of X relative to other goods, if one were to compensate the individual for the resulting change in his real income.

Consider an individual who chooses to work a 40-hour week at $10 per hour. He does not work any additional hours per week because he considers that each hour of leisure is worth at least $10, i.e. $10 (the hourly wage rate) is what he gives up for every hour of leisure he enjoys. If the wage rate were to increase to $20 per hour, would he work more or less than 40 hours per week? The answer depends on the relative strengths of the income and substitution effects.

If he were to continue working 40 hours per week, his weekly income would increase from $400 to $800. At the higher income he would buy more of all goods, including leisure. If he 'buys' more leisure, it means that he will reduce the number of hours worked per week; therefore, the income effect could lead to a reduction in hours worked. However, the price of leisure does not remain constant. The cost of an hour of leisure now rises to $20 per hour because that is what it costs the individual to 'buy' an hour of leisure. At a higher price for leisure, the individual will substitute relatively cheaper goods for leisure, i.e. buy less leisure. But buying less leisure is equivalent to working longer hours. Therefore, the substitution effect would lead to an increase in hours worked. Thus in this example the income effect (buy more leisure) and the substitution effect (buy less leisure) work in opposite directions. It is not possible to know *a priori* for any given individual which effect will predominate since it will depend upon an individual's preferences for leisure versus income. However, the effect on the average worker over time is clear: hours worked by the average individual have decreased over the past 100 years from over 60 hours to under 40 hours per week, and from this it appears that the income effect has outweighed the substitution effect.

The substitution effect is always negative, in other words the increase in the price of a good will always lead to the substitution of relatively cheaper goods for that good. The income effect may be positive or negative. For 'normal' goods the decrease in real income caused by an increase in the price of a good will lead to a smaller quantity being purchased and the increase in real income caused by a decrease in price will lead to a larger quantity being purchased, so that the income effect is positive since changes in real income and quantity purchased are of the same sign. Thus for normal goods the positive income effect reinforces the negative substitution effect and the demand curves for such goods always slope downwards from left to right. Thus, for instance, where part of a household's weekly budget is allocated to chicken, red meat and fish and the price of chicken falls, not only will more chicken be purchased because real income will have risen but more chicken will also be purchased as it is substituted for the now relatively more expensive meat and fish, i.e. the income and substitution effects reinforce each other leading to more chicken being purchased at lower prices and exhibiting the normally sloped demand curve for chicken.

An increase in real income may be caused by an increase in money income (budget) while the prices of goods remain constant, or by a decrease in the prices of goods, while income remains constant. An *inferior good* is one for which the quantity purchased falls when real income rises. As a family's money income rises and the prices of all goods remain the same, the family's real income will also rise. When this occurs, the purchases of cheap cuts of meat, for example, usually decline. Another example of an inferior good is low-cost housing since, with higher incomes, families switch to superior housing.

For inferior goods, the negative substitution effect of a price change counteracts and generally exceeds the income effect, i.e. the change in the relative price of good X leads to a change in the quantity of good X purchased, and this is greater than the change in the quantity of good X purchased because of the change in real income caused by the price of X changing. The demand curve is therefore still normally sloped. For example, assume margarine is an inferior good. While a decrease in the price of margarine may lead to a family substituting margarine for butter, the resulting increase in real income may cause more butter and less margarine to be purchased. The substitution effect, i.e. margarine for butter, as a rule outweighs the income effect, which for this good is negative and which causes less margarine to be purchased because of rising real income. Thus the demand curve for margarine will be normally sloped even though it is an inferior good, such that a lower price of margarine will be accompanied by a larger quantity purchased, *cet. par.*

There is one famous example where the income effect exceeds the substitution effect and where, over a range of prices and real income, a non-normal demand curve can be produced. This type of good is known as a *Giffen good*: at higher prices a larger quantity is purchased, though the range of prices for which this condition holds is limited.

In nineteenth-century Ireland, it is alleged that the bulk of family income was spent on potatoes because this was all the people could afford. However, when the potato crop was exceptionally large and the price of potatoes fell considerably, Irish peasants experienced a significant rise in real income. The same number of potatoes could be bought for much less and with the money left over Irish families could buy vegetables and meat. However, the purchase of other foodstuffs meant that fewer potatoes were required and thus their purchases could be reduced. Thus at lower prices fewer potatoes were purchased, and this was caused by the very large income effect.

There is no empirical evidence that such a good ever existed but students, being students, like exceptions, and if Victorian economist Sir Robert Giffen's hypothesis about the 1845 Irish potato famine helps clarify income and substitution effects, I am all for Sir Robert. With the exception of people who spend a large proportion of their incomes on only one good (alcoholics and drug users), it is very difficult to imagine any family in advanced economies today for whom changes in the price of any one good could have such a dramatic effect on their real income.

Movements along a typical demand curve indicate that higher quantities are purchased as price falls, i.e. the quantity purchased varies as price varies. Prices and quantities are, not surprisingly, known in this context as *variables*. All factors other than price that affect purchases and that we assume to be unchanged in constructing a demand curve are known as *parameters*. Given that the parameters are fixed, the quantity of a good purchased will therefore depend on the price of that good; quantity becomes the *dependent variable*. Because it is highly unlikely in most markets that any single individual's purchase of a good will affect the price of that good, price is known as the *independent variable*.

Thus the position and shape of a demand curve will depend upon the parameters; and, if any change occurs in any one of the parameters, the position and shape of the demand curve will change. We stated earlier that household tastes and preferences ultimately determine how resources are allocated in capitalist economies. In constructing an individual demand curve, tastes and preferences are considered to be parameters; in other words they are assumed to be unchanged in a given time period. Over time, however, tastes and preferences do change, and consequently the position and shape of demand curves shift. This can lead to price changes, or at least signals to producers, forcing them to redeploy resources if they wish to stay in business.

Let us now consider a parameter change using the example given in Table 3.2 and Figure 3.2. The parameter we have changed is income. It is one of the factors we held constant in constructing the initial demand curve D for a weekly income of $400. At that income level and with the price of salmon at $6 per kilo, our individual would have purchased 2 kilos. However, if the weekly income were $200 instead of $400 and the price of salmon still $6 per kilo, less than 1 kilo would have been purchased, i.e. the demand curve for this individual has shifted to the left. It has shifted because one of the parameters determining its position has shifted. The shift has nothing whatsoever to do with the price of salmon. When the price of salmon changes, the individual moves along his demand curve; only a change in one of the parameters determining the position of the curve will cause the curve to shift.

Table 3.2 Individual demand for salmon in a week (lower income)

Price ($)	Quantity (kilos)	Weekly income	Price ($)	Quantity (kilos)	Weekly income
10	0	400	10	0	200
8	1	400	8	0	200
6	2	400	5	1	200
4	3	400	4	1.4	200
2	4	400	2	2	200

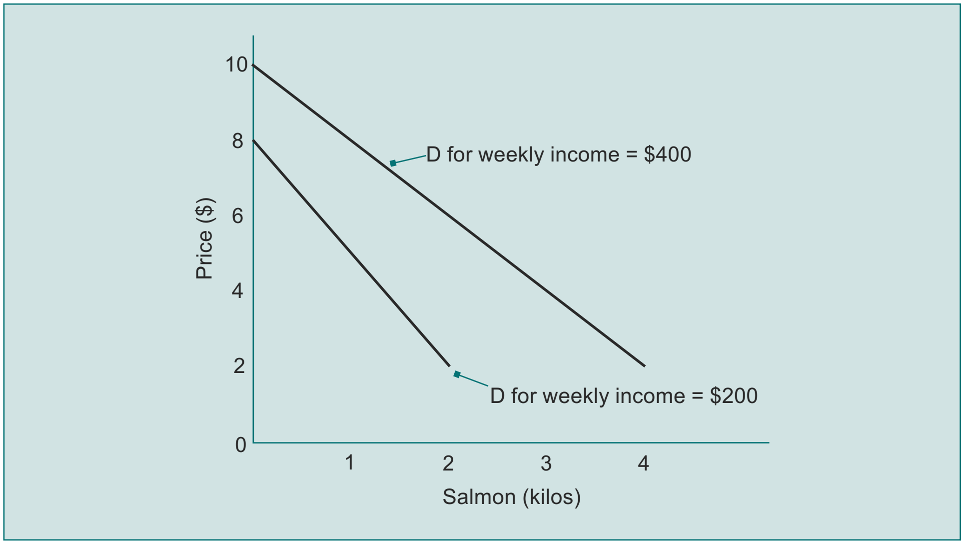


Figure 3.2 Individual demand curves for salmon for different levels of income

The following parameter changes could each shift the salmon demand curve, but in varying ways.

Parameter	Direction of shift
Dinner party for Americans who love salmon	right
Good barbecuing weather	right
Medical report relating heart disease to meat consumption	right
Lower price of chicken	left
Children go off to college	left
Mercury poisoning found in salmon	left
Gift of a salmon smoker	right
Holiday in India	left
Expectation of higher salmon prices *next* week	right

To simplify our analysis above, we have concentrated on a relatively brief time period, i.e. one week. For many households one week or one month is a convenient planning period for many household expenditures, because they are paid weekly or monthly, have a weekly or monthly budget constraint, and plan expenditures accordingly. Many household decisions, however, require to be analysed over a longer time period – for example, the decision to borrow money over a period of 25 years so as to purchase a house. Similarly the ability of students to borrow money and maintain a level of consumption above current income – and the willingness of lenders to lend money to such students – is based on expectations of both students and lenders that a student's future income will be sufficient to pay back the sum borrowed plus interest. However, the principles of utility maximisation given budget or income constraints remain in force.

Earlier we argued that utility maximisation dictates that

$$\frac{MU_a}{P_a} = \frac{MU_b}{P_b} = \cdots = \frac{MU_z}{P_z}$$

and that as an individual consumes more of a good in any given time period, total utility increases but marginal utility decreases. What then is the connection between utility theory and our downward-sloping demand curve? The Law of Diminishing Marginal Utility implicitly argues that an individual will only be willing to buy more of a good if the price falls.

Suppose you are in equilibrium and consequently maximising your utility in a two-good world in which you allocate your budget on wine (w) and sandwiches (s), with wine costing \$1 per glass and sandwiches costing \$0.50 each. Then

$$\frac{MU_w}{P_w} = \frac{MU_s}{P_s}$$

Further, suppose that at those prices

$$\frac{MU_w}{P_w} = \frac{MU_s}{P_s} = \frac{4}{1}$$

If the price of wine is reduced to \$0.50 per glass, then

$\frac{MU_w}{P_w}$ would become $\frac{8}{1}$

so that the last \$1 spent on wine will now yield approximately twice as much satisfaction since you can now buy two glasses instead of one. However,

$\frac{MU_s}{P_s}$ still equals $\frac{4}{1}$

and spending one dollar less on sandwiches reduces utility by 4 while spending one dollar more on wine increases utility by almost 8, giving a net gain of 4 (+8 − 4). Nevertheless, as you consume more wine and fewer sandwiches, the marginal utility of wine will decline and the marginal utility of sandwiches will increase. You will continue to reallocate expenditure until you are back in equilibrium, i.e. until

$\frac{MU_w}{P_w}$ again equals $\frac{MU_s}{P_s}$

Thus your demand curve for wine will be negatively inclined to the price axis. The same reasoning applies to the sandwiches, or any other good or service.

So far we have concentrated on individual demand. In markets where there are large numbers of firms (suppliers), the individual demand for any product can be assumed to have a negligible influence on both the price and quantity bought and sold of that good. Whether or not you buy a loaf of bread one day in your local supermarket does not significantly affect the sales or profits of the supermarket selling approximately 5000 loaves per day. Similarly, if you decide to go on vacation to a foreign country and change some of your currency into the currency of that country, you would not be surprised to observe that your purchase of foreign currency did not cause your domestic currency to fall in value nor cause the foreign currency to appreciate.

However, when we consider *total* demand in any given market we are adding together all of the individual demands. Total market demand is something about which companies are very concerned.

3.4 Market Demand

The market demand for a good is the sum of the demands for that good of the individuals who comprise the market. Some markets are highly localised – for example – the markets for baby sitters, plumbers, fresh bread, aspirin and hairdressing. At the other extreme, some markets are international – wheat, gold, crude oil, and vacations. Independent of the location or size of a market, a *market demand curve* is derived from individual demand curves. We shall choose a very simple example to show how a market demand curve is constructed.

Table 3.3 Individual and market demand for salmon (per week)

Price ($)	Quantity (kilos) demanded by individuals with			Market demand (kilos)
	High income	Medium income	Low income	
10	1	0	0	1
8	2	1	0	3
6	3	2	1	6
4	4	3	2	9
2	5	4	3	12

There are three individuals in the salmon market for which data are given in Table 3.3, namely High income (*H*) , Medium income (*M*) and Low income (*L*). At a price of $10 per kilo, only *H* buys salmon and then only 1 kilo. Thus at the price of $10 per kilo the market demand is 1 kilo. At a price of $8 per kilo, *H* would buy 2 kilos, *M* would buy 1 kilo, and *L* would still not be interested. The market demand, therefore, at $8 per kilo is 3 kilos. At $2 per kilo, *H* would buy 5 kilos, *M* would buy 4 kilos, and *L* would buy 3 kilos, yielding a market demand of 12 kilos.

These data have been plotted in Figure 3.3 and as can be seen, the market demand curve is constructed by adding together horizontally the individual demand curves. In this simple example, we have violated the condition of no single buyer being significant: each one is. But if we had thousands of salmon buyers in this market, the market demand curve would still be constructed in the same fashion but the influence of any one buyer would be insignificant.

The conditions that hold for the individual demand curve hold for the market demand curve. Again, price is the independent variable and quantity demanded is the dependent variable. Parameters such as taste and preference and income levels again determine the position of the market demand curve relative to the axes, as does the number of buyers on the market. Changes in any of these parameters will cause the demand curve to shift; at higher prices a smaller quantity is demanded, at lower prices a larger quantity is demanded.

The demand for goods and services can be classified by the responsiveness of a quantity demanded to changes in price. For some goods a 'small' change in price brings about a 'large' change in the quantity demand (however these are measured); for other goods the quantity demanded is relatively insensitive to changes in price. In addition, the degree of sensitivity can vary depending upon the point selected initially on the demand curve, i.e. a 10 per cent price reduction when the price is $10 may lead to a 20 per cent increase in the quantity demanded, whereas a 10 per cent price reduction for the same good when the price is $8 may only lead to a 12 per cent increase in the quantity demanded. This sensitivity of quantity demanded to price changes, which is known as *price elasticity of demand*, also affects total expenditure on the good in question, with implications (as we shall see) for consumers, firms and the government.

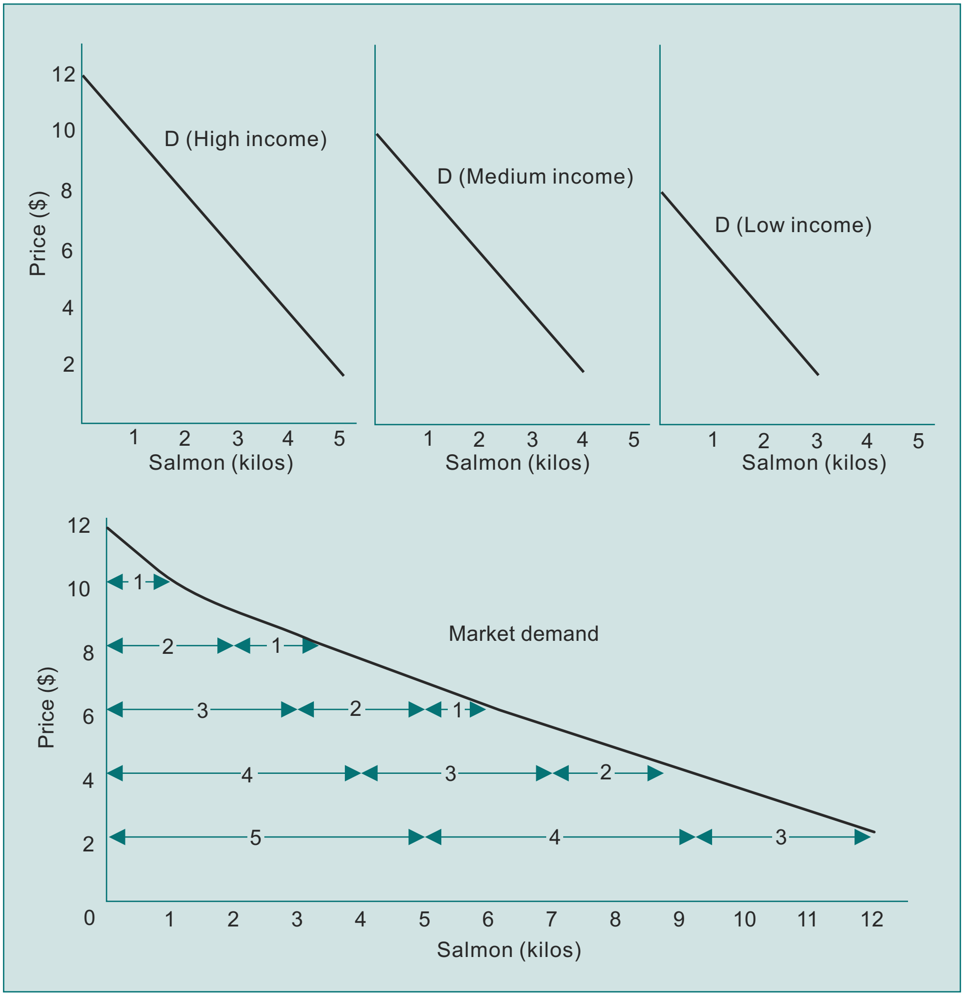


Figure 3.3 Individual and market demand for salmon

The three possible classifications of price elasticity of demand are:

(a) price inelasticity;
(b) price elasticity; and
(c) unitary price elasticity.

3.4.1 Price Inelasticity

A good is classified as price-inelastic when, as price changes, the proportional change in the quantity demanded is less than the proportional change in price. When the demand for a good is price-inelastic, a price increase will lead to an increase in expenditure on the good and a price decrease will lead to a decrease in expenditure on the good.

If we let P equal the initial price and Q equal the initial quantity demanded, and ΔP equal the change in price and ΔQ equal the resultant change in quantity demanded, then when

$$\frac{\Delta Q}{Q} < \frac{\Delta P}{P}$$

demand is price-inelastic. If a measure of price elasticity is denoted by E, the inelastic demand shown above can be expressed as:

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} < 1$$

This equation states that when E < 1, demand is price-inelastic. Total expenditure on the good (P × Q) increases as P increases and decreases as P decreases. To be mathematically correct, E is normally negative, since price and quantity changes move in opposite directions. For convenience, however, it is customary to ignore the negative sign.

As an example, let us consider the following data which refer to the quantity of petrol demanded by the residents of a town at two different prices during a typical day.

Price ($)	Quantity (litres)
1.00	100
1.10	99

In considering a price increase from $1.00 to $1.10, the quantity demanded decreases from 100 litres to 99 litres, i.e. the demand curve is normally sloped. We have:

$$\frac{\Delta Q}{Q} = \frac{100-99}{100} = \frac{1}{100}$$

and

$$\frac{\Delta P}{P} = \frac{1.10-1.00}{1.00} = \frac{1}{10}$$

and therefore

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} = \frac{\frac{1}{100}}{\frac{1}{10}} = \frac{1}{10} < 1$$

The value of E indicates that the demand for petrol over the price range $1.00–$1.10 is price-inelastic. And a price increase from $1.00 to $1.10 will cause total expenditure on petrol to increase from $100.00 ($1.00 × 100) to $108.90 ($1.10 × 99).

Now consider a price decrease from $1.10 to $1.00. In this case we have:

$$\frac{\Delta Q}{Q} = \frac{100-99}{99} = \frac{1}{99}$$

and

$$\frac{\Delta P}{P} = \frac{1.10-1.00}{1.10} = \frac{1}{11}$$

so that

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} = \frac{\frac{1}{99}}{\frac{1}{11}} = \frac{1}{9} < 1$$

The demand curve for petrol over the price range $1.00–$1.10 is still price-inelastic. (The difference in the values of price elasticity, $\frac{1}{10}$ and $\frac{1}{9}$, is caused by using different bases.) In considering total expenditure on petrol, the price decrease will cause total expenditure on petrol to decrease from $108.90 to $100.00.

If the owner of the petrol station were free to set the price of petrol, price elasticity of demand for petrol could be an important determinant of the price he sets. Indeed, the fact that the demand for oil in the world is price-inelastic typically leads to a rapid increase in the revenues flowing to the oil-exporting countries when world oil prices are raised.

3.4.2 Price Elasticity

A good is classified as price-elastic when, as price changes, the proportional change in the quantity demanded is greater than the proportional change in price. And when the demand for a good is price-elastic, a price increase will lead to a decrease in total expenditure on the good and a price decrease will lead to an increase in total expenditure on the good. That is, when

$$\frac{\Delta Q}{Q} > \frac{\Delta P}{P}$$

demand is price-elastic, and

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} > 1$$

For example, the following data refer to the quantity of roast beef demanded by diners at a certain restaurant in a normal day.

Price ($)	Quantity (portions)
1.00	100
1.10	80

In considering a price increase from $1.00 to $1.10, the price elasticity of demand for roast beef can be calculated as

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} = \frac{\frac{100-80}{100}}{\frac{1.10-1.00}{1.00}} = \frac{\frac{2}{10}}{\frac{1}{10}} = 2$$

Simultaneously with the price increase, total expenditure decreases from $100 ($1.00 × 100) to $88 ($1.10 × 80).

And in considering a price decrease from $1.10 to $1.00, we have

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} = \frac{\frac{2}{8}}{\frac{1}{11}} = 2.75$$

and total expenditure increases from $88 to $100 as price decreases. In both cases the value of E is greater than unity, indicating price-elastic demand.

3.4.3 Unitary Price Elasticity

A good is classified as being of unitary price elasticity when, as price changes, the proportional change in the quantity demanded exactly equals the proportional change in price. In these circumstances, total expenditure on the good remains constant for both price increases and price decreases.

In calculating the proportional changes in price and quantity, different numerical results will be obtained if the price and quantity before the change are used as a base compared with the price and quantity after the change. If the price of a glass of beer rises from 75 cents to $1, we consider that to be a 33 per cent increase; if the price of a litre of beer were to fall from \$1 to 75 cents we would consider that to be a 25 per cent decrease. Thus the calculation of elasticity using the formula would depend upon whether a price decrease or a price increase were being considered. When demand is highly price-elastic, or highly price-inelastic, and small changes only are being considered, the exact value of elasticity will not change significantly. If the price elasticity of demand is approximately unitary, however, a price increase or decrease could yield values slightly less than 1 or slightly greater than 1.

One way of resolving the problem of which price and quantity to use in calculating elasticity is to use the average of both prices and the average of both quantities in calculating the change in price. For example, the relative price change $\Delta P/P$ would become

$$\frac{\Delta P}{\frac{(P_1+P_2)}{2}} = \frac{P_1-P_2}{\frac{(P_1+P_2)}{2}}$$

Similarly, the relative quantity change $\Delta Q/Q$ would become

$$\frac{\Delta Q}{\frac{(Q_1+Q_2)}{2}} = \frac{Q_1-Q_2}{\frac{(Q_1+Q_2)}{2}}$$

and price elasticity of demand

$$E = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}} \text{ would become } E' = \frac{\frac{Q_1-Q_2}{\frac{(Q_1+Q_2)}{2}}}{\frac{P_1-P_2}{\frac{(P_1+P_2)}{2}}}$$

Thus in the beer example, the change in the price of the litre of beer is approximately 29 per cent whether we are considering a price increase or decrease. That is

$$\frac{\Delta P}{\frac{(P_1+P_2)}{2}} = \frac{P_1-P_2}{\frac{(P_1+P_2)}{2}} = \frac{1.00-0.75}{\frac{1.75}{2}} = \frac{0.25}{0.875} = 29\%$$

By using such a formula it can be shown that if a demand curve is of unitary price elasticity, total expenditure on the good remains constant for both price increases and price decreases.

Consider now the demand schedule and curve for beer as shown in Table 3.4 and Figure 3.4.

Table 3.4 Nightly demand and expenditure schedules for beer in a local bar

Price (\$)	Quantity	Total expenditure
1.00	100	100
0.80	125	100
0.50	200	100
0.40	250	100
0.20	500	100
0.10	1 000	100

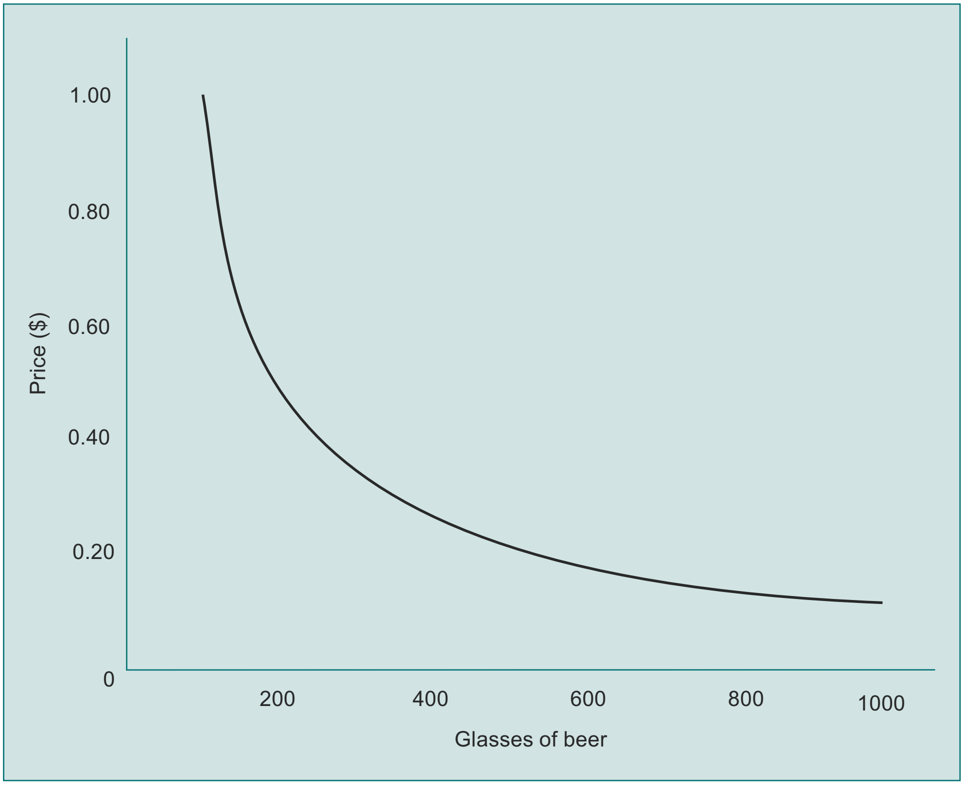


Figure 3.4 Demand curve for beer

In considering a price decrease from \$1.00 to \$0.80, the increase in the quantity demanded is 25 glasses (125 − 100). Using the standard formula, price elasticity of demand is calculated to be

$$E = \frac{\frac{125-100}{(100+125)\div 2}}{\frac{1.00-0.80}{(1.00+0.80)\div 2}} = \frac{\frac{25}{112.5}}{\frac{0.2}{0.9}} = \frac{\frac{1}{9}}{\frac{1}{9}} = 1$$

For a second incremental change along the demand curve, namely a price increase from \$0.10 to \$0.20, the decrease in the quantity demanded is 500 glasses (1000 − 500), and the price elasticity of demand is calculated to be

$$E = \frac{\frac{1000-500}{(1000+500)\div 2}}{\frac{0.20-0.10}{(0.20+0.10)\div 2}} = \frac{\frac{500}{750}}{\frac{0.1}{0.15}} = \frac{0.67}{0.67} = 1$$

3.4.4 Variation in Price Elasticity of Demand

Price elasticity of demand is not constant throughout the length of even a straight-line demand curve. At the top of the curve, price elasticity of demand is high (approaching infinity as the curve approaches the price axis), it decreases down the length of the curve, it passes through unity, and at the lower end of the curve, it becomes very low (approaching zero as the curve approaches the quantity axis).

The straight-line demand curve in Figure 3.5 shows that when price is \$20 nothing would be purchased and at zero price 100 units would be purchased.

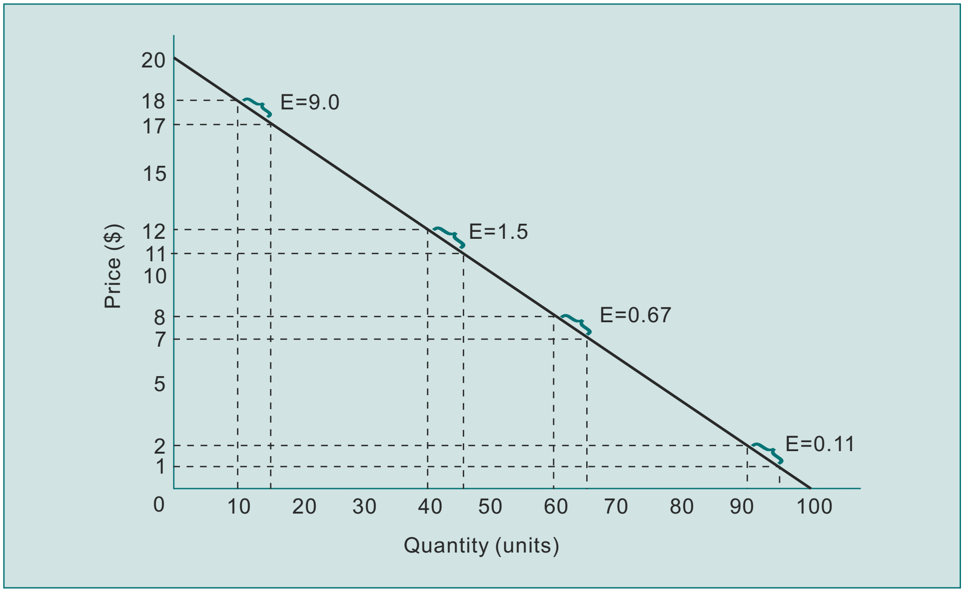


Figure 3.5 **Elasticity along a linear demand curve**

Table 3.5 contains various calculations of price elasticity of demand for that demand curve and shows that the value of E decreases as one moves along the linear demand curve from left to right.

Table 3.5 **Elasticities along a linear curve**

P	ΔP	ΔP/P	Q	ΔQ	ΔQ/Q	E=ΔQ/Q ÷ ΔP/P
18	1	1/18	10	5	1/2	9
17			15			
12	1	1/12	40	5	1/8	1.5
11			45			
8	1	1/8	60	5	1/12	0.67
7			65			
2	1	1/2	90	5	1/18	0.11
1			95			

It is clear that at some points on the demand curve the elasticity of demand is equal to one. This occurs in this example where P = 10, Q = 50. On a linear demand curve that touches both axes, as shown in Figure 3.6, the point of unitary price elasticity of demand is always the mid-point of that line. Over the range of the demand curve to the left of the mid-point, E > 1 approaches infinity as the quantity demanded approaches zero. Over the range of the demand curve to the right of the mid-point, E < 1 approaches zero as P approaches zero. Figure 3.6 shows these points.

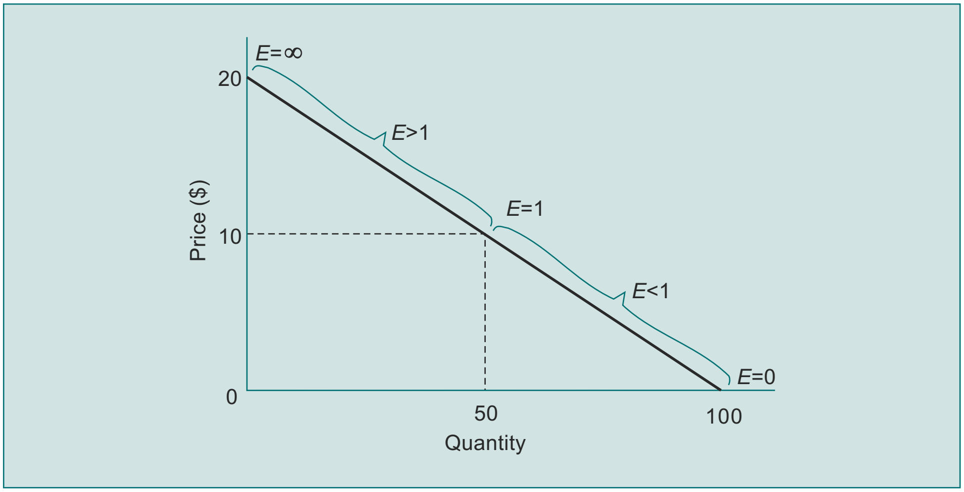


Figure 3.6 Linear demand curve

There are three special types of demand curve where price elasticity of demand does not vary along the length of the curve but instead remains constant. This happens when the whole demand curve is perfectly inelastic, perfectly elastic, or of unitary elasticity.

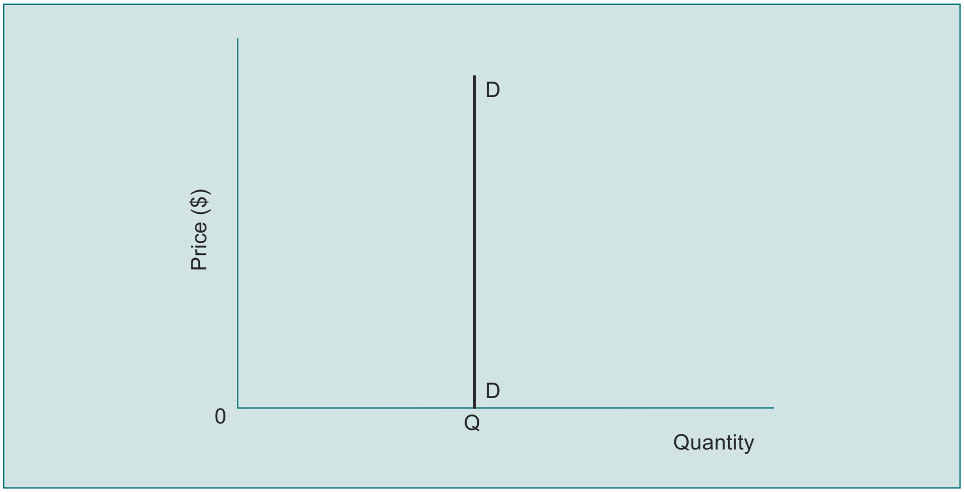


Figure 3.7 Perfectly inelastic demand curve

In Figure 3.7, the demand curve is a vertical line. The quantity demanded in this case remains the same (i.e. 0Q) at each price level. Price elasticity of demand is zero at every point on this curve, since no change in quantity demanded takes place, and in the formula for calculating price elasticity:

$$\frac{\Delta Q}{Q} = 0$$

This demand curve is said to be perfectly inelastic. This type of demand curve would apply only to a good or service that is independent of price. For an individual suffering from chronic heart disease, an operation involving the transplant of a new heart may be the only hope of survival; the individual's demand for such an operation may be represented by a perfectly inelastic demand curve.

In Figure 3.8, the demand curve D_1D_1 is a horizontal line. This type of demand curve will be met when we analyse the theory of the competitive firm. It could describe the demand for one farmer's wheat output, i.e. he can sell as much as he can produce at the price 0P, the going market price, but nothing at any price above 0P. He would never consider any price below 0P because the market will take all he can produce at 0P. In the formula for calculating price elasticity, $\frac{\Delta P}{P}$ is zero, and so price elasticity of demand is infinite. This demand curve is perfectly elastic.

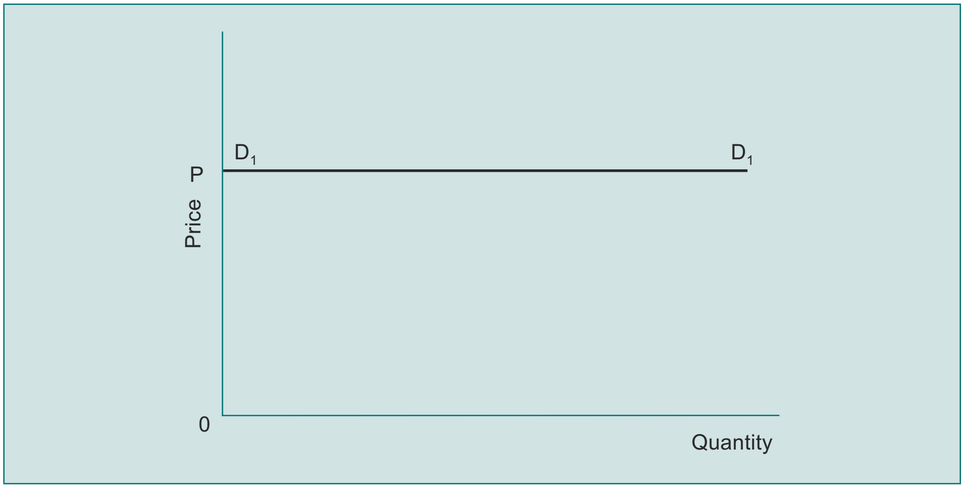


Figure 3.8 Perfectly elastic demand curve

The demand curve D_2D_2 shown in Figure 3.9 is said to be of unitary elasticity. The reason is that price elasticity of demand is equal to one at all points on the curve. Any given proportionate change in price will be matched by an equal proportionate change in quantity demanded, and total expenditure on the good will remain unchanged at each price.

It is misleading to judge the price elasticity of demand for a good by looking purely at the steepness of its demand curve. Everything depends on the scales used on the axes. For example, consider Figure 3.10. Demand curve D_1D_1 is steeper than demand curve DD, but it would be wrong to infer that D_1D_1 was a more inelastic demand curve than DD: in fact, curves DD and D_1D_1 are identical demand curves and only the quantity axis is represented differently. Since the steepness of the curves depends on the units denoted on the axes, and since quantities are measured at 100-unit intervals in Figure 3.10b compared with 50-unit intervals in Figure 3.10a, it follows that D_1D_1 will be steeper than DD. But in both cases, exactly the same quantities would be demanded at each price.

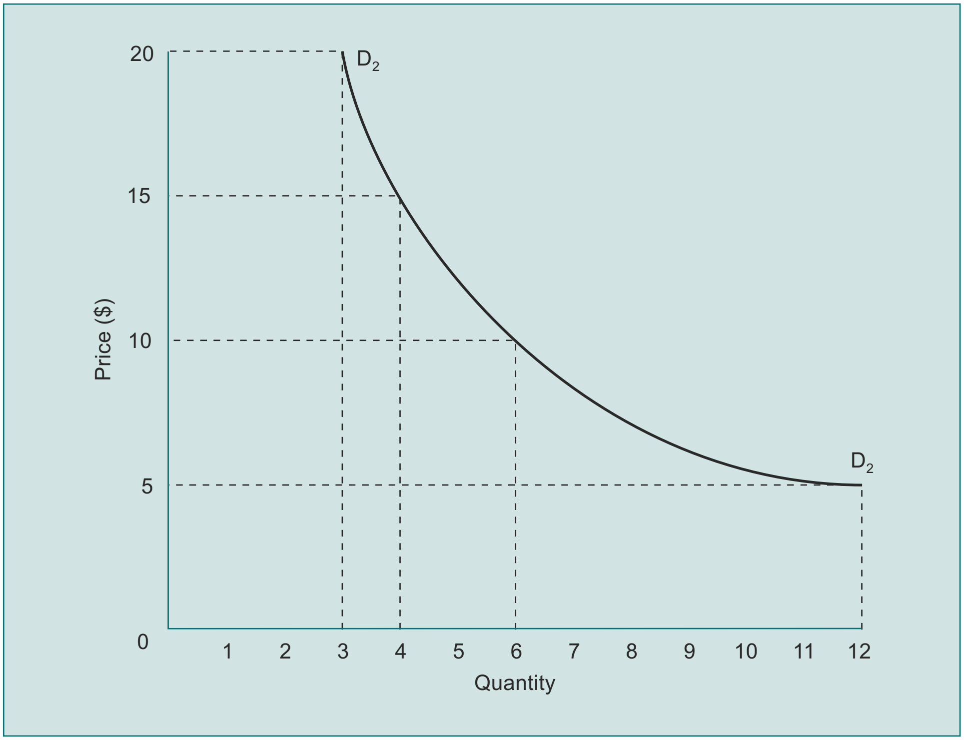


Figure 3.9 Demand curve of unitary elasticity

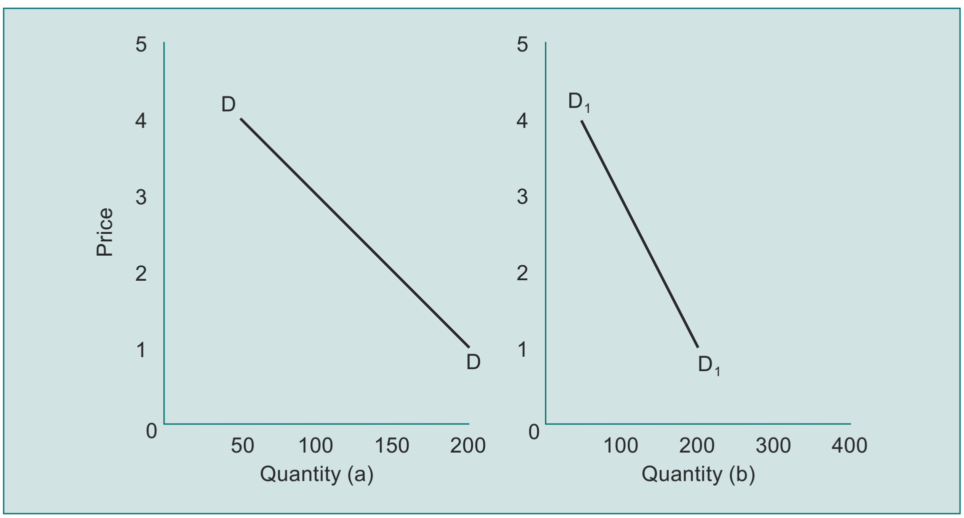


Figure 3.10 Drawing demand curves with varying axes

Goods have certain characteristics that determine whether their demand is price-elastic or price-inelastic. The more numerous and the closer are the substitutes for a good, the more elastic the demand for that good tends to be. Experience suggests that the demand for some brands of cigarette is price-elastic, since a small increase in the price of those brands leads to buyers switching purchases to other brands that are regarded as close substitutes. The demand for certain other cigarette brands may, however, be price-inelastic: individuals who buy such brands are highly loyal and would not smoke alternative brands even if the price of their normal brand were to increase slightly, *cet. par.* The demand for cigarettes as a general category of good is highly inelastic. Since few substitutes exist, a rise in the price of all brands of cigarette normally results in a relatively small decrease in the total quantity of cigarettes demanded.

The demand for goods that are necessities tends to be price-inelastic, whereas the demands for luxury goods tend to be price-elastic. The demand for salt is highly inelastic because salt is regarded by most people as a necessary part of their diet. Changes in the price of salt have little effect on the quantity demanded. In contrast, the fall in the price of package holidays in the past two decades has led to a significant increase in the quantity demanded, indicating price elasticity.

The demand for goods that account for small proportions of an individual's budget tends to be price-inelastic, whereas the demand for goods that account for a high proportion of an individual's budget tends to be price-elastic. Most people send a number of letters each year, but their total expenditure on stamps is only a small proportion of their income. The demand for stamps is price-inelastic since a small rise in the price of stamps leads to a relatively small decrease in the quantity of stamps demanded. Conversely, for low-income families, who spend a high proportion of their budget on food, a general rise in the price of food often leads to such families cutting back significantly on food purchases.

3.4.5 Other Elasticities of Demand

Other demand elasticities exist besides price elasticity of demand. A change in the price of one good (B) can affect the quantity demanded of another good (A). The measure of this responsiveness is known as *cross elasticity of demand* and is measured as

$$\frac{\frac{\Delta Q_A}{Q_A}}{\frac{\Delta P_B}{P_B}}$$

This ratio can vary from plus infinity to minus infinity. *Complementary goods* have a negative cross elasticity, i.e. an increase in the price of one good will lead to a reduction in the quantity demanded of the complement. *Substitute goods* have a positive cross elasticity, i.e. an increase in the price of one good leads to an increase in the quantity demanded of the substitute. If the demand for a good is independent, the cross elasticity will be zero. Suppose, for instance, that the price of butter increased by 10 per cent and it was observed that the demand for bread decreased by 20 per cent. The cross elasticity would be $-20/10 = -2$, and the negativity of this value indicates that bread and butter are complements.

As income changes, the amount of a good that is demanded may also change, *cet. par.* The measure of the responsiveness of quantity of a good demanded to changes in income is called *income elasticity*, where

$$\text{Income elasticity} = \frac{\frac{\Delta Q}{Q}}{\frac{\Delta \text{Income}}{\text{Income}}}$$

There are five important values for income elasticity:

(a) *negative*: the quantity of the good demanded decreases as income increases;
(b) *zero*: the quantity of the good demanded does not change as income increases;
(c) *between zero and one*: the quantity of the good demanded increases by a smaller proportion than the increase in income;

(d) *unitary*: the quantity of the good demanded increases in proportion to the increase in income; and
(e) *greater than one*: the quantity of the good demanded increases by a greater proportion than the increase in income.

It has been observed that as incomes rise, the expenditure on housing tends to rise by a greater proportion, i.e. the income elasticity of demand for housing is greater than one. This means that as people earn higher incomes they tend to spend larger proportions on housing. Goods with negative income elasticity are known as inferior goods, for example cheap cuts of meat.

3.5 The Theory of Consumer Behaviour and the Real World

Many of the concepts introduced in this module are used in the real world by businesses and government. It is important for companies (businesses and firms) to have some idea of likely demand for their products and services not only next month or next year but also, say, five years from now. In this way they can plan the number and location of factories or buildings, hire and train labour, and order the necessary plant and equipment.

The first step is to identify the characteristics of the type of people who consume the product in question. Are they young or old, on high income or low income, male or female, home owners or renters, car drivers or public transport users? The list can be extended indefinitely. Which characteristics are important will depend on the product in question: if you are in the retirement home business you won't be overly concerned with the spending patterns of young newlyweds; but if you are in the baby-food business, you will be concerned with the family plans of newlyweds.

Having identified major consuming groups, the next step is to estimate the elasticities we have discussed, namely price, income and cross elasticities (remembering both substitutes and complements). How are the incomes of the targeted populations likely to change? How are the prices of competing or complementary products likely to change? Are the targeted populations likely to increase or decrease, and by how much?

It is only when all such data have been collected and analysed that realistic estimates can be made as to whether the demand curve for the product will be shifting to the right, to the left, or remaining where it is. As someone formulating long-term strategy for a company, you would have to worry about what your competitors will be doing, but they undoubtedly will be making similar calculations about future demand.

Predicting demand, of course, is a hazardous game but most companies believe, with some justification, that such predictions are better than random guesses. Predictions may be wildly off the mark for a variety of reasons outside the control of the company. A world recession, by lowering incomes, can cause demand curves to move significantly to the left. A war can shift the demand curve to the right for many war-related products, but if accompanied by threats of terrorism can shift the demand for air travel sharply leftwards. Unanticipated changes in tastes, new inventions, and the imposition of tariffs and quotas all make demand estimation difficult. However, for companies to have the right investments in the right place at the right time requires forward planning, and sensible forward planning in turn requires estimates, however inaccurate, of future demands.

Another major user of the concepts discussed is government. As a supplier of certain goods and services, such as public transport, roads, schools, universities, police protection, fire services, hospitals, a judicial system and national defence, the government is in a position similar to that faced by private companies. It too has to make estimations of future demand and to include them in election manifestos. Indeed, one of the major differences between parties in power and those attempting to displace them are often the promises of increased quantity and quality of publicly provided goods and services without too much emphasis on how those expenditures will be financed.

However, there is one area of government activity for which there is no private company parallel, and that is the levying of taxes. The imposition of income taxes affects households' disposable incomes, i.e. how much of their earnings will remain free for spending after taxes have been deducted. A secondary, but not insignificant issue of income taxes, especially when they are high, is how such taxes will affect work effort. If you are living in a country with high *marginal tax rates* (i.e. where the proportion of income tax paid on the last dollar earned is substantial) and you are offered a higher-paid job that requires additional effort on your part, you may be unwilling to accept it since it may yield only a minor increase in your disposable income. Indeed, if you feel the tax structure too punitive, you may emigrate and thereby remove yourself as a valuable resource from your native land.

Taxes on goods have a dual effect: they reduce real income and they make the taxed goods relatively more expensive and consequently less attractive as purchases. The imposition of such taxes can also have a significant impact on the distribution of real income within a nation. It is well proved that excess consumption of alcohol and tobacco can have serious effects on people's health. It is also well known that households in low-income groups spend proportionately more of their incomes on tobacco and alcoholic beverages than people in higher-income groups. In addition, for many people in low-income groups the demand for tobacco and alcohol is price-inelastic. Consequently, raising taxes on tobacco and alcohol can significantly reduce the real incomes of people whose living standards are already low. Thus, despite the good intentions of legislators who wish to persuade people to give up the 'evils of tobacco and alcohol', imposing taxes may be counter-productive towards that end. In some countries 'bad' goods are prohibited by law, for example alcohol in Saudi Arabia. The manufacture and sale of alcohol was also banned in the USA in the early twentieth century; demand remained, however, and the Mafia stepped in as illegal suppliers. When prohibition was repealed in the USA in 1933, legal liquor returned, prices fell (but the Mafia remained). Today we face similar legislation with regard to drugs. Some nations have a liberal attitude; others are highly punitive. Many people argue that drugs, like alcohol and tobacco, are goods that people want and, independent of legislation, will therefore continue to be demanded and supplied, albeit illegally. As a consequence, they argue, the use of drugs should be allowed, according to consumer preferences. The appropriate role of government, they maintain, is providing information about the potential impact of these 'bads' on people's health, hopefully persuading them to change their tastes and preferences.

The government also has, in its armoury of tools, incentives (as well as tax disincentives) to change individuals' behaviour. These take the form of transfers and grants. Transfers include income supplements, food stamps, children's allowances, 'free' school meals, 'free' health services, and subsidised housing. Grants take the form of transfers of cash to local governments or other authorities, payments to university students, funds to universities and colleges to carry out research, and payments to businesses to locate their operations in specified areas. Implicitly all such schemes are either income supplements expanding budget constraints or reductions in the prices of specified goods and services. How effective such schemes are will be affected by the income and price elasticities we have been studying.

Do we need to understand the theory of consumer behaviour to be a rational consumer? The answer is 'no'. People learn by doing; people intuitively understand that consuming more of one good necessitates consuming less of another; people live with budget and time constraints – there is not enough income or time to enjoy all goods and services – and they realise that choices have to be made. What the theory does is to provide a good description of people's behaviour.

In addition, the theory is useful for making predictions about demand and provides useful tools for government legislators implementing tax–transfer policies to increase society's well-being. You don't have to be an economist to maximise utility from the way you spend your income on goods and services, just as you don't have to be a physics graduate to be a good snooker player.

Learning Summary

The keystone of the market mechanism is the determination of prices. To understand this mechanism requires you to appreciate the interaction between demand and supply. This module has introduced you to the first of these and you have learned to distinguish between movements along a demand curve and shifts in a demand curve, and the relevant factors involved in both.

The importance of the concept of elasticity is reflected in its relationship to total expenditure, a concept vital to both firms and governments in setting prices and taxes. You now understand the meaning of equilibrium conditions for maximising consumers' utility and how expenditure on goods and services changes as prices, incomes and other related factors change. Such an understanding is a necessary condition for grasping the meaning of 'economic efficiency' in a market economy (Module 6).

Review Questions

Multiple Choice Questions

3.1 The concept of demand is described by

I. how fully a good satisfies an individual's wants.

II. how much of a good an individual would be prepared to buy.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

3.2 In a household with a given food budget, factors that will affect how fully the family's wants for food are satisfied are

I. the prices of foodstuffs.
II. the prices of non-foodstuffs.
III. total household income.

Which of the following is correct?

A. I only.
B. I and II only.
C. I, II and III.
D. Not I, not II, not III.

3.3 A traveller arrives at an airport and discovers that the taxi fare to the city centre is $5.00. The additional economic variables that will influence the average individual's decision whether or not to hire a taxi are

I. the price of the airport bus to the city.
II. the individual's income.
III. the price of a new taxi.

Which of the following is correct?

A. I only.
B. I and II only.
C. I and III only.
D. I, II and III.

3.4 In some countries, such as Saudi Arabia, the sale of alcohol is prohibited. Which of the following is correct? It can be concluded that

A. in those countries no individual has any demand for alcohol.
B. law makers in those countries believe that the consumption of alcohol is harmful to society.
C. there are no benefits from drinking alcohol.
D. individuals collectively in those countries have no demand for alcohol.

3.5 A well-known football club is considering allowing old-age pensioners (people over 65 years of age) to attend home games for $2 instead of the normal ticket price of $10. If total ticket receipts were to increase as a result of adopting such a policy, which of the following is correct? It follows that

A. old-age pensioners, as a group, will have more income remaining to spend on other goods and services.
B. the demand for tickets by old-age pensioners is more price-elastic than the demand for tickets by all other age groups taken together.
C. total revenue from ticket sales would increase if all ticket prices were reduced to $2.
D. the increase in ticket revenue from new fans more than offsets the loss in revenue from old-age pensioners, who always attended games but now pay $2 instead of $10.

3.6 On a desert island, only two foods are available, fish and coconuts. On some days a castaway eats well; on other days he goes hungry. On any given day

I. the higher the marginal utilities of fish and coconuts the hungrier he is.

II. the lower the marginal utilities of fish and coconuts the hungrier he is.

III. when the marginal utilities are zero, his total utility is zero.

Which of the following is correct?

A. I only.

B. II only.

C. I and III only.

D. II and III only.

3.7 A person spending his income to yield maximum satisfaction buys hot dogs at 25 cents each and glasses of beer at 50 cents each. He derives 40 utils (units of utility) from the fourth and last hot dog he buys. How many utils does he derive from the last glass of beer he buys in the same period?

A. 20.

B. 40.

C. 80.

D. 160.

3.8 In a village in Ireland in the nineteenth century, all income was earned by producing linen and all family income was spent on food. When the price of meat fell, family F bought fewer potatoes and said it was better off. Family G said it was neither better nor worse off in the same circumstances. Assuming the 'law of diminishing marginal utility' to hold, it follows that

I. family F's marginal utility of meat must have fallen.

II. family G must not have consumed any meat either before or after the price change.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

3.9 A consumer's typical demand curve is downward sloping to the quantity axis because

I. at lower prices the consumer buys more of the good in question in place of other goods that are now relatively more expensive.

II. at lower prices the consumer can buy all she bought of the good at the higher price, and with the money left over she can still buy more.

Which of the following is correct?

A. I only

B. II only

C. Both I and II

D. Neither I nor II

3.10 An individual's demand curve for beer could shift because of

I. a change in the price of beer.
II. a change in the price of wine.
III. a change in the individual's income.

Which of the following is correct?

A. I only.
B. I and III only.
C. II and III only.
D. I, II and III.

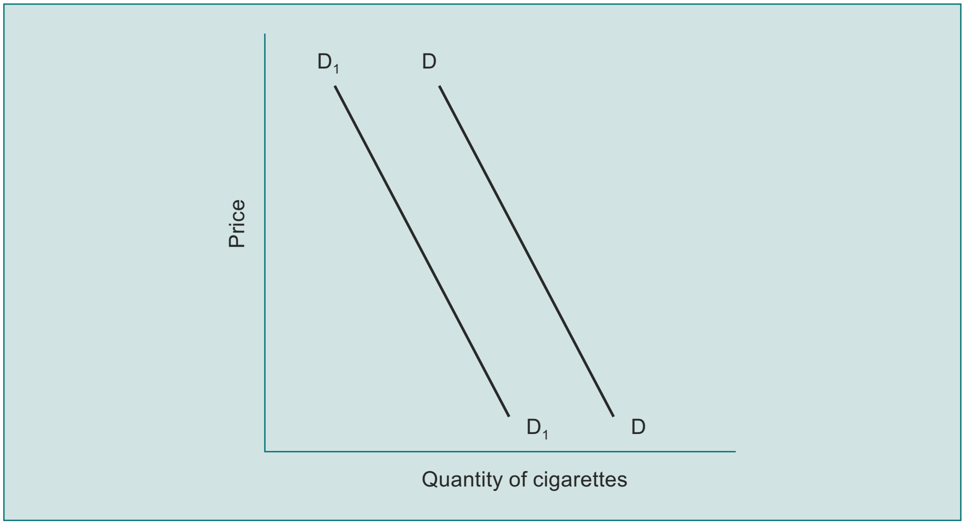


Figure 3.11 Demand curve for cigarettes

3.11 In Figure 3.11 the demand curve for cigarettes could have shifted from DD to D_1D_1 (assuming cigarettes not to be an inferior good) because of

I. the BMA's report suggesting a link between cigarette smoking and premature death.
II. a decrease in the price of cigars and pipe tobacco (cigar and pipe smoking being substitutes for cigarette smoking for many people).
III. an increase in smokers' income.

Which of the following is correct?

A. II only.
B. I and II only.
C. I and III only.
D. I, II and III.

3.12 If beef is a substitute for lamb and mint jelly is a complementary good to lamb, which of the following is correct?

An increase in the price of lamb, *cet. par.*, will lead to

A. a decrease in the quantity of beef demanded.
B. an increase in the quantity of beef demanded and an increase in the quantity of mint jelly demanded.
C. a decrease in the quantity of mint jelly demanded.
D. an increase in the quantity of mint jelly demanded and a decrease in the quantity of beef demanded.

3.13 In travelling about London, most people use either the underground train system or the bus (which are substitutes for one another). Suppose all underground train fares were doubled but bus fares remained unchanged. How would the underground train fare increase affect the total fare revenue?

A. It would increase for buses but might increase or decrease for the underground.
B. It would increase for the underground but might increase or decrease for buses.
C. It would increase for both underground and bus travel.
D. It would increase for the underground and remain unchanged for buses.

3.14 The local Community Centre operates so as to raise just enough revenue to cover its costs. Its revenue comes from a membership fee of $5 and a price of 50 cents for its nightly disco in the main hall. Attendance at the disco results in overcrowding of the hall. In order to reduce the crowding without increasing total receipts, if disco demand is elastic and membership demand is inelastic, the Centre should do which of the following:

A. raise both the membership fee and the disco price.
B. lower both the membership fee and the disco price.
C. raise the disco price and lower the membership fee.
D. lower the disco price and raise the membership fee.

Case Study 3.1: The Price of Potatoes

This case is an application of demand theory. Before you tackle it you should understand:

a. linear demand curves
b. total revenue
c. elasticity of demand
d. *ceteris paribus*
e. movements along a demand curve
f. shifts of a demand curve
g. what is meant by ‘estimating a demand curve’.

The case demonstrates:

- why it is important to have information about the shape of a demand curve when setting the price of a good;
- the consequences of price changes for demand curves of different elasticities; and
- that many issues in economics are resolved only when theories/hypotheses are subjected to empirical tests.

The Subsidy to Farmers

In a certain country, agriculture is subsidised by the government. The market for potatoes is run by its Potato Marketing Board, which acts as a buffer between potato growers and the free market. The Board advises each farmer on how much acreage he should allocate to potatoes and guarantees a price for all potatoes produced. If a farmer does not like the price offered, he can grow some other crop; the Board will then recommend to some other farmer that he increase potato acreage. In this way the Board has a fairly accurate idea of what the overall potato yield will be for each year, subject to variations in weather.

Each farmer must sell his potatoes to the Board. The Board in turn sells the potatoes to wholesalers, who in turn sell them to retailers, who in turn sell them to the public. Most farmers are in favour of this method of setting price and acreage because it protects them from free-market price fluctuations.

Last Year's Performance

Last year the Board had two million bags of potatoes that were produced on the basis of the recommendations and price offered the year before. It sold these potatoes to wholesalers for €2 per bag. The Board was rather pleased with its performance; its revenue of €4 million from the sale of potatoes more than covered its payments to farmers and its own running costs. In addition, no potatoes remained unsold and wholesalers did not come back asking for more potatoes. It appeared that the supply of, and demand for, potatoes was in balance.

How the Potato Marketing Board Got into Trouble

In planning this year's output, the Board increased both acreage and price offered to farmers. The result was that this year the Board will have three million bags of potatoes to sell. The Board decided to put the three million bags up for sale at the same price as last year. However, the Chief Economic Adviser to the Cabinet heard about this and pointed out that this would lead to one million bags of potatoes being unsold if demand conditions were similar to last year. Furthermore, the Chief Adviser recommended that, since the potatoes had now been produced, the Board should sell them at the price that maximised revenue from sales. The Cabinet agreed with his reasoning and instructed the Chairman of the Board to set the price this year to maximise the revenue from the sale of potatoes.

In some panic, the Chairman employed two teams of economists to find out how, if at all, the price should be altered this year to maximise revenue. The two teams were instructed to estimate the national demand curve for potatoes and to recommend what action should be taken on the price. On reading the reports of the two groups, the Chairman found to his surprise that they had come up with different answers. Group A recommended that the price should be lowered to €1.50 per bag; Group B recommended that the price be raised to €3 per bag. On examining the reports in more detail, the Chairman discovered that the critical difference between them was the demand curves that they had estimated. These two demand curves are shown in Figure 3.12.

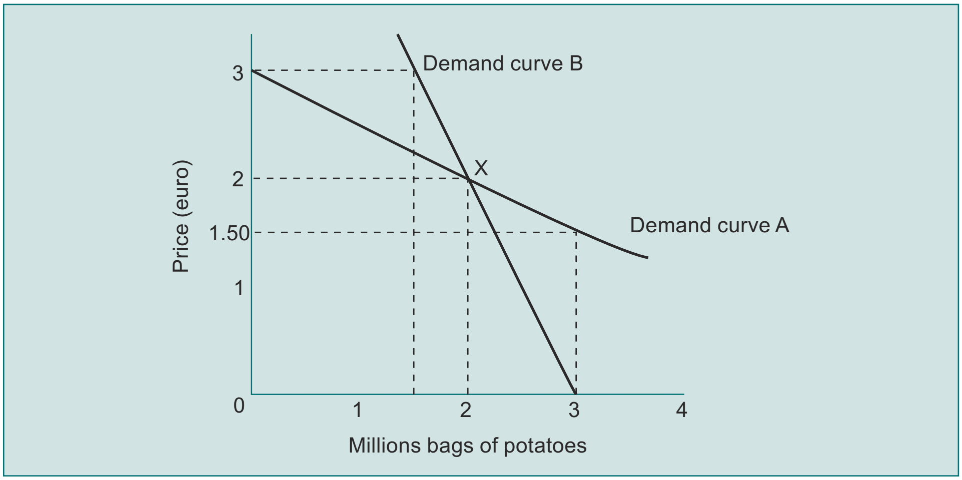


Figure 3.12 The demand for potatoes

The Chairman realised immediately that any price change he put into effect could be disastrous if he followed one team's advice and that team's demand curve was the wrong one. Furthermore, he noted that while the recommended price of €1.50 would result in all 3 million bags of potatoes being sold, the recommendation of Group B would result in only 1.5 million bags being sold.

1 Consider the following:

1. If Group A were correct, what would be the effect on revenue of their recommended reduction in the price?
2. If Group B were correct, what would be the effect on revenue of their recommended increase in the price?
3. What is the difference between Demand Curve A and Demand Curve B?
4. What would be the effect of reducing the price to €1.50 if Group A is wrong and Group B is right?
5. What would be the effect of increasing the price to €3 if Group B is wrong and Group A is right?

Case Study 3.2: Cash versus Vouchers

This case study is an application of the concepts of preferences and the maximisation of utility. Before you tackle it you should understand:

a. utility maximisation
b. preferences
c. value judgements
d. budget constraint.

The case demonstrates that well-meaning people can actually make others worse off than they would be if they were left to make their own decisions.

Cash versus Vouchers

In a certain European country there has been a significant rise in the total bill for Social Security payments. In the past few years, vigorous campaigning by pressure groups on behalf of the chronically sick, the elderly, the unemployed, the blind, widows, people injured at work, and other sections of the population has highlighted their problems and brought about increases in the size and scope of Social Security payments, i.e. sums of money paid by the government to people in the above (and other) categories.

Recently, however, a number of politicians made speeches in which they criticised the allegedly comfortable lifestyle of many of the disadvantaged groups. In one newspaper article entitled 'Luxury at the Taxpayers' Expense', one politician claimed that the general public was being 'taken for a ride' as the 'supposedly underprivileged members of society squander their Social Security payments on drinking, smoking, and gambling, rather than on necessities such as food and clothing, for which payments were really intended'.

In response to such criticisms, the government decided to conduct an experiment. A group of people who were receiving Social Security payments, and who were all identical in terms of background, household size, and monthly expenditure, was divided into two test samples. Each household in the first sample continued to receive the normal Social Security payment of €150 per month, while households in the second sample received a food voucher worth €100 and a clothing voucher worth €50 every month. These vouchers could be exchanged only for food and clothing respectively, could not be sold for cash, and could not be transferred to other households. All of the households in each sample also had an earned income of €100 per month. After a trial period of three months, the expenditure patterns of the two samples were compared, as shown in Table 3.6.

Table 3.6 Expenditure patterns of two sample groups of households

	Monthly expenditure (euro)	
	Sample 1 (Cash)	Sample 2 (Vouchers)
Food	80	100
Clothing	40	50
Other goods and services	130	100
Total	250	250

The supporters of the voucher system have issued a statement to the press claiming that the experiment justifies their position, since the households who received cash spent €30 less per month on food and clothing than the households receiving vouchers, and obviously squandered the rest of their Social Security payments on 'non-essentials'.

The government, however, is still not entirely convinced about the merits of the voucher system.

I The government has asked its Chief Economic Adviser to analyse the results of the experiment and comment on the statement issued by the supporters of the voucher system. In particular he has been asked to answer the following questions (which you also should consider):

1. Under what conditions would it make no difference whether a family were given cash or vouchers?
2. What evidence is there to suggest that, had the families who received vouchers been given cash, they would have bought a different bundle of the goods and services?
3. If giving vouchers affected how people spent their incomes, is it possible to conclude that they would be better off under one scheme rather than another?
4. Is it possible to conclude that poor families should be given either cash or vouchers?
5. Is it possible to conclude how much aid should be given to poor families in any form?

Module 4

Supply

Contents

4.1 Introduction

The bulk of goods and services we consume is produced by firms. Some of these firms, like the local fish-and-chip shop, are small family-owned entities producing only a few products (such as fish-and-chip suppers). Other firms, by contrast, are enormous, producing a large variety of products – for instance ICI and Unilever. Some large firms produce a range of one product, such as car manufacturers like Toyota, Mercedes-Benz, Ford and General Motors. Other large firms produce a range of related products, such as BP, Exxon Mobil and Shell producing oil, petrol and petroleum products; IBM, Hewlett-Packard and DEC producing computers, workstations and medical instrumentation.

Some of the goods and services we consume are produced by local and national governments: clean streets, police protection, national defence, elementary education, social services and, in some countries, health services.

All those suppliers of the goods and services we enjoy face three basic problems:

(a) How much output should be produced for each time period?
(b) What price should be charged for each unit of output?
(c) What is the most efficient way to produce that output?

The analysis of supply is in many ways similar to the analysis of demand, except that it is much more complex. To derive a *market supply curve*, we must first derive a *firm's supply curve*. To derive a firm's supply curve, we must first study *productivity* and *costs*. The first step is to analyse how much output will result from different amounts and combinations of factor inputs (resources) – this is the study of productivity. Then we must discover how much it will cost to produce such outputs – this is the study of costs. Having mastered productivity and costs, we shall derive the firm's supply curve, and by summing all firms' curves in a market we shall arrive at the market supply curve. In Module 5 we shall bring together the 'two legs of the scissors', the market demand curve and the market supply curve, and analyse how the price and quantity of output are determined.

But first, productivity and costs. You may find some of this boring, because it is. However, if you find yourself in a position in a company where there is constant pressure to match or outperform competitors, this will necessitate looking at costs, which in turn will necessitate looking at productivity; reducing costs in some cases may be achieved by obtaining better prices of factor inputs; in many cases it will be achieved by increased factor productivity.

4.2 Productivity

Let us choose a simple example of a firm in the ocean salmon-fishing business. This family firm owns one fishing boat and 18 fishing rods that fit into 18 rod holders distributed around the boat. On board there is a captain, who drives the boat, and a large supply of lines, hooks, bait, landing nets, gutting knives, storage tanks etc., in other words all the capital required to catch fish. Assume, for the present, none of these capital inputs can be changed – they are a given, and we are back to our *ceteris paribus* assumption. What *can* be varied is the number of fishermen hired for each fishing trip. Table 4.1 records the postulated salmon catch with different numbers of fishermen in our example.

Table 4.1 Salmon output for one boat and varying labour inputs

Input (man-days)	Total output (kilos of salmon)
0	0
1	7
2	18
3	36
4	62
5	99
6	143
7	191
8	235
9	271
10	297
11	314
12	324
13	331
14	335
15	337
16	338
17	339
18	340
19	340
20	340

Now consider Figure 4.1, which shows that with 10 fishermen on board, the maximum catch is 297 kilos of salmon. It also shows that other outputs are possible with 10 fishermen, for example 200 kilos or 100 kilos. The output will be 297 kilos when the firm operates in an engineering-efficient fashion. If the fishermen, however, 'goof off' half the time, don't check their baits and allow lines to become tangled, the maximum quantity of salmon will not be caught; the firm will not be operating in an engineering-efficient fashion. The boat will return to the dock with fewer salmon than it could have caught, it won't earn as much revenue from the sale of salmon as it could have, and consequently the trip will not have been as profitable to the firm as it could have been; it may even run at a loss.

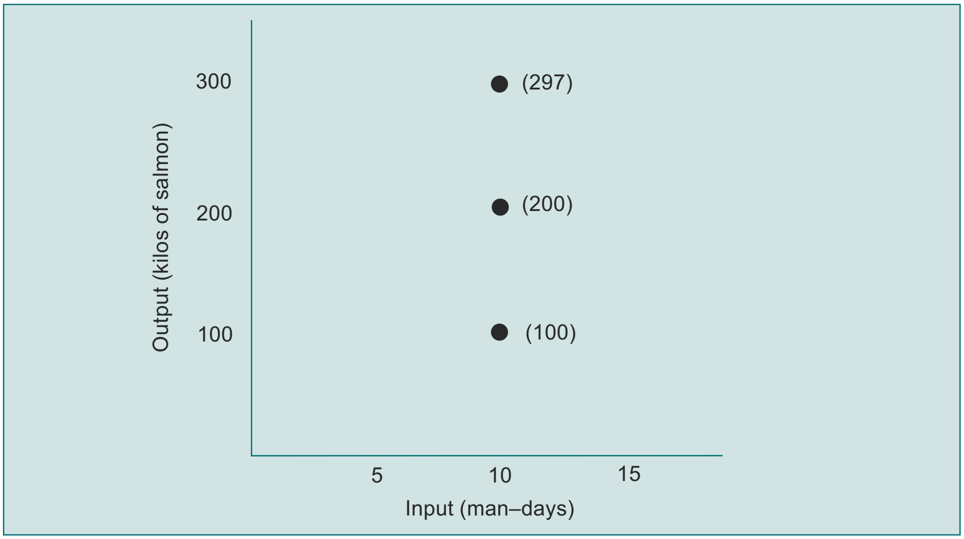


Figure 4.1 Ten fishermen fishing

What we are observing in Table 4.1 is the maximum salmon output that could be caught with varying labour inputs. Now, if you are an ocean fisherman or woman you might want to argue that catching salmon isn't that simple – you can have good days and bad days, but what is important is finding the fish. Fortunately, economists can overcome such trifling objections with a 'let us assume' statement. Therefore, let us assume the captain can always find the salmon shoal and consequently the amount of fish caught depends only on the number of fishermen aboard. Further, let us assume no 'goofing off' by the fishermen – they always perform in an engineering-efficient fashion. That's not too hard to swallow, is it? You may wonder why we keep talking about 'engineering efficiency'. As you will see, and hopefully remember from Module 1, there is a difference between engineering and economic efficiency. If no one wants to buy the salmon when the boat returns to harbour, economic resources will have been wasted: an output will have been produced that no one wants, and scarce resources with alternative employment opportunities will have been wasted despite the fact the fish were caught in a engineering-efficient fashion.

Let us return to Table 4.1. We observed that as labour inputs are added, total output increases until 18 fishermen are hired. The hiring of a 19th fisherman and then a 20th adds nothing to output. That's not surprising because of the limited number (18) of rods. The maximum amount of fish that can be caught is 340 kilos, and this is achieved when 18 fishermen are hired. It can also be seen from Table 4.1 that there is not a constant relationship between output and input.

The data from Table 4.1 are plotted in Figure 4.2. The *total product curve* traces out the *production frontier* for the firm, i.e. it shows the maximum possible output for different amounts of the varying factor input *ceteris paribus*. With a labour input of 10 man-days (10 fishermen) the output of 297 kilos lies on the frontier. An output of 200 kilos is possible, but such an output has not been produced in an engineering-efficient fashion and consequently lies within the frontier. An output of 320 kilos with 10 man-days is not feasible (given this particular boat) and consequently lies outwith the frontier.

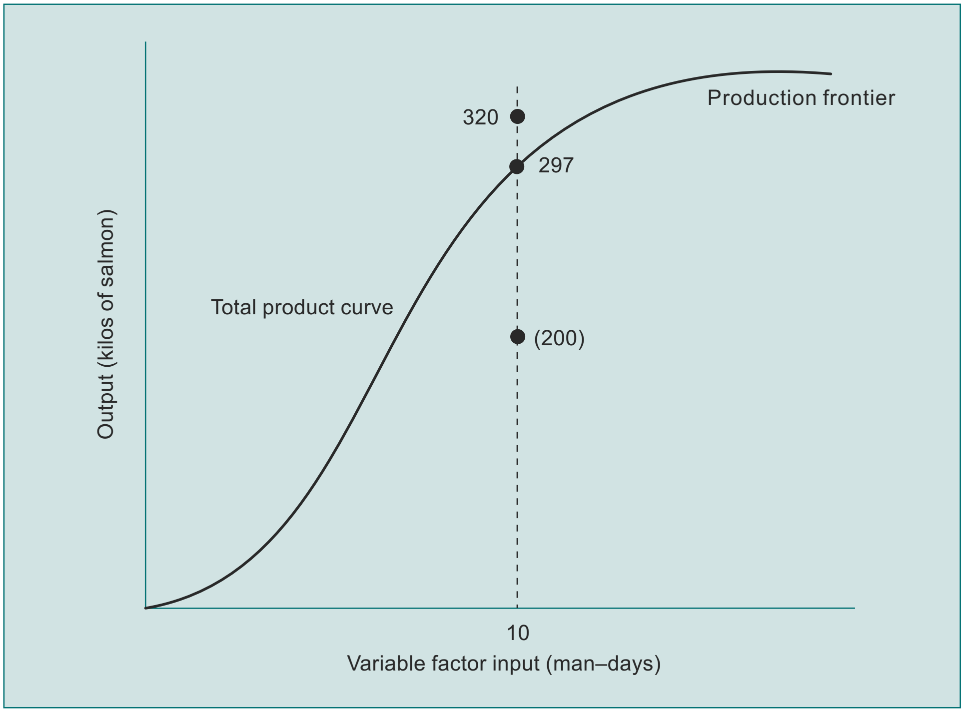


Figure 4.2 The production frontier

It is generally accepted in many areas of economics activity that production frontiers are shaped as shown in Figure 4.2. Up to a certain level of output, increases in the variable factor of production yield proportionately greater increases in output and after a certain level of output yield less than proportional increases in output. Possible explanations for the shape of the production frontier lie in the scope for specialisation of activities. In the fishing-boat example, if only one angler is employed he will have to perform all tasks: baiting hooks, putting out lines, netting fish, untangling lines etc., and so he will have very little time for fishing! As more fishermen are added, specialisation will occur. The scope for such specialisation is of course limited, and consequently as the number of fishermen keeps on increasing the number of additional fish caught falls, eventually becoming zero, so that the total product curve flattens off in its upper reaches.

This specialisation, or *division of labour*, can be observed in many business activities. Consider your local large supermarket with the butchers in the meat department, the bakers making bread and cakes, assistants putting goods on shelves and the checkout personnel and the security guards doing their jobs. If one employee attempted to perform all these tasks, sales would be close to zero. At the other extreme, a point will be reached when hiring additional checkout personnel adds nothing to sales.

The production frontier can shift only if the assumption of certain fixed inputs is relaxed. For instance, in our fishing-boat example, suppose we had a superior vessel – bigger, faster and containing electrical reels sensitised to the weight of fish hooked, so that fewer fish escaped when hooked. As you might imagine, this would also be a more costly vessel to purchase, a subject to which we shall return.

Figure 4.3 compares the production frontiers of the two vessels. For any given amount of the variable factor input (fishermen) catching salmon in an engineering-efficient fashion, more fish will be landed in the superior boat, i.e. the production frontier has shifted upwards. Conversely if you take a row-boat out to the fishing grounds on your own, with one or two friends, we could draw your production frontier[1] pretty close to the horizontal axis!

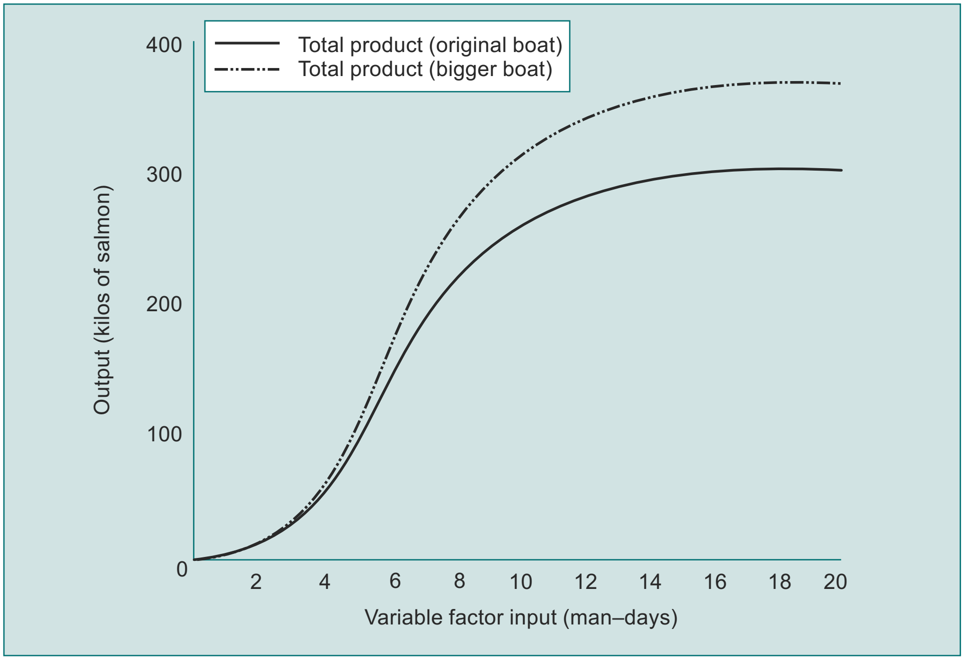


Figure 4.3 Shifting the production frontier

Using the data from Table 4.1, we can calculate two important measures: *average product* and *marginal product* of the variable factor of production (in this example, labour). This extension to Table 4.1's data is shown in Table 4.2.

[1] Readers who are familiar with production frontiers from previous economics courses may remember from their economics texts a description of production frontiers turning down after a maximum point was reached. The standard explanation is that the hiring of additional labour after a certain point – maximum output – is reached leads to the newly hired workers 'getting in the way' of other workers, causing total output to fall. This is a fallacious argument. Since the production frontier traces only maximum possible quantities of output from different amounts of labour inputs, the maximum output from, say, 18 workers can always be achieved with, say, 20 workers. Take our fishing-boat example. The owner knows from Table 4.1 that the maximum catch possible is 340 kilos of salmon, achieved when 18 fishermen are hired. Suppose some local labour agreement dictates that 20 fishermen must be hired; then the boat owner will hire 20 men but leave two on the dockside when he sets sail, so that if the 19th and 20th men would 'get in the way' of the other 18 if taken on board and reduce output below its maximum, then it is in the boat owner's interest to pay them – he has to because of the local government – but leave them on the dock.

Average product of the variable factor input (in this case, labour) is calculated by dividing total output/product (Q) by the number of units of the variable factor input (labour, L). Thus we have:

$$\mathrm{AP_L} = \frac{\mathrm{Q}}{\mathrm{L}}$$

When three fishermen are hired, the total catch is 36 kilos of salmon; and the average product per unit of labour, $\mathrm{AP_L}$, is therefore 12 kilos (36/3), on average each of the three fishermen catching 12 kilos of salmon. When ten fishermen are hired, the total output is 297 kilos and $\mathrm{AP_L}$ is 29.7 kilos (297/10). As can be seen from Table 4.2, $\mathrm{AP_L}$ increases to start with as labour input increases, reaching a maximum when nine men are hired. Thereafter $\mathrm{AP_L}$ declines as more labour is hired.

The marginal product of the variable factor input is the change in output (ΔQ) attributable to the change in the quantity of that input factor (ΔL) hired at a given level of employment of that factor. Thus we have

$$\mathrm{MP_L} = \frac{\Delta \mathrm{Q}}{\Delta \mathrm{L}}$$

We have already discussed in Module 1 the importance of marginal analysis in problems involving maximisation. In hiring the optimal amount of any variable factor input, the firm requires information on the marginal product of each factor input at each level of employment. The concept of *marginal productivity* is critical to efficiency in the theory of supply, as we shall see.

In Table 4.2 the hiring of the first fisherman results in 7 kilos of salmon caught. The addition of a second fisherman yields a total catch of 18 kilos, i.e. he adds 11 kilos to the catch. Thus $\mathrm{MP_L}$, when one fisherman is already employed equals:

$$\frac{\Delta \mathrm{Q}}{\Delta \mathrm{L}} = \frac{18-7}{1} = 11$$

Table 4.2 Total output: average and marginal product of labour

Input (L) (man-days)	Output (Q) (kilos of salmon)	Average Product of Labour (Q/L) (kilos of salmon)	Marginal Product of Labour (kilos of salmon (ΔQ/ΔL))
			0
0	0	–	
			7
1	7	7.0	
			11
2	18	9.0	
			18
3	36	12.0	
			26
4	62	15.5	
			37
5	99	19.8	
			44
6	143	23.8	
			48

Input (L) (man-days)	Output (Q) (kilos of salmon)	Average Product of Labour (Q/L) (kilos of salmon)	Marginal Product of Labour (kilos of salmon (ΔQ/ΔL))
7	191	27.3	
			44
8	235	29.4	
			36
9	271	30.1	
			26
10	297	29.7	
			17
11	314	28.5	
			11
12	324	27.0	
			7
13	331	25.5	
			4
14	335	23.9	
			2
15	337	22.5	
			1
16	338	21.1	
			1
17	339	19.9	
			0
18	340	18.9	
			0
19	340	17.9	
			0
20	340	17.0	

When seven fishermen are employed, the total catch is 191 kilos. When eight are employed, the catch is 235 kilos. Thus the eighth fisherman adds 44 kilos, and the value of MP_L, when seven fishermen are employed, is given by

$$\frac{\Delta Q}{\Delta L} = \frac{235-191}{1} = 44$$

When eight fishermen are employed the total catch is 235 kilos. This figure can also be reached by summing the marginal products of the fishermen up to the employment level, i.e. 235 = 7 + 11 + 18 + 26 + 37 + 44 + 48 + 44.

In our fishing-boat example, MP_L reaches a maximum at an employment level of seven fishermen, so that at that level of employment the addition of the seventh fisherman added 48 kilos, a number greater than any previous, or any additional, hire. When the addition of a further fisherman contributes nothing to output, $MP_L = 0$. This occurs when the nineteenth fisherman is hired – he is the one who stays behind!

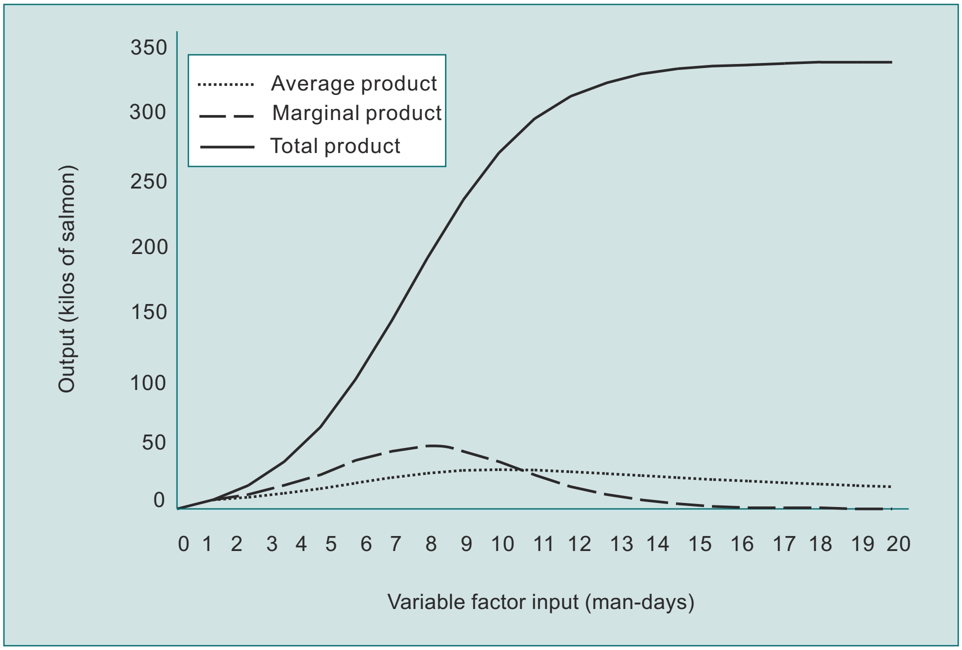


Figure 4.4 **Total, average and marginal product**

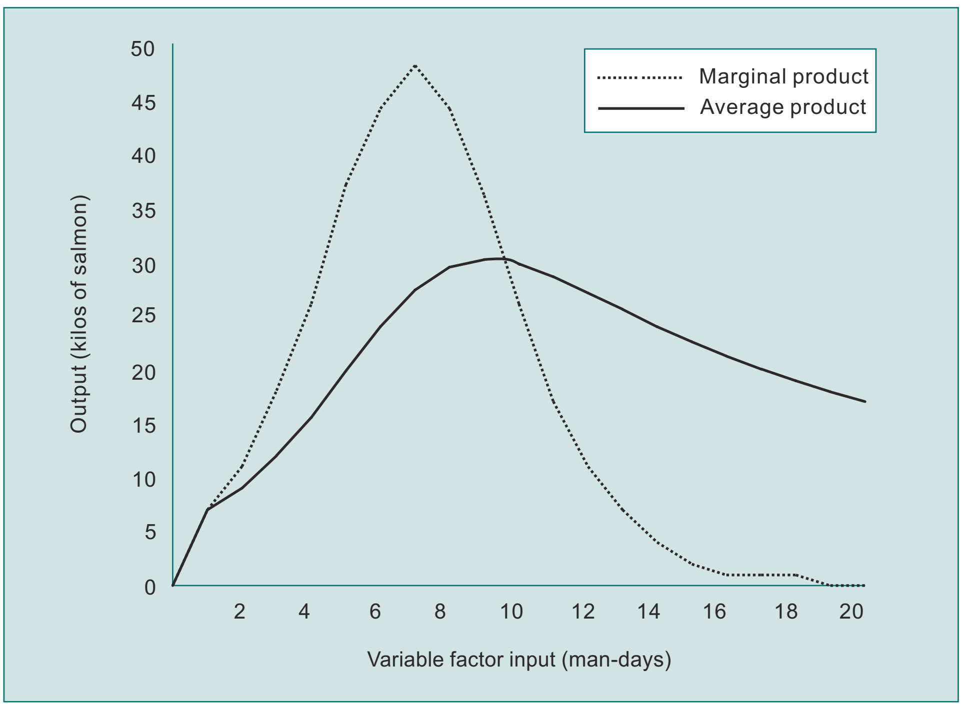


Figure 4.5 **Average and marginal product**

Figure 4.4 shows three curves: the average product and marginal product curves of labour and the total product curve from which they are derived. Figure 4.5 reproduces the average and marginal curves with the scale on the vertical axis altered.

In our analysis so far we have been dealing with physical units of output, namely kilos of salmon. For this reason the appropriate terms are *total physical product* (TP), *average physical product* (AP) and *marginal physical product* (MP). When output units are expressed in monetary terms, a different terminology will be used; thus unless otherwise stated TP, AP and MP will refer to physical units of output.

Because the AP and MP data and curves are derived from the TP data and curve, it follows that there are a number of relationships among these variables. The most important of these are:

1. When TP = 0, AP = 0
2. When AP is at its maximum, AP = MP
3. When MP is greater than AP, AP is increasing
4. When MP is less than AP, AP is decreasing
5. When the change in TP is 0, MP = 0

Point 1 is obvious, because by definition

$$AP_L = \frac{TP}{L}$$

where L = number of units of labour employed, and so if TP = 0 then

$$AP = \frac{0}{L} = 0$$

Point 5 is also obvious, for by definition

$$MP = \frac{\Delta TP}{\Delta L}$$

and so if ΔTP = 0 then

$$MP = \frac{0}{\Delta L} = 0$$

Points 2, 3 and 4 are not quite so obvious. For clarification consider the example of a basketball team consisting of five players, whose average height is 2 metres, and three reserves whose heights are respectively 2.6 m, 2.1 m and 1.4 m (the last being small but fast). As we add each player to the squad on the basis of height, the tallest first, then the next tallest, then finally the fast midget, what happens? Table 4.3 shows the results.

As the sixth player, whose height is 2.6 m, is added, the total height of all players summed increases from 10 m (5×2 m) to 12.6 m. The average height thereby rises from 2 m to 2.1 m, indicating that when the marginal height is greater than the average height, the average height will increase. When the seventh player, whose height is 2.1 m, is added, the existing average height will not change since it equals the marginal addition. When the eighth player, whose height is 1.4 m is added, the total height of the eight players increases to 16.1 m and the average falls to 2.01 m, showing how when the marginal height is less than the average height, the average height will decrease. These data are plotted in Figure 4.6.

Table 4.3 The basketball team

Players	Total height (m)	Average height (m)	Marginal height (m)
1–5	10	2.0	–
6	12.6	2.1	2.6
7	14.7	2.1	2.1
8	16.1	2.01	1.4

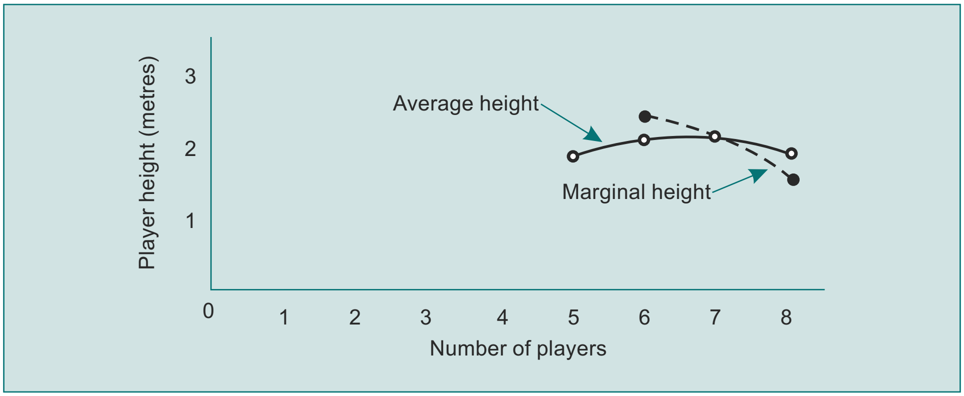


Figure 4.6 The basketball team

Now the relationship between average and marginal measures can be seen clearly. When the marginal exceeds the average, the average will increase; when the average is at a maximum, the average will equal the marginal; and when the marginal is less than the average, the average will decrease.

The marginal concept is of fundamental importance to decision making on the part of a firm. The rule that a profit-maximising firm must follow in hiring factor inputs is: continue to hire a factor of production up to the point at which the value of the marginal product of that factor just equals the unit cost of the factor. To see why this must be the case, let us return to our salmon fishing. Suppose the fishermen are paid in salmon and further suppose the going rate, or wage, is 20 kilos of salmon per day. At that rate, you, the boat owner, can hire as many fishermen as you want but no one will work for you at a lower rate. Why? Because any fisherman can always get the going rate of 20 kilos of salmon per day from any of the other boat owners who are your competitors.

Look again at Figure 4.4. How many fishermen will you hire? The answer is 10. The solution is found by using marginal analysis. Start off with any number of hires, for example eight. Does it pay to hire the eighth fisherman? Yes. Why? He will add 44 kilos of salmon to the catch. He will cost you 20 kilos, the going wage rate. Therefore, he is hired. What about a ninth? Again, yes, he will add 36 kilos and cost 20. A tenth? Sure. He will add 26 kilos and also cost 20. Now the eleventh. He will add 17 kilos to the catch but will cost 20 kilos. Therefore, although total output will increase by hiring an eleventh fisherman, the marginal cost (20 kilos) exceeds the marginal benefit or output (17 kilos). Consequently, hiring an eleventh man adds more to costs than it does to output or revenues and in this case involves a loss of 3 kilos (20 − 17).

Now note, and note well, that the hiring of 11 men will result in a total catch of 314 kilos and a total cost of 220 kilos, yielding a ‘net catch’ to you of 94 kilos. But this is not the maximum ‘net catch’ you could achieve. Hiring 10 fishermen results in a total catch of 297 kilos and a total cost of 200 kilos (10 × 20 kilos) yielding a net catch to you of 97 kilos. Thus you are 3 kilos better off by hiring 10 rather than 11 fishermen. You could estimate the net catch for all possible units of labour inputs (fishermen hired) and you should reach the same results as we have discovered using marginal analysis; the use of marginal analysis avoids all those tedious calculations.

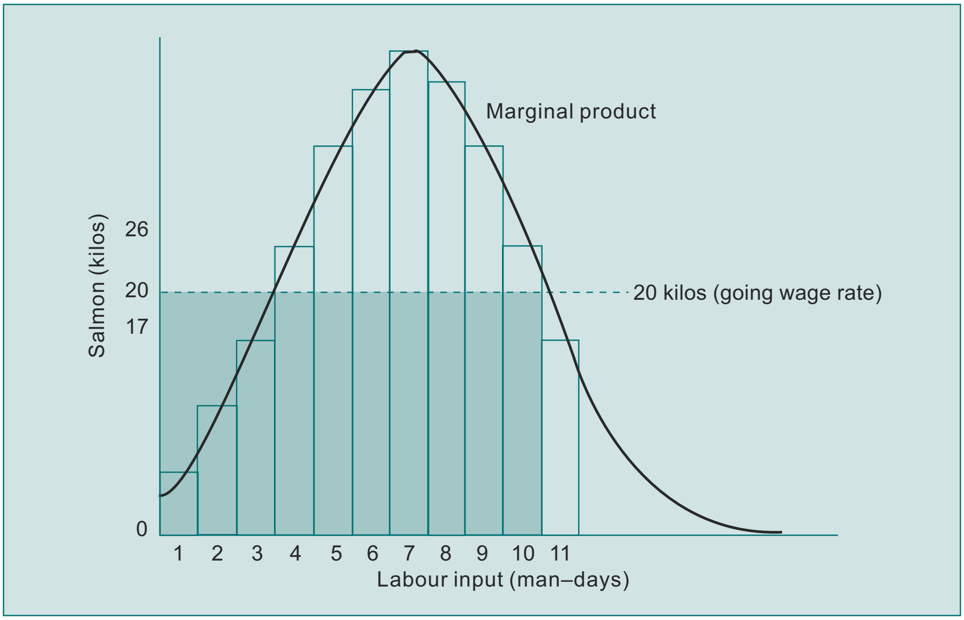


Figure 4.7 Hiring resources using marginal analysis

Maximising behaviour can be shown diagrammatically (*see* Figure 4.7). The heavily shaded area shows the labour cost of hiring 10 fishermen, i.e. 200 kilos of salmon. The marginal product curve and its intersection, when decreasing, with the marginal cost line (20 kilos) shows why the eleventh fisherman should not be hired; at that level of employment the marginal product is 17 kilos and the marginal cost is 20 kilos; stop hiring when 10 fishermen are on board.

The area under the marginal product curve up to 10 units of labour, i.e. the sum of the first 10 rectangles, is total product and consists of the salmon caught when one fisherman is hired, plus the salmon caught when the second fisherman is hired + … + the salmon caught when the tenth fisherman is hired. These 10 rectangles sum to 297 kilos of salmon. If we subtract the cost of the 10 fishermen, namely the 200 kilos shown by the heavily shaded area, from the area under the curve up to and including the tenth labour unit, we are left with 97 kilos of salmon. If 200 kilos of salmon is the return to the labour input required to catch them, the difference of 97 kilos of salmon is the return to all other factors of production also required to catch salmon.

The next step in investigating the concept of Firm supply in the short run requires the conversion of productivity data to cost data.

4.3 Costs

The total cost of producing any good equals the amount of each factor of production employed, multiplied by the price (or cost) of each unit of the factors employed. With all resources classified as either capital or labour, total cost of production equals the sum of {the number of units of capital employed times the price of that unit} plus the sum of {the number of units of a type of labour (workers), employed times the price of a unit of that labour (the wage rate)}.

$$\begin{aligned} \mathrm{TC} &= \mathrm{CP_C} + \mathrm{LP_L} \\ &= \mathrm{C_1P_C} + \mathrm{C_2P_C} + \cdots + \mathrm{C_nP_{Cn}} + \mathrm{L_1P_L} + \mathrm{L_2P_L} + \cdots + \mathrm{L_nP_{Ln}} \end{aligned}$$

This generalised equation indicates that output may require many different capital inputs and many different labour inputs.

Now we require a small, but not insignificant digression, involving another economist's assumption. As you will see, this is a quite reasonable assumption; unfortunately it is not always accepted by accountants and financial vice-presidents in companies, to their cost.

The time period within which firms make decisions present a complication that requires careful analysis. For analytical and practical considerations, it is useful to consider a firm in two distinct time periods. This distinction differs from the need to state the rate at which a firm produces output, e.g. the salmon catch per day, per week or per month. The two time periods are defined with reference to the ability of the firm to alter the quantities of the different factor inputs required to produce output.

When a firm decides to 'step up production' or 'slow down production', there are certain factors of production that cannot be altered practically in quantity. Such factors are known as *fixed factors of production*. Those factors whose numbers can be altered 'immediately' are known as *variable factors of production*. Normally capital items such as factories and machine tools are assumed to be the fixed factors of production and labour, e.g. man-days, is assumed to be a variable factor of production. The time period during which certain factors of production cannot be altered in terms of quantity hired is known as '*the short run*'. The time period during which all factor inputs can be varied is known as '*the long run*'.

During wars, many firms producing armaments have been asked by governments to increase their output. Many have been able to do so by hiring more labour so as to operate two or even three shifts instead of one. Their factories, instead of working an 8-hour day, worked 16- or 24-hour days, and even weekend shifts would become common. In other words, additional labour, i.e. more inputs of the variable factor of production labour, was hired and worked in given factories with given machines and tools, i.e. where capital as the fixed factor of production remained unaltered in quantity.

There can be circumstances, of course, when labour is not the variable factor and in which capital is not the fixed factor of production. Imagine a firm producing newspapers, which has entered into a binding agreement with a trade union in which it has to employ an agreed number of employees. Further, suppose that the firm is not constrained by the union in the number and type of computers and printing machines that it can rent. Thus an increase in demand for newspapers can be met with the fixed labour force by renting faster and more technologically advanced machines. In such an example, labour would be the fixed factor and capital the variable factor of production.

End of digression, but not of the implications of having fixed and variable input factors. Remember

$$\mathrm{TC} = \mathrm{CP_C} + \mathrm{LP_L}$$

If capital is assumed to be the fixed factor and labour the variable factor, we can rewrite the equation to signify that total cost now equals fixed cost and variable cost:

$$\mathrm{TC} = \mathrm{FC} + \mathrm{VC}$$

Back to our fishing boat. The rods, nets, reels etc. (capital) will comprise our fixed cost and the fishermen (labour) will constitute our variable cost. More simplifying assumptions are in order. The boat owner paid $14 600 for the boat. The boat will last for one year, after which it will sink and the rods and equipment disintegrate. It will go fishing each of the 365 days in the year, yielding a daily fixed cost of $40 ($14 600/365). The going wage rate for fishermen is $10 per day. None can be hired at a rate less than $10 per day and any boat owner can hire as many men as he wants at $10 per day. What we can now do is convert our productivity data into cost data, and this is done in Table 4.4.

Table 4.4 **Total cost of catching salmon**

Total product (kilos of salmon)	Fixed cost ($)	Unit of labour (man-days)	Wage rate $ (/day)	Variable cost ($)	Total cost = Fixed cost + Variable cost ($)
0	40	0	10	0	40
7	40	1	10	10	50
18	40	2	10	20	60
36	40	3	10	30	70
62	40	4	10	40	80
99	40	5	10	50	90
143	40	6	10	60	100
191	40	7	10	70	110
235	40	8	10	80	120
271	40	9	10	90	130
297	40	10	10	100	140
314	40	11	10	110	150
324	40	12	10	120	160
331	40	13	10	130	170
335	40	14	10	140	180
337	40	15	10	150	190
338	40	16	10	160	200
339	40	17	10	170	210
340	40	18	10	180	220
340	40	19	10	190	230
340	40	20	10	200	240

The table shows the total cost associated with each level of output. For example, 297 kilos of salmon cost a total of $140 to catch. The cost is composed of the fixed cost of $40 plus a variable cost of $100, i.e.

$$\begin{aligned} \mathrm{TC} &= \mathrm{FC} + \mathrm{VC} \\ &= \$40 + \$100 = \$140 \end{aligned}$$

Since an increase in output requires an increase in factor inputs, the total cost of production will increase as output increases. However, since a given increase in the amount of the variable factor input does not necessarily result in a proportional increase in output, the relationship between cost of production and output will be positive but not necessarily fixed.

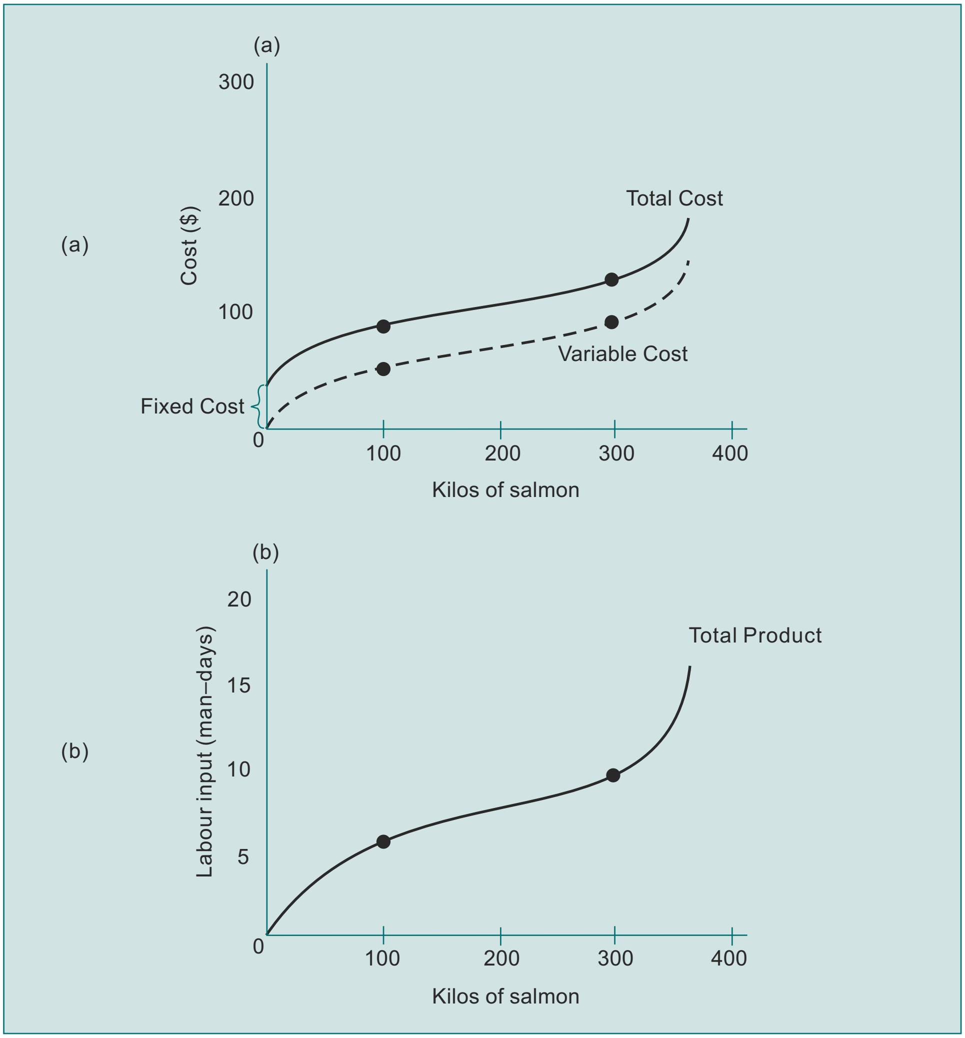


Figure 4.8 Total and variable costs related to labour productivity of salmon fishing

In Figure 4.8 output and cost data from Table 4.4 are plotted. The relationship between the various curves can be traced. For example, given our fixed factor input (the fishing boat), five units of labour input produce a catch of 99 kilos. Figure 4.8a shows that the cost of catching 99 kilos is $90, which can be broken down into fixed cost of $40 (the boat) and variable cost of $50 (five fishermen at $10 per man per day). Similarly, 10 units of labour input produces a catch of 297 kilos. The concomitant total cost is $140, i.e.

$$\begin{aligned} TC &= FC + VC \\ &= \$40 + \$100 = \$140 \end{aligned}$$

Through this example we can observe a very important relationship: the shape of the total cost and variable cost curves is determined by the shape of the total product curve. Thus given certain factor input prices, total cost of output is a function of the productivity of factor inputs.

We can now derive other relationships.

Average total cost (ATC) of output for each level of production (Q), also known as the *average cost of production*, is calculated by dividing total cost (TC) by output (Q). Thus:

$$\mathrm{ATC} = \frac{\mathrm{TC}}{\mathrm{Q}}$$

Similarly, average variable cost (AVC) is calculated by dividing total variable cost (VC) by output. Thus:

$$\mathrm{AVC} = \frac{\mathrm{VC}}{\mathrm{Q}}$$

Finally, average fixed cost (AFC) can be found by dividing fixed cost (FC) by output. Thus:

$$\mathrm{AFC} = \frac{\mathrm{FC}}{\mathrm{Q}}$$

And since TC = FC + VC, it follows that

$$\mathrm{ATC} = \mathrm{AFC} + \mathrm{AVC}$$

Using the data in Table 4.4 we can calculate ATC, AVC and AFC for each level of output. This is shown in Table 4.5.

Table 4.5 Average total, average variable and average fixed costs for salmon fishing

Output	Average total	Average variable	Average fixed cost
0	–	–	–
7	7.14	1.43	5.71
18	3.33	1.11	2.22
36	1.94	0.83	1.11
62	1.29	0.65	0.65
99	0.91	0.51	0.40
143	0.70	0.42	0.28
191	0.58	0.37	0.21
235	0.51	0.34	0.17
271	0.48	0.33	0.15
297	0.47	0.34	0.13
314	0.48	0.35	0.13
324	0.49	0.37	0.12
331	0.51	0.39	0.12
335	0.54	0.42	0.12
337	0.56	0.44	0.12
338	0.59	0.47	0.12
339	0.62	0.50	0.12
340	0.65	0.53	0.12
340	0.68	0.56	0.12
340	0.71	0.59	0.12

As output increases, AFC decreases since we are dividing a fixed amount ($40) by an ever increasing number of units of output. Thus the greater is output, the smaller will be AFC and consequently the smaller will be the difference between ATC and AVC. In Figure 4.9 these data have been plotted in conjunction with the productivity data from Table 4.2.

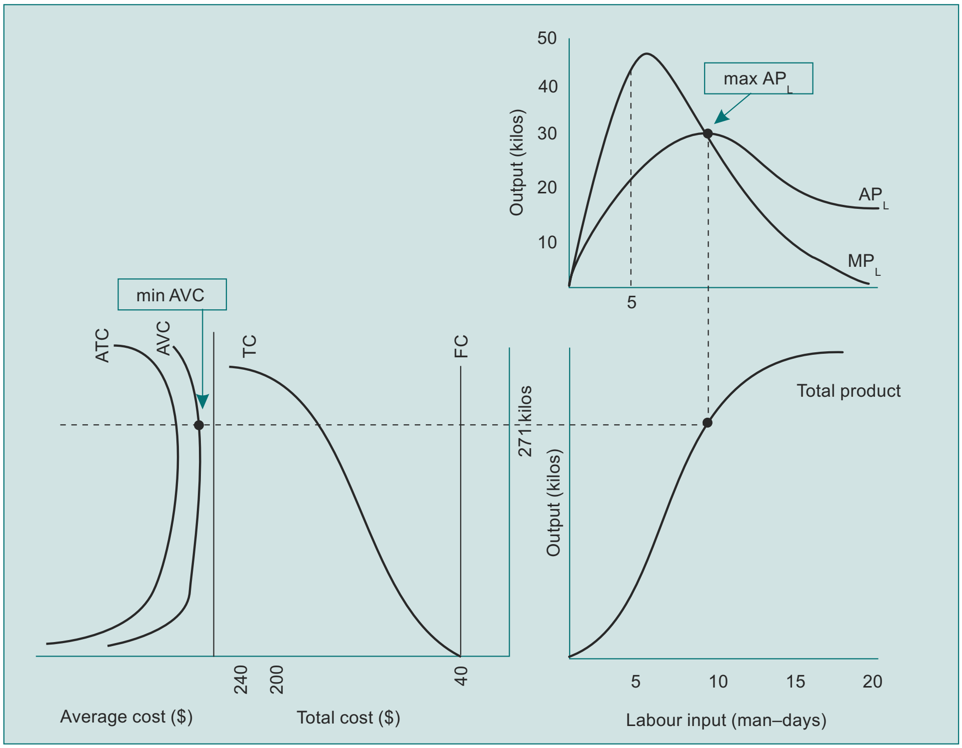


Figure 4.9 Average cost and average productivity curves for salmon fishing

The shape of the AVC curve is inversely related to the AP_L curve. When AVC is at its minimum, i.e. at 271 kilos of output, AP_L is at its maximum, i.e. 30.1 kilos. Thus the shape of the AVC curve is determined by the shape of the AP_L curve. Given the cost of factor inputs, the average variable cost of output is determined by the average productivity of variable factor inputs.

Marginal cost is the cost of producing an additional unit of output. In Table 4.4 we see that the total cost of catching 337 kilos of salmon is $190. We also see that the total cost of catching 338 kilos of salmon is $200, i.e. the marginal cost of a kilo of salmon when 337 kilos are being caught is $10, i.e. the cost of the 338th kilo. For all levels of output up to 337 kilos, however, the marginal cost of salmon, i.e. the cost of catching one additional kilo, cannot be observed directly; instead, the data show the number of additional kilos of salmon caught for each additional fisherman hired. An approximation to the marginal cost of salmon can be calculated by dividing the increase in cost of hiring one additional fisherman by the weight of extra salmon caught. For example, the hiring of the tenth fisherman leads to an increased output of 26 kilos of salmon. The cost of the tenth fisherman is $10 and therefore the cost of each additional kilo is $0.385, i.e. $10/26, on average. It follows that if we had nine fishermen on board and could hire $\frac{1}{26}$th of a fisherman – he is out swimming and climbs on board and fishes for 20 minutes, is paid $0.385 and catches one salmon weighing 1 kilo – we would achieve the same result. But don't worry, we are not going to assume 'swimming fishermen' ever again in this course!

The marginal cost of salmon is calculated on the foregoing basis and shown in Table 4.6.

Table 4.6 Calculating marginal cost

Total output (kilos)	Labour input (man-days)	Wage rate ($/man/day)	Increase in output (kilos)	Marginal cost ($)
				–
0	0	10	–	
				1.43
7	1	10	7	
				0.91
18	2	10	11	
				0.56
36	3	10	18	
				0.38
62	4	10	26	
				0.27
99	5	10	37	
				0.23
143	6	10	44	
				0.21
191	7	10	48	
				0.23
235	8	10	44	
				0.28
271	9	10	36	
				0.38
297	10	10	26	
				0.59
314	11	10	17	
				1.00
324	12	10	10	
				1.43
331	13	10	7	
				2.50
335	14	10	4	
				5.00
337	15	10	2	
				10.00
338	16	10	1	
				10.00
339	17	10	1	
				10.00
340	18	10	1	

The marginal cost is 'high' at low levels of output, declines to a minimum of \$0.21 at an output level of 191 kilos, and after that it increases. In Figure 4.10 the marginal cost data from Table 4.6 have been plotted together with the marginal productivity data from Table 4.2. The curves show quite clearly the inverse relationship between marginal cost and marginal productivity. When marginal cost is at its minimum, the marginal product of labour is at its maximum. In other words, the cost of producing additional units of output, given the prices of factor inputs, is determined by the marginal productivity of the variable factor inputs.

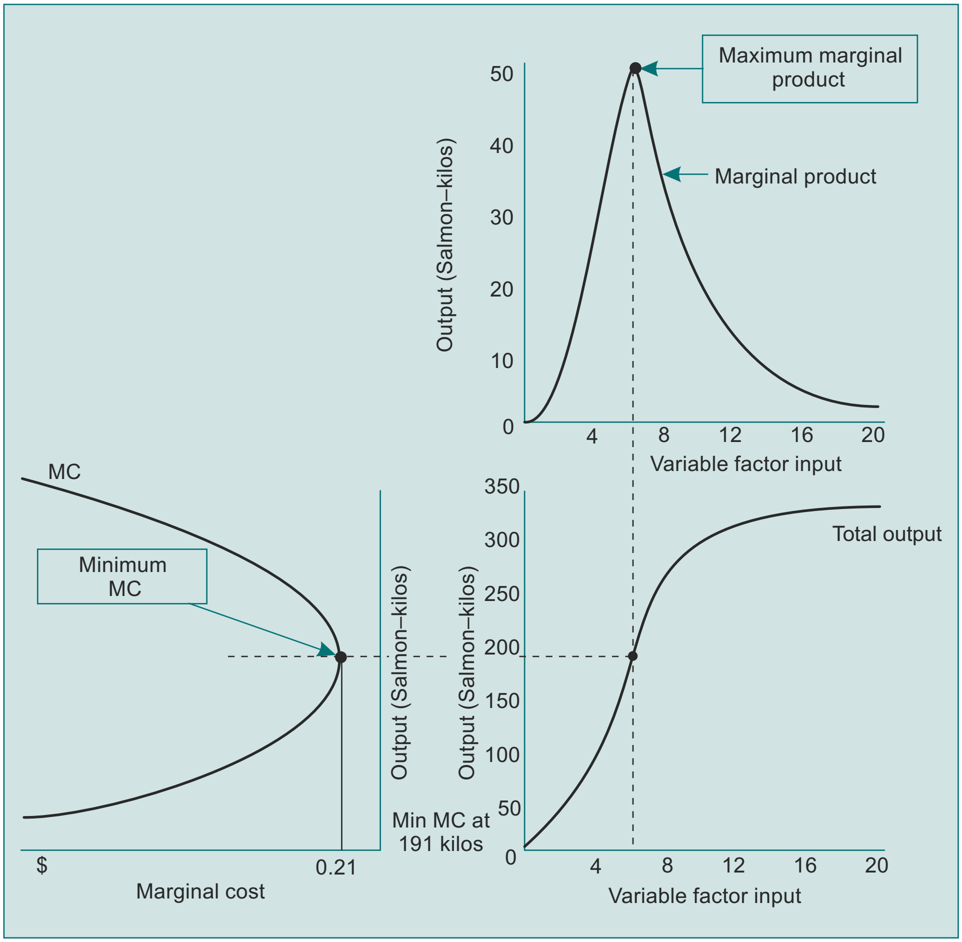


Figure 4.10 Marginal cost and productivity

A little algebra may help show these inverse relationships.

$$AVC = \frac{TVC}{Q} \quad \text{and } TVC = L \cdot W$$

$$AP_L = \frac{Q}{L}$$

$$\therefore Q = AP_L \cdot L$$

$$\therefore \text{AVC} = \frac{\text{L}\cdot\text{W}}{\text{Q}} = \frac{\text{L}\cdot\text{W}}{\text{AP}_\text{L}\cdot\text{L}} = \frac{\text{W}}{\text{AP}_\text{L}}$$

Since W is fixed (given), AVC varies inversely with AP_L, i.e. when AVC is minimum AP_L is maximum.

Second,

$$\text{MC} = \frac{\Delta\text{TVC}}{\Delta\text{Q}} = \frac{\Delta(\text{L}\cdot\text{W})}{\Delta\text{Q}}$$

$$\text{MP}_\text{L} = \frac{\Delta\text{Q}}{\Delta\text{L}}$$

$$\therefore \Delta\text{Q} = \text{MP}_\text{L}\cdot\Delta\text{L}$$

$$\therefore \text{MC} = \frac{\Delta\text{L}\cdot\text{W}}{\Delta\text{Q}} = \frac{\Delta\text{L}\cdot\text{W}}{\text{MP}_\text{L}\cdot\Delta\text{L}} = \frac{\text{W}}{\text{MP}_\text{L}}$$

Since W is fixed (given), MC varies inversely with MP_L, i.e. when MC is minimum, MP_L is maximum.

Just as we identified important relationships among total product, average product and marginal product of variable factor outputs, so we can identify matching relationships among output and costs of output.

1. When TP = 0, TVC = 0 and TC = FC
2. When ATC is minimum, ATC = MC
3. When AVC is minimum, AVC = MC
4. When MC is greater than ATC (AVC), ATC (AVC) is increasing
5. When MC is less than ATC (AVC), ATC (AVC) is decreasing

Armed with all these productivity and cost data, we can now turn to firm supply and market supply. We shall also consider the two time periods, namely the short run and the long run.

4.4 Firm Supply in the Short Run

The productivity and cost relationships we have studied provide the basis for determining how much output a firm would be prepared to supply at each and every price. To simplify the analysis in both the short run and long run we shall make four assumptions.

(a) The firm produces only one good.
(b) The market price of the good in question is outside the control of the firm; it can sell as much as it can produce at the going market price and consequently has no incentive to offer the good at a price lower than the market price. Think of yourself going on holiday in Europe and offering your currency for euro. You have to accept the going exchange rate, i.e. the price of euro in terms of your currency; your being in the market or not being in the market does not affect the price. In other words, each firm is so small relative to the total market that its influence on price is completely insignificant. Such an assumption, of course, is not true for all firms in the world economy. IBM, for example, has had some say in the price of computers, and if IBM were to disappear from the world scene both the going price and the quantity of computers sold would alter. (We shall deal with corporate giants in a future module.)

(c) The price or cost of factor inputs is outside the control of any one firm. In other words, in both product and factor markets each firm is a price taker. It accepts the going market prices and makes decisions on the basis of its activities not being able to affect prices in either the product or factor markets.
(d) The firm's goal is profit maximisation. Profit (π) is maximised when total revenue (TR) minus total cost (TC) is maximised.

Given all these reasonable assumptions, let us return to our fishing boat. The firm (boat owner) produces only one good, salmon. The going market price for salmon is \$0.70 per kilo – that's for whole fish, ungutted. The costs we looked at in Table 4.4 have not changed; the fixed cost of the boat is still \$40 per day and the going wage of fishermen is still \$10/man/day. The firm's goal is profit maximisation, so that the boat owner wants to hire the number of fishermen who will yield the catch that maximises the difference between TR and TC.

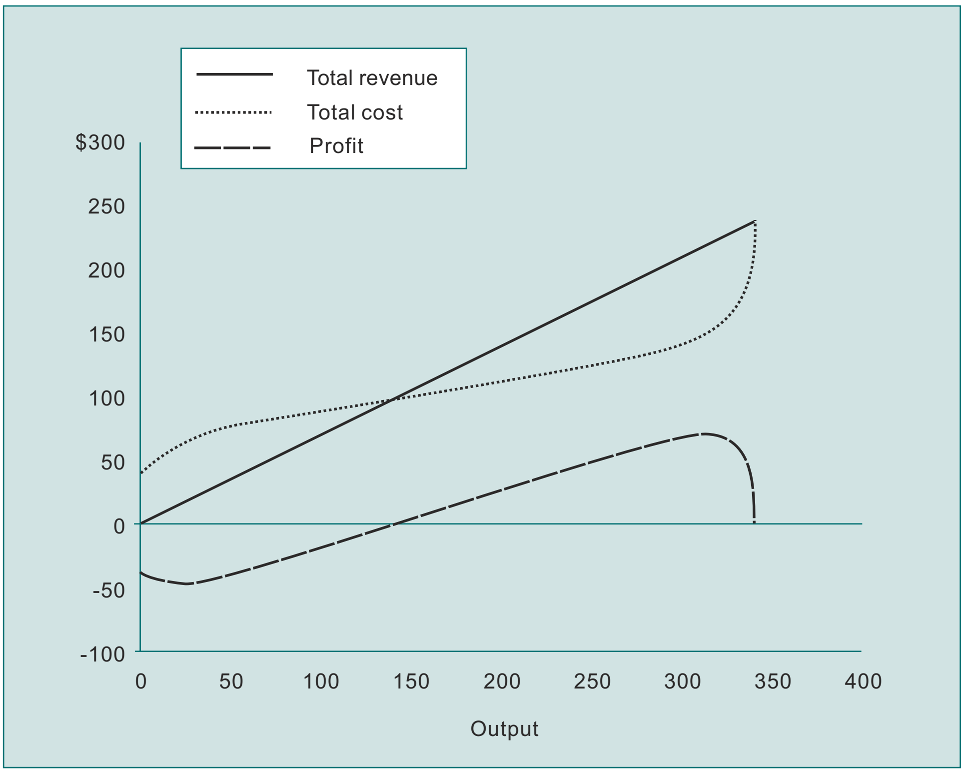


Figure 4.11 Total cost, revenue and profit for salmon fishing

Table 4.7 and Figure 4.11 show TR, TC and π for different levels of output and, concomitantly, for different numbers of fishermen hired. The value of π is at a maximum when total output is 314 kilos of salmon. Total revenue equals the quantity (Q) of output produced times the price (P) of each unit of output. Total cost, you may recall, equals fixed cost (FC) plus variable cost (VC). At the given level of output, we have

$$\begin{aligned}
\mathrm{TR} &= \mathrm{P} \times \mathrm{Q} = \$0.70 \times 314 = \$220 \\
\mathrm{TC} &= \mathrm{FC} + \mathrm{VC} \\
&= \$40 + (\$10 \times 11) \\
&= \$40 + \$110 = \$150 \\
\pi &= \mathrm{TR} - \mathrm{TC} \\
&= \$220 - \$150 \\
&= \$70
\end{aligned}$$

Table 4.7 Total cost, revenue and profit for salmon fishing

Output (kilos)	Total cost ($)	Total revenue ($)	Profit ($)
0	40	0	−40
7	50	5	−45
18	60	13	−47
36	70	25	−45
62	80	44	−36
99	90	69	−21
143	100	100	0
191	110	134	24
235	120	165	45
271	130	190	60
297	140	208	68
314	150	220	70
324	160	227	67
331	170	232	62
335	180	234	54
337	190	236	46
338	200	237	37
339	210	237	27
340	220	238	18
340	230	238	8
340	240	238	−2

A catch of 314 kilos is not the only output level at which the firm makes a profit. As can be easily seen from both Table 4.7 and Figure 4.11, π is positive as soon as the sixth fisherman is hired, i.e. as soon as total output is equal to or greater than 143 kilos. Profit remains positive until the twentieth fisherman is hired, after which a loss is incurred. Thus as long as between 6 and 19 fishermen are hired, the firm makes a profit; but only one unique combination of factor inputs (one boat plus 11 fishermen) and only one resulting level of output (314 kilos) results in maximum profit.

At only two levels of output does the firm break even, when $\pi = 0$. This occurs when six fishermen are hired and when between 19 and 20 fishermen are hired. The resultant catch for six fishermen is 143 kilos, yielding approximately $100 in total revenue. Total cost at this output level is $100 ($40 + $60) and consequently total profit is zero.

There are two other ways to calculate the profit-maximising level of output. The first is to consider average revenue and average total cost. Average revenue is calculated by dividing total revenue by output. If each kilo of salmon sells for the same price ($0.70), average revenue will be constant and equal to price. For example, if we had 100 kilos of salmon and the price is $0.70 per kilo, total revenue from the sale of the 100 kilos will be $70 ($0.70 × 100). Average revenue, i.e. the revenue from each kilo, will be $0.70, i.e. $70/100, which is of course the price per kilo.

Table 4.8 and Figure 4.12 show average cost (total and variable) and average revenue for the salmon-fishing firm. So as to be able to concentrate on the important parts of the curves, the average costs for very low outputs are omitted. To be profitable the firm must operate at levels where TR > TC or where AR > ATC, i.e. where $\frac{TR}{Q} > \frac{TC}{Q}$. In Table 4.8 such output levels occur between 143 and 340 kilos.

Table 4.8 Average revenue and average costs for salmon fishing

Total output (kilos)	Average variable cost ($)	Average total cost ($)	Average revenue ($)
0	–	0	0.70
7	1.43	7.14	0.70
18	1.11	3.33	0.70
36	0.83	1.94	0.70
62	0.65	1.29	0.70
99	0.51	0.91	0.70
143	0.42	0.70	0.70
191	0.37	0.58	0.70
235	0.34	0.51	0.70
271	0.33	0.48	0.70
297	0.34	0.47	0.70
314	0.35	0.48	0.70
324	0.37	0.49	0.70
331	0.39	0.51	0.70
335	0.42	0.54	0.70
337	0.44	0.56	0.70
338	0.47	0.59	0.70
339	0.50	0.62	0.70
340	0.53	0.65	0.70
340	0.56	0.68	0.70
340	0.59	0.71	0.70

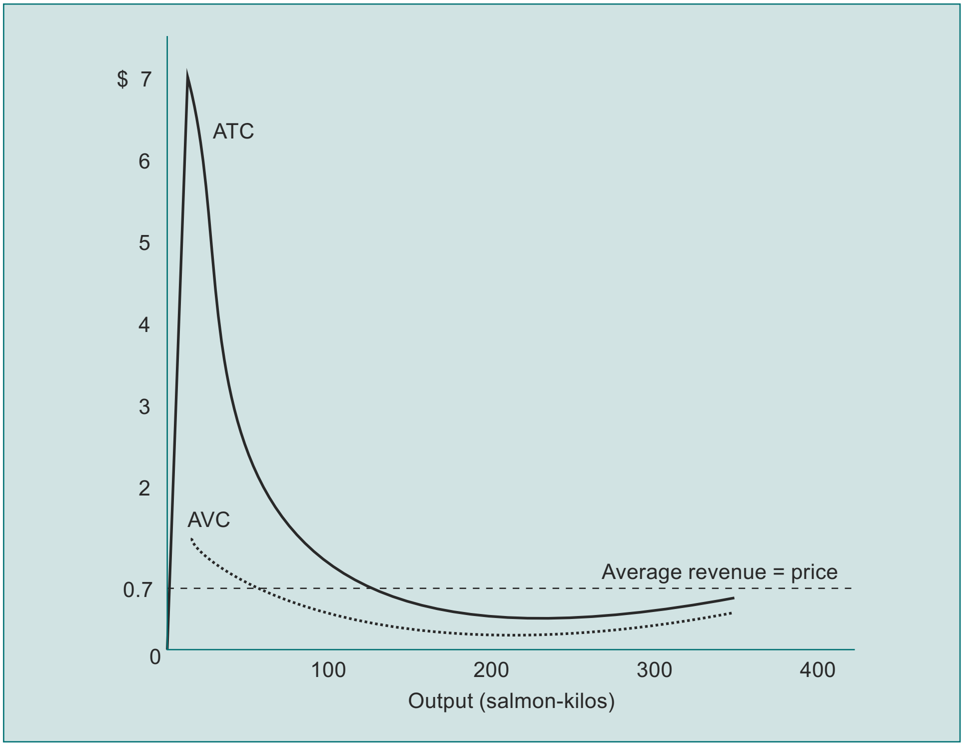


Figure 4.12 Average revenue and costs for salmon fishing

Since π = TR − TC, we have

$$\frac{\pi}{Q} = \frac{TR}{Q} - \frac{TC}{Q}$$

Thus the profit per kilo of salmon caught and sold in our example will be at a maximum when AR − ATC is at a maximum. This will occur when 297 kilos of salmon are caught because at that level of output ATC is at a minimum ($0.47) and AR is constant ($0.70), yielding a profit per kilo of $0.23 and a total profit equal to $0.23 × 297 kilos = $68.

But maximising π per kilo does not necessarily imply maximising total π. Before we figure out why this is the case – and not unexpectedly – we shall use marginal analysis in relation to a very important lesson that can be learned from Figure 4.12.

Suppose the boat owner, because of some local labour dispute, could hire no more than five fishermen and as a result could catch only 99 kilos of salmon: would it 'pay him' to go fishing? Both Table 4.7 and Figure 4.12 suggest no. Table 4.7 shows he would suffer a loss of $21 if he could have only five men and catch only 99 kilos of salmon. We can see the loss again in Figure 4.12; at 99 kilos of output ATC > AR, the loss per kilo being the difference between the curves at an output level of 99 kilos and the total loss being the difference × 99 kilos, i.e. $21.

But what happens if the boat owner doesn't go fishing? In this case he will lose $40, the daily fixed cost of the boat. Thus he would be better off going fishing with five fishermen than staying on shore; if he does, he covers his variable cost of $50 ($10 × 5) and from the total revenue of $69 from selling 99 kilos of salmon at $0.70 per kilo, he will have $19 to set against his fixed cost of $40.

Even although he cannot cover his total cost, the boat owner should stay in business, i.e. keep going fishing, in the short run, as long as he can at least cover his variable cost.

Of course, if the price of salmon fell to \$0.30 per kilo, a price below the minimum average variable cost of \$0.33, the boat owner would not go to sea and would then suffer a loss of at least \$40 (his fixed cost). At a price of \$0.30 per kilo, hiring any number of fishermen will involve a loss greater than \$40.

The final, and – dare we say it – the best way to find the profit-maximising output level is to employ the technique of marginal analysis. The profit-maximising output level occurs where marginal revenue (MR) equals marginal cost (MC). Marginal revenue is the revenue obtained from selling one additional unit of output – in our example one kilo of salmon. Given a selling price of \$0.70 per kilo, the marginal revenue (MR) must equal that price. Marginal cost (MC) is the cost to the firm of producing one additional unit of output, i.e. 1 kilo of salmon.

When MR > MC, the firm will expand output because each additional unit produced will add more to total revenue than it will to total cost and consequently profit will be increased or loss decreased. When MR < MC the firm will reduce output because the decrease in total revenue will be less than the decrease in total cost, and consequently profit will be increased or loss decreased. (Note that in this analysis MC must be increasing.) When the firm reaches an output level at which MR = MC, profit will be at a maximum. Any increase or decrease in output will cause π to decrease.

To drive home the very important message let us analyse the data in Table 4.9. Let's choose any given output level – say 235 kilos – and apply marginal analysis. Does it pay the boat owner to expand output? Remember the rule that if MR for additional output exceeds MC for producing that output, then it is worthwhile doing so, i.e. π will increase. From Table 4.9 we can see that increasing output from 235 kilos to 271 kilos requires hiring one additional fisherman at a cost of \$10. Thus the cost of the additional 36 kilos (271 − 235) is \$10 and the marginal cost of each kilo is \$0.28 (\$10/36). However, the MR of salmon is constant, equal to the selling price of \$0.70. Thus the additional 36 kilos of salmon will add more to revenue than they will to cost to the tune of \$15: \$25 (\$190 − \$165) is the increase in total revenue and \$10 (\$130 − \$120) is the increase in total cost.

Table 4.9 Marginal revenue, marginal cost and profit maximisation

Total output (kilos)	Total revenue (\$)	Total cost (\$)	Marginal revenue (\$)	Marginal cost (\$)	Profit (\$)
			–	–	
0	0	40			−40
			0.70	1.43	
7	5	50			−45
			0.70	0.91	
18	13	60			−47
			0.70	0.56	
36	25	70			−45
			0.70	0.38	
62	44	80			−36
			0.70	0.27	

Total output (kilos)	Total revenue ($)	Total cost ($)	Marginal revenue ($)	Marginal cost ($)	Profit ($)
99	69	90			−21
			0.70	0.23	
143	100	100			0
			0.70	0.21	
191	134	110			24
			0.70	0.23	
235	165	120			45
			0.70	0.28	
271	190	130			60
			0.70	0.38	
297	208	140			68
			0.70	0.59	
314	220	150			70
			0.70	1.00	
324	227	160			67
			0.70	1.43	
331	232	170			62
			0.70	2.50	
335	234	180			54
			0.70	5.00	
337	236	190			46
			0.70	10.00	
338	237	200			37
			0.70	10.00	
339	237	210			27
			0.70	10.00	
340	238	220			18
			0.70	10.00	
340	238	230			8
			0.70	10.00	
340	238	240			−2

Will profit be increased even further if output is increased yet again? The answer is once more, ‘yes’, for we only have to look at MR and MC. The value of MR remains constant at $0.70; MC increases from $0.28 to $0.38; but since $0.38 is less than $0.70, the expansion is worthwhile. What about increasing total output from 314 to 324 kilos? Now the answer is ‘no’. The MC of each kilo now increases to $1.00 and as a consequence a loss of $0.30 ($1.00 − $0.70) would be incurred on each of the additional 10 kilos, lowering profit below its maximum figure of $70. As output continues to expand, the excess of MC over MR increases, causing profit to decrease.

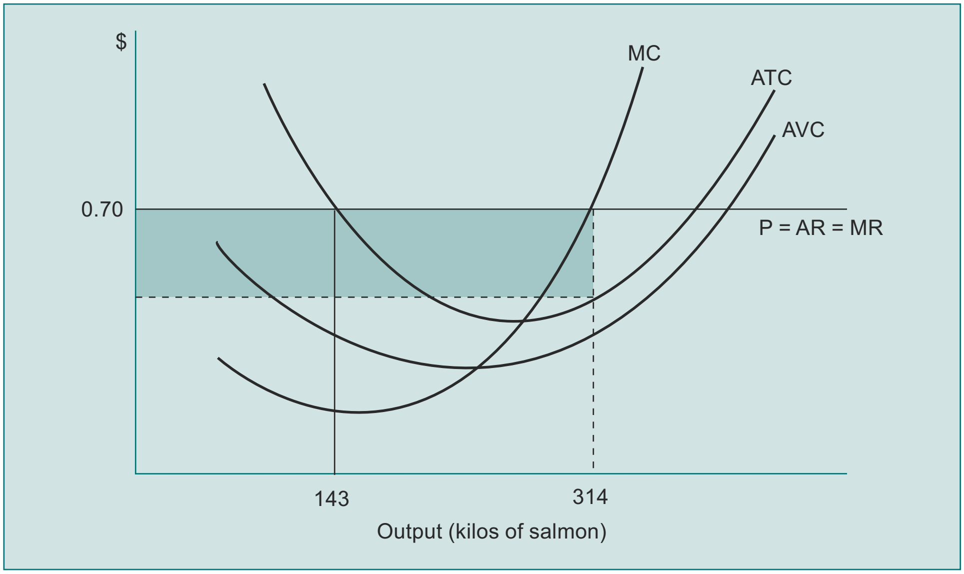


Figure 4.13 Marginal revenue, marginal cost and profit maximisation

This profit-maximising rule can be seen in Figure 4.13, which is a repeat of part of Figure 4.12 with the addition of the MC curve. MC intersects MR at an output level of 314 kilos. To the right of this point, profit would still be positive because AR > ATC. Similarly, to the left of this point AR > ATC for output levels greater than 143 kilos. The heavily shaded area represents maximum profit at an output level of 314; i.e. at the output level where MC = MR. Total profit equals AR − ATC (profit per kilo sold) times 314 (number of kilos sold).

Where does the profit go? To the boat owner. He is the entrepreneur who bought the boat and who takes the risk that salmon prices might fall; he is the entrepreneur who ensures that all the rods and equipment are in sound working order; it is he who hires the fishermen and ensures that they work efficiently. Such efforts and talents have to be rewarded or the entrepreneur will seek employment with a firm or another area of activity where he can exercise his entrepreneurial talent.

Thus the profit we have spoken about must be sufficient to keep him in the salmon fishing industry. What is sufficient? When he cannot do better elsewhere. Thus strictly speaking an element of cost should be a salary or payment to the boat owner/entrepreneur just sufficient to keep his talents in the fishing industry, i.e. just greater than his opportunity costs. Henceforth the salary or payment will be included in a firm's costs like any other factor input.

With the tools now at your command, deriving firm supply in the short run is a relatively straightforward task. Consider Figure 4.14, where we represent a hypothetical firm operating in an engineering-efficient fashion (i.e. on its production frontier), and again we assume a single product firm, a profit maximiser, and a price taker in both the product and factor markets. The question we want to be answered is: what quantities of output would the firm be willing to produce in the short run at each and every price? Being in the short run, the firm cannot escape its fixed cost whether or not it produces any output. We have chosen a set of prices $P_1 \dots P_6$ and wish to derive the quantities of output that the firm would be willing to produce at each of these prices.

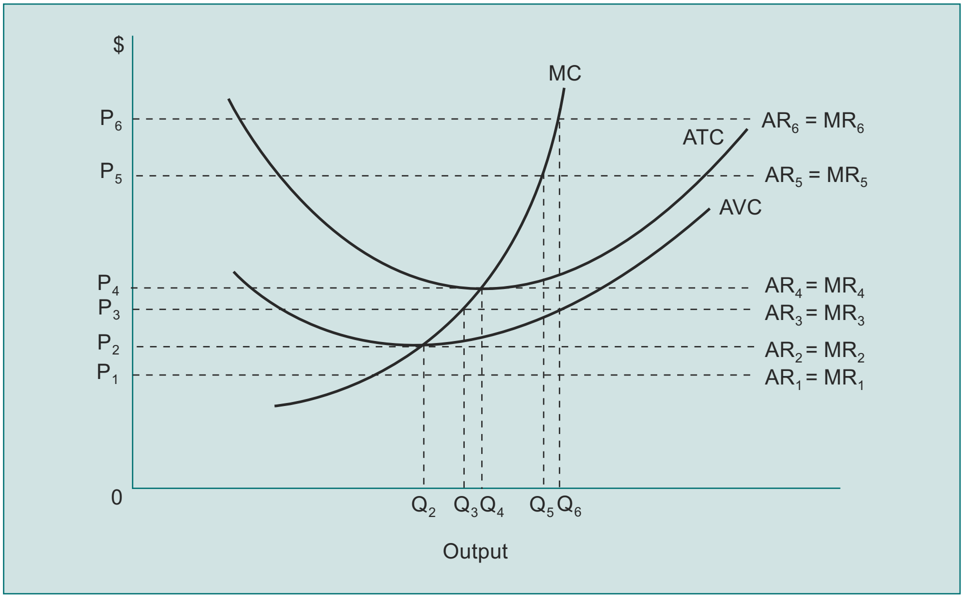


Figure 4.14 Firm supply in the short run (1)

Consider P_1. Because P_1 is constant, $P_1 = AR_1 = MR_1$. How much output would the firm produce if the price were P_1? The answer is zero. Why? At that price $AR_1 < AVC$; as a consequence, the firm cannot cover its variable cost in producing output. By producing zero, the only cost incurred is fixed cost. The production of any output would incur this fixed cost plus some variable cost. Therefore, loss minimisation dictates that the firm produces zero output at a price of P_1.

What about P_2? At this price there is one level of output, and only one, where variable cost is covered. At output level $0Q_2$, AR = AVC. At all other levels of output, $AVC > AR_2$. Thus the firm is indifferent whether it produces zero or $0Q_2$; in both cases the loss incurred equals fixed cost.

At a price of P_3 the firm is still not making a profit because $AR_3 < ATC$, but by producing an output of $0Q_3$ the firm is able to cover its variable cost ($AR_3 > AVC$) and make some contribution towards fixed cost. But why produce $0Q_3$ at a price of P_3; why not $0Q_2$? If it were to produce $0Q_2$, the firm would observe that, at that level of output, marginal revenue (MR_3) was greater than MC. Thus producing additional output, up to $0Q_3$, would add more to revenue than it would to cost. Being a profit maximiser, the firm should increase output up to $0Q_3$. Why not up to $0Q_4$? For any output level greater than $0Q_3$, $MC > MR_3$ and consequently the production of additional output would add more to cost than it would to revenue.

At a price of $0P_4$, AR_4 = ATC at an output level of $0Q_4$. At a price of $0P_4$, $0Q_4$ is the optimal output because $MC = MR_4$. Since AR_4 = ATC, the firm breaks even, i.e. covers both its fixed and variable costs. No level of output, other than $0Q_4$ at a price of P_4 permits the firm to break even; at all other levels of output, ATC > AR.

In a similar fashion we can see that if the price were P_5, the firm would produce an output level of $0Q_5$ because at that output level MR_5 = MC. Finally if the price were P_6, the firm would produce $0Q_6$. What profit would the firm make if the price were P_6 and it produced

$0Q_6$? On each unit of output the profit would be AR_6 minus ATC at $0Q_6$. Thus total profit would equal unit profit × $0Q_6$ (the number of units produced).

In Figure 4.15 we can now plot the firm's short-run supply curve. Figure 4.14 and Figure 4.15 have the same main title – 'Firm supply in the short run'. This is because one curve is common to both. That curve is the firm's marginal cost curve (MC) above the minimum point on the AVC curve of Figure 4.14, and this represents the firm's short-run supply curve.

What would happen to the firm's short-run supply curve for prices greater than P_6? Would the firm keep on producing higher levels of output? The answer is 'no'. At the same level of output the short-run supply curve would become a vertical line, i.e. no additional output would be forthcoming no matter what the market price. Why is this so? To answer that question we have to go back to our marginal productivity curve. Remember our fishing boat example? After 18 fishermen had been hired, the marginal productivity of labour became zero, i.e. hiring additional fishermen added zero to the catch. Thus if Q_6 in Figure 4.15 represented 340 kilos of salmon, the catch when 18 fishermen were on board (or 19 or 20), the firm's short-run supply curve in Figure 4.16 would become a vertical line at output $0Q_6$, i.e. hiring more fishermen would only add to cost; it would produce no additional salmon, irrespective of the price of salmon.

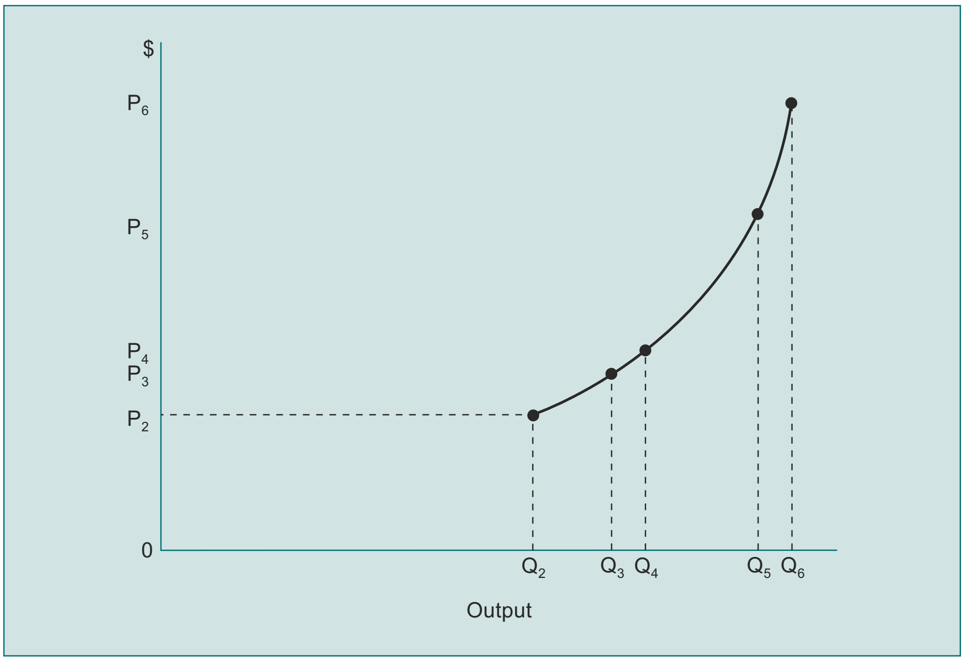


Figure 4.15 Firm supply in the short run (2)

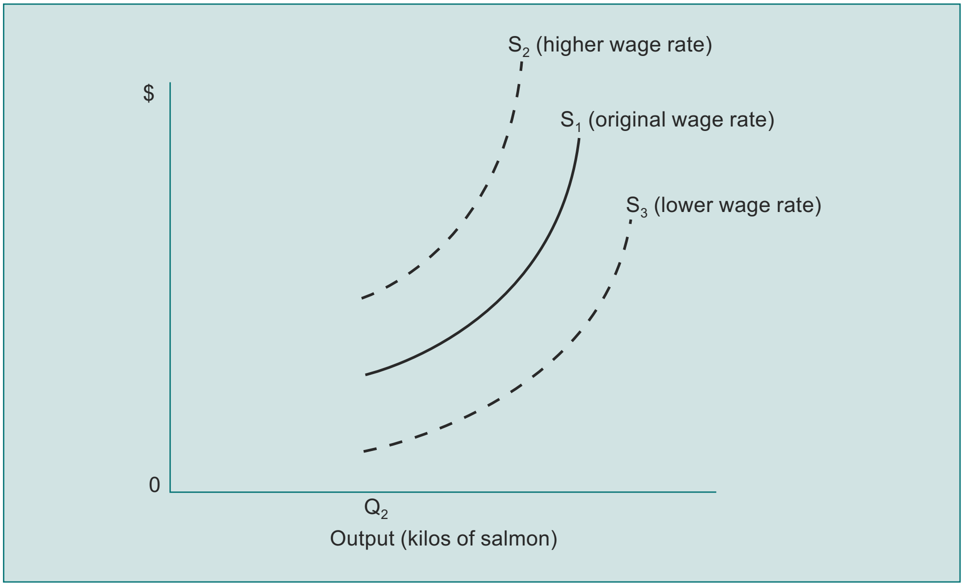


Figure 4.16 Shifting the short-run supply curve

Given the production frontier of any firm, i.e. fixed capital input plus a given state of technology, the only factor that can cause the firm's short-run supply curve to shift is the cost of the variable factor input. If, in the fishing boat example, the going wage rate of fishermen were to increase, the marginal cost of catching salmon would increase, i.e. the marginal cost curve would shift upwards, and consequently the firm's short-run supply curve would shift upwards. Conversely, if the wage rate were to decrease, the firm's short-run supply curve would shift downwards. These shifts are shown as curves S_2 and S_3 in Figure 4.16. At the higher wage rate (S_2) the minimum AVC will increase; at the lower wage rate (S_3) the minimum AVC will decrease. Each supply curve will originate where MC = minimum AVC.

Having mastered the firm's short-run supply curve, producing the industry short-run supply curve is a straightforward matter. The long-run supply curve of the firm is slightly more complex; the industry long-run supply curve is significantly more complex. However, when we put demand and supply together, economics, you will be pleased to learn, becomes quite interesting. It's like learning to ski: the basics can be quite painful and boring, but once the basics have been mastered skiing can be exciting. Before the fun, however, a little more pain!

4.5 Market Supply

4.5.1 Market Supply in the Short Run

On the supply side, a market is composed of a number of competitive firms. The *market supply* is the aggregate of the amounts that all firms in a given market would be willing to supply at each price. Similar to the construction of the market demand curve, the *market supply curve* is found by adding together the supply curves of all the firms in a given market. Some markets are supplied by local firms, e.g. the plumbing market, the babysitting market and the newspaper-delivery market. Other markets are supplied by international firms, e.g. the computer market, the car market and the rare antiques market. To simplify the analysis we shall assume we are discussing a local market in which there is, in the short run, a very large, but finite, number of firms. The very large number of firms is necessary to guarantee that no individual firm can influence the price of the good in question. For example, if the going rate for babysitting is $5.00 per hour and your usual babysitter goes off to college, you can always find another babysitter for $5.00 per hour but you cannot find a babysitter for less than $5.00 per hour – that's the going rate.

How do we add together the short-run supply curves (the marginal cost curves above the minimal points on the AVC curves) to obtain the market short-run supply curve? Consider only three firms, each of which owns one salmon-fishing boat: A owns a large sophisticated boat with electric reels and 40 rod holders; B owns an average-sized, 5-year-old vessel with 20 rod holders; and C owns a small old boat with 12 rod holders. Each of their short-run supply curves are shown in Figure 4.17, as well as the aggregate supply curve of the three boats.

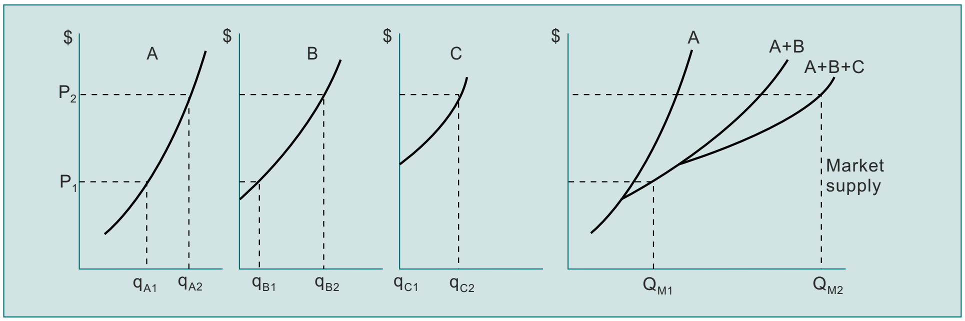


Figure 4.17 Derivation of the short-run market supply curve

From Figure 4.17 it can be seen that

- at a price of P_1, A would supply q_A;
- at a price of P_1, B would supply q_B; and
- at a price of P_1, C would supply $q_C = 0$.

Therefore at a price of P_1 the three boats taken together would supply the 'market' quantity of Q_M equal to $q_A + q_B + q_C$ (the last item being zero). Similarly, at a price of P_2 the three boats would supply the market with an output of Q_M equal to $q_A + q_B + q_C$. For *n* existing boats (a large number), we can add together the quantity each boat would be willing to supply at each and every price and thus derive the short-run market supply curve.

4.5.2 Shifting Market Supply Curve

A change in the price of the variable factor input will, as we have seen, shift each firm's short-run supply curve – remember that this is because a change in the variable factor input price will affect each firm's marginal cost curve. It follows that the market short-run supply curve, the aggregate of the firms' short-run supply curves, must also shift.

Suppose a recession hits the area in which the fishing fleet is based. The unemployment rate rises and wages, on average, decrease. Some unemployed workers who enjoy fishing in their leisure time attempt to find employment on the fishing boats. Suppose this causes the going daily wage rate for hiring fishermen to fall from \$10.00 per day to \$8.00 per day. This will cause the minimum value of AVC for each fishing boat to decrease and also cause the MC curve (the short-run supply curve) of each boat to shift to the right. The market short-run supply curve, the summation of all firms' short-run supply curves, will also, therefore, shift to the right. These movements are captured in Figure 4.18.

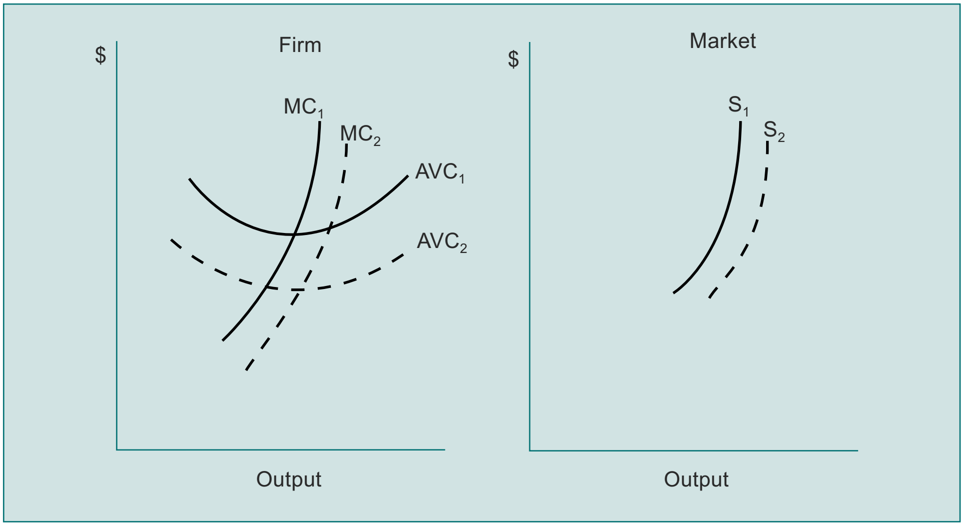


Figure 4.18 Shift in firm and market short-run supply curves

- S_1 is the summation of all the MC_n; and
- S_2 is the summation of all the MC_n, with *n* equal to the number of firms in the market.

Similarly, an increase in the going wage rate for fishermen would cause S_1 to shift to the left. Why? Because AVC_1 would rise, MC_1 would shift to the left, and consequently fewer kilos of fish would be offered at each and every price by each firm and thereby by all firms in the market taken together.

4.5.3 Equilibrium of the Firm

In the short run, namely that time period when some factor inputs and consequently some costs are fixed, the equilibrium output of the firm will be that output at which marginal revenue equals marginal cost. This will be an equilibrium level of output in the sense that while other levels of output may yield a profit, no other output level yields as high a profit. Thus the firm has no incentive to move from that level of output and can be viewed as *in equilibrium* in the short run.

However, being in equilibrium in the short run does not imply the firm is in long-run equilibrium. Indeed, the firm may never reach long-run equilibrium. Why? In the real world, change occurs constantly. New goods and services are constantly being discovered and invented, consumer tastes and preferences are constantly – albeit slowly – changing, research and development is finding new uses for natural resources, and technological change, whereby more efficient ways are found of producing goods and services, is, fortunately, always with us. As an example, consider the changes that have occurred in the transportation business in crossing the Atlantic in the past 150 years. In the early to mid-nineteenth century, if you wanted to travel from Europe to America you had little choice: your sole means of transportation was a sailing vessel – small, cramped, unsafe, and with a large margin of error on estimated arrival time. Nobody spent a summer vacation on a two-week trip to Florida. Today you can fly from London to New York in the morning and be back in London for dinner the next night. All of this has been made possible, and is economically viable, because of changes in consumers' tastes and preferences, research and development, new uses of natural resources, and technological innovations.

In Module 1, in discussing scarcity and resource allocation, we recognised that whenever a choice is made, something has to be given up – a cost has to be incurred. The best alternative given up when a good is produced is known as its *opportunity cost*, i.e. the best alternative forgone. Thus there is a cost to produce any good over and above the costs that have to be expended for the fixed and variable factor inputs. This cost is the profit that could be earned by using the resources in producing the best alternative good or service. For example, if our fishing-boat owner analysed all possible uses of his boat and decided that, after fishing, the next most profitable employment of his boat was passenger pleasure cruising, then the profit that he could have made pleasure cruising is the boat owner's opportunity cost.

In the long run, when there are no fixed resources and consequently no fixed costs, if a resource owner cannot earn his opportunity cost in a given industry he will shift his resources out of that industry. In other words, in the long run, if a resource owner finds that he can earn a higher profit in the production of a different good or service from the one in which he currently employs resources, he will move resources to the more profitable employment. Thus in our fishing-boat example, the day will arrive – probably at the end of the one year that was the useful life we assigned to the fishing boat – when the boat owner will evaluate the expected return from staying in the salmon-fishing industry against all realistic alternatives.

To recognise a resource owner's opportunity cost as a cost of doing business, we shall henceforth include that opportunity cost as part of a firm's fixed cost. Thus when a firm is just breaking even, i.e. when total cost equals total revenue, the resource owner is earning a sufficient return to prevent his resources from leaving the industry; in such a circumstance we would say the firm is earning 'normal profit'. If profit in excess of what is required to keep current resources in the industry exists, then new resources, i.e. new firms and/or expanded existing firms, will be attracted into the industry in the long run. To the long run we now turn!

4.5.4 The Firm in the Long Run

In the long run, no factors of production are fixed. The firm is in the planning stage with no prior commitments. Our boat owner's vessel, rods, reels, lines, hooks and nets have all worked perfectly until the end of their useful lives, at which point they disintegrate instantaneously and simultaneously. Alternatively, let us consider the case where you are a young entrepreneur with your new MBA and you decide to go into the supermarket business in some chosen city. You have a greenfield site – but how big should you make your supermarket? That's your long-run decision but, remember, as soon as you build that supermarket you are right back on the short run with a fixed capital resource.

The rule for profit maximisation in the long run is, with one important modification, the rule for profit maximisation in the short run: marginal revenue must equal marginal cost. The modification is that the marginal cost must be the *long-run marginal cost*.

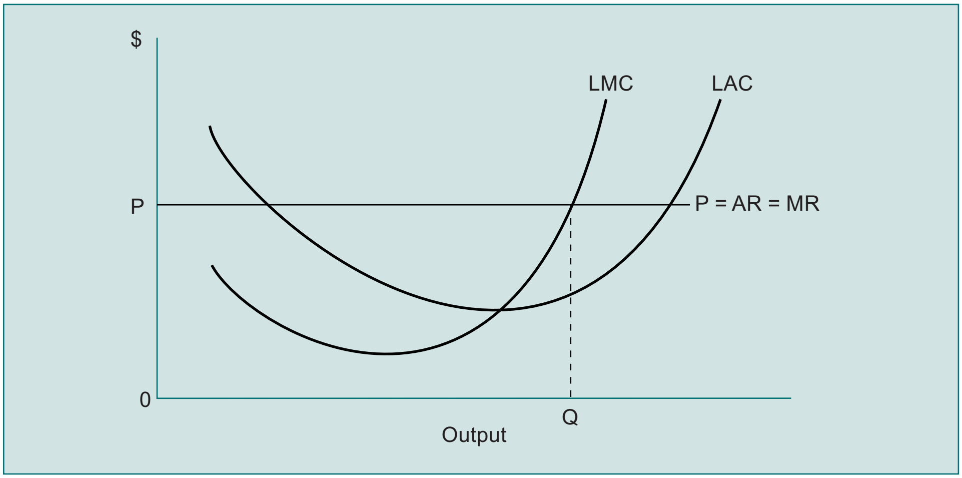


Figure 4.19 Long-run equilibrium

Consider Figure 4.19. The firm is in long-run equilibrium producing an output level of Q. The logic of the short run applies in the long run. The firm is again a price taker and consequently P = AR = MR. Profit-maximising behaviour dictates that MR must equal, in this case, long-run marginal cost (LMC). For output levels less than Q, producing additional output adds more to revenue than it does to total cost because MR > LMC. For output levels greater than Q, the reverse is true; each additional unit of output adds more to total cost than it does to total revenue. Total profit is positive for all outputs at which AR > ATC but for only one output level will profit be at a maximum, i.e. at output Q where MR = LMC.

But there is another important message to be derived from Figure 4.19. At output Q the most efficient/least cost combination of all resources required to produce Q by a firm is found on the long-run average cost (LAC) curve, for that is what the LAC traces out, namely the lowest average total cost for each and every level of output for a firm when no resources are fixed. Thus if Figure 4.19 represents our salmon-fishing industry and if the price of salmon were P per kilo and were going to remain there, the boat owner could calculate what size and type of vessel and what number of fishermen would yield maximum profit. All he has to do is figure out where price (MR) equals LMC and both the optimal boat and the optimal number of fishermen can be deduced. But, you might say, that cannot be a long-run equilibrium price of salmon. If each firm is making above-normal profit, resources will move into the industry and the price of salmon must fall. That's quite correct, but the purpose of the current exercise is not to determine the long-run equilibrium price of salmon. What we are doing here is explaining what entrepreneurs in the salmon-fishing industry would do if the price of salmon were to be P. If the price were to be $P per kilo, our boat owner would produce Q from a unique combination of a boat of certain specifications and a certain number of fishermen. We can superimpose this optimally sized boat (capital) in the long-run equilibrium figure, and this we have done in Figure 4.20.

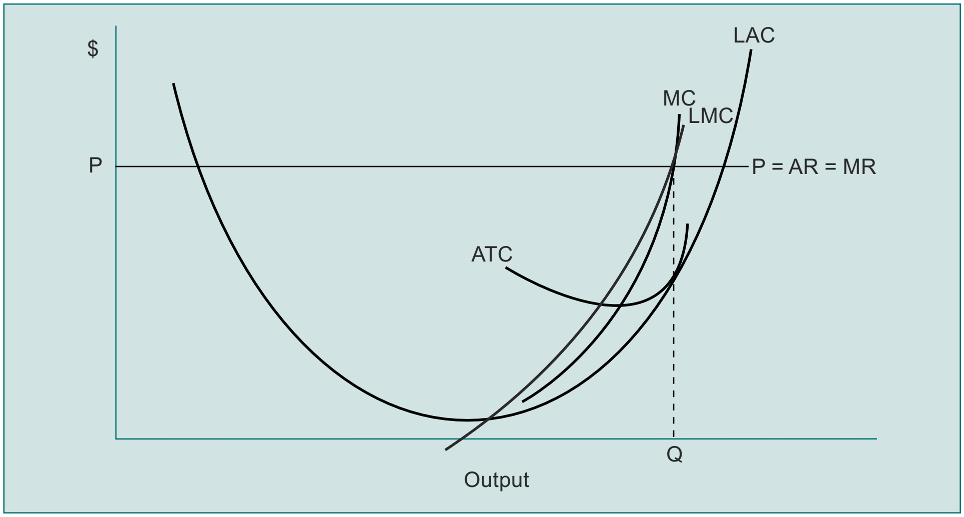


Figure 4.20 Long-run and short-run equilibrium

The firm represented in Figure 4.20, with plant (boat) size represented by ATC, is in both short-run and long-run equilibrium. It is in long-run equilibrium because MR = LMC at output Q. It is in short-run equilibrium because MR = MC at output Q.

As soon as the boat has been acquired, the firm is back in the short run because fixed cost has been incurred. If the price of salmon now changes, the firm will adjust its variable factor inputs, i.e. fishermen – it cannot 'adjust' the boat because that's a fixed cost – so that the new price (MR) equals MC. No longer will MR = LMC and consequently no longer will the firm be in long-run equilibrium. Thus it is only when the price of the good in question, the state of technology, and all factor input prices remain unaltered that the firm can remain in both short-run and long-run equilibrium. In the real world, of course, the chances of these other things remaining unchanged (our *ceteris paribus* assumption), are essentially zero. Long-run equilibrium, however, despite all these uncertainties, is the position for which profit-maximising firms must aim despite the fact they may never attain it.

It also follows from Figure 4.20 that if price ever falls below the minimum point on the long-run average cost curve (LAC), the firm would produce zero output. Why? The reason is simple: the firm could not make a profit as opportunity cost would not be covered and resources could be more profitably employed in some other area of business activity.

4.5.5 Market Supply in the Long Run

You will recall that we derived the short-run market supply curve by horizontal simulation of the existing firms' short-run supply (MC) curves. This we cannot do in the long run because there are no 'existing' firms in the long run.

Figure 4.21 shows again a short-run market supply curve and one representative firm. Suppose we have *n* firms in this market in the short run, and further suppose they are all identical, i.e. they all have identical AVC, ATC and consequently MC curves. For any price less than P_1, firm output will be zero because variable costs cannot be covered. Consequently, market supply (S) will be zero for prices less than P_1. At P_1, a firm would cover its variable cost but be indifferent between producing q_1 units or zero. Since a firm will produce output at a price of $P_1 + .0001P_1$, let us assume that a firm will actually produce $0q_1$ units at a price of P_1. Thus the market supply at a price of P_1 will be $q_1 \times n$ (number of firms). And we know that $q_1 \times n = Q_1$.

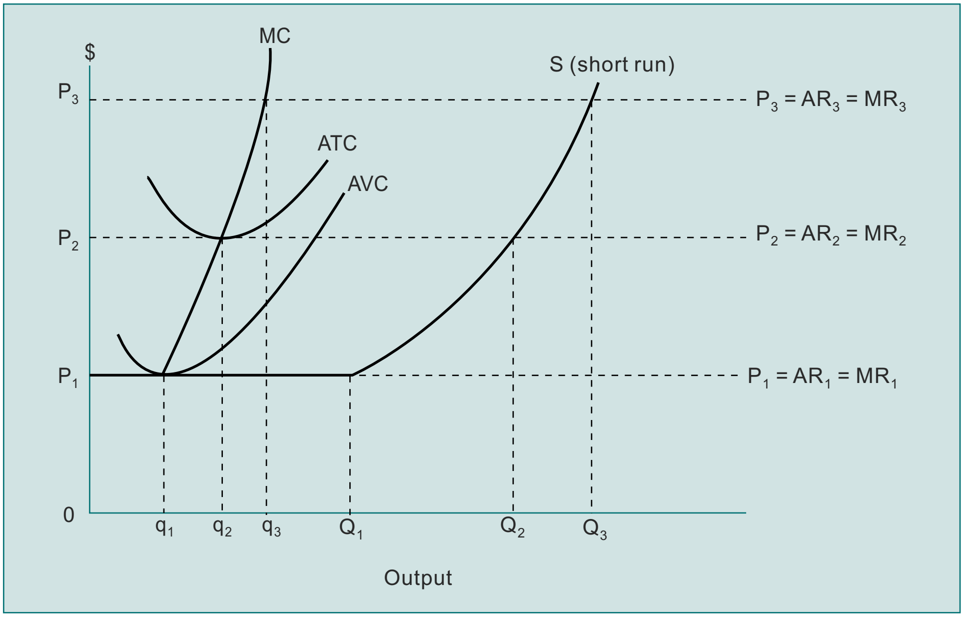


Figure 4.21 Firm and market supply in the short run

At a price of P_2, a firm will produce q_2 as the output where MC = MR_2 and similarly to the above, market supply will be Q_2. At a price of P_3, the market supply will be Q_3, and so on. Market supply S, then, is the sum of the *n* firms' marginal cost curves.

Suppose the going price is P_3. What will happen in the long run? Resource owners will observe above-normal returns being made in this industry and will divert resources into the industry. There are two cases to consider in these circumstances: the first is when the movement of resources into the industry does not increase the prices of the factors of production employed in the industry; and the second is when factor prices rise.

First, if factor prices do not increase as new firms move in, the cost of producing a unit of output will remain unchanged. Firms will continue to move into an industry as long as they can cover their average total costs (including, remember, their opportunity cost). Because factor prices do not increase with the influx of new firms, the average total cost of production will remain constant and consequently the long-run industry supply curve will be a horizontal line. This is shown in Figure 4.22.

At any price below P_2, no output will be supplied in the long run because no firm can cover its average total cost. At a price of P_2, firms can cover all costs and earn normal returns. At prices above P_2, above-normal return (excess profit) will exist. Since, as new firms enter, the prices of factor inputs remain unchanged, the unit cost of production will also remain unchanged, resulting in a horizontal industry supply curve at a price of P_2 (equal to the minimum average total cost of production). Thus if there were 80 identical firms in the industry, each would be willing to produce q_2 in the long run at a price of P_2, industry supply would be Q_{80}. If there were 100 identical firms instead of 80 and the price were still P_2, total industry supply would be Q_{100}. Each firm would still produce q_2.

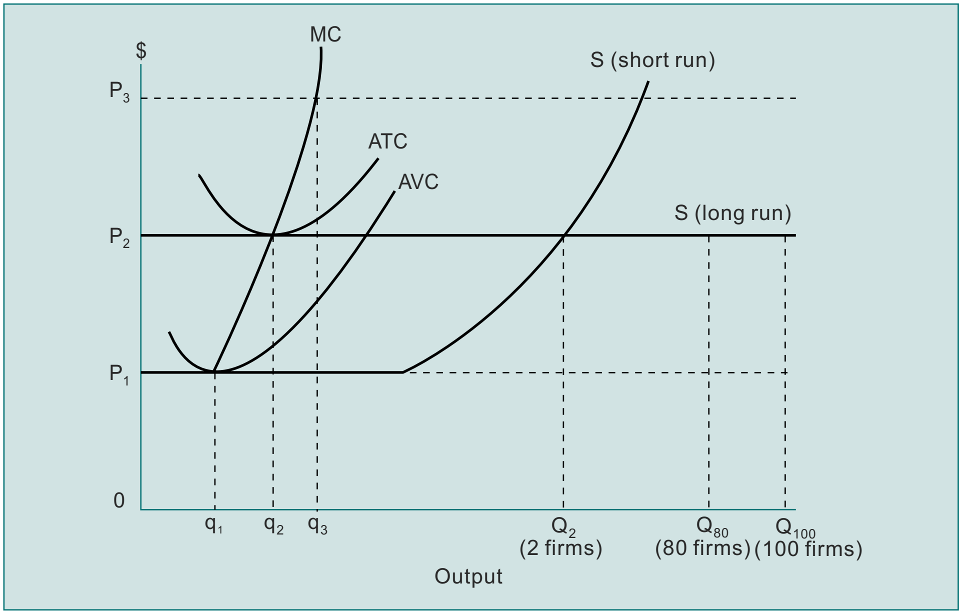


Figure 4.22 Long-run supply curve

How realistic is the assumption of no increase in factor input prices as an industry expands? The answer depends on the size of the expansion relative to the factor markets involved, for all industries compete in the same markets for factor inputs. Steel is used by the car industry, the shipbuilding industry and the construction industry, to name but a few; the private home construction industry uses steel but not a lot. It is quite conceivable, therefore, to imagine a significant increase in home construction having a negligible impact on the price of steel. In contrast, a major expansion in the shipbuilding industry might well be expected to cause the price of steel to rise.

If we assume that additional factor inputs for an industry will only be forthcoming at higher prices, then the long-run supply curve of that industry will no longer be a horizontal curve, but will be positively inclined to the price axis. Such a situation is represented in Figure 4.23.

Suppose we start off with *n* firms in the industry each producing q_2: industry output will be Q_2. Another *R* firms now move into the industry, causing factor input prices to rise. Each firm's average total cost curve (ATC_1) increases to ATC_2. In the long run each firm will only be willing to supply q_2 at a price of P_3. However, there is now a larger market supply; we now have $n + R$ firms each producing q_2, but since average total cost for each firm has risen, it requires a price of P_3 to attract an industry output of Q_3 in the long run. Thus in the long run, because a greater supply is forthcoming only at a higher price, the long-run supply curve will be upward sloping, positively inclined towards the price axis.

It is not impossible to envisage a situation in which an increase in demand for factor inputs causes a fall in their unit costs. Many suppliers of factor inputs – for instance electronic companies supplying personal computers to industry and publishing companies supplying textbooks to schools and universities – give quantity discounts, so that the larger the quantity bought, the lower is the unit price. If such discounts were not offset by other factor input prices rising as an industry output expanded, the long-run supply curve of such an industry would be a downward sloping curve, negatively inclined to the price axis. But since such a case tends to be the exception rather than the rule, it will be ignored here.

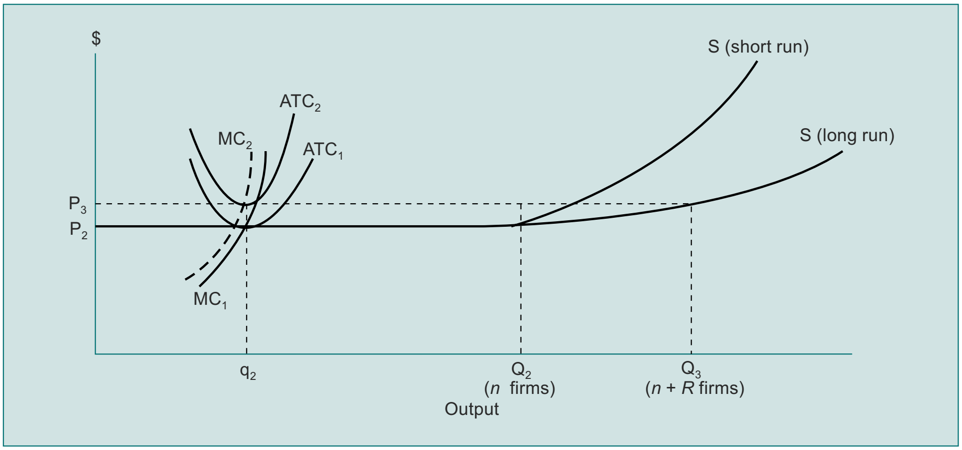


Figure 4.23 Long-run supply curve (increasing factor prices)

4.6 Real-World Applications

4.6.1 Estimating Production Frontiers

In the real world, estimating production frontiers is not an easy task. The data that are collected in plants or companies on inputs and outputs relate to actual performance and therefore do not necessarily record maximum output achievable from given inputs. Many companies offer bonus schemes to employees or groups of employees who can suggest ways of increasing output from given resources. What such companies are implicitly doing is attempting to move closer to their production frontiers.

Several years ago a well-known car producer set up identical automobile plants in different European countries. What was observed is shown in Figure 4.24.

Consider the two observations on the lower left. With the same plant in both countries, country A is producing fewer cars in the same time period but employing more labour. The two observations on the upper right show the situation when both plants were operating double shifts (i.e. employing night staff). Again, country A produces fewer cars but still employs more men than country B. Now, we do not know whether country B is on the production frontier – it isn't according to the production frontier we have drawn in, and what we do know is that country A is definitely not on the production frontier and is definitely operating in a less engineering-efficient fashion than the country B plant. Note carefully that such a comparison is possible only when the capital inputs and all other significant factors are comparable in both plants.

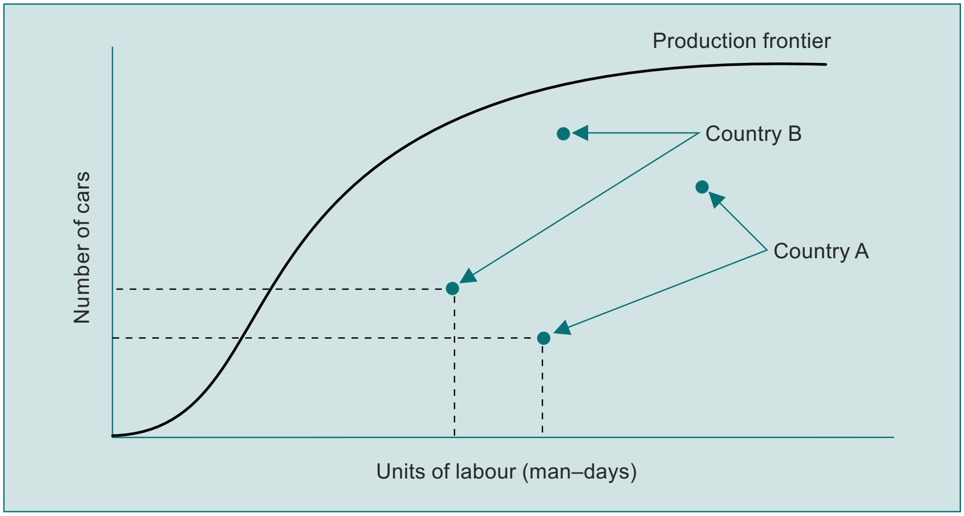


Figure 4.24 Output comparisons in car plants

4.6.2 Labour Productivity

From Figure 4.24 it is possible to calculate average labour productivity, i.e. output per unit of labour for a specified time period (per day, per month or per annum, etc.). When such rates differ by plant, given identical fixed factor inputs, multinational companies with plants in different countries will take such facts into account when deciding on new plant locations. Such comparisons, however, in the real world cause considerable controversy between management and organised labour. Restrictive labour practices, overmanning, weak management, inadequate supervision, poor communication between workers and managers, and inferior capital equipment all contribute to labour productivity being lower than it otherwise could be. This results in costs being higher than they otherwise could be, and, in the highly competitive international market, this in turn results in lost sales.

Ignoring different international capital endowments, estimates are prepared annually for the more developed economies for total annual output of the economy divided by labour input, and growth rates per annum for labour productivity can then be estimated. For the major industrialised nations for the last three decades, annual labour productivity growth has averaged just over 3 per cent.

Another important implication of what we have studied in this module lies in the equating of the marginal product of labour to the wage rate in competitive labour markets. If each worker is paid the value of his marginal product, and since for most people the primary source of income arises through the sale of labour services, how well off a person is will depend upon his contribution to output. Individuals whose marginal products are low, namely the unskilled, tend to have very low incomes. Individuals whose marginal products are considered to be zero, or close to zero, by firms will have no jobs. Low productivity poses problems not only for households whose members are primarily in such a category but also for society as a whole in 'taking care' of such a segment of society. As we shall see when studying macroeconomics, a principal cause of poverty is unemployment: when unemployment rates fall, poverty levels fall. However, no matter how low the unemployment rate becomes, those individuals with low or zero labour productivity remain with us – and are probably unemployed.

4.6.3 Factor Returns and Scale Returns

From the shape of the production frontiers and concomitant short-run average cost curves, it is obvious that the relationship between changes in factor inputs and the resultant changes in output is not constant. Were it constant, production frontiers would be straight lines. The relationship between changes in output and changes in the variable factor input can take three forms: they can be increasing, constant and diminishing returns.

Increasing returns to the variable factor input result when an increase in the variable factor input results in an increase in total output that is proportionally greater than the increase in the factor input; increasing returns occur when

$$\frac{\Delta Q}{Q} > \frac{\Delta L}{L}$$

where Q is the output and L is the variable factor input.

Constant returns occur when

$$\frac{\Delta Q}{Q} = \frac{\Delta L}{L}$$

and diminishing returns occur when

$$\frac{\Delta Q}{Q} < \frac{\Delta L}{L}$$

Returns to factor inputs must be distinguished from *returns to scale*. Returns to factor inputs refer to varying one factor and holding all other inputs constant, and so it is a short-run phenomenon. Returns to scale refer to a change in output resulting from a change in all factor inputs, so that it is a long-run phenomenon.

Increasing returns to scale occur when, as all factor inputs are increased, the resulting increase in output is proportionally greater than the increase in the factor inputs. Increasing returns to scale occur when

$$\frac{\Delta Q}{Q} > \frac{\Delta (L,C)}{L,C}$$

where Q is the output, L and C are both variable factor inputs.

Constant returns to scale occur when

$$\frac{\Delta Q}{Q} = \frac{\Delta (L,C)}{L,C}$$

and diminishing returns to scale occur when

$$\frac{\Delta Q}{Q} < \frac{\Delta (L,C)}{L,C}$$

There is no relationship between returns to factors of production and returns to scale.

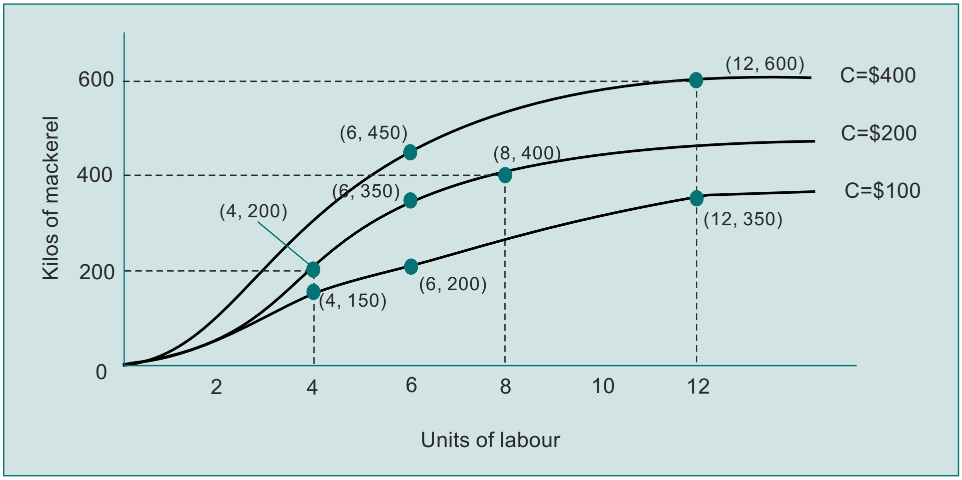


Figure 4.25 Factor returns: returns to scale

Figure 4.25 shows examples of these concepts. The figure contains three production frontiers representing different capital inputs of $100, $200 and $400 (three different fishing boats, say – we are mackerel fishing now; salmon is out of season). Labour inputs are man-days.

Choosing the $100 capital input (lowest curve in Figure 4.25), we can see that a doubling of labour input (the variable factor of production) from 6 to 12 results in an increase in output from 200 to 350 kilos. Therefore

$$\frac{\Delta Q}{Q} = \frac{350-200}{200} = 75\%$$

$$\frac{\Delta L}{L} = \frac{12-6}{6} = 100\%$$

Thus over the range of output there are diminishing returns to the variable factor input. Choosing the $200 capital input, a 100 per cent increase in labour input from 4 to 8 results in an increase of output from 200 to 400 kilos (a 100 per cent increase), indicating constant factor returns.

A doubling of capital input from $100 to $200 and at the same time a doubling of labour input from 4 to 8 increase output from 150 kilos to 400 kilos, indicating increasing returns to scale. A doubling of capital input from $200 to $400 and at the same time a doubling of labour input from 6 to 12 results in an increase of output from 350 to 600 kilos, a 71% increase, indicating diminishing returns to scale.

4.6.4 Supply Elasticity

Price elasticity of supply (normally called *supply elasticity*) is a measure of the responsiveness of quantity of output supplied to a change in price. In the short run we would measure elasticities at different points on the industry short-run supply curve; in the long run these elasticities would be measured along the long-run industry supply curve. Elasticity of supply is defined algebraically as

$$\frac{\frac{\Delta Q}{Q}}{\frac{\Delta P}{P}}$$

Supply is said to be elastic when:

$$\frac{\Delta Q}{Q} > \frac{\Delta P}{P}$$

and inelastic when

$$\frac{\Delta Q}{Q} < \frac{\Delta P}{P}$$

and of unitary elasticity when

$$\frac{\Delta Q}{Q} = \frac{\Delta P}{P}$$

The supply of some goods, of course, is completely inelastic: there is only one *Mona Lisa*; when two given straight roads intersect there are only four corners; and at Wimbledon tennis championships there are a fixed number of Centre Court seats. Excepting all these goods and services of fixed supply, long-run elasticities are normally greater than short-run elasticities. Back to our fishing example: at the end of the day when all boats return and the fish are counted, they are in fixed supply, i.e. a higher price will not bring forth more fish that day. The following day, however, more fishermen will be hired in response to the higher price, and the supply of fish should increase. The ultimate limit on supply in the short run will be the number of boats. In the long run, no such constraint exists and if prices remain high, resources will move into the fishing industry and the supply of fish will continue to increase.

Learning Summary

You should by now understand thoroughly the relationship between productivity of factor inputs and cost of goods produced and that underlying every cost curve is a productivity curve. You can derive average and marginal productivity curves from total production functions, and marginal and average cost curves from total cost functions. You are able to use marginal analysis to find profit-maximising outputs at different prices and profit-maximising input combinations for different levels of output.

You can now explain the difference between the short and long run, understand the significance of fixed costs versus variable costs, know the meaning of equilibrium for the firm in both the short and long run, and derive market supply curves. You can distinguish between returns to factor input and returns to scale.

Review Questions

Multiple Choice Questions

4.1 Which of the following is correct?

When one good is being produced using the minimum amount of resources

A. economic efficiency will prevail in the economy as a whole.
B. engineering efficiency will prevail in the production of that good.
C. all resources in the economy will be employed.
D. an increase in the production of that good must lead to a reduction in the production of other goods.

4.2 The economy is *not* producing its maximum output of goods if

I. goods are not being produced in an engineering-efficient manner but all resources are employed in production.
II. goods are being produced in an engineering-efficient manner but not all resources are employed in production.
III. goods are not being produced in an engineering-efficient manner and not all resources are employed in production.

Which of the following is correct?

A. I only.
B. II only.
C. I or II only.
D. I or II or III.

4.3 Economic efficiency is achieved in an economy when

I. all goods and services are produced using the least possible amount of resources.
II. all resources are employed.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

4.4 If you were to observe a firm supplying none of a particular good at its existing price, it could justifiably be concluded that

I. the good is incapable of making a profit for the firm at any price.
II. the firm expects the price to fall in the future.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

4.5 A firm that produces a range of cosmetics products moves from Britain to Germany. In comparing monthly sales, it discovers that in Germany it supplies a greater proportion of high-cost items than it did in Britain. The factors that could *not* have caused this change in the firm's sales are

I. higher income levels in Germany compared with Britain.
II. the length of the period being compared.
III. different relative prices of cosmetic products in the two countries.

Which of the following is correct?

A. I only.
B. II only.
C. II and III only.
D. I, II and III.

4.6 What is the correct definition of the term 'short run' as used in the theory of supply?

A. The time period between the decision to produce a good and its final appearance on the market.
B. A time period of less than one year in which more than one factor of production can be altered.
C. A time period of up to three years in which any factor of production can be altered.
D. The time period in which a firm cannot increase or decrease the quantity it hires of at least one factor of production.

4.7 In the theory of supply, what does the term 'long run' refer to when used in connection with a firm's decision making?

A. The number of years over which a firm can expect to produce with unaltered factors of production.
B. Any period of time between three and five years in which a firm can increase or decrease all the factors of production it employs.
C. Any period of time over five years in which a firm can increase or decrease all the factors of production it employs.
D. The time period in which a firm can increase or decrease all the factors of production it employs.

4.8 A firm must make maximum profits when

I. it raises the maximum amount of revenue from selling its products.
II. its products are produced at the lowest possible costs per unit of output.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

Questions 4.9 and 4.10 are based on Figure 4.26.

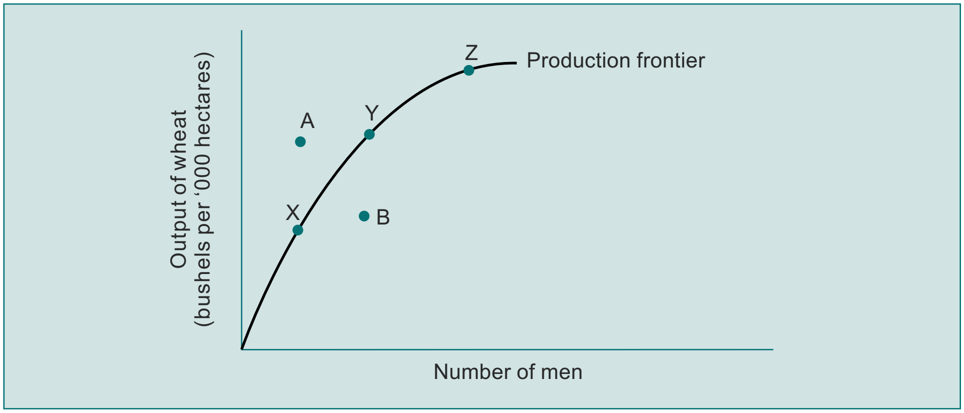


Figure 4.26

4.9 Figure 4.26 provides information about

I. engineering-inefficient outputs.

II. economic efficiency, since wheat is always in demand because of world food shortages.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

4.10 From Figure 4.26 it can be concluded that

I. points X, Y and Z are the only attainable outputs.

II. B is an attainable output.

III. A is an unattainable output even in the long run.

Which of the following is correct?

A. I only.

B. II only.

C. III only.

D. Not I, nor II nor III.

4.11 Which of the following is correct?

The output of a good (i.e. the quantity produced) divided by the number of units of a variable factor used in the production of that good is a definition of:

A. marginal product of that factor.

B. average product of that factor.

C. total cost of production.

D. variable cost of production.

4.12 A hop grower in Kent noted that, within limits, the larger the number of workers he hired the greater was the output of hops from his fields. This means that for the number of workers being employed by the grower the marginal product of labour is necessarily

A. higher than the average product of farm labour.
B. lower than the average product of farm labour.
C. greater than zero.
D. less than zero.

4.13 Which of the following is correct?

For a firm, in the long run

A. fixed costs tend to be greater than variable costs.
B. variable costs tend to be greater than fixed costs.
C. all costs are fixed.
D. all costs are variable costs.

4.14 If a perfectly competitive firm is maximising profit in the long run, price equals

I. short-run marginal cost.
II. long-run marginal cost.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

4.15 The long-run supply curve of a competitive industry is found by adding

I. the long-run marginal cost curves of all firms in the industry.
II. the long-run marginal cost curves above the average cost curves of all firms in the industry.
III. the long-run average cost curves of all firms in the industry.

Which of the following is correct?

A. I only.
B. II only.
C. III only.
D. None of the above.

4.16 If an increase in the demand for aircraft causes the price of aluminium to rise, how will this, other things being unchanged, affect the industry that produces aluminium beer cans?

A. Its output will increase.
B. Its long-run supply curve will shift to the right.
C. Its long-run supply curve will shift to the left.
D. Its long-run supply curve will not be affected.

4.17 If a firm is producing output at a point where diminishing returns have set in, then

I. each additional unit of output will be more expensive to produce.
II. each additional unit of output will require increasing amounts of the variable factor of production.
III. the marginal product of the variable factor of production decreases as the quantity used increases.

Which of the following is correct?

A. I only.
B. I and II only.
C. I and III only.
D. I, II and III.

4.18 A recent economics graduate was hired by a highly competitive restaurant and was assigned the task of improving efficiency in the restaurant. It was felt that an inefficient combination of two variable factor inputs was being employed. The economist was able to make the following calculations where each lunch sells for $2.00:

	Marginal product	**Cost per unit of factor input**
Factor A	6 lunches	$3.00
Factor B	3 lunches	$1.00

The economist correctly diagnosed the problem and advised the restaurant owner on the appropriate course of action. That course was to:

A. hire more A and more B.
B. hire more A but no more B.
C. hire more B but no more A.
D. hire less B and less A.

4.19 The higher the

I. marginal product of a factor input, the lower the marginal cost of output.
II. average product of a factor input, the lower the average cost of output.
III. costs of fixed factor inputs, the lower the marginal cost of output.

Which of the following is correct?

A. I only.
B. III only.
C. I and II only.
D. I, II and III.

4.20 For a firm with given plant and equipment, which of the following will always decrease as output increases?

A. Total variable cost.
B. Total fixed cost.
C. Variable cost per unit.
D. Fixed cost per unit.

4.21 In one of the largest manufacturing firms, an agreement was recently reached between management and unions that wage rates would increase proportionately with average labour productivity (the output per employee). Last year the firm increased its plant and equipment, retained all of its labour force fully employed, and enjoyed a 6 per cent increase in output. The unions have gone on strike because management has refused to pay any wage increase.

According to the agreement, which of the following is correct?

A. The unions should not receive a wage increase because the 6 per cent increase in output is attributable to the increase in plant and equipment.

B. The unions should receive a wage increase because average labour productivity has increased.

C. The unions should receive a wage increase only if the firm increases its plant and equipment by less than 6 per cent.

D. Whether the unions should receive an increase according to the agreement cannot be determined from the given information.

4.22 A family will be away from its house for six months. The monthly mortgage payment on the house is \$130. Local utilities, to be paid by the owner, are \$30 per month if the house is occupied, otherwise zero. If the family wished to minimise its losses (or maximise its gains) from the house (assuming wear and tear, etc. to be zero whether or not the house is occupied), it should let for as much as the market will bear as long as monthly rent is above the following:

A. 0.

B. \$30.

C. \$130.

D. \$160.

4.23 The short-run supply curve of an industry would shift to the right as a result of

I. an increase in the number of firms in the industry.

II. an increase in the cost of the variable factor of production.

III. a decrease in the cost of the variable factor of production.

Which of the following is correct?

A. I only.

B. III only.

C. I and II only.

D. I and III only.

Case Study 4.1: How to Produce Switches

This case is concerned with decision making in a firm that has the option of buying different types of machinery to perform a given task. Before you tackle it you should understand:

a. total cost
b. variable cost
c. fixed cost
d. marginal cost
e. average cost
f. profit maximisation.

The case demonstrates that information on costs is not sufficient to make decisions on investments. A great deal depends on what is expected to happen to demand in the future.

Switch-Production Options

A firm that produces electrical components has decided to break into the market for electrical switches. Studies on both the cost of producing switches and the potential market have been carried out. At a board meeting, the information on costs was presented with a view to deciding which type of machine should be purchased for the production of switches. It had been decided that the best two machines on the market were a small type, renting for €1000 per week, and a larger one renting for €5000 per week. The smaller machine could be used by up to four workers, and each worker would produce five switches per week; the larger machine could be used by up to eight workers, each producing 15 switches per week. The average cost and output of each type of machine working at full capacity is shown in Table 4.10.

Table 4.10 Cost of production (euro)

Total cost	Machine cost	Wage cost	Output	Average cost	No. of machines	No. of workers
1 400	1 000	400	20	70	1	4

Machine cost = 1000

Workers per machine = 4

Wage rate = 100

Output per worker = 5

Total cost	Machine cost	Wage cost	Output	Average cost	No. of machines	No. of workers
5 800	5 000	800	120	48	1	8

Machine cost = 5000

Workers per machine = 8

Wage rate = 100

Output per worker = 15

The Cost Accountant explained that the firm would be committed to one type of machine only, because different skills and facilities were required to operate the different machines. Thus it would *not* be possible to buy some large and some small machines. Some board members jumped to the conclusion that the larger machine should be purchased because the average cost was much lower at €48 per switch compared with €70 per switch for the smaller machine. However, the Managing Director pointed out that the average cost would vary depending on how fully utilised the machines were and instructed the Cost Accountant to calculate the average cost for a wide range of output. The Cost Accountant later presented the data in Table 4.11 and Table 4.12, showing the average costs and levels of output for different numbers of workers and machines of the small and large types respectively.

It became immediately clear that the selection of the appropriate machine depended on the expected level of sales.

Table 4.11 **Cost of production as output varies (small machine) (euro)**

Total cost	Machine cost	Wage cost	Output	Average cost	No. of machines	No. of workers
1 100	1 000	100	5	220	1	1
1 200	1 000	200	10	120	1	2
1 300	1 000	300	15	87	1	3
1 400	1 000	400	20	70	1	4
2 500	2 000	500	25	100	2	5
2 600	2 000	600	30	87	2	6
2 700	2 000	700	35	77	2	7
2 800	2 000	800	40	70	2	8
3 900	3 000	900	45	87	3	9
4 000	3 000	1 000	50	80	3	10
4 100	3 000	1 100	55	75	3	11
4 200	3 000	1 200	60	70	3	12
5 300	4 000	1 300	65	82	4	13
5 400	4 000	1 400	70	77	4	14
5 500	4 000	1 500	75	73	4	15
5 600	4 000	1 600	80	70	4	16
6 700	5 000	1 700	85	79	5	17
6 800	5 000	1 800	90	76	5	18
6 900	5 000	1 900	95	73	5	19
7 000	5 000	2 000	100	70	5	20
8 100	6 000	2 100	105	77	6	21
8 200	6 000	2 200	110	75	6	22
8 300	6 000	2 300	115	72	6	23
8 400	6 000	2 400	120	70	6	24
9 500	7 000	2 500	125	76	7	25
9 600	7 000	2 600	130	74	7	26
9 700	7 000	2 700	135	72	7	27
9 800	7 000	2 800	140	70	7	28
10 900	8 000	2 900	145	75	8	29
11 000	8 000	3 000	150	73	8	30
11 100	8 000	3 100	155	72	8	31
11 200	8 000	3 200	160	70	8	32
12 300	9 000	3 300	165	75	9	33
12 400	9 000	3 400	170	73	9	34
12 500	9 000	3 500	175	71	9	35
12 600	9 000	3 600	180	70	9	36
13 700	10 000	3 700	185	74	10	37

13 800	10 000	3 800	190	73	10	38
13 900	10 000	3 900	195	71	10	39
14 000	10 000	4 000	200	70	10	40
15 100	11 000	4 100	205	74	11	41
15 200	11 000	4 200	210	72	11	42
15 300	11 000	4 300	215	71	11	43
15 400	11 000	4 400	220	70	11	44
16 500	12 000	4 500	225	73	12	45
16 600	12 000	4 600	230	72	12	46
16 700	12 000	4 700	235	71	12	47
16 800	12 000	4 800	240	70	12	48

Machine cost = 1000; Workers per machine = 4; Wage rate = 100; Output per worker = 5.

Table 4.12 Cost of production as output varies (large machine) (euro)

Total cost	Machine cost	Wage cost	Output	Average cost	No. of machines	No. of workers
5 100	5 000	100	15	340	1	1
5 200	5 000	200	30	173	1	2
5 300	5 000	300	45	118	1	3
5 400	5 000	400	60	90	1	4
5 500	5 000	500	75	73	1	5
5 600	5 000	600	90	62	1	6
5 700	5 000	700	105	54	1	7
5 800	5 000	800	120	48	1	8
10 900	10 000	900	135	81	2	9
11 000	10 000	1 000	150	73	2	10
11 100	10 000	1 100	165	67	2	11
11 200	10 000	1 200	180	62	2	12
11 300	10 000	1 300	195	58	2	13
11 400	10 000	1 400	210	54	2	14
11 500	10 000	1 500	225	51	2	15
11 600	10 000	1 600	240	48	2	16
16 700	15 000	1 700	255	65	3	17
16 800	15 000	1 800	270	62	3	18
16 900	15 000	1 900	285	59	3	19
17 000	15 000	2 000	300	57	3	20
17 100	15 000	2 100	315	54	3	21
17 200	15 000	2 200	330	52	3	22
17 300	15 000	2 300	345	50	3	23
17 400	15 000	2 400	360	48	3	24

Machine cost = 5000; Workers per machine = 8; Wage rate = 100; Output per worker = 15.

The following questions need consideration.

1. Why does average cost have a wave-like pattern for both machines?
2. If it was estimated that between 45 and 60 switches would be sold per week, which machine should be rented?
3. If it was estimated that between 135 and 150 switches could be sold per week, which machine should be rented?
4. If it was considered likely that more than 165 switches would be sold per week, which machine should be rented?
5. If it was thought that there might be a 50 per cent increase in wage costs by the time production began, what general effect would this have on the relative attraction of the two types of machine?

Case Study 4.2: The Price of Butter

This case is an application of the theory of supply. Before you tackle it you should understand:

a. linear supply curves
b. total cost
c. elasticity of supply
d. *ceteris paribus*
e. movements along a supply curve
f. shifts of a supply curve
g. what is meant by 'estimating a supply curve'.

The case demonstrates:

- why it is important to have information about the shape of the supply curve when setting the price offered to producers;
- the consequences of price changes for supply curves of different elasticities;
- that many issues in economics are resolved only when theories and hypotheses are subjected to empirical tests.

The Subsidy to Farmers

A country in the EU is currently producing four million kilos of butter per year. The price paid to producers by the EU Commissioners is $2 per kilo. This price is unaffected by the subsidised price charged to consumers on the open market, which tends to vary between $1.00 and $1.50 per kilo.

The Attempt to Increase Supply

It has been decided that the amount of butter produced should be substantially increased, and that this should be done by increasing the price offered to producers from $2 to $3 per kilo. To determine the likely effect of this price increase on the supply of butter, the EU Commissioners asked the Confederation of Milk Producers to carry out a study to estimate the supply curve for butter. In response to objections that the Milk Producers might not be impartial, an independent team of economists was also employed to carry out the same study.

The two reports produced conflicting findings. The Milk Producers – Group A – estimated that the supply of butter would increase from the current four million kilos per year to nine million kilos in response to the price increase. This, they claimed, meant that the price increase was well worthwhile because a 50 per cent price increase would lead to a 125 per cent increase in the output of butter. The economists – Group B – estimated that the output of butter would increase by only one million kilos per year, to five million kilos. Furthermore, they claimed, the additional one million kilos of butter would cost almost as much to the EU as the existing four million kilos cost at the moment.

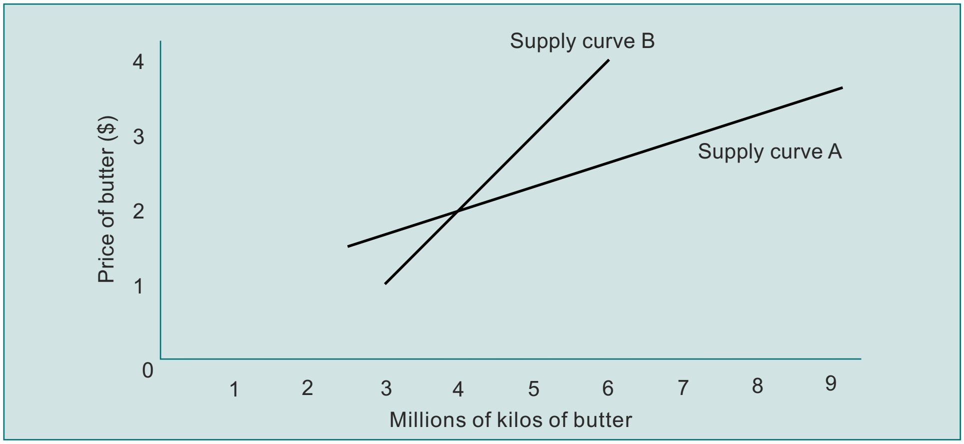


Figure 4.27 The supply of butter

The Chair of the Commission pondered the two reports, and in particular the two versions of the supply curve, which are shown in Figure 4.27.

1 The following points in particular taxed the Commission in its deliberations; what is your view?

1. If Group A is correct, what would be the cost of the additional five million kilos of butter? What would be the cost per kilo?
2. If Group B is correct, what would be the cost of the additional one million kilos of butter? What would be the cost per kilo?
3. What is the difference between the two supply curves?
4. Why does it cost more than $3 per kilo for the additional output of butter in both cases?
5. If Group B is correct and the price is increased to $3 per kilo, how confident would you be that the supply would turn out to be precisely five million kilos of butter next year?

Case Study 4.3: Production Designs for a Jewellery Manufacturer

This case is concerned with finding the profit-maximising level of output for goods with different costs and prices. Before you tackle it you should understand:

a. fixed cost, average fixed cost
b. variable cost, average variable cost
c. marginal cost
d. profit-maximising output
e. loss-minimising output
f. the reason for ceasing production in the short run.

The case demonstrates that:

- in the short run it is necessary to alter output as market conditions change in order to maximise profits;
- despite the fact that a good may be making a loss it may still be rational to carry on production in the short run; and
- in the short run it may be necessary to cease production of certain loss-making outputs.

Production Decisions

The Easycharm Manufacturing Company produces four products: agate bracelets, turquoise necklaces, opal rings, and amethyst brooches. Its objective is to produce and sell the quantity of each product that maximises profit. Current market prices of the products are given in Table 4.13.

Table 4.13 The Easycharm Manufacturing Company's products and prices

Product	Price ($)
Bracelets	4.00
Necklaces	300.00
Rings	15.00
Brooches	8.00

Retail outlets are prepared to take at these prices as many bracelets, necklaces, rings and brooches as the Easycharm Manufacturing Company can produce. Total cost of production for different quantities of each product are given in Table 4.14.

Table 4.14 Total costs for different levels of output

Bracelets		Necklaces		Rings		Brooches	
Output per day	Total cost ($)	Output per day	Total cost ($)	Output per day	Total cost ($)	Output per day	Total cost ($)
0	100	0	800	0	300	0	0
100	200	1	900	10*	400	10	100
200	350	2	1 025	20	540	20	210
300*	550	3	1 200	30	700	30*	330
400	850	4	1 500	40	900	40	490
500	1 300	5*	1 900	50	1 500	50	700

* Current daily production

1 Madame Pauline, the Managing Director, is worried about the financial position of the company and doubts whether the goal of profit maximisation is being achieved. You have been employed as a consultant by the Marketing Director and given the following brief:

> Given the continuation of current market prices and production schedules, please advise me which, if any, of our activities should be expanded, contracted or eliminated in the short run.

Case Study 4.4: The Empty House

This case is an application of the concepts of fixed and variable costs. Before you tackle it you should understand:

a. fixed cost
b. variable cost
c. opportunity cost.

The case demonstrates that the 'common sense' method of deciding what rent to charge for an empty house is usually wrong. Sometimes a rather surprising answer can emerge from the application of economic analysis.

Letting-out Decisions

A surgeon at a local hospital has been offered a one-year visiting appointment in the USA. The visiting appointment, which carries a very handsome salary, was offered to the surgeon after she had completed pioneering research on kidney transplants. Her hospital is prepared to grant her leave of absence, without pay, for one year. She is considering whether or not to let her house while she is in the USA, and has compiled the figures shown in Table 4.15, which she considers relevant in helping her to decide whether to let the house, and also to decide the minimum weekly rent she would be prepared to accept rather than leave the house unoccupied. The surgeon also wishes the following conditions observed:

1. Anyone renting the house pays for the gas and electricity used.
2. Anyone renting the house will take care of the dog for free.
3. The house will not be left unattended for long periods since this would necessitate the services of a security firm. The surgeon has stated that the minimum she would be willing to accept in rent for the house is $30 per week, equal to the mortgage payments.

Table 4.15 The surgeon's housing cost estimates

Item	Average house outlays per week ($)	
Mortgage repayments		30.00
Electricity	(if house empty)	0.00
	(if house occupied)	6.00
Gas	(if house empty)	0.00
	(if house occupied)	20.00
Gardener	(if house empty)	10.00
	(if house occupied)	0.00
Depreciation	(if house empty)	5.00
	(if house occupied)	10.00
Insurance		3.00
Rates		15.00
Security firm	(if house empty)	10.00
	(if house occupied)	0.00
Dog care	(if house empty)	8.00
	(if house occupied)	0.00

1 Assuming that no other factors affect the decision to let, what is your calculation of the minimum amount the surgeon should be willing to accept for letting her house? Set out your reasoning so that the surgeon will understand it.

Module 5

The Market

Contents

5.1 Introduction

A market exists when individuals, households and firms who want to buy a good or service are in contact with firms or institutions willing to supply that good or service. Some markets are highly localised; others are international. For example, the village babysitting market is small (few transactions because there are only a few babies in the village) and highly localised (all babysitters are teenage neighbours). In contrast, coffee, cocoa, sugar, shipping, cars, copper, steel and wheat are bought and sold at the wholesale level in international markets by dealers from many countries throughout the world.

Markets exist for resources as well as for goods and services. Most major cities have newspapers containing a property section and a stock market. The property section contains information not only on houses, flats, and commercial property for sale, but also on properties for rent. Sometimes a 'fixed' price for a piece of property is advertised, i.e. the first person to offer that price secures the property; sometimes an 'offers over' price is stipulated with a closing date for offers. Sometimes offers are invited, counter-offers are made, and then yet more counter-offers until a price is agreed between buyer and seller. Furthermore, the property section can contain 'accommodation/property wanted' advertisements accompanied by price ranges. The stock market is concerned with the buying and selling of 'pieces' of companies, denoted by shares. The major companies in the world are quoted on the international stock markets and the numbers of shares traded in these companies each day are recorded and published, as are the prices at which the exchanges took place. 'Bid' and 'offer' prices (as the buying and selling prices are known) are also published. With the advent of sophisticated computers, data on world transactions are recorded and available as they are made.

In markets in which exchange takes place, the competition (either implicit (e.g. a supermarket) or explicit (the stock market)) among buyers to buy and among sellers to sell, determines the price or prices and the quantities that are bought and sold. Market prices provide information to buyers attempting to maximise utility from the allocation of their budgets and also to suppliers attempting to maximise profit from the production and sale of goods, services and resources.

While the decisions of suppliers determine which resources are employed and consequently which goods and services are produced and how much of each good and service is produced, economic efficiency requires that resources are allocated to the production of those goods and services that satisfy society's wants as fully as possible. Prices indicate to suppliers society's preferences for different goods and services; and to survive in a competitive world, it is to these prices that suppliers must respond in their decision-making processes. Thus, ultimately, prices determine how resources are allocated. For example, the dramatic decrease in the price of air travel in the past three decades has been a major contributing factor to the number of people seeking vacations abroad, especially in countries offering sun and surf. This in turn has spawned construction of massive hotel and apartment complexes near previously deserted beaches and has attracted resources to vacation-related activities – eating, drinking, dancing.

In a market economy, a system of interrelated markets for goods and services determines how resources are allocated, and it operates without any centralised planning mechanism or central decision-making body. In such economies, millions of households daily buy various quantities of food and drink, including bread, cheese, meat, milk, coffee, tea, tins of fruit, fresh fruit, frozen vegetables, fresh vegetables, butter, margarine, fish, beer, whisky, Alka Seltzer and aspirin. Retail outlets of every size, from small corner grocery stores to giant supermarket chains, regularly estimate how much of each kind of food and drink they will sell, and place orders with wholesalers, who in turn place orders that filter down to the primary suppliers. These suppliers base production plans for food, drink etc. on estimated consumer demand and allocate resources to the supply of various kinds of food and drink.

The relative quantities of food and drink bought by households are not decided by some central plan, nor are the resources allocated to the production of food and drink decided by a planning mechanism. Instead, the pattern of resource allocation is the outcome of millions of households expressing their preferences for various kinds of food and drink in markets and the decisions of independent suppliers to hire resources to produce the food and drink to satisfy these households. By these means, in the best restaurants in the major capitalist cities in the world, you can take a party of friends and with a high degree of probability be able to order Russian caviar, Scottish smoked salmon, well-aged sirloin steak, fresh salad and fine Bordeaux wines.

The magic of the price mechanism makes such a meal possible at almost any time. If you are having that meal in New York, spare a thought for the number of interrelated markets involved in delivering your wild smoked salmon; it was caught in a Scottish river by fishermen wearing waders, a waterproof coat covering a tweed jacket over a woollen sweater, cotton shirt and cotton underwear. He was smoking Virginia tobacco in a brier pipe and he shaved that morning with an electric razor, having cleaned his teeth, taken a bath and brushed his hair. He caught the salmon on a 15-foot rod that had a steel and brass reel, 50 metres of nylon line and a steel hook to which was tied a fly made from duck feathers. The salmon was transported by road to a smokehouse, where it was gutted, smoked, packaged, refrigerated and trucked to London. It was then flown to New York on a 747 aircraft (guess how many markets were involved in making the 747) and eventually served to you in the restaurant at a price of $15. All these markets just to satisfy your demand for a piece of Scottish smoked salmon and costing you only $15! That is the magic of the price mechanism.

5.2 Market Supply and Demand

Let us now see how the price of a good and the quantity exchanged are determined in a market. A market demand curve relates the quantities of a good that individuals would be prepared to buy at different prices; a market supply curve relates the quantities of a good suppliers would be prepared to sell at different prices. A necessary condition for exchange to take place in a market is that there must be at least one common price at which suppliers would be prepared to sell a quantity of the good and at which individuals would be prepared to buy a quantity of the good.

Table 5.1 contains data relating to hypothetical demand and supply schedules for sandwiches in a British city. The data in Table 5.1 are also expressed diagrammatically in Figure 5.1.

Table 5.1 Market demand and supply schedules for sandwiches in a British city

Price (£)	Quantity demanded ('000 per week)	Quantity supplied ('000 per week)
1.10	180	0
1.20	150	30
1.30	120	60
1.40	90	90
1.50	60	120
1.60	30	150
1.70	0	180

Exchange could take place in this market at any price above £1.10 and below £1.70. This is because at least one consumer would be prepared to buy a sandwich and at least one supplier would be prepared to sell a sandwich at all prices within this range. Exchange would be possible at a price of £1.60, for example, since consumers would be prepared to buy 30 000 sandwiches per week and firms would be willing to supply 150 000 sandwiches per week at that price. At a price of £1.20, exchange would also be possible since consumers would be willing to buy 150 000 sandwiches per week and suppliers would be willing to sell 30 000 sandwiches per week at that price. It would not be possible, however, for exchange to take place at £1.10, because although individuals would be willing to buy 180 000 sandwiches per week at that price, no firms would be prepared to supply sandwiches at £1.10 each. Similarly, exchange could not take place at £1.70 because no individuals would be willing to buy sandwiches at that price, although firms would be prepared to supply 180 000 per week at that price.

The reason that exchange takes place in a market is that both buyers and sellers benefit. Assuming rational behaviour, any consumer who is prepared to buy a unit of a good at a certain price must prefer that good to cash; otherwise there would be no incentive for him to sacrifice cash in order to acquire the good. Similarly, any producer who is prepared to supply a unit of a good at a certain price must prefer cash to the good, otherwise there would be no incentive for him to participate in exchange. In the market for sandwiches illustrated in Table 5.1 and Figure 5.1, consumers would be prepared to buy 120 000 sandwiches at a price of £1.30, and at that price producers would be willing to supply 60 000 sandwiches. Both consumers and producers would gain through exchange. Consumers who wish to buy at £1.30 would gain more satisfaction from a sandwich than from £1.30. Producers who would be prepared to sell at £1.30 would prefer the £1.30 received from the sale of each sandwich to the sandwiches themselves.

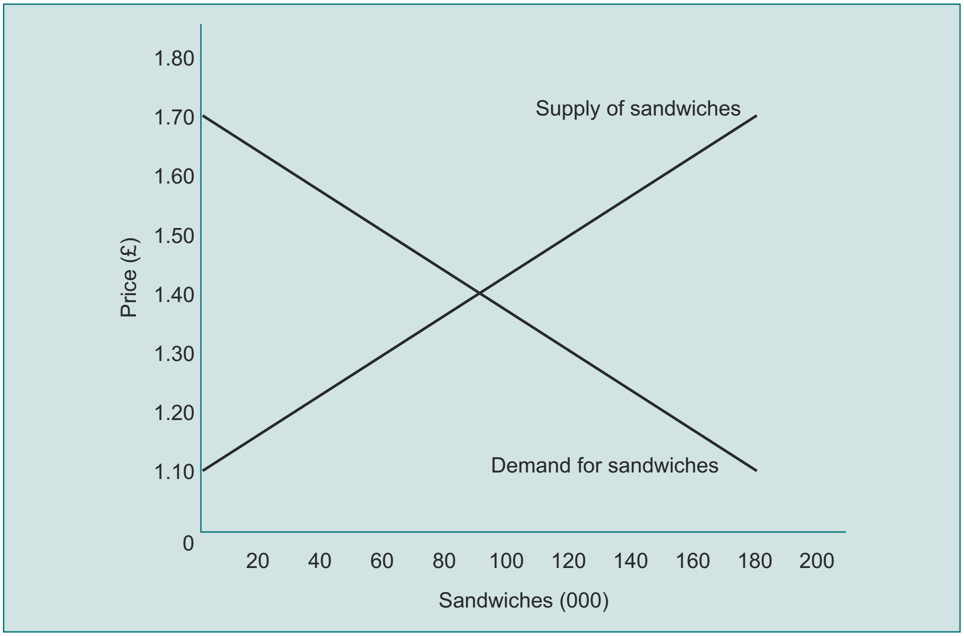


Figure 5.1 Market demand and supply curves for sandwiches in a British city

Even where a market demand schedule for, and a market supply schedule of, a good exist, exchange would be impossible if there were no common price at which at least one consumer was prepared to buy a unit of the good and a least one producer was willing to supply a unit of that good. Suppose the electronics industry developed a new micro-television set with a one-inch square screen to be worn on the wrist, and suppose that the market demand and supply schedules are as shown in Table 5.2.

Table 5.2 Market demand and supply schedules for micro-TV sets

Price ($)	Quantity demanded (per week)	Quantity supplied (per week)
500	200	0
1 000	100	0
1 500	0	0
2 000	0	0
2 500	0	100
3 000	0	200
3 500	0	300

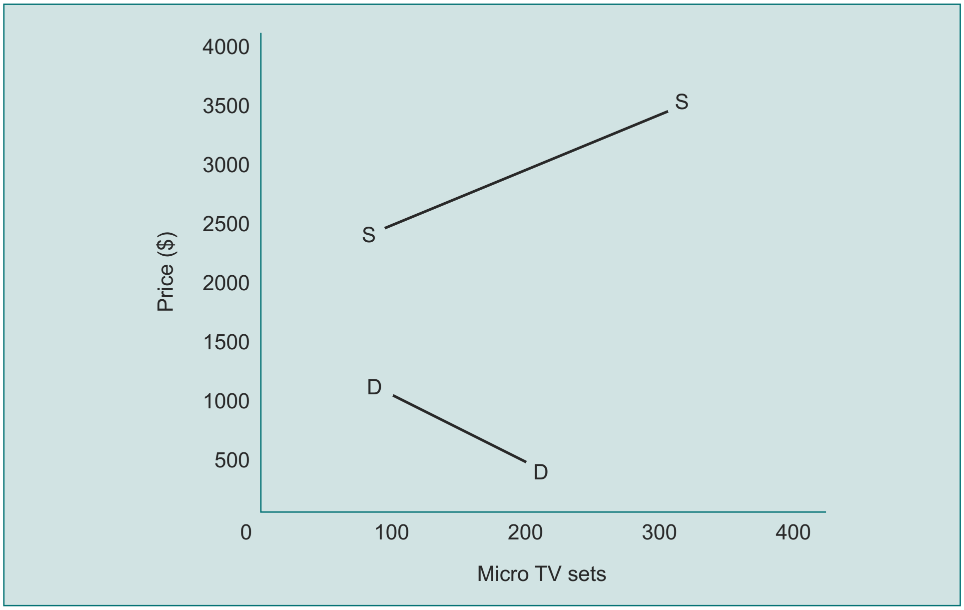


Figure 5.2 Market demand and supply curves for micro-TV sets

Exchange would not be possible in this market because there is no common price at which both consumers and producers would be prepared to buy and sell the micro-TV sets. Consumers would be willing to buy certain quantities of these sets at prices of $1000 and below; producers would only be prepared to supply the TV sets at prices of $2500 and above. As shown in Figure 5.2 (a graphical representation of the Table 5.2 data), there is no price at which exchange would occur.

If at all positive prices in a market the quantity of a good supplied exceeds the quantity demanded, the price of the good will be zero. When such market circumstances occur, the good is referred to as a *free good*. Goods such as fresh air are often classified as free goods when consumers can have as much of them as they want without having to pay for them. As shown in Figure 5.3, the market for air in the Scottish Highlands may be represented by the supply curve SS, and the demand curve by DD. The supply curve eventually slopes upwards because at various positive prices, individuals would be prepared to supply different quantities of air, e.g., in the form of air-conditioning.

The quantity of air available at a price of zero is $0Q_2$. The largest quantity of air people desire is $0Q_1$ at a price of zero. Thus at a price of zero the quantity supplied exceeds the quantity demanded.

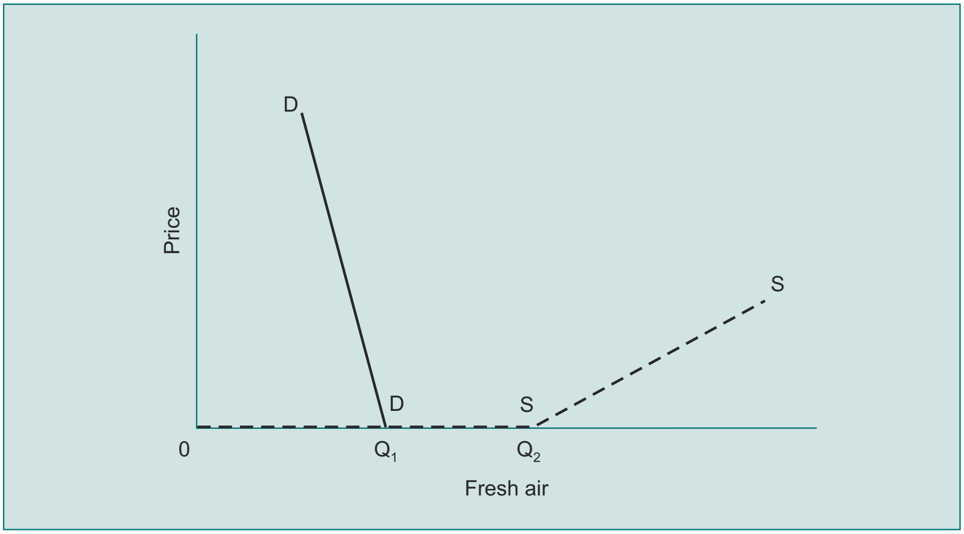


Figure 5.3 Demand for and supply of air in the Scottish Highlands

Given these demand and supply schedules, a zero price for air would exist in the market because, at that price, people can have as much air as they want. Goods such as air and water that are traditionally regarded as free may, however, have particular markets in which exchange at a positive price can take place. In Los Angeles, for example, fresh air is not always a free good but it can be made available through air-conditioning. The market for air-conditioning would be represented by a diagram similar to the sandwich market of Figure 5.1. As shown in Figure 5.4, consumers would be willing to buy different quantities of air-conditioning systems at different prices and producers would be prepared to supply different quantities of air-conditioning systems at different prices. In the air-conditioning market there would be a range of prices at which exchange could take place.

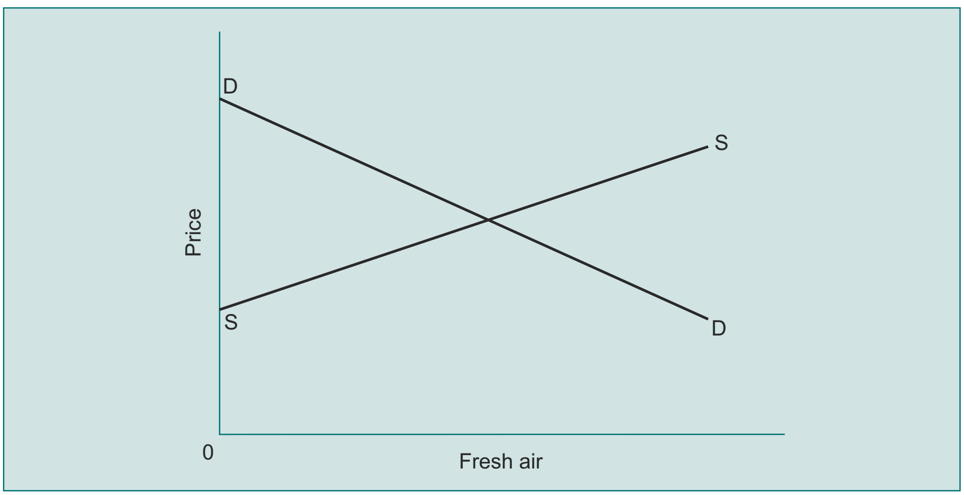


Figure 5.4 Market demand and supply curves for air-conditioning

Excess supply exists in a market when, at a given price, the quantity of a good firms are prepared to supply exceeds the quantity consumers are prepared to buy. Figure 5.5 illustrates a situation of excess supply in the market for sandwiches.

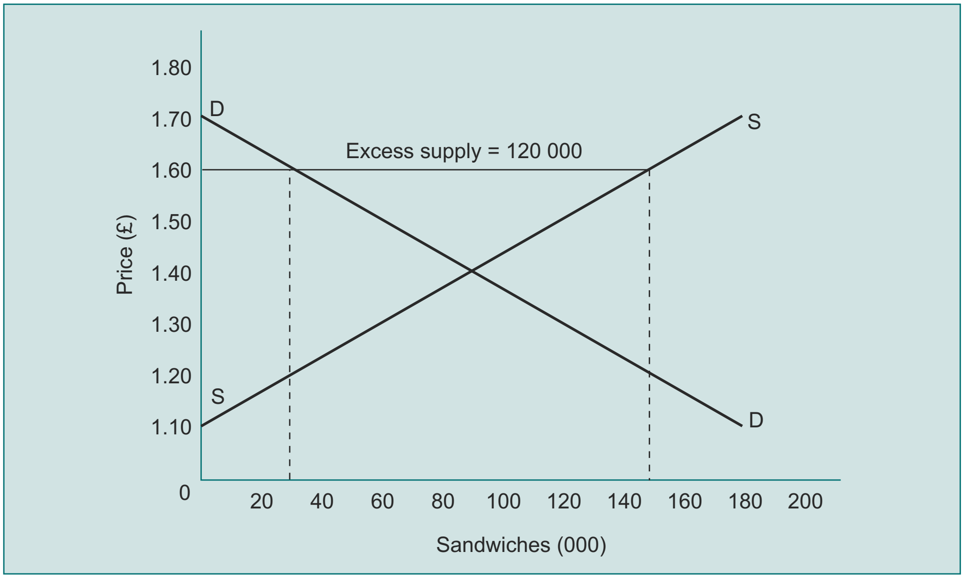


Figure 5.5 Excess supply in a market

At a price of £1.60 the sandwich suppliers would be willing to sell 150 000 sandwiches per week whereas consumers would be willing to buy only 30 000 sandwiches per week. Therefore, at a price of £1.60 excess supply of 120 000 sandwiches per week would exist.

Excess demand exists in a market when, at a given price, the quantity of a good consumers are prepared to buy exceeds the quantity that firms are prepared to supply. Figure 5.6 illustrates a situation of excess demand in the market for sandwiches.

At a price of £1.20 consumers would wish to buy 150 000 sandwiches per week but suppliers would be willing to sell only 30 000 sandwiches per week at that price. Hence excess demand of 120 000 sandwiches per week would exist if the price were £1.20.

At the unique price in a market at which the quantity of a good that consumers are willing to buy equals the quantity of the good that firms are willing to sell, neither excess demand nor excess supply exists. All individuals who are willing to buy as much of the good as they want at that price are able to do so, and all firms who are willing to sell as much of the good as they wish at that price are able to do so. Figure 5.7 illustrates the price in the market for sandwiches at which neither excess demand nor excess supply exists, for at a price of £1.40 consumers would be prepared to buy 90 000 sandwiches per week and suppliers would be willing to sell exactly the same quantity. Every individual sandwich producer who is willing to supply sandwiches at £1.40 each would be able to sell them to consumers, and every individual wishing to buy sandwiches at £1.40 each would be able to do so.

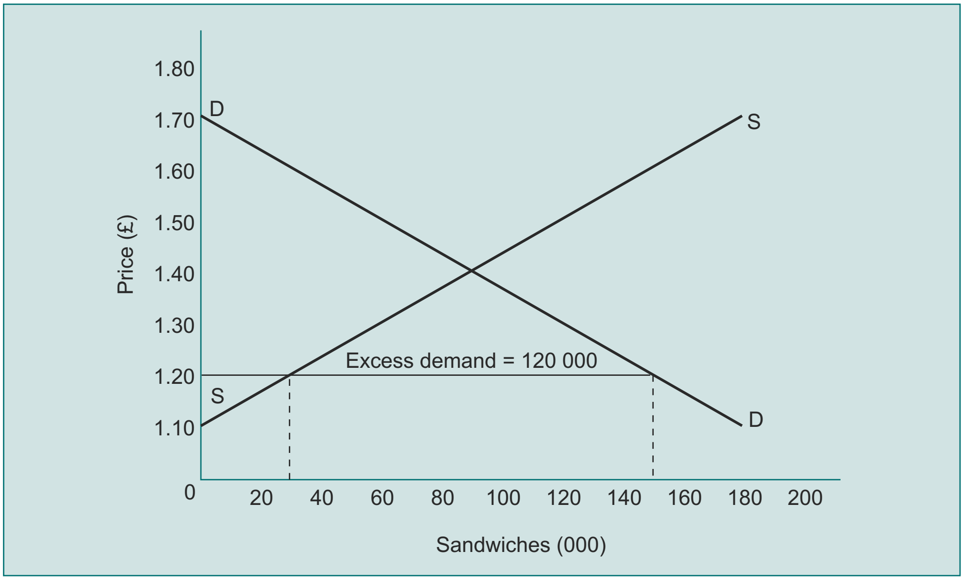


Figure 5.6 **Excess demand in a market**

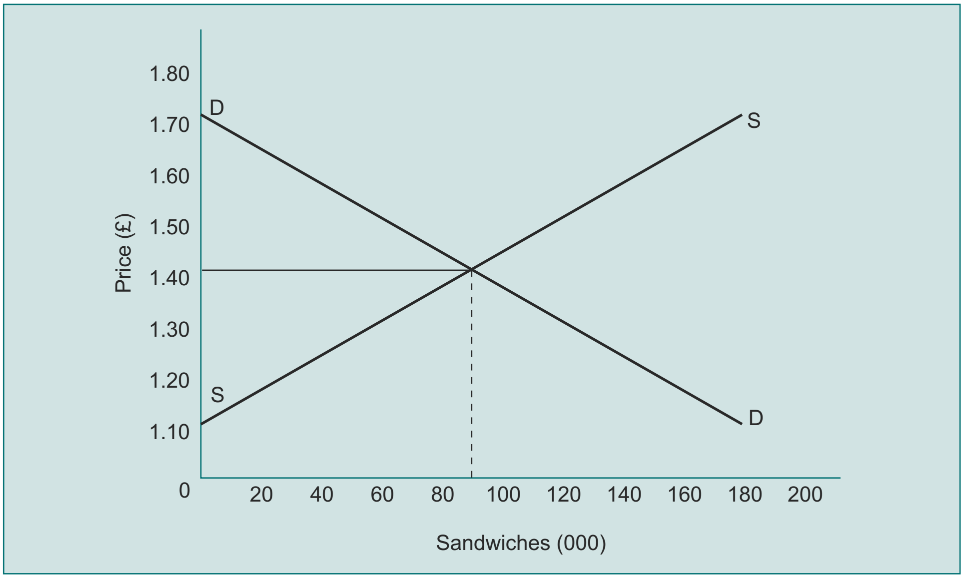


Figure 5.7 **Sandwich market with neither excess demand nor excess supply**

5.3 The Operation of Markets

If a price causes excess demand or excess supply to exist in a market, forces in the market will change the price of the good and the quantity bought and sold. These forces will eventually eliminate any excess demand or excess supply. If, at a given price, excess supply were to exist in a market, competition among suppliers would force down the price of the good. Simultaneously, at lower prices a larger quantity of the good will be demanded. Thus at lower prices the excess supply will be reduced. Eventually, a price would be reached at which excess supply will be zero.

Figure 5.8 again illustrates the market for sandwiches in a British city. It is possible for at least one sandwich to be offered for sale and to be demanded at any price above £1.10 and below £1.70. This is because at least one supplier is prepared to sell a sandwich for at least £1.10 while at least one buyer is prepared to pay up to £1.70 for a sandwich. If one buyer and one seller agreed to terms of exchange within this price range, and if the terms of exchange were known to all other potential buyers and sellers in the market, they would act on the information provided.

If one sandwich were sold to a consumer for £1.50 and as a result all suppliers attempted to sell sandwiches for £1.50 each, some suppliers would discover that they could not find customers at that price. This is because excess supply would exist at £1.50 since 60 000 more sandwiches would be offered for sale than consumers would wish to buy. Some suppliers who could not find individuals wishing to buy sandwiches at £1.50 would be willing to lower their asking price below £1.50, because the supply curve indicates that some suppliers would be prepared to sell sandwiches at prices less than £1.50. But this act of lowering the price would provide information to all other buyers and sellers in the market. Buyers who are prepared to pay £1.50 or more for a sandwich will seek those suppliers offering sandwiches for sale at less than £1.50. Additionally, at lower prices some consumers will buy more sandwiches. Others who would not buy at the higher prices will now be prepared to purchase sandwiches.

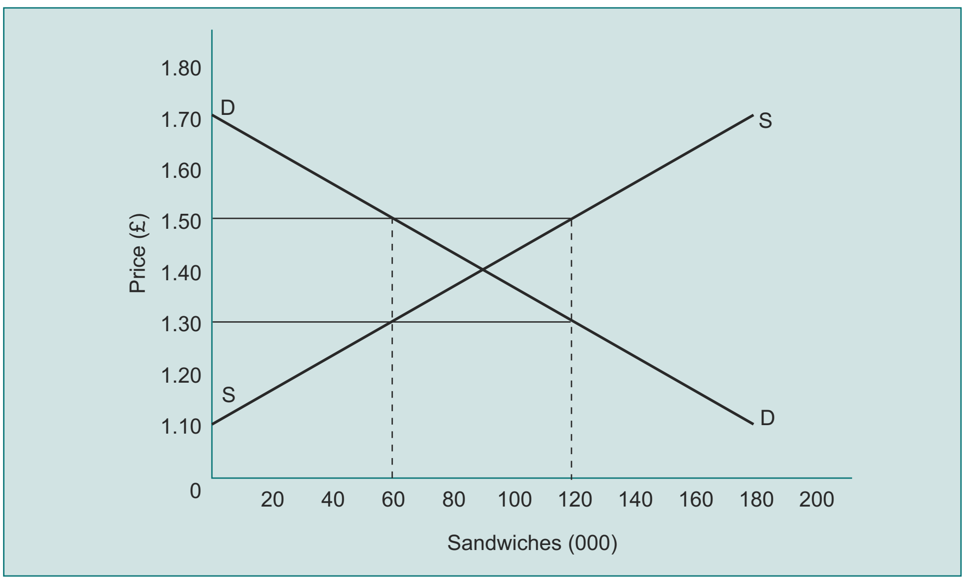


Figure 5.8 **The market for sandwiches in a British city**

At £1.40 each, the quantity consumers would be prepared to buy is 90 000 per week, and this exactly equals the quantity that firms would be prepared to supply. The excess supply that originally existed at £1.50 would therefore be eliminated.

If, at a given price, excess demand were to exist in a market, competition between consumers would force up the price of a good. Simultaneously, at higher prices, suppliers would be willing to sell a greater quantity of the good. Thus at higher prices excess demand would be reduced. Eventually, a price would be reached at which excess demand will be eliminated. In our example in Figure 5.8, if the original exchange between one buyer and seller in the market for sandwiches were agreed at £1.30, and as a result all sellers decided to ask for a price of £1.30 each for sandwiches, some potential buyers would find it impossible to obtain them. This is because excess demand would exist at a price of £1.30, since 60 000 more sandwiches would be desired by consumers than firms would be prepared to supply. Those individuals who could not obtain sandwiches would be willing to bid up the price of sandwiches because, as indicated by the demand curve, some consumers would be prepared to pay prices above £1.30 to obtain sandwiches. But those individuals who pay more than £1.30 each for sandwiches would provide information to all other buyers and sellers in the market. Suppliers who are prepared to sell sandwiches for £1.30 or less will seek those consumers who are prepared to pay more than £1.30 to obtain sandwiches. Additionally, at higher prices some suppliers will supply more sandwiches. Other suppliers who would not supply sandwiches at lower prices will now be prepared to supply sandwiches.

At £1.40 each, the quantity consumers in total would be prepared to buy would exactly equal the quantity firms would be willing to supply, i.e. 90 000 per week. The excess demand which originally existed at £1.30 would therefore be eliminated.

When the forces that act to eliminate excess demand or excess supply exactly offset each other, a market is defined as being *in equilibrium*. The price at which equilibrium occurs in a market is the price that allows all individuals who wish to buy a good to obtain as much as they wish at that price, and enables all producers who wish to sell the good to sell as much as they wish at that price. The price at which the intentions of all potential buyers and sellers are exactly matched is called the *equilibrium price*, and the identical quantity of the good which the two sides wish to buy and sell at that price is called the *equilibrium quantity*.

In Figure 5.9 the equilibrium price of sandwiches in this market is £1.40, since all buyers who wish to purchase sandwiches at £1.40 each are able to do so and all firms who wish to sell at £1.40 each are able to find buyers. Diagrammatically, the equilibrium price is established where the demand and supply curves intersect, i.e. at point E. At the price of £1.40 the market is just cleared since both excess demand and excess supply are zero. The quantity bought and sold at £1.40 is 90 000 per week and there are no unsatisfied buyers or sellers. In this market, 90 000 sandwiches per week is the equilibrium quantity.

When exchange takes place in a market, gains accrue to both consumers and producers. If such gains did not exist, there would be no incentive for exchange to take place. A demand curve relates the quantities of a good individuals would buy at different prices, and it also relates the maximum prices that individuals would be willing to pay for different quantities of a good. However, if individuals can obtain units of a good at a price that is less than they would be prepared to pay, they would derive an additional gain from exchange. This additional gain is called *consumer surplus* and can be measured by the difference between the price individuals actually pay for units of a good and the higher prices they would be prepared to pay.

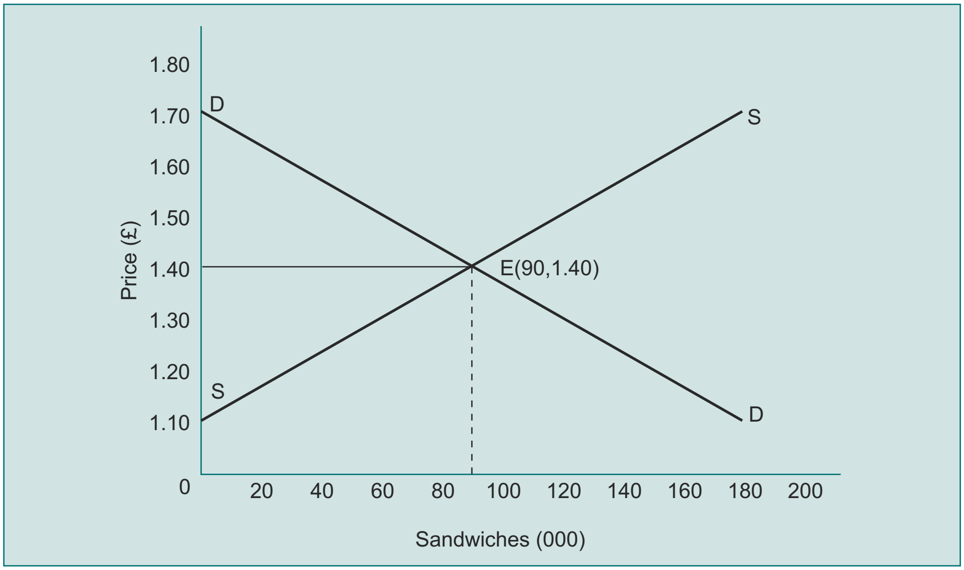


Figure 5.9 Equilibrium price and quantity in a market

Figure 5.10 again illustrates the market for sandwiches in a British city, and highlights the foregoing point. The equilibrium price of sandwiches is £1.40 and the equilibrium quantity is 90 000 per week. This means that a price of £1.40 is paid by every individual for every sandwich bought. But according to the downward sloping demand curve, some consumers would have been prepared to pay more for sandwiches.

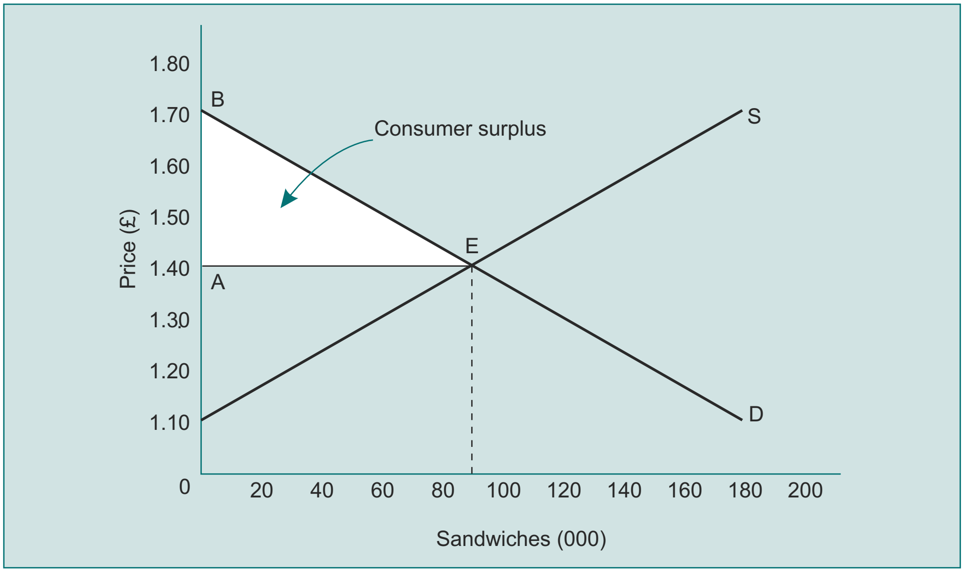


Figure 5.10 Measurement of consumer surplus

For example, one individual would be prepared to pay up to £1.70 for the first sandwich, but he has to pay only £1.40. He has a consumer surplus of £0.30. Similarly, for successive quantities up to 90 000 per week there is a difference between the price of £1.40 that consumers pay for sandwiches and the price they would be willing to pay. This difference, as described above, is called consumer surplus, and is illustrated graphically for each quantity of sandwiches by the vertical distance between the line AE (the price actually paid, i.e. £1.40) and the demand curve (the prices individuals would be willing to pay for different quantities). Hence for successive quantities of sandwiches up to 90 000 per week, consumer surplus is denoted by a series of vertical lines between AE and the demand curve. At the equilibrium price of £1.40 consumer surplus is measured by the unshaded triangle BAE.

A supply curve relates the various quantities of a good that suppliers would be prepared to sell at different prices, and it also relates the minimum prices at which suppliers would be prepared to sell different quantities of a good. If suppliers can sell units of a good at a price that is higher than they would be prepared to accept, they derive an additional gain from exchange. This additional gain is called *producer surplus*, and can be measured by the difference between the price suppliers actually receive for selling units of a good and the lower prices they would be prepared to accept.

Figure 5.11 illustrates how producer surplus is measured. Assuming that the market is in equilibrium, a price of £1.40 for sandwiches is received by every supplier for each of the 90 000 sandwiches sold. But the upward sloping supply curve indicates that suppliers would be willing to sell sandwiches for prices less than £1.40.

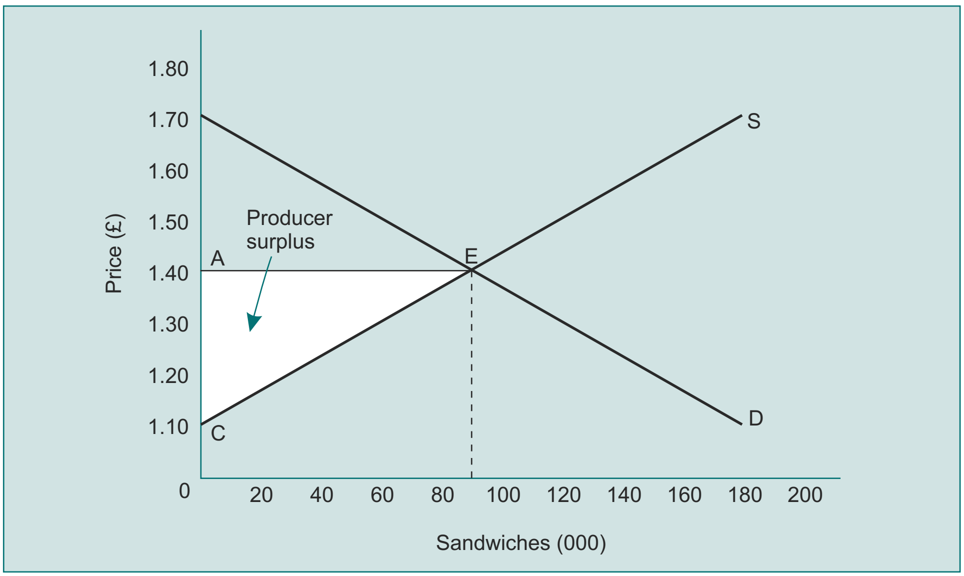


Figure 5.11 Measurement of producer surplus

For example, one supplier would be prepared to sell a sandwich for just over £1.10, but in fact receives £1.40. He has a producer surplus of £0.30. Similarly, for successive quantities of sandwiches up to 90 000 there is a difference between the price suppliers actually receive and the prices they would be prepared to accept. This difference is called producer surplus and is illustrated graphically for each quantity of sandwiches by the vertical distance between the line AE (the price received by sellers for each sandwich, i.e. £1.40) and the supply curve (the prices suppliers would be prepared to accept for different quantities). Hence for successive quantities of sandwiches up to 90 000, producer surplus is denoted by a series of vertical lines between AE and the supply curve. At the equilibrium price of £1.40, producer surplus is measured by the unshaded triangle CAE.

An example of how both consumer and producer surpluses provide the motivating forces in a market to achieve equilibrium is provided by Myron Joseph's role-playing wheat game. This educational experiment is based on a hypothetical wheat market in which a class of economics students adopts the role of dealers, half of whom buy and half of whom sell wheat. In the experiment, a card is given by the teacher to each student, requiring him or her to buy or sell 1000 units of wheat for not more or less than a certain price, as shown in Table 5.3.

Table 5.3 **Distribution of instruction cards**

Price ($)	Number of buyers (not more than the price)	Number of sellers (not less than the price)
2.80	4	–
2.60	4	2
2.40	4	2
2.20	4	2
2.00	4	2
1.80	4	2
1.60	2	4
1.40	2	4
1.20	2	4
1.00	2	4

Students are told that when the fictional market opens, they should adopt the role of dealers for clients and attempt to do the best they can in a transaction. Once the game commences, students are free to contact each other to discuss terms of exchange. Each time an exchange is agreed, the teacher makes the price known to all students, and those students who have completed a deal may then pick up fresh cards and start again. After a certain time, the market is closed and the whole class can then see the range of prices at which exchange actually took place and the number of transactions that were recorded at each price. Table 5.4 shows the results from one experiment conducted by Joseph.

Table 5.4 **Distribution of transactions**

Price ($)	Number of transactions
2.60	1
2.50	0
2.40	1
2.30	0
2.20	6
2.10	1
2.00	15
1.90	25
1.80	27
1.70	20
1.60	16
1.50	13

Price ($)	Number of transactions
1.40	5
1.30	1
1.20	3
1.10	0
1.00	1

According to Joseph, the initial exchanges were agreed at a wide range of prices. But then there was a tendency for prices to move towards a theoretical equilibrium level of $1.80. By the end of the game virtually every exchange was agreed at or close to $1.80.

This game illustrates the role of *information* in markets, since the disparity in prices soon disappeared once the terms of each exchange became known to all buyers and sellers. Any exchange at prices well above $1.80 (which would have enlarged the area of producer surplus if all deals were made at that price) resulted in fervent competition among sellers, which in turn drove prices down. Conversely, any exchanges arranged at prices well below $1.80 (which would have extended the area of consumer surplus if all exchanges had taken place at that price) resulted in keen competition among buyers, which in turn drove prices up. Eventually, equilibrium would be reached.

5.4 Changes in Market Equilibrium

If a market were in equilibrium and if any of the conditions determining demand or supply were to change, forces would establish a different equilibrium price and/or a different equilibrium quantity. Assuming that the supply conditions of a good remain unchanged, an increase in demand would lead to an increase in equilibrium price and an increase in equilibrium quantity, while a decrease in demand would lead to a decrease in equilibrium price and a decrease in equilibrium quantity.

Figure 5.12 illustrates again the market for sandwiches in a British city, in which the original equilibrium price and quantity are £1.40 and 90 000 sandwiches per week. If the price of all restaurant meals were to rise, and if this led to consumers substituting sandwiches for such meals, this would be defined as an increase in demand for sandwiches and would be represented by a shift of the demand curve from DD to D'D'. Note that there is a difference between an increase in demand and an increase in the quantity demanded. The former refers to a *shift in the position of the whole curve* caused by a change in a factor other than the price of the good in question. The latter refers to an increase in the quantity that would be demanded if the price were to be lower, other factors remaining unchanged, and is represented by a *movement along a demand curve*.

Assuming no change in supply conditions, at the original equilibrium price of £1.40 there would be excess demand for sandwiches, given the demand curve in position D'D'. Suppliers would still wish to sell 90 000 sandwiches per week but consumers would be prepared to buy 150 000 sandwiches per week at a price of £1.40.

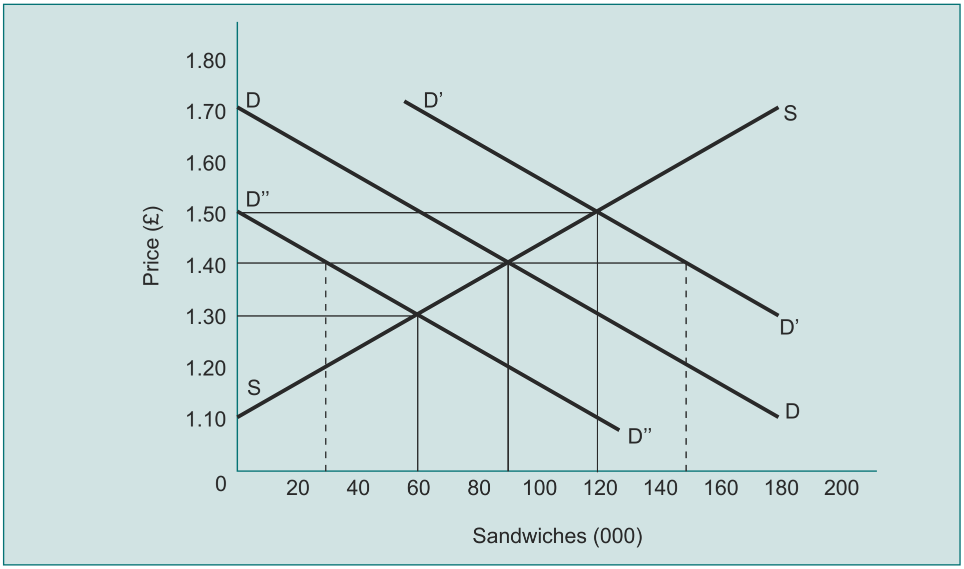


Figure 5.12 A change in demand with supply conditions constant

Upward pressure would be exerted on price until a new equilibrium point were established at £1.50, where D'D' intersects the supply curve. The new equilibrium price of £1.50 and the new equilibrium quantity of 120 000 sandwiches per week would both be higher than the original equilibrium price and equilibrium quantity.

Conversely, if a health risk were to arise in connection with eating sandwiches, and if consumers wished to substitute other products for sandwiches and buy fewer sandwiches at every price, the decrease in demand for sandwiches would be represented by a shift of the demand curve from DD to D"D". At the original equilibrium price of £1.40 there would be an excess supply of 60 000 sandwiches per week, assuming supply conditions remained constant. Consumers would now wish to buy only 30 000 sandwiches per week whereas suppliers would be prepared to sell 90 000 per week.

Downward pressure would be exerted on price until a new equilibrium point were established at £1.30, where D"D" intersects the supply curve. The new equilibrium price of £1.30 and the new equilibrium quantity of 60 000 sandwiches per week would both be lower than the original equilibrium price and equilibrium quantity.

In each of the above examples, the new equilibrium price and equilibrium quantity would be determined as a result of a shift in the demand curve and a movement along an unchanged supply curve. In similar fashion, with the demand conditions for a good remaining unchanged, an increase in supply would lead to a fall in equilibrium price and an increase in equilibrium quantity, while a decrease in supply would lead to a rise in equilibrium price and a fall in equilibrium quantity.

Suppose that in the market for sandwiches, suppliers were able to use cheaper imports of cooked meats to produce sandwiches and as a result they would be willing to sell more sandwiches at every price. This increase in supply is represented by a shift of the supply curve from SS to S'S', as shown in Figure 5.13. Note that there is a difference between an increase in supply and an increase in the quantity supplied. The former refers to a *shift in the whole position of the supply curve*, caused by a change in a factor other than the price of the good in question. The latter refers to an increase in the quantity which would be supplied if the price were to be higher, other factors remaining unchanged, and is represented by a *movement along a supply curve*.

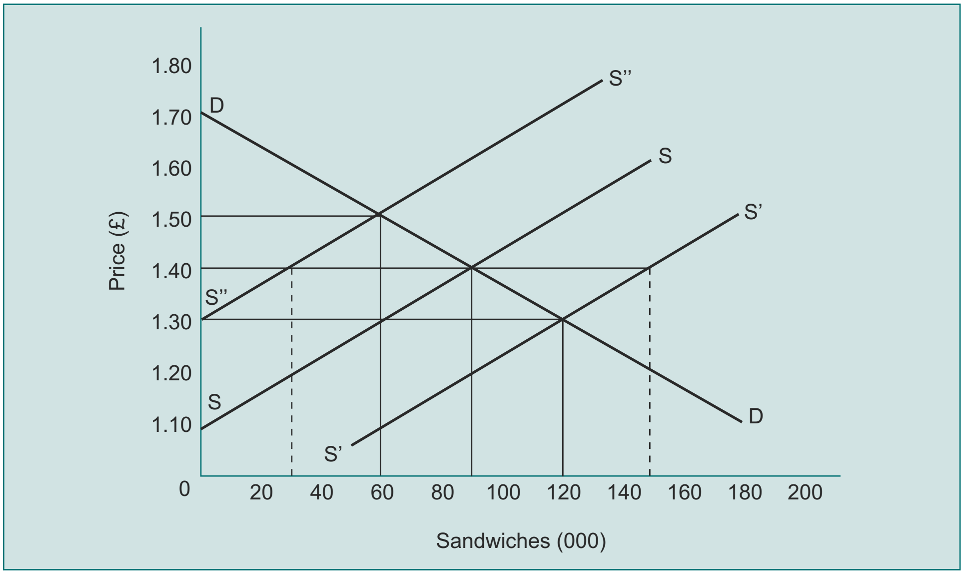


Figure 5.13 A change in supply with demand conditions constant

Assuming no change in the demand for sandwiches, there would be excess supply at the original equilibrium price of £1.40 since consumers would still wish to buy 90 000 sandwiches per week whereas suppliers would be prepared to sell 150 000 sandwiches per week at £1.40 each. This would generate downward pressure on price until a new equilibrium price of £1.30 were reached, where the supply curve S'S' intersects the demand curve. As a result of the increase in supply, the new equilibrium price would be lower than the original equilibrium price, and the new equilibrium quantity of 120 000 sandwiches per week would be greater than the original equilibrium quantity of 90 000 per week.

Conversely, if sandwich suppliers were to experience a rise in production costs following an increase in wages paid to all workers employed in producing sandwiches, and were prepared to supply fewer sandwiches at every price, this would be represented by a shift of the supply curve from SS to S"S". Assuming no change in the demand curve, there would be excess demand for sandwiches at the original equilibrium price of £1.40 since consumers would still wish to buy 90 000 per week whereas suppliers would be prepared to sell only 30 000 per week. This would generate upward pressure on price until a new equilibrium position were reached. In this market, it would occur at a price of £1.50 and a quantity of 60 000 sandwiches, where the supply curve S"S" intersects the demand curve.

As a result of the decrease in supply, the new equilibrium price would be higher than the original equilibrium price and the new equilibrium quantity of 60 000 sandwiches per week would be lower than the original equilibrium quantity of 90 000 per week.

In each of the above examples, the new equilibrium price and equilibrium quantity would be determined as a result of a shift in the supply curve and a movement along an unchanged demand curve. If there were an increase in demand for a good and simultaneously an increase in supply of the good, the equilibrium quantity would rise. The equilibrium price of the good, however, could rise, fall or remain the same, depending on the extent to which demand increased relative to supply.

In the market for sandwiches represented in Figure 5.14, suppose the demand for sandwiches were to increase as a result of a rise in the prices of sandwich substitutes (hamburgers, beans on toast, etc.) and the supply of sandwiches were also to increase as a result of a fall in production costs. The increase in demand is shown by the shift in the demand curve from DD to D'D' and the increase in supply is indicated by the shift of the supply curve from SS to S'S'. As a result of these simultaneous changes, the equilibrium price of sandwiches would fall from £1.40 to £1.30, as indicated by the intersection of D'D' and S'S'. The new equilibrium quantity would increase from 90 000 per week to 160 000 per week.

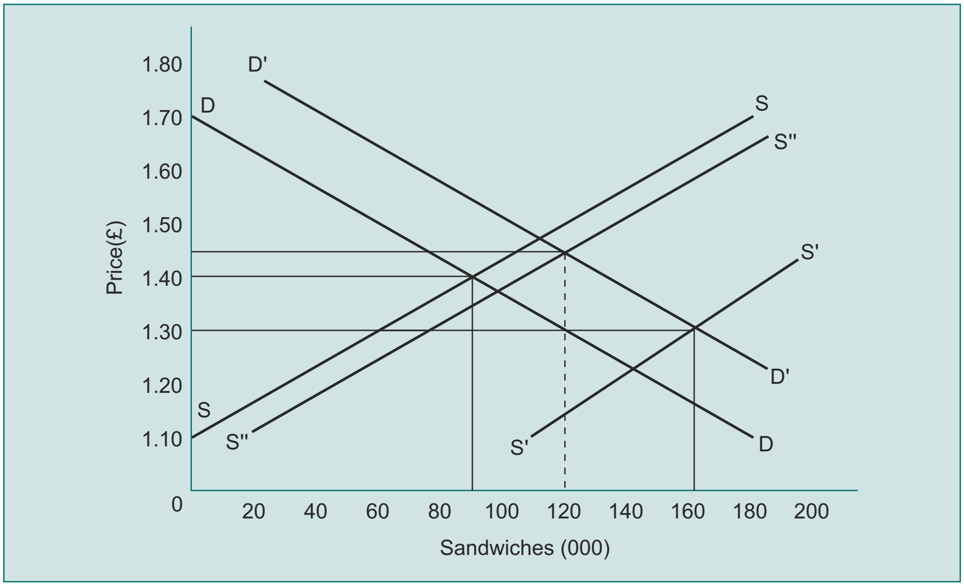


Figure 5.14 An increase in demand and an increase in supply

While the equilibrium quantity will always increase as a result of these increases in supply and demand, equilibrium price might rise or even remain at the original level instead of decreasing. If the supply curve were to shift only slightly from SS to S"S", the new equilibrium price and quantity would be £1.45 and 120 000 per week respectively, so that equilibrium quantity would increase and equilibrium price would increase also. In addition, it would be possible for both curves to shift just sufficiently to maintain equilibrium price at £1.40.

If the demand for a good were to increase and the supply of the good were to decrease simultaneously, the equilibrium price of the good would rise. The equilibrium quantity of the good might increase, decrease or remain the same, depending on the extent of the relative changes in demand and supply.

In the market for sandwiches, suppose that consumers wish to buy more at every price as a result of an increase in the prices of all substitutes and that, simultaneously, sandwich suppliers were to experience a rise in production costs that they reflected in their selling prices. The impact of these changes in supply and demand is illustrated in Figure 5.15.

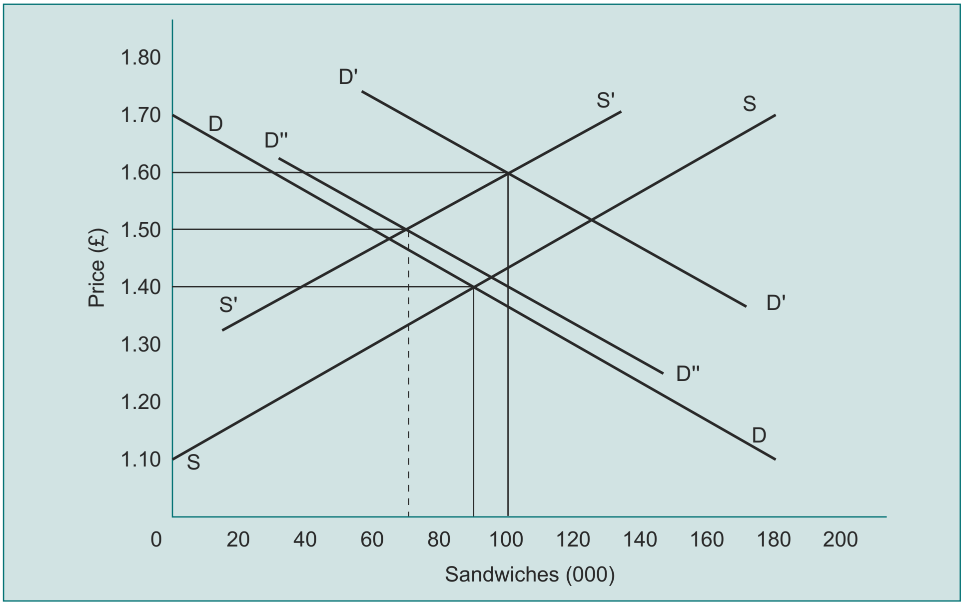


Figure 5.15 An increase in demand and a decrease in supply

The demand curve would shift from DD to D'D' and the supply curve would shift from SS to S'S'. As a result, the new equilibrium price and quantity are established where D'D' and S'S' intersect. The new equilibrium price would be £1.60 and the new equilibrium quantity would be 100 000 per week, both of which are higher than the original equilibrium levels.

While equilibrium price would always rise as a result of these two simultaneous changes in supply and demand, equilibrium quantity might fall, or remain at the same level, instead of increasing. If the demand curve were to shift only slightly from DD to D"D", the new equilibrium price and quantity would be £1.50 and 70 000 sandwiches per week. So although equilibrium price would still increase, equilibrium quantity would fall from its original level. Also, it would be possible for both curves to shift just sufficiently to maintain the equilibrium quantity at 90 000 per week.

If the supply of a good were to increase and the demand for the good were to decrease simultaneously, the equilibrium price of the good would fall but the equilibrium quantity might increase, decrease or remain at the same level.

Suppose that suppliers discovered a more efficient method of producing sandwiches and that, simultaneously, consumers learned that there might be a health risk involved in eating sandwiches. The effect of these changes in the market for sandwiches is illustrated in Figure 5.16. The increase in supply is denoted by the shift in the supply curve from SS to S'S' and the decrease in demand by a shift in the demand curve from DD to D'D'. The point at which D'D' and S'S' intersect denotes the new position of equilibrium in the market, i.e. a price of £1.10 and a quantity of 110 000 sandwiches per week bought and sold.

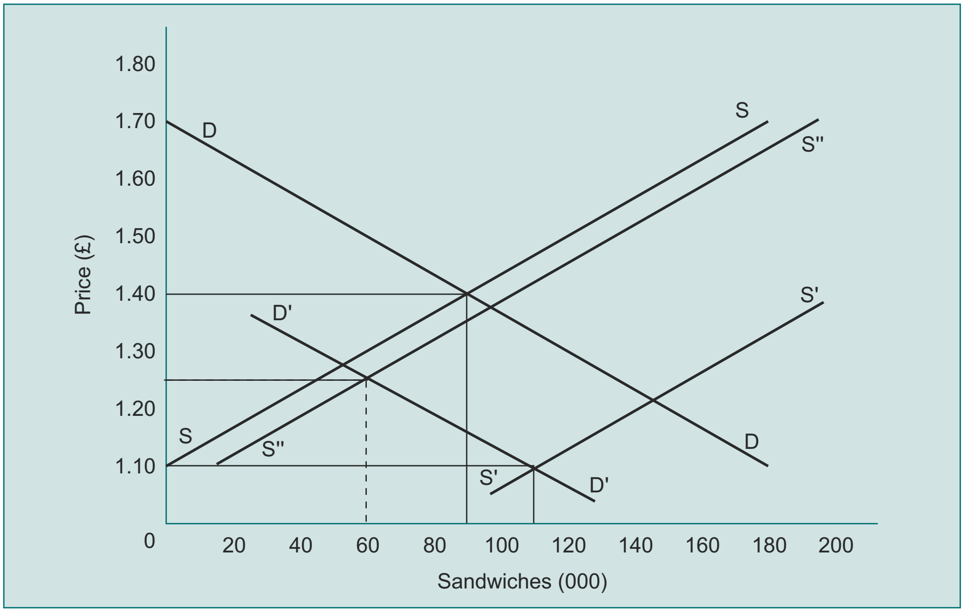


Figure 5.16 An increase in supply and a decrease in demand

Although equilibrium price of the good would always fall as a result of these two changes in supply and demand, equilibrium quantity might fall or remain unchanged instead of increasing. If the supply curve were to shift from SS to S''S'' instead of to S'S', the new equilibrium price and quantity would decrease also. It would be possible for both curves to shift just sufficiently to maintain the equilibrium quantity at 90 000 per week.

If the demand for a good and the supply of a good were both to decrease, the equilibrium quantity of the good would fall, but equilibrium price might increase, decrease or remain at its original level.

Suppose that some consumers substituted sausage rolls for sandwiches and, simultaneously, sandwich producers experienced a rise in costs. The impact of these two changes is shown in Figure 5.17. The decrease in demand is represented by a shift in the demand curve from DD to D'D' and the decrease in supply by a shift from SS to S'S'. As a result, the equilibrium price would fall from £1.40 to £1.35 and the equilibrium quantity would fall from 90 000 sandwiches per week to 40 000 per week.

Although equilibrium quantity would always fall as a result of a simultaneous decrease in supply and demand, equilibrium price might increase or remain unchanged instead of decreasing. If the demand curve shifted only slightly from DD to D''D'' instead of to D'D', the new equilibrium price would be £1.45, which is higher than the original equilibrium price. And it would be possible for both curves to shift just enough to maintain the equilibrium price at £1.40.

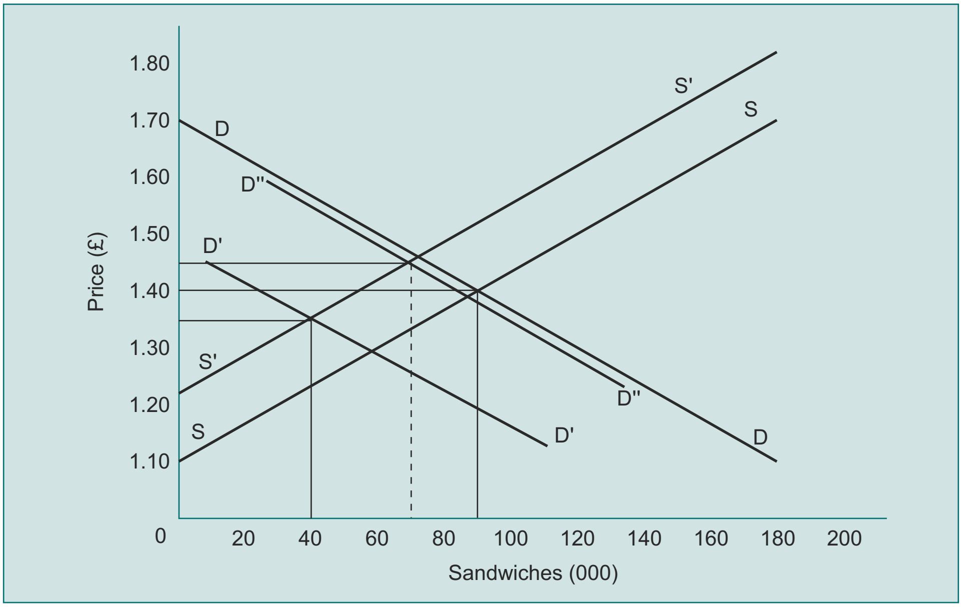


Figure 5.17 A decrease in demand and a decrease in supply

5.5 Intervention in the Market

Market intervention or market regulation occurs when a non-market force causes the price and quantity of a good bought and sold in a competitive market to be different from the price/quantity combination that would occur if the market were allowed to operate freely.

Tickets for major sporting events such as the Wimbledon tennis tournament or the American Superbowl are priced below the level that would occur if the forces of supply and demand were allowed to operate freely. The result is that more people wish to buy tickets at the prevailing price than there are tickets available. Conversely, minimum-wage laws in some countries, which stipulate that employees must be paid at least a stipulated minimum per hour, provide an example of the price in a resource market that is set above the equilibrium level, where more people are willing to work at the stipulated minimum than there are jobs available.

5.5.1 Price Ceilings

When the price of a good is fixed below the equilibrium level, a *price ceiling* is said to exist in the market for the good. Price alone will be an inadequate mechanism for allocating the available supply of the good among potential buyers, and some other allocative mechanism (such as first-come-first-served, reliance on suppliers' preferences, or rationing) may be employed.

When world-famous pop groups stage a concert, the price of the tickets is often below the equilibrium price that would be established if tickets were sold without restriction in a competitive market. Suppose Figure 5.18 illustrates the market for a pop concert in a concert hall where the capacity is 2000 seats. Supply is therefore represented by the completely inelastic supply curve SS. If the demand curve for tickets were represented by DD, the equilibrium price would be $8 per ticket. But if the organisers of the concert decided to fix the price of tickets at $4 each, excess demand for tickets would exist. In fact, at a price of $4 twice as many people would wish to see the concert as there are seats available, since at $4 each 4000 tickets would be demanded.

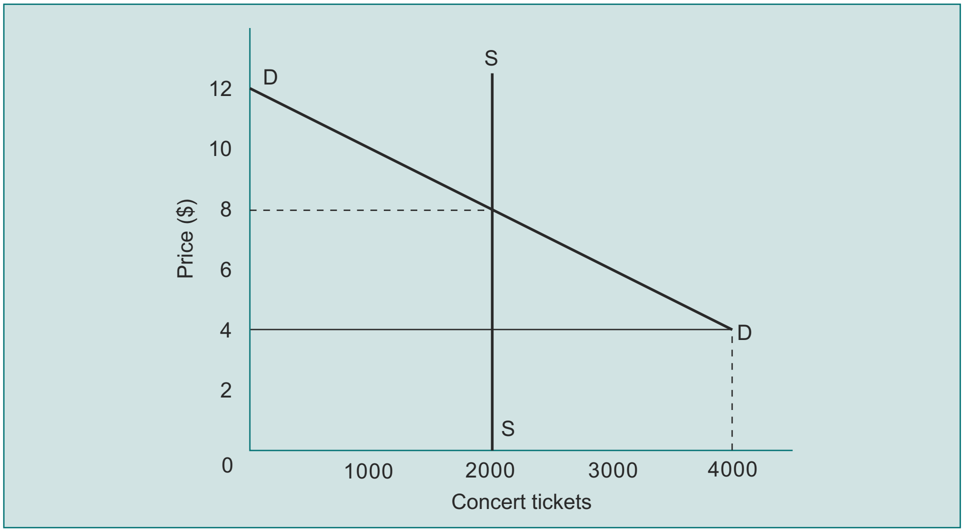


Figure 5.18 A price ceiling in a pop-concert market

Some way would have to be found of allocating the available 2000 tickets among potential buyers. One obvious method would be to allocate tickets to the first 2000 buyers who appeared; this is why large queues are often seen when tickets for major events go on sale. Alternatively, the concert organisers might allocate some of the 2000 tickets according to their own preferences, e.g. to personal friends, disc jockeys and other show-business personalities. As a third option, a system of rationing might be used whereby each individual would be allowed to purchase only a limited number of tickets.

When price ceilings are in operation and excess demand exists, a black market, in which units of a good are exchanged for prices higher than the fixed level, tends to develop. Black markets, despite being illegal in many countries, serve the function of establishing equilibrium prices and quantities in a market. If the black market is 'perfect', the higher price will be the equilibrium price, where all buyers who want a ticket at that price can obtain one and any ticket owner who wishes to sell a ticket at that price can do so. The motivation for exchange at black-market prices exists because, as shown by the downward-sloping demand curve, some individuals are prepared to pay much higher prices than others to obtain units of a good, and there is no guarantee that those individuals who are willing to pay higher prices actually receive any units of the good by the initial allocation.

In the hypothetical market for pop concert tickets shown in Figure 5.18, there is at least one individual who is prepared to pay up to $12 for a ticket, and at least one individual who would be willing to pay the fixed price of $4 but no more. Suppose the latter individual actually obtains a ticket for $4 by turning up early when the tickets go on sale, but the former individual arrives too late and discovers that the tickets have all been sold. Exchange would be possible between these two individuals because the person with the ticket would prefer any sum of money above $4 to the ticket, whereas the person without a ticket would be prepared to pay up to $12 to obtain a ticket. Thus exchange of a ticket for a price between

$4 and $12 can occur because both individuals would gain by the exchange. In this one-to-one situation, the price of the ticket is indeterminate – it will lie between $4 and $12, and will depend ultimately on the bargaining skills of the buyer and seller. However, if two similar buyers and one seller constituted the market, the price of the ticket would be $12 and, conversely, if one buyer and two similar sellers constituted the market, the price of the ticket would be $4.01.

5.5.2 Price Floors

When the price of a good is fixed above the equilibrium level, a *price floor* is said to exist in the market for the good. At the fixed price, there would be excess supply of the good and some method other than price would have to be found for disposing of the surplus.

For example, the prices of many agricultural goods in the European Union (EU) are fixed above the equilibrium level. Figure 5.19 illustrates the implications of such price floors in a wheat market. Suppose that DD and SS represent the demand and supply curves for wheat and that the EU Commission were to set a price floor at $2.50 per hectolitre, which is higher than the equilibrium price of $2.00 shown at point E. The quantity of wheat that farmers would be prepared to supply at $2.50 is 100 million hectolitres per year, whereas the quantity of wheat that consumers would be prepared to buy at the regulated price is 70 million hectolitres per year. Hence a situation of excess supply of 30 million hectolitres per year would exist with a $2.50 unit price floor.

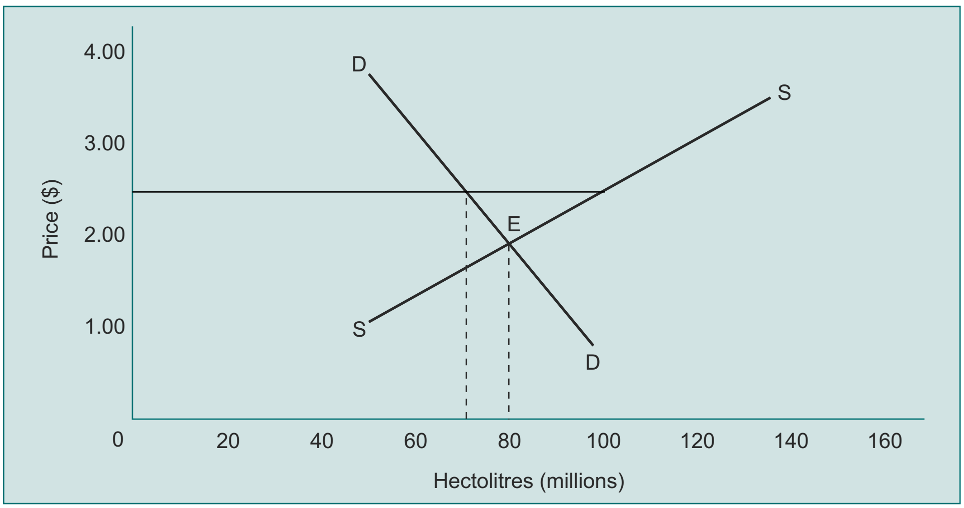


Figure 5.19 A price floor in a wheat market

In imposing a price floor, the regulating agency faces the problem of disposing of the excess supply. One solution would be to destroy the surplus. While this is obviously a waste of resources, some European countries have resorted to this course of action for many agricultural goods. Society as a whole suffers because fewer goods are available, but if the demand for these agricultural goods is relatively price-inelastic, the income of the farmers, i.e. revenue from the sale of these agricultural goods, will be higher, the higher the price. Indeed, such an argument is often used to justify these schemes. Alternatively, a government could buy the surplus, stockpile it and release it on to the market at a time when crop production was very low. Such a scheme could help stabilise price over time.

One further choice available to a government is to pay farmers not to produce the excess supply. In Figure 5.19 this would involve the government trying to restrict the quantity of wheat supplied to 70 million hectolitres per year. This policy has been tried by governments, and many farmers have been paid by their government not to produce!

One further example of price floors, this time in a resource market, is the law on minimum wages that operates in many countries. Suppose this law makes it illegal for firms to hire labour at less than $4.30 per hour. Figure 5.20 illustrates a hypothetical market for low-skilled labour. According to the market demand and supply curves, the equilibrium wage rate would be $4.00 per hour. At that rate, 200 people each working 40 hours per week would be employed, so that the quantity of labour offered and hired at $4.00 per hour would be equivalent to 8000 hours per week. But if a legal minimum wage of $4.30 were in effect, an excess supply of low-skilled labour would exist. Individuals would wish to work 8800 hours per week, i.e. 220 people would wish to be employed at a rate of $4.30 per hour. Firms, however, would wish to employ individuals to work only 7200 hours per week, i.e. they would wish to employ only 180 people at a rate of $4.30 per hour. There would be an excess supply of labour amounting to 40 people, comprising 20 extra people who would wish to work at a rate of $4.30 compared with $4.00, and 20 people originally employed at $4.00 but whom firms would no longer be willing to hire at $4.30. The 180 people who remain employed, however, would be better off than before.

Many economists have tried to discover the impact of minimum-wage legislation on employment, and studies have shown that certain sections of the labour force, such as teenagers and low-skilled workers, are most affected by minimum-wage laws. The greater the gap between the going wage rate for low-skilled groups and a newly imposed minimum-wage rate, the greater the increase in unemployment among those low-skilled groups.

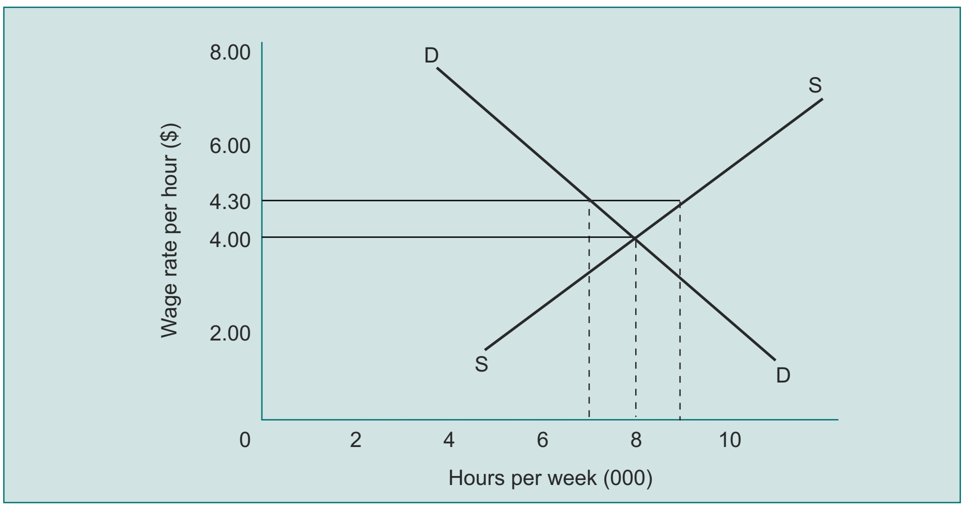


Figure 5.20 Effect of a minimum wage in a competitive labour market

Controversy rages about minimum-wage laws. Some people benefit; others suffer, but there is little disagreement that such laws cause unemployment among various groups of workers and, consequently, make total output lower than it otherwise would be.

Nowadays it is argued that the minimum-wage law in the USA does not seriously disrupt labour market operations because the minimum wage set by law is close to the equilibrium market wage for low-skilled labour. Were the legal minimum wage of France, which is substantially higher than that in the USA, to be imposed in the USA, it would result in an estimated 30 million US workers becoming unemployed.

In Britain, there is considerable doubt about whether women are actually paid less than men. Some jobs are male dominated; others female dominated. If women were paid less than men for doing the same job, why would firms ever hire men? In many industries and professions the differences between men's and women's pay can be explained by differences in the types of jobs performed, hours worked, relocating, travelling, family choices and motivation. In those markets where females were formerly paid lower wage rates than men for identical jobs, the economic effects of equal pay legislation are similar to those arising from minimum-wage legislation. Such legislation benefits those females who remain employed at the legislated wage rate, but makes worse off those females who become unemployed.

5.5.3 Taxes and Subsidies

Taxes and subsidies also affect market prices and quantities. The equilibrium price and equilibrium quantity of a good may be changed by the imposition of a tax on each unit of the good bought and sold.

Table 5.5 Effect of a tax on the supply schedule of a good

Price ($)	Quantity supplied per week	Price plus tax ($)
1	10 000	2
2	20 000	3
3	30 000	4
4	40 000	5
5	50 000	6

Table 5.5 gives data relating to the supply schedules in the market for tennis balls before and after the imposition of a tax of $1 per ball. The imposition of a tax affects the market supply curve through its impact on the marginal cost curve of each producer, who would have to pay to the government $1 in tax for each tennis ball sold. Hence suppliers would be willing to supply the same quantities at prices $1 higher than the original prices. This change in suppliers' intentions can be represented by a shift of the supply curve for tennis balls from SS to S'S', as shown in Figure 5.21.

The shift of the supply curve merely indicates a change in sellers' intentions. To discover what actually happens to the equilibrium price and quantity of tennis balls, it is necessary to compare the point of intersection of the supply and demand curves for tennis balls before and after the tax. As shown in Figure 5.22, the equilibrium price of tennis balls would rise from $3.00 to $3.75, and the equilibrium quantity would fall from 30 000 to 26 000.

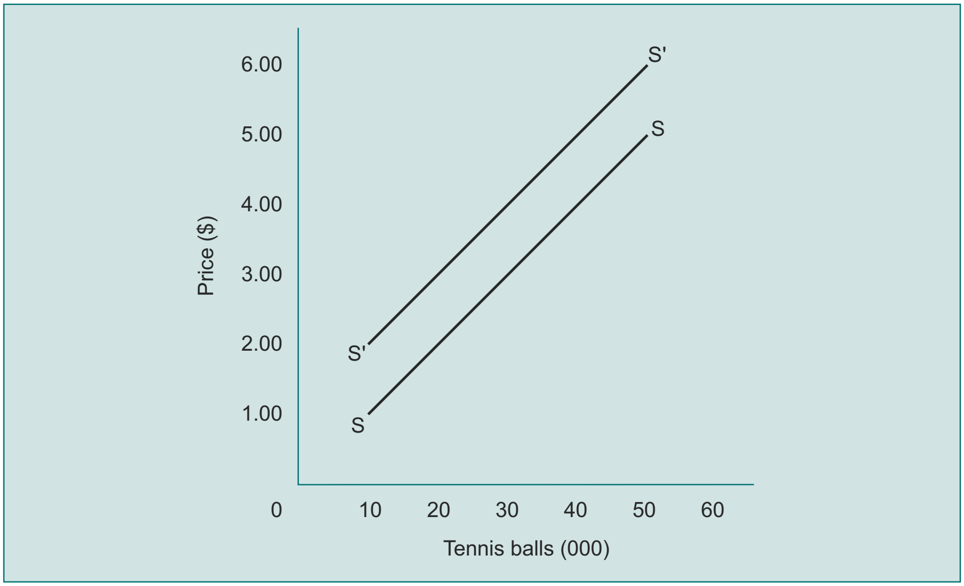


Figure 5.21 Shift of supply curve due to tax

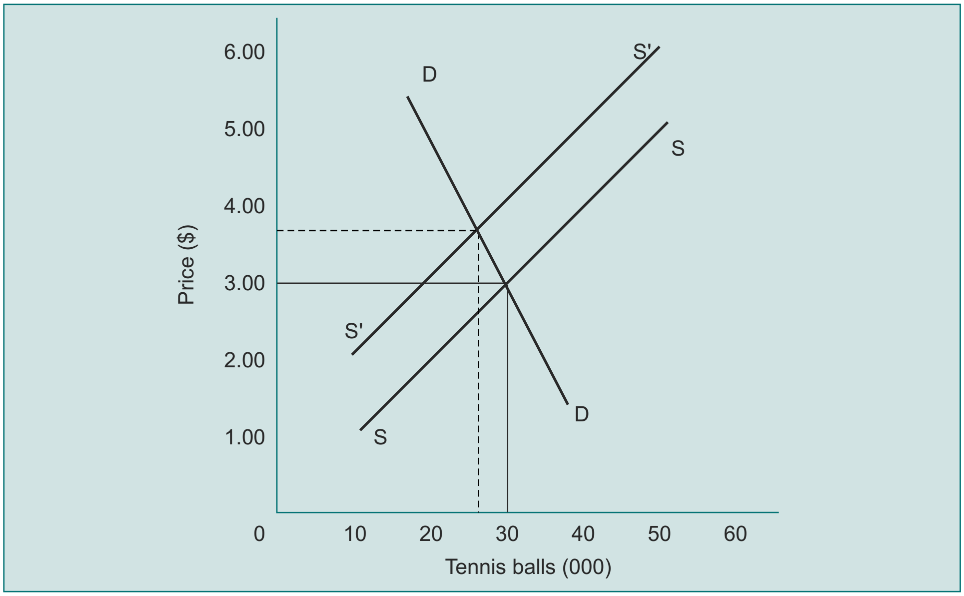


Figure 5.22 Effect of tax on equilibrium price and quantity

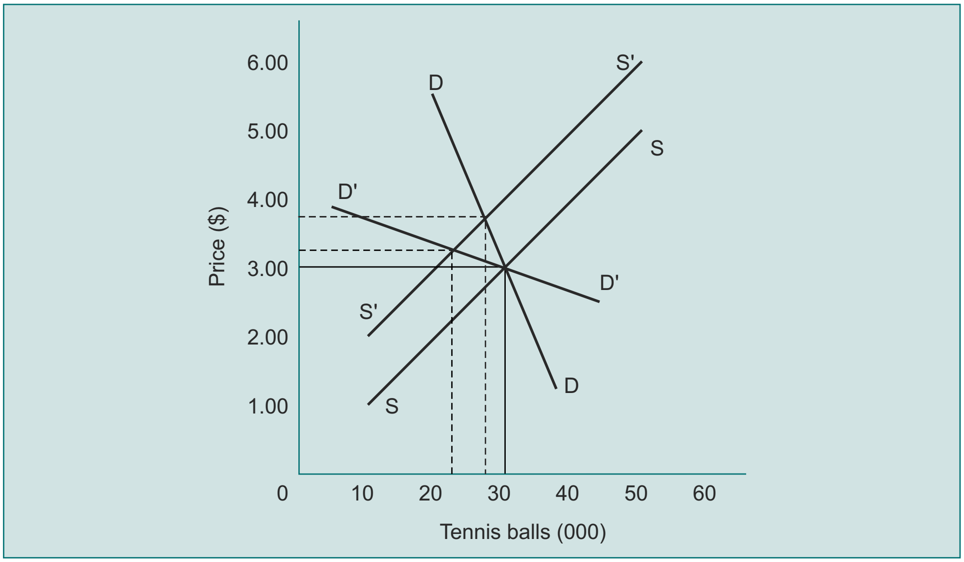


Figure 5.23 Effect of tax on equilibrium price and quantity with different demand elasticities

If a tax were to be imposed on every unit of a good bought and sold, the extent to which equilibrium price and quantity would change, and the incidence of the tax between consumers and producers, would depend on the relative elasticities of supply and demand.

If the demand for tennis balls were represented by DD as illustrated in Figure 5.23, then for every dollar paid to the government in tax, three-quarters would fall on consumers because of the price rise from $3.00 to $3.75. The remaining one-quarter would fall on producers, whose revenue from the sale of each tennis ball would fall by 25 cents, i.e. from $3.00 to $2.75 ($3.75 − $1.00 tax = $2.75).

But if the demand for tennis balls were represented by the more elastic demand curve D'D', the price after a tax of $1 per tennis ball was imposed would rise to $3.25, and the quantity bought and sold would fall to 23 000. For every $1 paid to the government in tax, only one-quarter would fall on consumers through the rise in price from $3.00 to $3.25. The remaining three-quarters would fall on producers, whose revenue from the sale of each ball would now be reduced by 75 cents, i.e. from $3.00 to $2.25 ($3.25 − $1.00 tax = $2.25).

If a tax were imposed on every unit of a good bought and sold, there would be a loss of both consumer surplus and producer surplus. Figure 5.24 again illustrates the effect of a tax on the market for tennis balls. At the pre-tax equilibrium price of $3.00, consumer surplus is denoted by the area DPE. At the post-tax equilibrium price of $3.75, consumer surplus is denoted by the smaller area DP'Y. Hence the loss of consumer surplus due to the imposition of the tax is denoted by the area P'PEY. At the pre-tax equilibrium price of $3.00, producer surplus is denoted by the area 0PE. The post-tax price, net of tax, received by producers is $2.75. At this price, producer surplus is denoted by the smaller area 0P''Z. Hence the loss of producer surplus due to the imposition of the tax is denoted by P''PEZ.

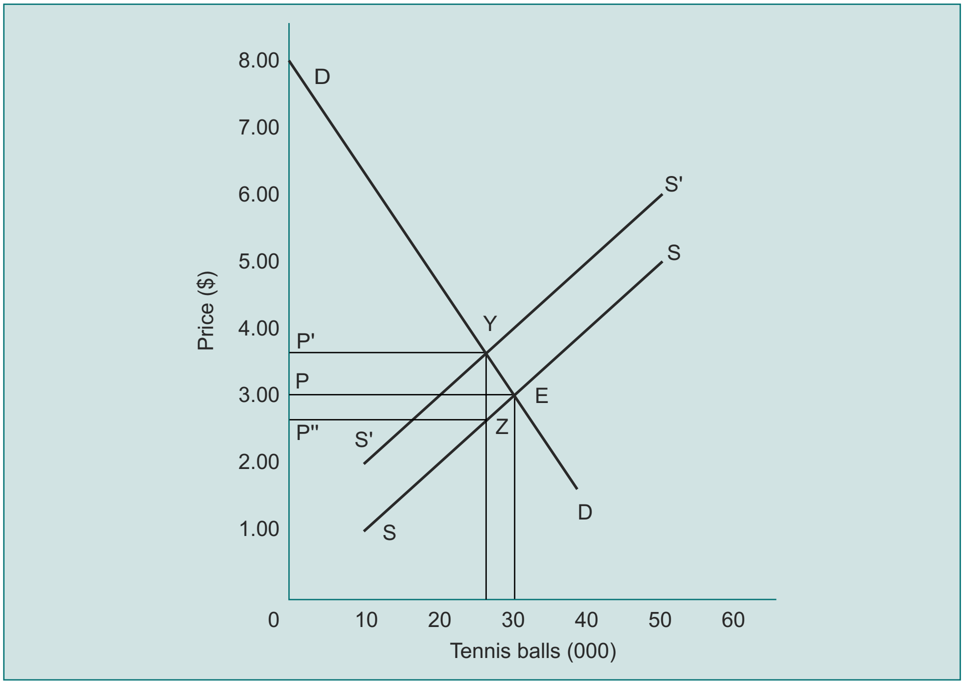


Figure 5.24 Effect of a tax on consumer surplus and producer surplus

The imposition of a tax, by raising the price paid by individuals for the good and lowering the price received by suppliers from the sale of the good, reduces the volume of transactions and thus reduces the gains from exchange.

5.6 Dynamic Adjustments in the Market

When conditions of demand and/or supply change in a market, a new equilibrium price or equilibrium quantity, or both, will be established. This method of analysing market changes is called *comparative statics* since it involves comparing equilibrium points before and after certain demand or supply changes occur. But the analysis does not stipulate the time necessary for a new equilibrium to be established, nor does it trace the exact process by which a market adjusts from one equilibrium point to another.

It was shown earlier that an increase in demand for sandwiches would lead to an increase in the equilibrium price of sandwiches, say from £1.40 to £1.50. The analysis did not specify whether the price of sandwiches would move up from £1.40 to £1.50 in one jump or more gradually until equilibrium is restored, or even whether price would suddenly rise above £1.50 immediately following the demand increase and then fall to the new equilibrium level. When the demand for a good changes, price will reach different levels in different time periods depending upon differences in elasticities of supply. If a change in demand were to occur while a market was in progress, and the supply curve for a good were completely inelastic, indicating that producers were unable to adjust the quantity supplied no matter what the price in the 'market period', the equilibrium price would change, but there would be no change in equilibrium quantity. For instance, if there were a sudden increase in demand in the market for salmon in a given day (the market period), suppliers would be unable to bring any more salmon to the market that day. This effect is illustrated in Figure 5.25.

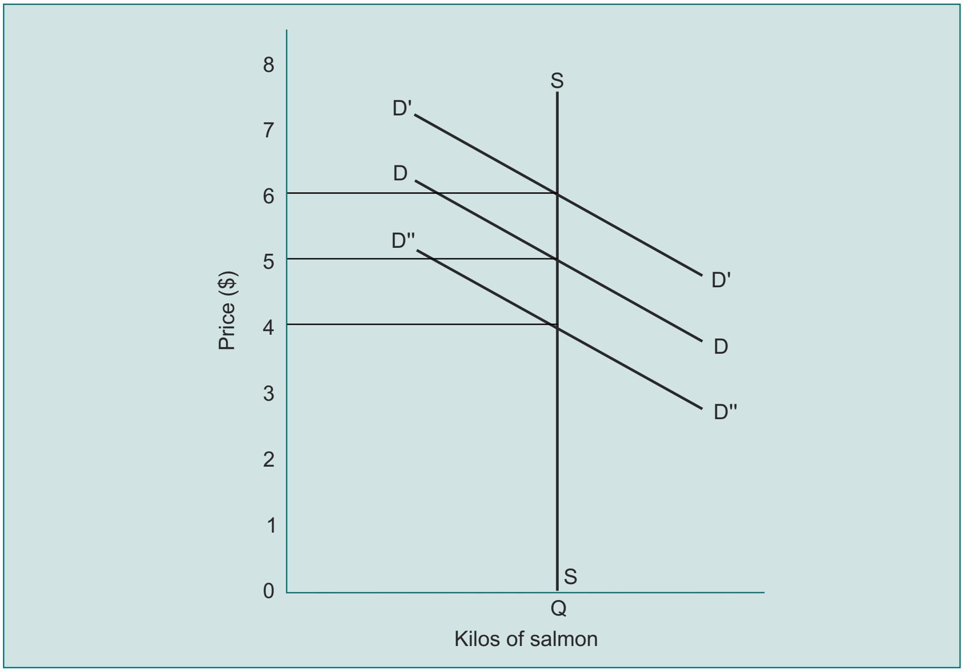


Figure 5.25 An increase in demand in the market period

The market is initially in equilibrium with a price of $5 per kilo and 0Q kilos of salmon bought and sold. But in this initial market period, a rise in demand for salmon, as shown by a shift in the demand curve from DD to D'D' would result solely in a rise in price from $5 to $6 per kilo. The equilibrium quantity of 0Q kilos would remain the same because in this time period the supply curve SS is completely inelastic because the time period is too short to catch more salmon and bring them to market. Conversely, if there were to be a sudden decrease in demand for salmon, as indicated by the shift of the demand curve from DD to D''D'' in the market period, equilibrium price would fall from $5 to $4 per kilo, but equilibrium quantity of 0Q kilos would remain unchanged.

5.6.1 The Effect of Changes in the Short Run

In the short run, the supply of a good would be less inelastic than in the market period because suppliers would be able to adjust the quantity hired of the variable factors of production. Comparing the short run with the market period, given a change in demand, equilibrium price would change, but not by as much as in the market period. Equilibrium quantity would change.

Example

Following an increase in demand, in the short-run period the boat owners would employ more labour in order to catch more salmon. The supply curve for this period would be represented by S'S', which is less inelastic than the supply curve SS in the market period. The short-run period is illustrated in Figure 5.26.

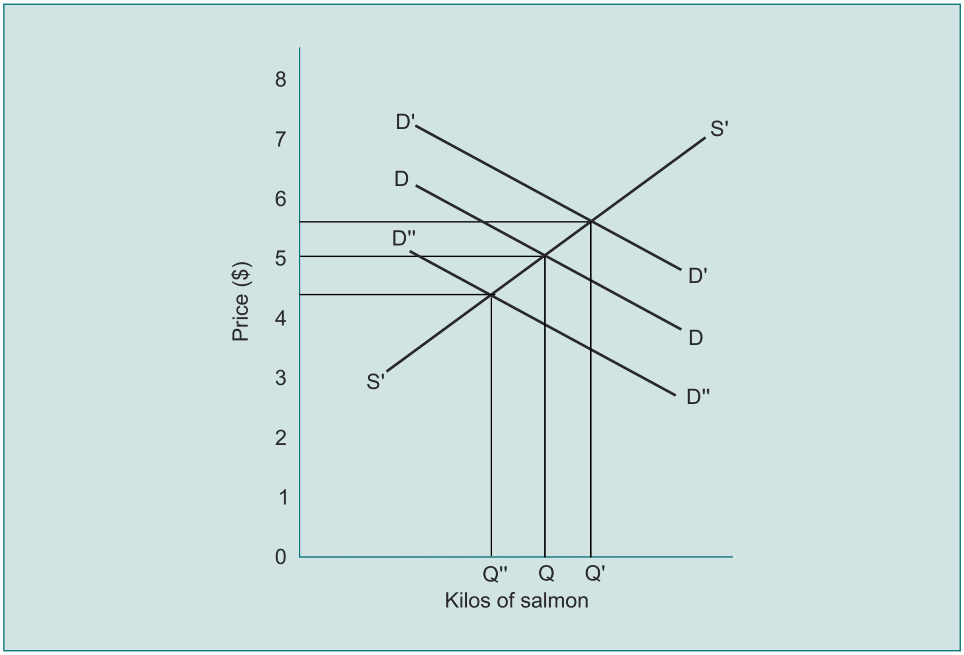


Figure 5.26 An increase in demand in the short run

The equilibrium price in the short run would be \$5.75 per kilo, as shown by the intersection of D'D' and the short-run period supply curve S'S'. This price is higher than the initial equilibrium price of \$5.00 but lower than the market period equilibrium price of \$6.00. Equilibrium quantity would also increase from 0Q to 0Q' kilos of salmon.

Similarly, following a decrease in demand, as indicated by a shift of the demand curve from DD to D''D'', equilibrium price would fall from \$5.00 to \$4.25 per kilo, which is not quite as large a fall as occurred in the market period. Equilibrium quantity would also fall from 0Q to 0Q''.

5.6.2 The Effect of Changes in the Long Run

In the long run, the supply of a good would be more elastic than in the short run because producers would be able to alter the quantities hired of all factors of production and firms would be able to enter or leave the industry. Equilibrium price would differ from its original level but not by as much as in the short-run period, provided the long-run supply curve were upward sloping. If the long-run supply curve were completely elastic, price would return to its original level. Equilibrium quantity would change by more in the long run than in the short run.

Example

Following an increase in demand for salmon, in the long run the suppliers would be able to order more fishing boats as well as employing more variable factors, and new suppliers would be able to enter the market. Assuming that costs were to rise as firms expanded and as new firms entered the market, the long-run supply curve would slope upwards but would be more elastic than in the short run, as shown in Figure 5.27.

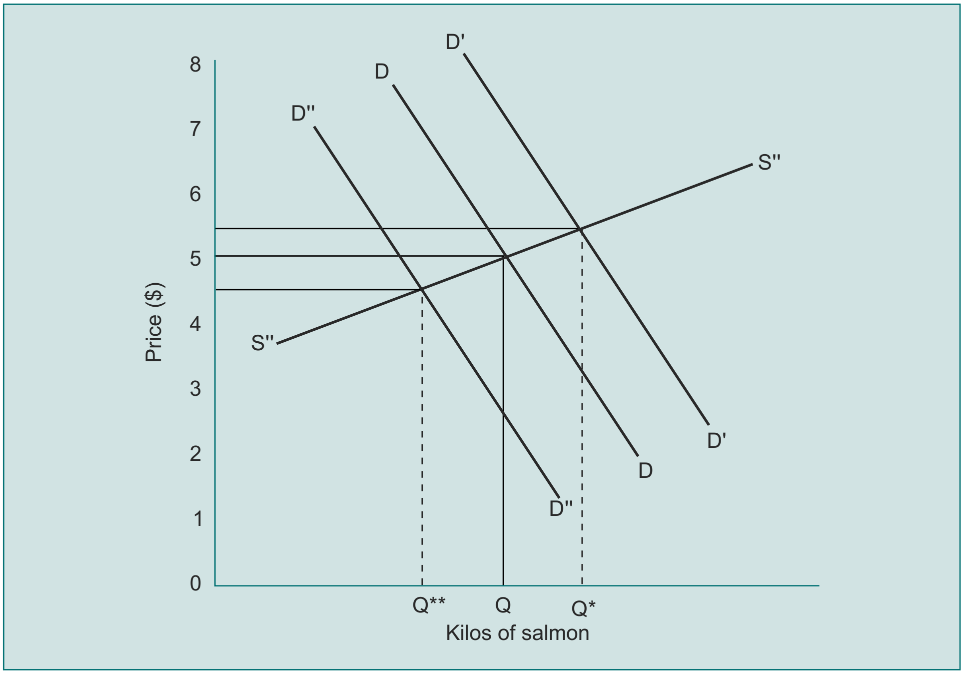


Figure 5.27 An increase in demand in the long run with an upward sloping supply curve
The equilibrium price in the long run would be $5.50 per kilo, determined by the intersection of D'D' with the long run supply curve S''S''. This price is higher than the original equilibrium level but lower than equilibrium price in the short run. Equilibrium quantity would be 0Q* which is greater than the equilibrium quantity in the short run. Similarly, following a decrease in demand for salmon, as represented by a shift of the demand curve from DD to D''D'', the equilibrium price would fall to $4.50, which is higher than its level in the short run. The equilibrium quantity would be 0Q** which is lower than the short-run equilibrium quantity.

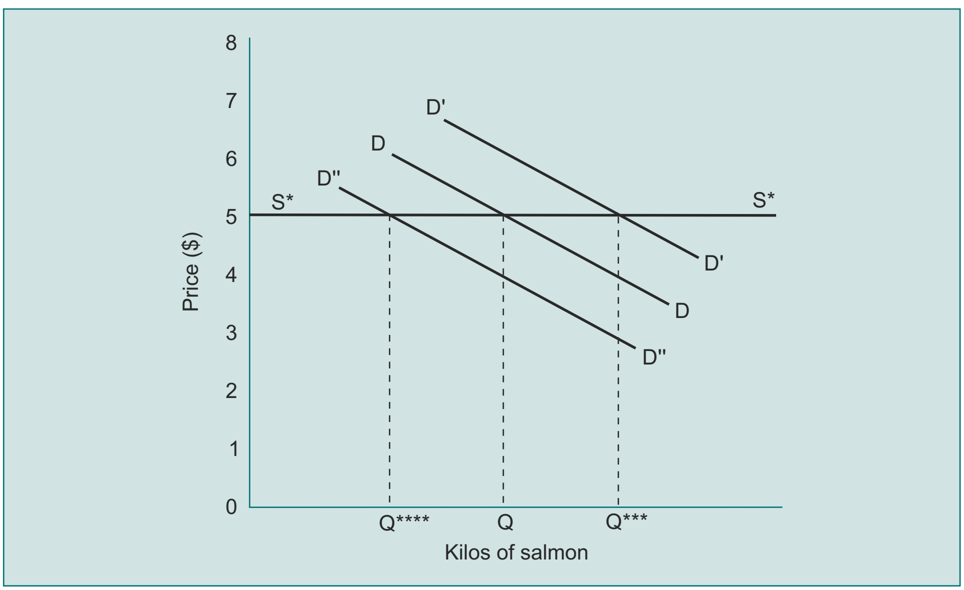


Figure 5.28 An increase in demand in the long run with a horizontal supply curve

If, however, costs were to be unaffected by firms entering the market, the long-run supply curve would be a horizontal line S*S* as shown in Figure 5.28. As a result of a change in demand, equilibrium price would be restored to its original level of $5.00 per kilo. Equilibrium quantity, however, would change by more than in the short run, rising to 0Q*** as a result of an increase in demand, and falling to 0Q**** as a result of a decrease in demand.

5.6.3 Cyclical Patterns in Markets

In the markets for some goods and resources whose supply adjusts only after a time lag, the prices of those goods and resources tend to fluctuate in a cyclical fashion over time. For example, for many years the price of pigs in Britain oscillated from year to year, sometimes reaching relatively high levels and sometimes relatively low levels. Given the time required from breeding to fattening, the quantity supplied by pig farmers could respond to market changes only after many months had elapsed. Similar patterns have occurred in the markets for resources such as physicists in the US. Their salaries became substantially higher than those earned in similar highly qualified occupational groups, but then declined in relation to comparable salaries two decades later.

Markets that operate in such a cyclical fashion are often analysed using what is known as the *cobweb model*. Although it uses conventional tools of supply and demand, it is different in that the market being analysed operates in discrete blocks of time, and the supply curve contains a time lag, where one period's prices affect the *next* period's quantities.

Example

In Figure 5.29, suppose the curves DD and SS represent a hypothetical demand curve for, and supply curve of, pigs. If the market were initially in equilibrium at E, then the quantity of pigs that farmers would plan to produce would just match consumers' demands for pork. But suppose instead that there has been widespread disease amongst pigs, causing the quantity that farmers could supply to fall to $0Q_1$, which is below the equilibrium level.

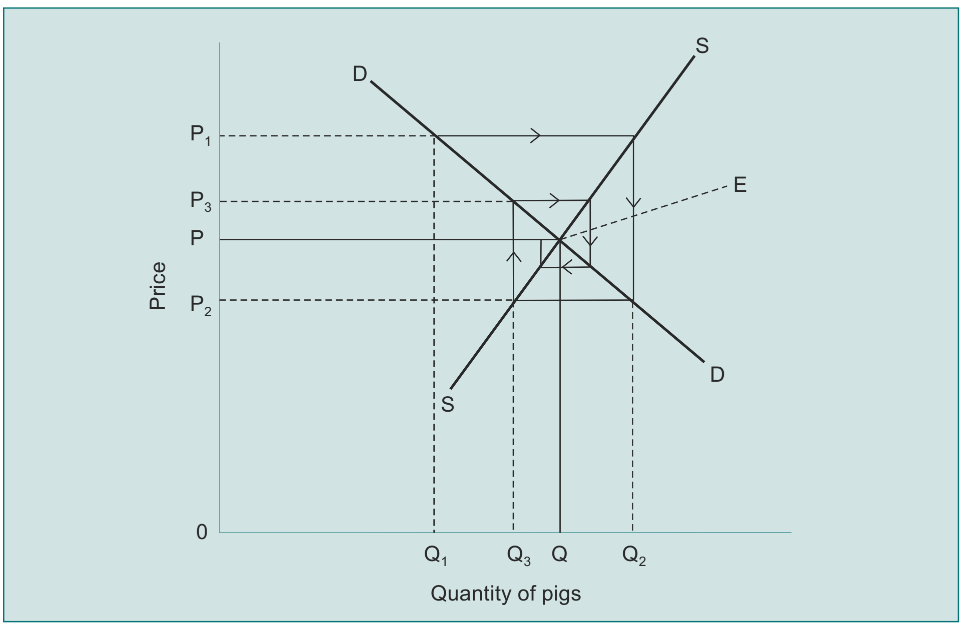


Figure 5.29 The cobweb model – convergence towards equilibrium

The price that $0Q_1$ will fetch on the market is $0P_1$, i.e. the price that consumers, as shown by the demand curve DD, are prepared to pay for $0Q_1$ units. But the relatively high price in this time period will induce farmers to increase substantially their planned supply of pigs for the following year. If farmers use price $0P_1$ as a basis for determining next year's supply of pigs, they would bring $0Q_2$ units to the market. But the price that clears the market of $0Q_2$ is $0P_2$, i.e. the price according to curve DD which consumers would pay for $0Q_2$ pigs. This price is below the equilibrium level, whereas the previous price of $0P_1$ was above the equilibrium level.

But in response to this relatively low market price, farmers will cut back on production plans and be prepared to supply next year only $0Q_3$ at a price of $0P_2$ When $0Q_3$ pigs arrive at the market, the clearing price will be $0P_3$, which is once more above the equilibrium level. And so the process of cyclical price movement goes on. But in each time period, price moves closer to the equilibrium level until, as shown in Figure 5.29, price P is finally reached. Over a period of time the market has converged towards equilibrium.

However, in some markets, using the cobweb model, cyclical movements in price do not converge towards equilibrium but move further away from equilibrium.

Example

In Figure 5.30, if price were initially at $0P_1$, in each succeeding time period price would move to $0P_2$, $0P_3$, and so on, gradually diverging from the equilibrium level.

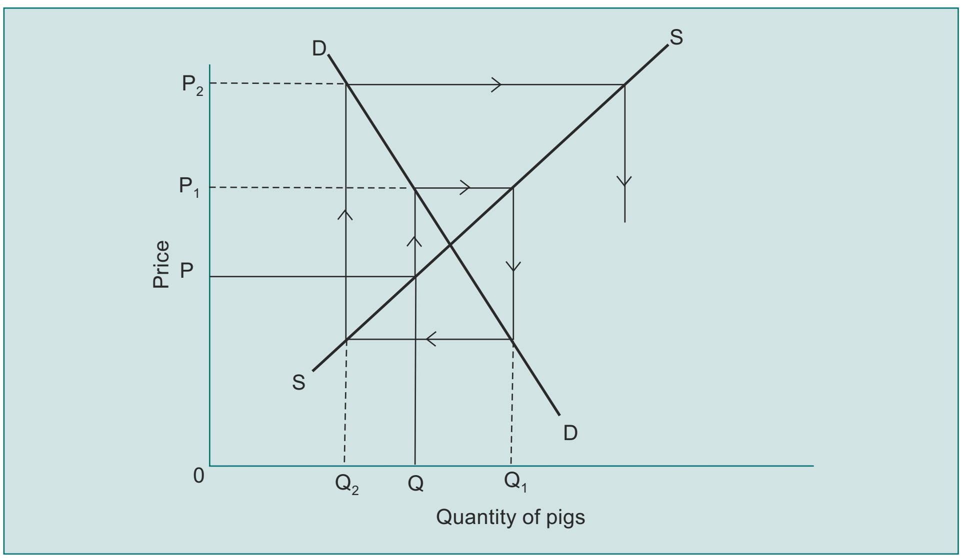


Figure 5.30 Divergence from equilibrium

In use of the cobweb model, whether a market will converge or diverge will depend upon the relative slopes of the demand and supply curves. When the supply curve is steeper than the demand curve, price will converge towards equilibrium. When the demand curve is steeper than the supply curve, price will diverge from equilibrium.

In Figure 5.29, the supply curve SS is steeper than the demand curve DD, and price converged towards the equilibrium level of 0P from a situation in which the market was not in equilibrium. As price changes, suppliers adjust quantity less than consumers do, and this has a dampening effect on the cyclical movements around equilibrium. Conversely, in Figure 5.30, the demand curve DD has a steeper slope than the supply curve SS, and from an original disequilibrium situation represented by price $0P_1$, the price in successive time periods diverged even further from equilibrium. In these circumstances, buyers adjust the quantity demanded less than producers adjust quantity supplied, and so the cyclical price movements diverge increasingly from the equilibrium level.

The cobweb model of dynamic adjustment depends on the further assumption that producers never learn from experience and adjust their production plans accordingly. It also depends on the assumption that speculators, i.e. people who may buy quantities of the good at low prices and then resell later at high prices, have no role to play in the market. It is, though, highly unlikely in the real world that, over a period of years, producers and consumers would learn nothing about the way in which the market was behaving and adjust their behaviour accordingly. Pig farmers would realise that basing next year's production plans on this year's prices and on the expectation that these prices would persist was an unrealistic assumption. Despite this 'learning-by-doing' effect, the cobweb model provides a realistic explanation of some markets, such as the world shipbuilding market, which throughout its history has been subject to 'feasts' and 'famines'.

Learning Summary

You are now in a position to understand how the forces of demand and supply in competitive markets determine equilibrium prices and equilibrium quantities of goods bought and sold. You can explain how equilibrium prices and quantities respond to changes in demand and supply. (You will meet the concept of equilibrium again and again as you make progress through the course.) You are also able to explain how price ceilings and price floors can lead to black markets and surpluses respectively in controlled markets. You understand the impact of sales taxes on both prices and quantities. You can explain why some markets adjust only after a time lag and why such adjustments can cause the price of the goods in question to fluctuate in a cyclical fashion.

Review Questions

Multiple Choice Questions

5.1 Which of the following is correct?

A market for a good, service or resource exists when

A. there is a specific location where trade can take place.

B. individuals have something that they can sell.

C. individuals wish to buy something.

D. potential buyers are in communication with potential sellers.

5.2 A local store deals in second-hand musical instruments. It offered for sale three pianos at $1000 each but nobody wished to buy the pianos at that price. Over a period of months the store manager gradually reduced the asking price for the pianos but he was still unable to find buyers. Eventually a representative of a music college called at the store and offered to take the pianos at zero price on condition that the store paid the college for their removal. The store manager agreed.

Which of the graphs in Figure 5.31 represents the market for the pianos?

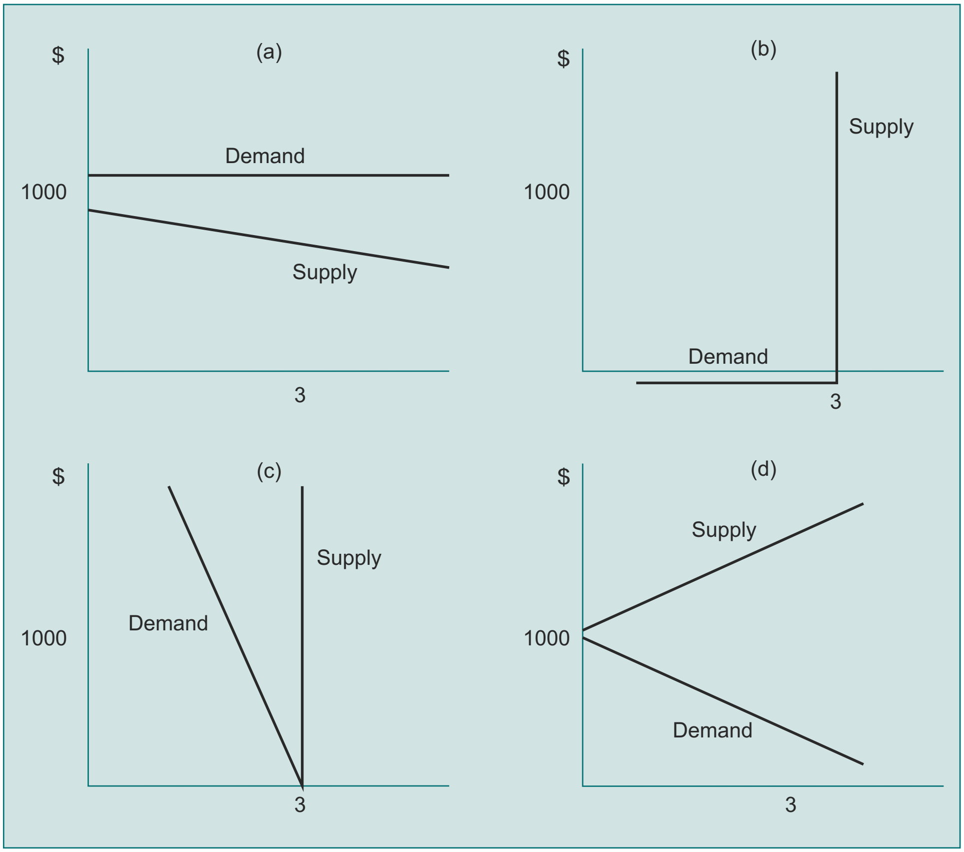


Figure 5.31

A. Figure (a) only.
B. Figure (b) only.
C. Figure (c) only.
D. Figure (d) only.

5.3 Suppose the market for a newly developed gold-plated pencil were represented by the demand and supply schedules below:

Price ($)	Quantity demanded per day	Quantity supplied per day
600	0	1 000
500	0	700
400	0	350
300	0	0
250	0	0
200	50	0
150	100	0
100	150	0
50	250	0

Which of the following is correct?

A. Equilibrium quantity would lie between 50 and 350 pencils per day.

B. No exchange would occur.

C. Suppliers would have to lower their prices to $200 since that is the maximum price consumers are willing to pay.

D. Consumers would have to pay $400 to obtain pencils since that is the minimum price at which suppliers are prepared to sell.

5.4 The price system reacts to excess demand for a good in the short run by

A. lowering the price and profits of firms producing the good.

B. raising the price and producer profits.

C. lowering the price but increasing producer profits.

D. raising the price but lowering producer profits.

5.5 When influenza vaccine first became available in the US, the government set the price equal to the cost of production. At that price, output was insufficient to fulfil orders and the government regulated the distribution of the vaccine. Which of the following is correct if the vaccine had been sold privately without government intervention?

A. The price would have been higher.

B. The price would have been lower.

C. The price would have been the same.

D. Whether the price would have been higher or lower cannot be determined from the information given.

5.6 If the equilibrium price prevails in the market for a good

I. no consumers would buy units of the good if the price were higher.

II. excess demand and excess supply are both zero.

III. no producers would sell units of the good if the price were lower.

Which of the following is correct?

A. I only.

B. II only.

C. I and III only.

D. I, II and III.

5.7 A city has decided to build 5000 dwelling units and to lease them to low-income residents at a rental below cost and the going rate in the private market. Other things (such as population) being the same, which of the following effects would you expect this to have on the market for private housing?

A. A decrease in rent, followed later by a decrease in the quantity supplied.
B. A decrease in tenants, followed later by an increase in rents.
C. An increase in rents, followed later by an increase in the quantity supplied.
D. No effect, because the low-income persons who will be eligible for the 5000 city-built dwelling units cannot afford acceptable private housing.

5.8 Assume the market for golf clubs to be initially in equilibrium. If there were a considerable increase in the number of people wishing to take up golf and, simultaneously, a decrease in every golf-club manufacturer's production costs due to the use of a new metal compound, which of the following would be correct?

A. The equilibrium price of golf clubs might rise, fall or remain at the same level, but the equilibrium quantity would increase.
B. The equilibrium price of golf clubs would rise, but the equilibrium quantity might increase, decrease or remain at the same level.
C. The equilibrium price of golf clubs would fall, and the equilibrium quantity would increase.
D. The equilibrium price of golf clubs would fall, but the equilibrium quantity might increase, decrease or remain at the same level.

5.9 Suppose a government wants to raise money by means of a tax levied on each unit of a good bought and sold. The government wants the incidence of the tax to fall mainly on buyers, not producers. It also wants to disturb the quantity bought and sold as little as possible. Which of the following demand or supply curves best meets these requirements?

A. Highly inelastic supply.
B. Highly elastic demand.
C. Highly elastic supply.
D. Highly inelastic demand.

5.10 If a government were to increase the sales tax on new cars, what would be the effect on the equilibrium price and equilibrium quantity of second-hand cars?

A. Both increased equilibrium price and equilibrium quantity.
B. Equilibrium price would decrease but equilibrium quantity would increase.
C. Equilibrium price would increase but equilibrium quantity might increase, remain the same, or decrease.
D. There would be no effect, since the second-hand market is independent of the new car market and the sales tax applies only to new cars.

5.11 Assume the market for fresh salmon to be in equilibrium. If the demand for fresh salmon were suddenly to increase, which of the following would happen in the market (immediate) time period?

A. *Equilibrium price*: no change; *Equilibrium quantity*: no change
B. *Equilibrium price*: increase; *Equilibrium quantity*: increase
C. *Equilibrium price*: increase; *Equilibrium quantity*: no change
D. *Equilibrium price*: no change; *Equilibrium quantity*: decrease

5.12 Suppose that the market for tomatoes were initially in equilibrium, but that costs of production for every supplier were to rise as a result of a significant increase in wages paid to all workers in the industry. Which of the following would occur?

A. *Equilibrium price*: increase; *Equilibrium quantity*: increase
B. *Equilibrium price*: increase; *Equilibrium quantity*: decrease
C. *Equilibrium price*: decrease; *Equilibrium quantity*: increase
D. *Equilibrium price*: decrease; *Equilibrium quantity*: decrease

Case Study 5.1: Housing in Southern California

This case is an application of the theory of price determination. Before you tackle it you should understand:

a. shifts of the supply curve
b. shifts of the demand curve
c. equilibrium price and quantity
d. maximum price
e. lottery
f. substitution between goods
g. income distribution.

The case demonstrates that changing demand and supply conditions can have a substantial effect on equilibrium price and quantity, the extent of the effect depending on the shape of the demand and supply curves.

House Sales by Lottery

The following extracts are taken from a report entitled 'South Californian Bubble', which appeared in *The Economist*:

> South California is in the grips of a property boom quite out of step with the rest of the country. A frenzy of speculation has pushed up new house prices by 2 per cent a month for the past two years – 50 per cent faster than the national average …
>
> The market seems to be expanding as people move south, as post-war babies begin rearing their own families, and as the burgeoning numbers of divorced couples search for two homes rather than one; indeed, the number of single adults desiring to live alone is one of the most remarkable features of California's housing boom …
>
> … to make things worse, the recent worldwide inflation in materials' prices has significantly pushed up building costs, and builders also fear they may soon run out of land on which it is permissible to build…
>
> The hottest area for building and for speculation is Orange County, a populous, largely residential area south of Los Angeles, where the demand for homes is so great that lotteries have become commonplace. There are, for example, 5500 names on the waiting list for the next 100 houses to be built in one popular tract. Elsewhere, a construction company recently held a lottery to select those among 1000 applicants who might buy 191 houses priced at $115 000.

1 Now consider the following questions:

1. There has been both an increase in production costs and an increase in the demand for new houses. What effect will this have on the price of new houses and the quantity sold on the market?

2. What will be the effect of the increased demand for new houses on the market for second-hand houses?
3. Compare the effects on economic efficiency and income distribution of putting 191 houses up for sale at $115 000 and allocating them by using a lottery with the effects were the price to be determined on the free market.

Case Study 5.2: Obtaining Wimbledon Tennis Tickets

This case is an application of the theory of price determination. Before you tackle it you should understand:

a. the determination of equilibrium price
b. the effect of a maximum price on the quantity demanded and the quantity supplied
c. a price-inelastic supply curve
d. total revenue.

The case demonstrates:

- the effect on total revenue of setting a maximum price rather than allowing the market to determine price;
- that an attempt to make one group better off can result in making another group worse off; and
- that price fixing can reduce the efficiency with which the market system allocates resources.

Wimbledon Ticket Allocations

If you are a tennis enthusiast there are several ways to obtain Wimbledon tennis tournament tickets. You may:

(a) enter your name in the pubic ballot;
(b) enter a ballot of an association affiliated to the Lawn Tennis Association (LTA);
(c) become, or have your company become, a tennis sponsor and negotiate Wimbledon tickets from the LTA in return for sponsoring, say, a UK regional tournament or youth coaching;
(d) purchase a ticket from a debenture holder, for although this number is small it is the only legal 'market' source of tickets;
(e) buy a ticket from an agency specialising in sporting events, which normally means buying a package including travel, hotels, entertainment and tickets; or
(f) queue at the gates of Wimbledon for the daily gate allocation.

The above methods either have a low probability of success or are expensive. A much more reliable way is to contact a ticket tout and negotiate a price. Of course, if you enter the public ballot and are lucky enough to obtain a ticket for which you will pay face value, or obtain tickets by any other means, you yourself can become a ticket tout. The LTA is not fond of touts and it attempts to limit the resale of tickets by stipulating resale conditions, violation of which can lead to ejection from Wimbledon by the 'illegal' purchaser as a trespasser and removal of ticket-allocation privileges of the person acquiring the ticket originally.

Wimbledon is obviously losing a small fortune by adopting its particular pricing policy. However, Wimbledon officials argue that setting the price at a low level allows enthusiasts to obtain a ticket who would not be able to afford the free-market price.

An economist has been asked to comment on the method of allocating Wimbledon tickets. She starts by pointing out that all the bodies involved in the distribution of Wimbledon tickets are completely superfluous. She proposes that the Wimbledon officials should estimate the equilibrium price of tickets and sell them at that price; they may not get prices right first time, but they will learn by doing so and anyone who really wants a ticket will be able to obtain one at the prevailing price. After all, she points out, the markets for Rolls-Royce cars and gold bars work well; why shouldn't a Wimbledon ticket market work equally well? Furthermore, the economist claims that while some people may benefit from cheap tickets, many more people lose out and as a result the present method of allocating tickets is not efficient.

1 In evaluating the economist's conclusions, the following need to be considered.

1. What would be the effect on (a) the price of tickets, and (b) the demand for tickets if the proposal were adopted?
2. Who would benefit and who would lose from the proposal?
3. Wimbledon officials argue that price fixing makes it possible for enthusiasts who cannot afford the free-market price to attend. Do you think this argument is correct?
4. Is there a case for increasing the penalties for ticket touting on the basis of its effect on economic efficiency?

When answering these questions you will bear in mind that there are several different types of ticket to Wimbledon and that some days, such as Finals Days, are more popular than others. However, to avoid giving a very complicated answer, you should assume that there is only one kind of ticket and that there is no difference in the demand from one day to the next.

Module 6

Economic Efficiency

Contents

6.1 Introduction

Within legal limits, in free-market or capitalist economies, households are free to buy the goods and services they want and free to withhold or offer their resources – for many, their labour services – to whichever firm they choose. Firms, in turn, are free to decide which goods and services to produce and which resources to hire and fire. A market economy, however, does not suddenly appear in the world; it evolves slowly over time, and in the history of the human race it is a relatively recent phenomenon. If we go back a mere 300 years, we find there was no market economy as we have it today. There were markets then – indeed there were markets thousands of years ago – but most of them were small and local. The driving force that led to the establishment of larger markets – and eventually market economies – was people's desire to become better off.

The idea of the pioneering families in early times being self-sufficient, building their own homes, cultivating crops, hunting and fishing, and making their own clothes, has been romantically captured by Hollywood films – but, in reality, such life was 'nasty, brutish and short'. It became obvious to such pioneers that they could become better off by specialising in doing things they were good at and exchanging goods with other people who were better at doing other things. Such exchange required markets.

Consider first the advantages of exchange. Adam and Eve live on a small island; Eve is a very good swimmer and each day she spears five fish, which she eats. Adam cannot swim but can climb date-palms and each day he collects four baskets of dates, which he consumes. Adam and Eve have identical utility functions for fish and dates, as shown in Figure 6.1, and in these functions we shall assume that a measure of satisfaction is a util.

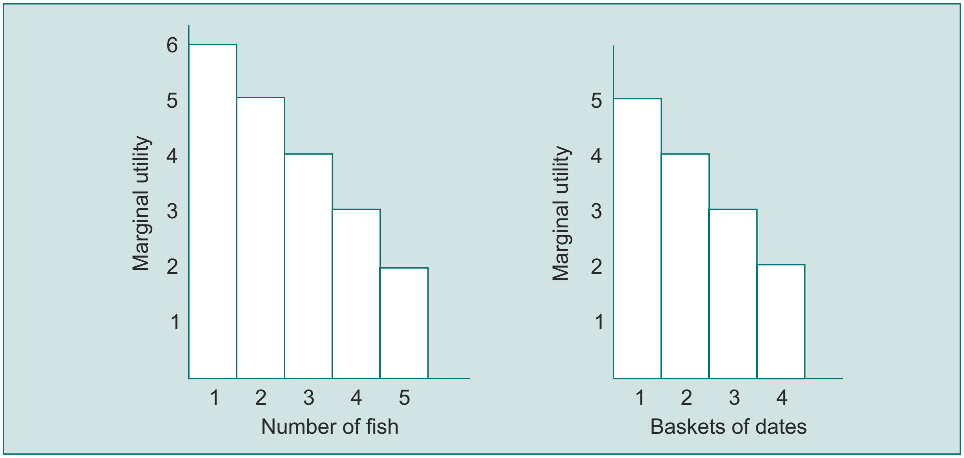


Figure 6.1 Marginal utilities of fish and dates

The principle of diminishing marginal utility exists. Eve receives six utils of satisfaction from consuming her first fish, five utils from the second and so on. Thus her total daily enjoyment for consuming her five fish is 20 utils. Adam, with an identical utility function for fish, would also receive 20 utils for consuming five fish – if he had them; but since he does not, his satisfaction is derived solely from consuming dates. He enjoys 14 utils from his consumption of four baskets of dates, i.e. 5 + 4 + 3 + 2.

Now Adam and Eve get together. Adam offers Eve one basket of dates. What is the marginal cost to Adam of this action? It is two utils: the satisfaction he derives from the last basket of dates consumed. The gain in utility to Eve is five utils: the satisfaction gained from her first basket of dates. Eve in return offers Adam one fish. She loses two utils, Adam gains six utils. Thus the net result of the exchange is Adam gaining four utils and Eve gains three utils. Both have become better off without any increase in the total amount of fish and dates. Does it pay Adam and Eve to indulge in further exchange? Yes. By giving up a second basket of dates for a second fish Adam loses three utils on the dates but gains five on the fish; Eve loses three utils on the second fish but gains four utils on the second basket of dates.

Now suppose Eve teaches Adam how to swim and spear fish. Unfortunately he is not very good at it and only catches one fish every second day. Likewise Eve, it is discovered, suffers from vertigo and because she tends to fall off the palm trees she can only collect half a basket of dates per day. To maximise utility, should each both fish and climb date-palms? The answer is ‘no’: total output of fish and dates will be greater if each specialises, so that Eve just fishes and Adam just collects dates. This is *division of labour*. Total utility will be greater with trade or exchange.

How much division of labour occurs in society in the real world is limited by the extent of the market. In small, isolated communities there is often only one doctor, who provides all medical attention. As markets expand, doctors specialise in children’s diseases (pediatricians), in skin diseases (dermatologists), in brain diseases (neurologists), in mind disorders (psychiatrists), in heart diseases (cardiologists), and in blood diseases (haematologists). As specialisation increases, so do the quality and quantity of the goods and services produced; and as a consequence, the materialistic well-being of society is enhanced.

The opportunity for individuals to specialise in the production of goods and services and exchange them for other goods and services in markets has a substantial impact on the supply of and demand for each good and service. In Module 5 we analysed the way in which demand

and supply interact to determine the equilibrium price and quantity in a market for an individual good. In doing so we analysed the basic mechanics of a market economy, but the equilibrium determined in any single market is only a partial equilibrium because it ignores the interdependence among markets. We saw that what happens in the smoked-salmon market has an effect on the fishing-reel market, which has an effect on the steel market, which in turn has an effect on the iron-ore market – these are some of the interdependencies that are well-nigh impossible to trace out in full because there are millions of such interdependencies. However, a competitive market or capitalist economy, in bringing about a general equilibrium of prices and quantities in all markets, automatically takes these interdependencies into account. It is possible for competitive markets to perform this complex task for two reasons. First, no single individual, firm or government has to make all the decisions: households and firms all pursue their own interests, and the utility, profit and general equilibrium solution that is produced is the result of all those individual decisions taken together. Second, markets give out signals that convey to the many individual decision makers the information necessary to plan their expenditure or production. These signals are, of course, prices.

The decisions of individual households and firms acting freely and independently are reflected in prices and outputs in the many competitive markets that exist in a capitalist economy. Equilibrium is reached in all markets when the combination of goods and services demanded is just equal to the combination of goods and services supplied.

Does this provide us with an efficient solution to the problem of resource allocation? To answer this question we should recall our earlier discussion about efficiency. Technological engineering efficiency occurs when a given output of a good is produced using the least possible amount of resources. If all goods and services in the economy were so produced, the economy as a whole would be technologically efficient. Is there a tendency for a capitalist economy to be technologically efficient? If markets are freely competitive, the answer is 'yes': self-interest, the desire to be profitable, forces firms in a competitive environment to adopt the most cost-effective ways to produce goods and services, for if they do not, they will be forced to do so or forced out of business by more efficient firms. Thus a competitive market economy will tend to be technologically efficient – we shall face the behaviour of firms operating in a non-competitive environment in Module 7.

Before an economy can be called economically efficient, it must be more than technologically efficient. Economic efficiency requires more than producing a combination of goods and services using the least amount of resources; it must also produce the combination of goods and services that satisfies consumers' wants as fully as possible. Will a competitive market economy tend to be economically efficient as well as technologically efficient? To answer that question, we must pull together most of what we have analysed to date.

6.2 The Marginal Equivalency Conditions

A consumer will maximise utility from a given income when that income is allocated to goods and services so that the marginal utility of each good and service divided by its price is equal. That is:

$$\frac{MU_A}{P_A} = \frac{MU_B}{P_B} = \cdots = \frac{MU_N}{P_N}$$

For example, in a sandwich/coffee world in which the price of a sandwich is four times the price of a cup of coffee, a utility-maximising consumer will allocate his income to coffee and sandwiches such that the last dollar spent on each will yield equal satisfaction, i.e. the last sandwich consumed will yield four times the utility of the last cup of coffee consumed.

A competitive firm will maximise profit by producing that level of output at which price equals marginal cost. That is:

$$P_A = MC_A$$

For example, if the market price of pencils is \$0.10 each and if a competitive pencil manufacturer is producing a level of output where the marginal cost is \$0.08, he will not be maximising profits. The production of additional pencils until the marginal cost rose to \$0.10 would add more to revenue than it would to cost and consequently add to profit or diminish loss.

Assuming a two-good world (goods A and B) composed of utility-maximising consumers and profit-maximising competitive firms, the following conditions must hold:

$$\frac{MU_A}{P_A} = \frac{MU_B}{P_B} \tag{6.1}$$

and

$$P_A = MC_A \quad \text{and} \quad P_B = MC_B \tag{6.2}$$

Substituting the MCs of Equation 6.2 for the Ps in Equation 6.1 results in

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B} \tag{6.3}$$

Because MU_A is the utility gained from consumption of the last unit of good A and MC_A is the dollar resources required to produce that unit,

$$\frac{MU_A}{MC_A}$$

is the utility gained by a consumer from the last dollar's worth of resources in the production of good A; similarly,

$$\frac{MU_B}{MC_B}$$

is the utility gained by a consumer from the last dollar's worth of resources in the production of good B. Thus Equation 6.3 signifies that for each utility-maximising consumer in a two-good world in which the goods are produced by profit-maximising competitive firms, the utility gained from the last dollar's worth of resources in the production of good A will equal the utility gained from the last dollar's worth of resources in the production of good B.

Example

If a student attempting to maximise total exam points has eight hours of study time remaining to allocate to History and Economics and allocates four hours to each, he will not maximise total points unless that last hour allocated to History adds the same increase in points as it would if it were allocated to Economics. Suppose that the third hour allocated to History adds 10 points and the fourth hour adds a further six points, whereas the fourth hour allocated to Economics adds 15 points and a fifth hour would add 10 points. To maximise total points, the student should reallocate his fourth hour of History study to Economics. Total History points will decrease by six points but total Economics points will increase by 10 points, yielding a net gain of four points. The last hour allocated to each will yield the same increment to score, i.e. 10 points.

Any allocation of resources that does not produce

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B}$$

results in a non-equilibrium situation. Forces will cause a resource reallocation until the ratios of marginal utility to marginal costs are equal for all goods and services.

If, on the other hand, the marginal utility-to-cost relationship is such that

$$\frac{MU_A}{MC_A} > \frac{MU_B}{MC_B}$$

there will be an incentive for society to transfer resources from industry B to industry A. Fewer resources in industry B will reduce the quantity of B produced, with a concomitant decrease in total utility from good B since less B will be consumed. However, the marginal utility of B will increase because of the assumption of diminishing marginal utility. More resources in industry A will increase the quantity of A produced, with a concomitant increase in total utility from good A since more A will be consumed. However, the marginal utility of A will decrease because of the assumption of diminishing marginal utility. Since

$$\frac{MU_A}{MC_A} > \frac{MU_B}{MC_B}$$

the reallocation of resources at the margin from industry B to industry A will result in a net gain of total utility.

Example

To return to our sandwich (A) and coffee (B) world: if

$$\frac{MU_A}{MC_A} = \frac{4}{1} \quad \text{and} \quad \frac{MU_B}{MC_B} = \frac{2}{1}$$

in other words the last dollar's worth of resources in the production of sandwiches yields twice the utility of the last dollar's worth of resources in the production of coffee, society would experience a net gain in utility by moving resources out of coffee into sandwiches. The net result of reallocating $1 of resources from coffee into sandwiches will be a gain in utility of two, i.e. a loss in utility of two attributable to one fewer unit of coffee being produced but a gain in utility of four attributable to one more sandwich being produced. Because society has less coffee, its marginal utility will increase.

Suppose after the reallocation of $1 of resources from industry B to industry A the new ratios are:

$$\frac{MU_A}{MC_A} = \frac{3.5}{1} \quad \text{and} \quad \frac{MU_B}{MC_B} = \frac{2.5}{1}$$

Since

$$\frac{MU_A}{MC_A} \quad \text{still exceeds} \quad \frac{MU_B}{MC_B}$$

the incentive remains for further resource reallocation. Only when

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B}$$

e.g.

$$\frac{MU_A}{MC_A} = \frac{3.15}{1} = \frac{MU_B}{MC_B} = \frac{3.15}{1}$$

will there be no incentive for further reallocation since further reallocation would lead to net loss in total utility.

When

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B}$$

total utility will be maximised and society will experience the highest utility possible given its limited resources. Moreover, since for each good MC = price, each good will be being produced in a technologically efficient (engineering-efficient) manner.

Thus when the *marginal equivalency* condition holds, i.e. when

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B}$$

economic efficiency will prevail in the economy. Each good will be produced using the least amount of resources and the goods that are produced will be produced in those quantities that yield the highest possible level of satisfaction.

Thus a competitive market economy tends to be economically efficient as well as technologically efficient, producing that combination of goods and services most in accordance with consumers' preferences and also producing these goods and services with the smallest quantities of resources. We use the expression 'tends to be' because we cannot state unequivocally that the competitive market economy *will* be economically efficient. There are several reasons for this. First, new goods and services are constantly appearing and, as a result, many firms never reach long-run equilibrium. Rapidly changing technology can also make the attainment of long-run equilibrium well-nigh impossible: as soon as you buy the latest personal computer, a faster and more efficient machine appears on the market. Changes in population, changes in consumer tastes and preferences, changes in the supply and prices of different factors of production all force firms to alter their production and their production techniques. Those are not unwelcome changes from the viewpoint of society – rapid technological change means more goods and services coming on stream, but it also makes the attainment of consumer and firm equilibrium less likely.

There are other areas of economic activity where competitive markets will not tend to produce economic efficiency. Before we turn to those areas let us review our typical competitive firm in both the short and long run in the general equilibrium setting.

6.3 Resource Allocation and Profit-Maximising Behaviour in the Short Run

In a competitive market, the equilibrium price of a commodity and the equilibrium quantity exchanged between buyers and sellers will be determined by the intersection of the demand and supply curves.

In Table 6.1 the market demand and supply schedules for sandwiches in a British city are set out. In Figure 6.2, the demand and supply curves derived from the data in Table 6.1 are shown. The equilibrium price of sandwiches is £1.40, and the equilibrium quantity exchanged per time period is 90 000 sandwiches.

Table 6.1 Market demand for and supply of sandwiches in a British city

Price (£)	Quantity demanded ('000 per week)	Quantity supplied ('000 per week)
1.10	180	0
1.20	150	30
1.30	120	60
1.40	90	90
1.50	60	120
1.60	30	150
1.70	0	180

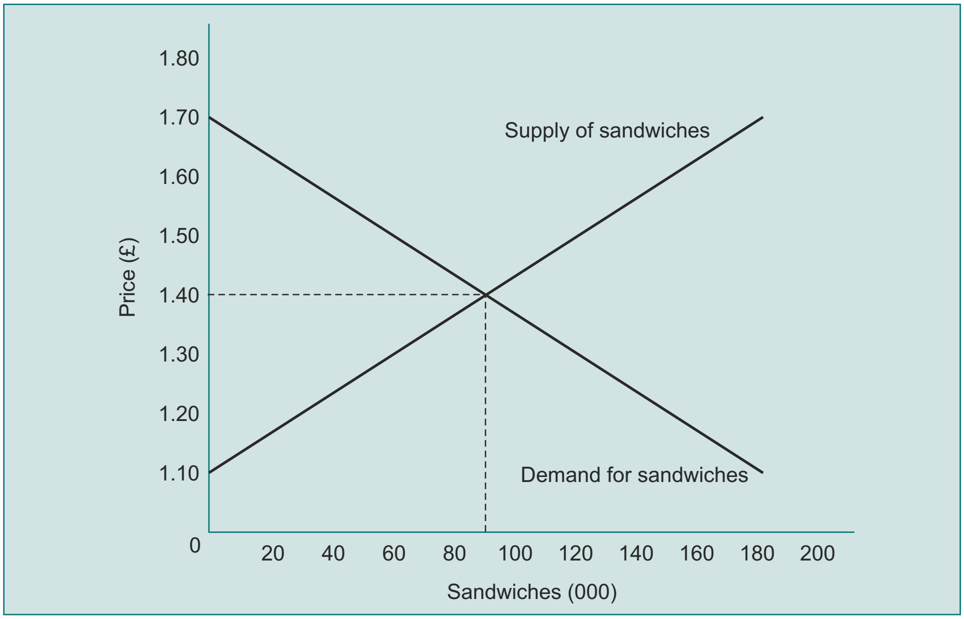


Figure 6.2 Market demand and supply curves for sandwiches in a British city

The market demand curve for a commodity is the summation of individual demand curves and the short-run market supply curve is the summation of individual supply curves (marginal cost curves above the minimum points of the average variable cost curves) of the firms in the market. Each consumer and each firm is a price taker, i.e. we assume that in any large city the price of a given quality of sandwiches is determined by the demand for, and the supply of, sandwiches, and if on some given day one sandwich stall were to close and/or if one family decided to buy or not buy sandwiches, the price of sandwiches would not be affected by such actions.

In the short run, a firm will produce output using both fixed and variable factor inputs. The firm in the short run will not be able to alter the amount of the fixed factor input, but it will be able to employ any number of units of the variable factor at the going price of that factor, i.e. in varying the amount of output in the short run the firm will employ different amounts of the variable factor in cooperation with a given amount of the fixed factor of production. For example, if a sandwich-stall operator wished, on a given day, to increase the number of sandwiches produced, he would be constrained by certain factors, for instance the size of his stall and the size and number of preparation places in his stall; but he could hire more workers and buy more raw materials, such as bread, butter, and fillings.

The profit-maximising firm in the short run will produce output until the marginal cost equals price. Figure 6.3 shows the market for sandwiches and the output produced in the short run by one firm. The firm is a price taker and its actions do not affect the equilibrium price of P_e. To maximise profit, the firm will produce up to the point at which $P_e = MC$. This occurs at a firm output level of $0q_e$.

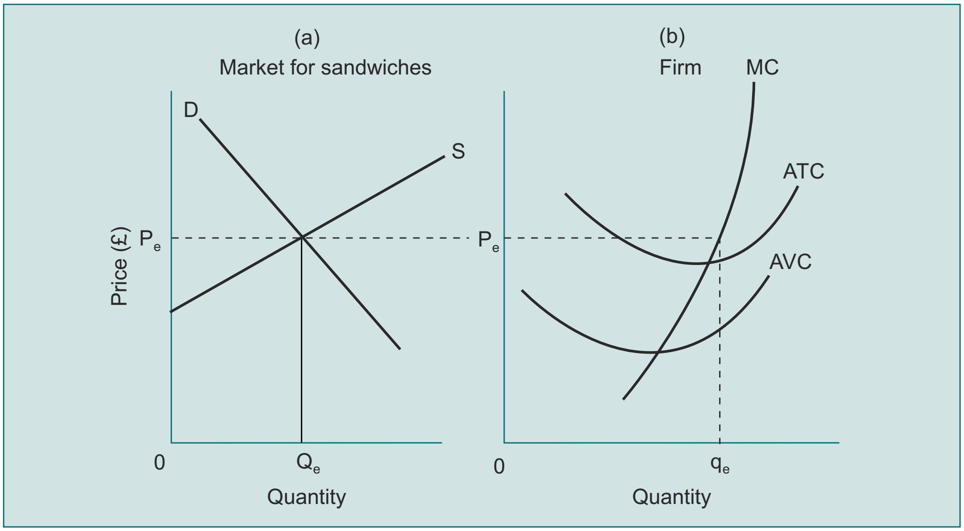


Figure 6.3 Market for sandwiches and a firm in short-run equilibrium

In a market composed of competitive profit-maximising firms, the price of a product will equal short-run marginal cost for each firm. Consequently, the price of a product will reflect the marginal cost to society of producing that product in the short run, i.e. given fixed factor inputs. In Figure 6.3 the number of sandwiches produced and sold in the market will be $0Q_e$, which will consist of the summation of the values of $0q_e$ of all firms in the market. Note that each firm need not produce identical amounts in the short run: each firm, while producing up to the point at which p = MC, does not guarantee identical output since different firms may have different amounts of fixed factors and consequently different MCs or supply curves.

Since commodities are produced by competitive profit-maximising firms, for each commodity in the short run, price will equal marginal cost, and since utility-maximising consumers will allocate their incomes until the ratio of marginal utilities to prices is equal for all commodities, the marginal equivalency condition will prevail and economic efficiency will exist in the economy in the short run. Thus the price of each commodity will reflect the marginal cost to society of that commodity, and societal utility could not be increased by any resource reallocation.

The market for all commodities will be similar to our sandwich market. The price of coffee will equal the marginal cost of coffee, just as the price of sandwiches will equal the marginal cost of sandwiches. Consumers will allocate their incomes on coffee and sandwiches until the ratios of marginal utilities to prices are equal. And the marginal equivalency condition will hold for all commodities, i.e.

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B} = \cdots = \frac{MU_Z}{MC_Z}$$

If, in the short run, something occurs that causes the equilibrium prices and quantities of commodities to alter, the utility-maximising behaviour of consumers and the profit-maximising behaviour of firms will establish new equilibrium prices and quantities. The marginal equivalency condition will be re-established and economic efficiency will again prevail.

Example

If, in a beer/coffee world represented by normally shaped demand and supply curves, a dramatic cold spell of weather causes the demand curve for beer to shift to the left and the demand curve for hot coffee to shift to the right, the price of beer and the quantity exchanged will decrease and the price of hot coffee and quantity exchanged will increase. Variable factors of production will be reallocated from the beer market to the coffee market until prices equal marginal costs for each firm in the markets. At the same time, consumers will reallocate their expenditures until the ratios of marginal utilities to prices for beer and coffee are again equal. Societal utility will once more be at a maximum and any further resource and income reallocation would only decrease total utility.

6.4 Resource Allocation and Profit-Maximising Behaviour in the Long Run

For economic efficiency to prevail in the long run in a market economy, the marginal equivalency condition must continue to hold, i.e.

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B} = \cdots = \frac{MU_Z}{MC_Z}$$

However, in the long run the relevant marginal cost of each commodity is the long-run marginal cost.

In the sandwiches example depicted in Figure 6.3, where diagram (b) is reproduced here as Figure 6.4, the firm is in short-run equilibrium producing output $0q_e$ with price = MC (short run). However, the firm is not in long-run equilibrium, for above-normal profits exist. Consequently the market for sandwiches is not in long-run equilibrium. It follows that the economy is not in long-run equilibrium.

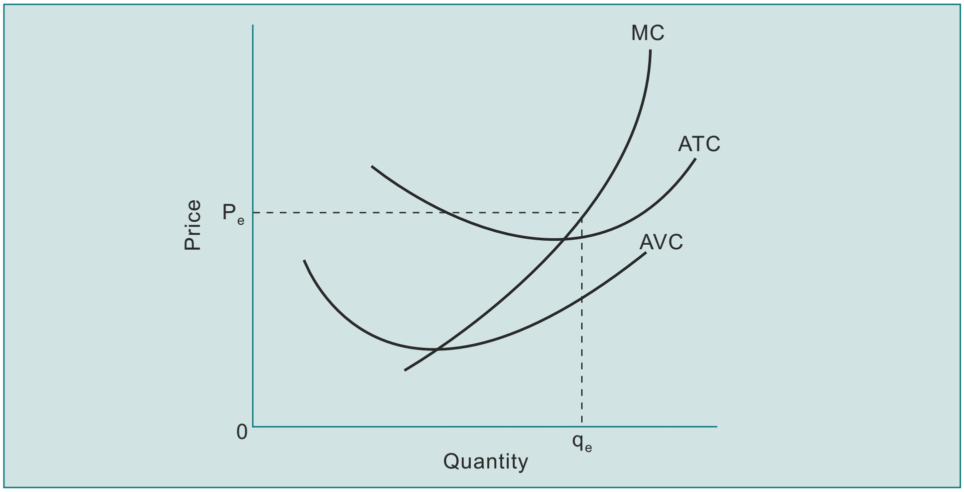


Figure 6.4 A firm in short-run equilibrium

For a firm to be in long-run equilibrium the price of the commodity being produced must equal long-run marginal cost.

Example

Figure 6.5 shows the market for sandwiches and the output produced in the long run by one firm. The firm, as in the short run, is a price taker. In long-run equilibrium the firm's plant size will be that depicted by the ATC curve. Its output will be $0q_e$. Price will equal long-run marginal cost (LMC). Price will also equal MC, minimum ATC and minimum LAC, and there will therefore be the conjunction of curves shown in Figure 6.5(b).

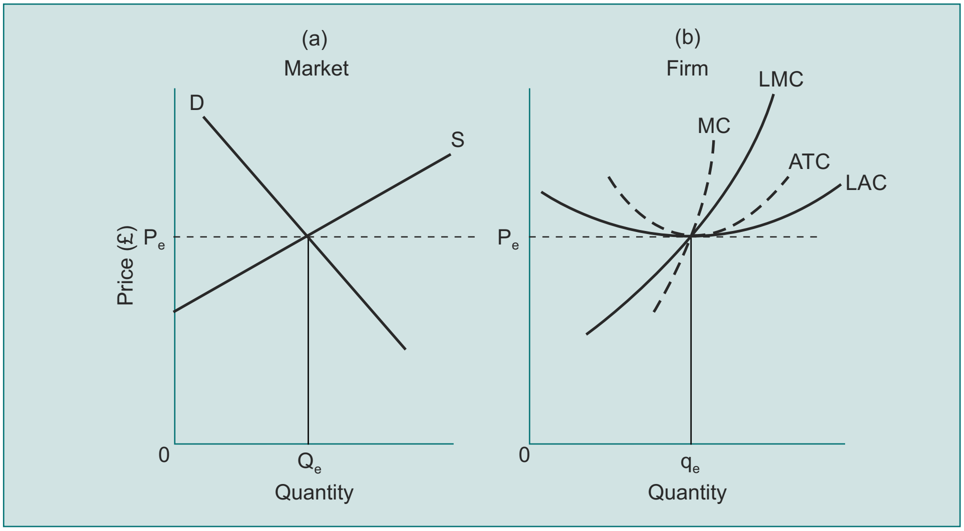


Figure 6.5 **Market for sandwiches and a firm in long-run equilibrium**

For an economy to be in long-run equilibrium, the prices of all commodities produced must equal their long-run marginal costs. And, when all firms in a market economy are in long-run equilibrium, the price of each commodity will reflect the marginal cost of producing that commodity.

For an economy to be in long-run equilibrium, the ratio of marginal utilities to prices for all commodities must be equal, i.e. no consumer in the economy could increase total utility by reallocating income from one commodity to another. Expressed algebraically, when

$$\frac{MU_A}{MC_A(\text{long run})} = \frac{MU_B}{MC_B(\text{long run})} = \cdots = \frac{MU_Z}{MC_Z(\text{long run})}$$

utility will be maximised and each firm will be earning normal profits by producing output where LAC and ATC are minimised, and price equals LMC and MC. Any reallocation of resources would reduce utility.

It follows from the shape of the long-run average cost curve of Figure 6.5 that there is only one optimally sized firm, i.e. the firm with plant size represented by ATC. Figure 6.6 exhibits a firm whose plant size is too small (ATC_1) and one whose plant size is too large (ATC_2). Both firms would suffer losses and in the long run would be forced to adopt plant size ATC or go out of business.

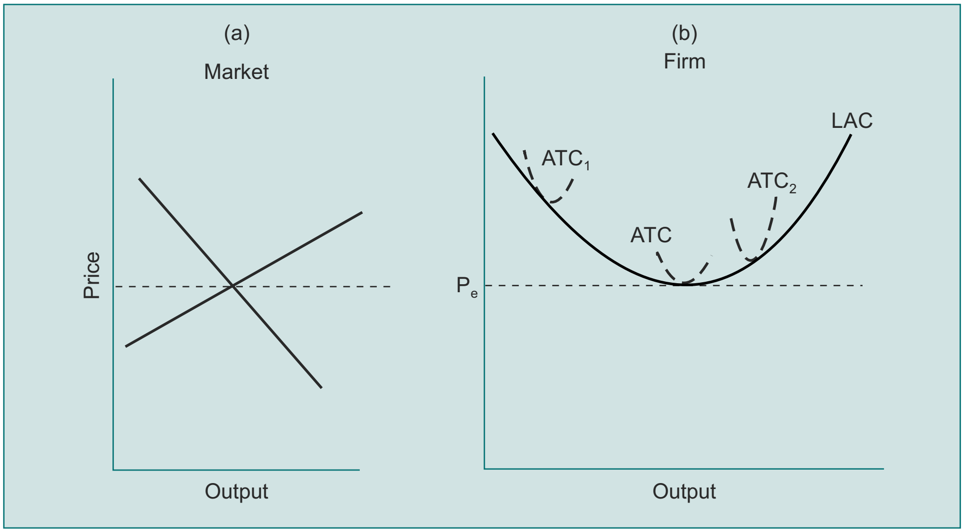


Figure 6.6 Non-optimally sized plants

Only if the long-run average cost curve were 'flat-bottomed' over a range of output could different-sized plants exist in the long run. Figure 6.7 exhibits such a situation. In those circumstances, firms represented by ATC_1 and ATC_2, producing output levels of $0q_{e1}$ and $0q_{e2}$ respectively, are both in short-run and long-run equilibrium.

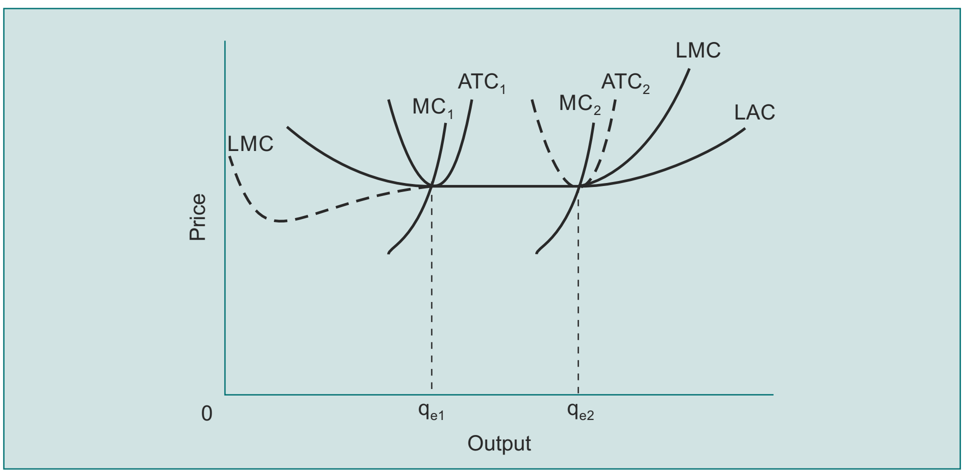


Figure 6.7 Flat-bottomed long-run average cost curve

6.5 The Wonderful World of Adam Smith versus the Real World

In the eighteenth century, Adam Smith formalised the main elements of Modules 1–6 and showed how the utility-seeking behaviour of consumers and profit-maximising behaviour of firms would tend to lead to an efficient solution to the fundamental problem faced by a nation of unlimited wants and scarce resources. He pointed out that no large group of planners, a scarce resource in itself, was needed to reach a solution that would tend to be economically efficient. He pointed out also, however, that market forces would not guarantee economic efficiency under all circumstances, that collective action or planning would be necessary in certain areas of economic activity, and that while market forces in many areas would tend to produce efficient resource allocation, there were many forces at work that would prevent the ideal from being reached. His conclusions still hold: it cannot be stated unequivocally that a competitive market economy will be efficient and will produce that combination of commodities most in accordance with consumers' wishes using the smallest possible quantity of resources. There are several reasons.

New goods and services are constantly appearing in the real world, just as old goods and services are constantly disappearing. This situation can prevent firms from reaching long-run equilibrium. For example, the invention of the car signalled the end of the life of horse-drawn transport just as the steamship spelt the death-knell for the sailing vessel as the principal means of transporting cargos and passengers at sea.

Changes in factor supplies and factor prices can make efficient long-term planning for firms difficult and may prohibit the achievement of long-term equilibrium. For example, volatile oil prices have in recent years led to problems for companies involved in the production of oil. As prices increased steadily through the early 2000s, many oil and gas corporations expanded into unconventional resources where oil is harder and more expensive to extract. Prime examples are the Canadian Oil Sands projects, where the raw product is extremely viscous and impure. These projects require an oil price of around $80 per barrel to be considered viable due to the extraction and processing techniques involved. Investment in such projects grew rapidly as the oil price increased steadily from 2000 onwards. The oil price peaked at $147 per barrel in July 2008 but six months later had fallen to around $45. This left companies with operations in unconventional resources in a situation where they were not covering all their costs – a loss in the short run. By the end of 2009 the oil price had increased again to around $80 per barrel and the projects were once again profitable. In such areas it is nearly impossible for a firm to reach long-run equilibrium.

Furthermore, although firms will attempt to reach long-run equilibrium in order to maximise profit, rapid technological change within an industry can make existing plant and equipment (fixed factors) obsolete and prevent the attainment of long-run equilibrium. The arrival of the jet engine necessitated larger runways at airports, and the increasing demand for air travel led to reorganisation of existing airports and construction of new ones. Because of the time lags involved in recognising such needs, making the necessary plans, and carrying out the work, many airports were completed in time to discover that further technological change resulting from the Boeing 747 jumbo jets made the new airports ineffective and that further change was immediately required in order to attain economic efficiency. The same has happened again recently with the introduction of the Airbus A380, also known as the 'Superjumbo'.

Technological change, by reducing the input requirements for the production of a good, makes possible a larger amount of total output in a society, with a concomitant increase in total utility. However, the more rapid the technological change in an industry, the smaller the probability of firms achieving long-run equilibrium.

Example

The rate of technological change in the electronics industry has been extremely rapid since the 1970s. Since 2000 the mobile phone has decreased steadily in price despite rapidly growing demand. Engineering breakthroughs have increased the capabilities of portable technology to the point that it is possible to carry around a well-powered, constantly connected personal computer in one's pocket. New firms have entered the industry, and old firms unable to adapt have struggled or left the industry altogether. While such rapid and dramatic changes continue to occur, the achievement of long-run equilibrium is impossible.

Changes in consumers' tastes and preferences, changes in population and the movement of people and firms from one geographical area to another cause profit-maximising firms to move resources from one industry to another and from one geographical area to another. The more rapid the changes, the less the chance firms have to reach long-run equilibrium – although profit-maximising behaviour dictates that they attempt to do so.

While some of the reasons for not achieving economic efficiency in the long run are concerned with the changing nature of many factors in the real world, there are other reasons why a market system will not achieve economic efficiency in both the short run and long run. Modules 7, 8 and 9 identify the factors that can prevent the achievement of economic efficiency and analyse collective action designed to overcome such factors.

Learning Summary

The meaning of economic efficiency is captured in the marginal equivalency conditions. In understanding these conditions you have become familiar with the utility-maximising behaviour of consumers and the profit-maximising behaviour of firms and you have seen how market prices are the common factor in establishing the marginal equivalency conditions.

You can distinguish between profit-maximising behaviour of firms in both the short run and the long run and are able to show the implications for resource allocation. You also realise the differences between the real world and the world of economic models and the fact that, while market forces tend to produce an efficient allocation of resources, there is no guarantee that economic efficiency will ever be attained. You are aware of the real-world factors that prevent the achievement of economic efficiency and how such factors affect society's well-being in the long run.

Review Questions

Multiple Choice Questions

6.1 In long-run, purely competitive, industry equilibrium

I. no firms are making either more or less than a 'normal' profit.
II. no firms wish either to enter or leave the industry.
III. price equals minimum long-run average cost.

Which of the following is correct?

A. I only.
B. I and II only.
C. I and III only.
D. I, II and III.

6.2 The main reason that a particularly efficient firm in a purely competitive industry cannot continue indefinitely to make more than normal profits is that

A. other firms will adjust their behaviour.

B. resources will leave the industry.

C. the extra profits will make the firm complacent, and it will lose its efficiency.

D. the government will tax away the extra profits.

6.3 Suppose the furniture industry is perfectly competitive and has been in long-run equilibrium. Then a rise in consumer demand for furniture takes place. Which of the following is the most likely sequence of events?

A. A quick rise in price, which will reduce demand, thereby returning the price to its original level.

B. A quick expansion of output as new firms enter, and then a slow rise of price to a higher equilibrium level.

C. A quick rise of price, increasing profits, and then a slow fall of price as new firms enter.

D. Decreased profits as firms are forced to produce beyond capacity, and then the exit of unprofitable firms.

6.4 Which of the following is correct?

The term 'general equilibrium' for a purely competitive economy implies that:

A. things are generally stable, though a few changes may be taking place.

B. no economic unit in the entire economy wants to change its behaviour.

C. industries as a whole are in equilibrium, but individual firms are not.

D. consumers in the economy are all well off.

6.5 Which of the following is correct?

The short-run industry supply curve in a perfectly competitive industry is:

A. a horizontal line.

B. inelastic in the long run.

C. the summation of firms' marginal cost curves above the AVC curves.

D. determined by the intersection of marginal revenue and marginal cost curves.

6.6 In order to make the most efficient use of a city's trains and buses, how should train and bus fares vary during the day?

A. They should be relatively low during rush hour to reduce cost for the maximum number of people.

B. They should be relatively low during rush hour to transport as many people as possible at that time.

C. They should be relatively high during rush hour to allocate the limited space to the rush hour travellers who value it most highly.

D. They should be the same at all times to avoid travellers altering their schedules because of price differences.

6.7 At going wage rates for military service, a country has been forced to resort to compulsory military service (conscription) to meet personnel requirements. As compared with a system that achieved the required personnel by relying on prices determined in a free market for military service, the conscription system results in which of the following:

A. a lower income for conscripts and a higher income for the civilian population.
B. a higher income for conscripts and a lower income for the civilian population.
C. a higher income for both conscripts and the civilian population.
D. a lower income for both conscripts and the civilian population.

6.8 'If at the going price there is excess demand, competitive firms will shift their supply curves to the right, causing prices to rise and quantity to increase until an equilibrium is reached.' Which of the following is correct with regard to competitive markets?

The above statement is:

A. essentially correct in describing how equilibrium is reached.
B. correct regarding how suppliers behave, but ignores buyers' reactions.
C. incorrect because the supply curve will not shift.
D. incorrect, because it confuses shifting supply curves with shifting demand curves.

6.9 In a private enterprise (free-market) economy, one of the principal

I. effects of competition is to force prices to the lowest level consistent with normal profits.
II. functions of profits is to indicate to the government where wages are too low.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

6.10 If, in a competitive industry, market prices and quantity are in equilibrium,

I. no buyer wants any more at that price.
II. sellers would not be willing to supply more even at a higher price.
III. there is no excess demand or excess supply.

Which of the following is correct?

A. I only.
B. I and III only.
C. I, II and III.
D. Not I or II or III.

6.11 Suppose the supply schedule of, and demand schedule for, babysitters in a town are as follows:

Price ($)	Number demanded	Number supplied
1.00	10	0
2.00	8	6
3.00	7	8
4.00	4	10

Which of the following predictions is *least* likely to come true?

A. If the price starts out at $2, it will probably go higher.

B. The equilibrium price will be somewhere between $2 and $3.

C. If the price starts at $1, the demand schedule will shift down (to the left) because at that price nothing will be supplied.

D. If the price starts out at $4, some suppliers will begin to offer 'bargains'.

Case Study 6.1: The Economics of Army Conscription

This case is an application of the theory of markets to assessing the efficiency of different methods of pricing. Before you tackle it you should understand:

a. equilibrium price
b. marginal benefit
c. opportunity cost
d. marginal cost
e. national welfare
f. economic efficiency.

The case demonstrates that an attempt to provide a service by directing resources using the power of law can have serious implications for the allocation of resources and the efficiency with which the economy operates.

Army Conscription in Britain

Today Britain has a volunteer army in the same sense that British car makers have voluntary employees. The employees do not volunteer their services free, but they have the freedom to apply for jobs at going wage rates. As far as the army is concerned, however, this was not always the case and is not the case in several other countries today.

During the 1950s, Britain had an army that was composed partly of volunteers and partly of conscripted men. The conscripts (mostly able-bodied men under the age of 26 years) were forced to serve in the army for a period of two years at wage rates significantly below what they could have earned in the private sector of the economy. Furthermore, those conscripted could not buy their way out, i.e. they could not pay someone else to take their place. Historically this has not always been the case; in the nineteenth century, conscripts were allowed to send replacements, the price (bribe?) being negotiated between the conscript and the replacement in the open market.

In Britain in the 1950s many people argued that the conscription system was both unfair and economically inefficient. The army, it was argued, should compete for labour like any firm and should pay the going wage rate for the labour it needed. In 1960 the conscription system was abandoned, and since then Britain has had an all-volunteer army.

1 Assume you are a 1950s economist and have been asked to analyse the situation. Prepare a case, paying particular attention to the following points.

1. If the aim is economic efficiency, how large should the army be?
2. How does the volunteer system allocate resources?
3. What is the equilibrium wage rate and number of soldiers under the volunteer system, compared with conscription?
4. Who benefits and who loses by operating a conscription rather than a volunteer system?
5. What effect would the operation of the nineteenth-century system have on the welfare of society?
6. How economically efficient is the volunteer system compared with the conscript system?
7. Does the volunteer system ensure large enough forces during a war?

Module 7

Organisation of Industries

Contents

7.1 Introduction

So far, we have analysed markets in which competition prevails, that is markets in which many consumers compete with each other on the demand side and in which many firms compete with each other on the supply side. The price of a good or service is determined by the forces of demand and supply. In such competitive markets, each buyer and seller is a price taker, such that an individual consumer or firm joining or leaving the market has no significant effect on the price of the commodity in question. This is the world of perfect competition.

There are underlying assumptions in such markets. On the demand side it is assumed that consumers are rational, utility-maximising individuals who know their own tastes and preferences and have perfect information on prices and other characteristics of all goods and services. On the supply side, it is assumed that all firms in a given market produce a homogeneous (identical) good, are profit maximisers, face no restriction in moving into or out of an industry, and have perfect information on the opportunity cost of all resources. You don't have to live long in this world to realise that perfect competition does not describe all markets. There are only a few firms in your country producing cars and there are generally fewer than a few companies producing telephone and postal services.

Table 7.1 summarises the types and characteristics of industries existing in all capitalist economies. In this module we shall analyse the markets in which each of those types of firm operates, and the implications for economic efficiency.

Table 7.1 Organisation of industries

Industry characteristic	Number of firms	Type of product	Barriers to entry	Control over price	Degree of concentration	Example
Perfect competition	Very large	Homogeneous	None	None	Zero	Many agricultural products
Monopolistic/ imperfect competition	Large	Differentiated	None/few	Some	Low	Restaurants, clothing stores
Oligopoly	Small	Homogeneous	Scale	Substantial	High	Cars, chemicals, oil
Monopoly	One	Unique	Scale or legal	Complete	100%	Public utilities

7.2 Perfect Competition

Let us begin with the type of firm we have analysed to date: those in *perfect competition*.

The number of firms in a perfectly competitive industry is so large that each individual firm's actions has a negligible effect on price of the commodity and quantity exchanged in the market. For instance, the wheat industry is composed of a large number of wheat farmers. Although some wheat farmers are larger than others in terms of land cultivated and wheat produced, no individual wheat producer – not even the largest – accounts for a significant proportion of the total quantity of wheat produced in the world. The withdrawal or addition of one wheat farmer would have a negligible impact on the market price of wheat and on the quantity exchanged.

All firms in a perfectly competitive industry produce an identical good. There is no real or imagined difference between each unit of output produced by one firm and each unit of output produced by all other firms. Wheat, for example, is bought and sold in large markets, often national and international in scale, and buyers cannot distinguish wheat by farm of origin.

There are no restrictions on resource owners moving into or out of the industry. Any individual who aspires to become a wheat farmer, for instance, can do so provided he commands the necessary resources to enter the industry.

Firms attempt to maximise profit both in the short run and in the long run. Each individual wheat farmer, whether making a short-run or long-run decision, will hire that combination of factor inputs that produces that level of output calculated to generate maximum profit.

Resource owners have information on the opportunity cost of their resources. It is assumed that the profit-maximising wheat farmer, for example, must know the return he would receive on the resources he commands not only if they were employed in different areas of agriculture but also if employed in any other activity, e.g. the use of the land as a golf course instead of a wheat farm.

Each firm in perfect competition faces a perfectly elastic demand curve, meaning that the firm can sell all the output it wishes to produce at a single price (the market price).

Let us return to our salmon-fishing example from Module 4. We shall assume that the salmon-fishing industry is perfectly competitive and that the equilibrium price of salmon in the market is \$0.70 per kilo. Figure 7.1 illustrates the market for salmon and the demand curve facing an individual supplier of salmon.

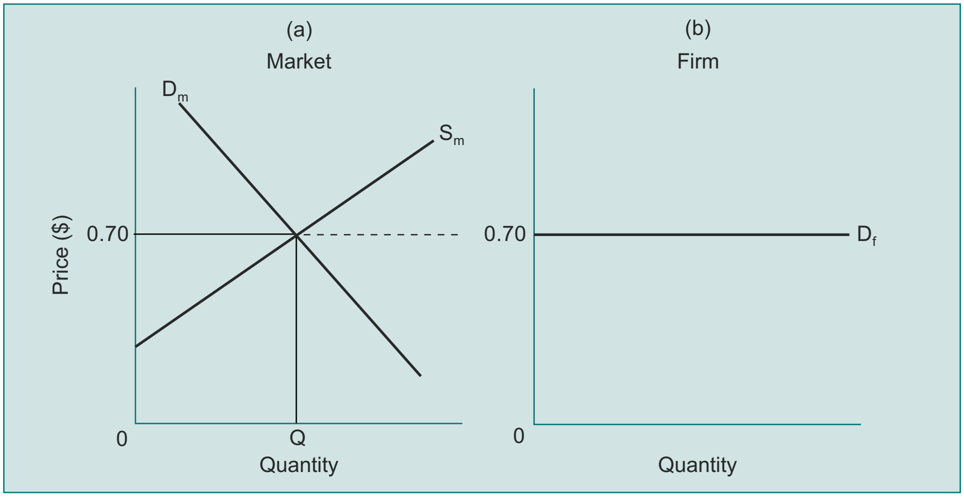


Figure 7.1 **The market for salmon and the demand curve facing a firm**

Figure 7.1(a) shows how, with normally sloped market demand (D_m) and supply (S_m) curves, the equilibrium price of salmon would be \$0.70 per kilo and the equilibrium quantity 0Q kilos. Since the catch of each individual firm is so small in relation to the total market output, \$0.70 per kilo is the price at which it is able to sell all the salmon it catches. This is shown in Figure 7.1(b), in which D_f denotes the individual firm's demand curve. The demand curve faced by the firm in perfect competition is, therefore, perfectly elastic. Since the firm is a price taker, the revenue it receives from each unit of output must be the same, in other words average revenue is constant and equal to price. Since the revenue obtained by the firm from each additional unit of output must also be the same, marginal revenue is also constant and equal to price. Consequently, the demand schedule faced by the firm is both its average revenue schedule and its marginal revenue schedule. Table 7.2 shows such a schedule for our salmon-fishing example.

Table 7.2 **Salmon-fishing firm's total, average and marginal revenue schedules under perfect competition**

Output (Q) (kilos)	Price (P) per kilo (\$)	Total revenue (TR) Q×P(\$)	Average revenue AR = TR/Q = P	Marginal revenue MR = ΔTR/ΔQ
0	0.70	0	–	–
7	0.70	5	0.70	0.70
18	0.70	13	0.70	0.70
36	0.70	25	0.70	0.70
62	0.70	44	0.70	0.70

Output (Q) (kilos)	Price (P) per kilo ($)	Total revenue (TR) Q×P($)	Average revenue AR = TR/Q = P	Marginal revenue MR = ΔTR/ΔQ
99	0.70	69	0.70	0.70
143	0.70	100	0.70	0.70
191	0.70	134	0.70	0.70
235	0.70	165	0.70	0.70
271	0.70	190	0.70	0.70
297	0.70	208	0.70	0.70
314	0.70	220	0.70	0.70
324	0.70	227	0.70	0.70
331	0.70	232	0.70	0.70
335	0.70	234	0.70	0.70
337	0.70	236	0.70	0.70
338	0.70	237	0.70	0.70
339	0.70	237	0.70	0.70
340	0.70	238	0.70	0.70

$$\text{Average Revenue (AR)} = \frac{\text{Total Revenue}}{\text{Quantity}}$$

but since Total Revenue = Quantity × Price, we must have

$$\text{AR} = \frac{\text{Quantity}\times\text{Price}}{\text{Quantity}} = \text{Price}$$

Similarly,

$$\text{Marginal revenue (MR)} = \frac{\Delta\text{TR}}{\Delta\text{Q}}$$

and since Total Revenue = Price × Quantity, we must have

$$\text{TR} = \text{P} \times \text{Q}$$

$$\text{MR} = \frac{\Delta\text{Q}\times\text{P}}{\Delta\text{Q}} = \text{P}$$

Diagrammatically, the demand curve facing a perfectly competitive firm is therefore also its average revenue curve and its marginal revenue curve, as shown in Figure 7.2.

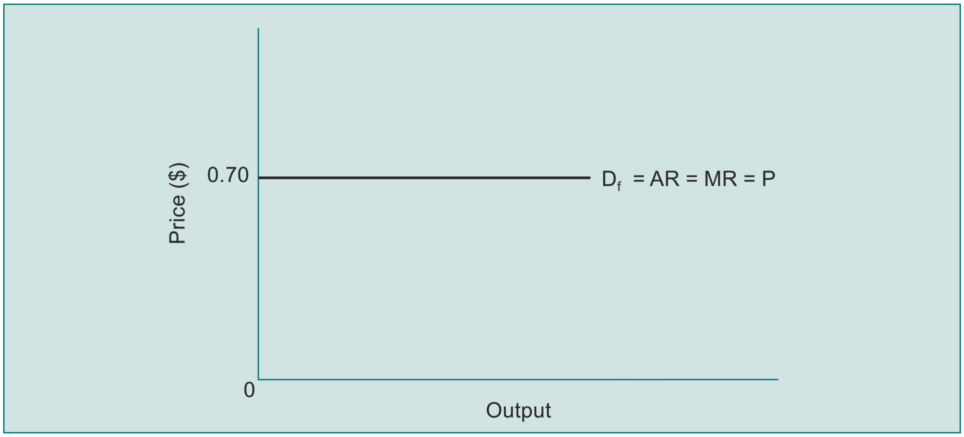


Figure 7.2 **Demand, average revenue and marginal revenue curve for a perfectly competitive firm**

Table 7.3 **Firm's cost data (salmon-fishing firm and perfect competition)**

Output	Fixed cost (FC) ($)	Varia-ble cost (VC) ($)	Total cost (TC) ($)	Average variable cost VC/Q($)	Average total cost TC/Q($)	Marginal cost ΔTC/ΔQ($)
0	40	0	40	0	0	–
7	40	10	50	1.43	7.14	1.43
18	40	20	60	1.11	3.33	0.91
36	40	30	70	0.83	1.94	0.56
62	40	40	80	0.65	1.29	0.38
99	40	50	90	0.51	0.91	0.27
143	40	60	100	0.42	0.70	0.23
191	40	70	110	0.37	0.58	0.21
235	40	80	120	0.34	0.51	0.23
271	40	90	130	0.33	0.48	0.28
297	40	100	140	0.34	0.47	0.38
314	40	110	150	0.35	0.48	0.59
324	40	120	160	0.37	0.49	0.93
331	40	130	170	0.39	0.51	1.43
335	40	140	180	0.42	0.54	2.50
337	40	150	190	0.44	0.56	5.00
338	40	160	200	0.47	0.59	10.00
339	40	170	210	0.50	0.62	10.00
340	40	180	220	0.53	0.65	10.00
340	40	190	230	0.56	0.68	∞
340	40	200	240	0.59	0.71	∞

To the revenue data in Table 7.2 and Figure 7.2 we can add the cost data for our salmon-fishing boat from Module 4, which are reproduced in Table 7.3. The level of output that ensures maximum profit can be found both by comparing total revenue and total cost, and marginal revenue and marginal cost, as illustrated in Figure 7.3. Figure 7.3 shows the firm's total and average revenue and cost curves together.

By use of the data given in Table 7.2 and Table 7.3 as represented diagramatically in Figure 7.3, the largest difference between total revenue and total cost, i.e. maximum profit, occurs at an output level of 314 kilos. By definition, at this output level marginal revenue equals marginal cost (point E in Figure 7.3). At this output level, however, the firm is earning above-normal profit since price (average revenue) exceeds average cost. The price at which each kilo of salmon is sold is measured by XE, whereas the average cost of producing 314 kilos of salmon is XC. Hence the firm is earning above-normal profit of EC on each kilo of salmon sold. The total amount of above-normal profit obtained by producing 314 kilos of salmon is denoted by the unshaded rectangular area AECB and is equal to the quantity of salmon sold (0X = AE) × profit per unit (EC).

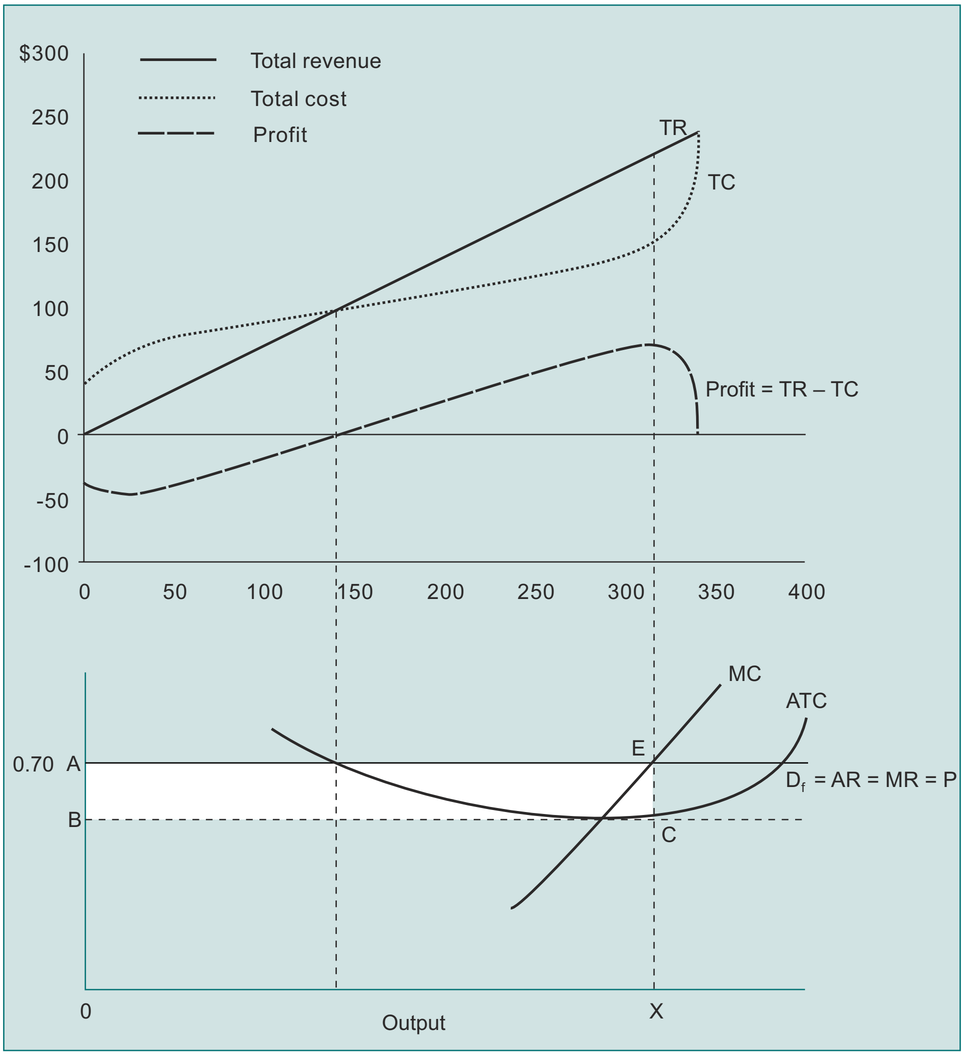


Figure 7.3 **Revenue, cost and profit**

Since the cost curves include normal returns to factors, if above-normal profit exists in an industry, resources will flow into that industry. Conversely, if less than normal profit exists in an industry, resources will flow out of that industry. In a market economy, the existence of different rates of profit in different industries acts as an incentive for resources to transfer out of low and declining profit sectors and into high-profit sectors. In the past three decades, for example, because of high profit earned in the electronics industry, many firms have expanded their output and many new firms have entered that industry. Conversely, after the initially high profit earned in the early days of the nineteenth-century US 'gold rush', resources moved out of the industry as profit fell.

Given above-normal or less than normal profit in an industry, resources will flow into (out of) that industry in the expectation that such profit will continue to be earned by resources in that industry. As resources move in (out), firms as a whole will be prepared to supply a greater (lower) quantity at each price. *Cet. par.*, the equilibrium price of the good will fall (rise) and the equilibrium quantity increase (decrease). This movement of resources will continue until the forces causing it are neutralised, at which point no resources can earn a higher return by moving from one industry to another. When this occurs, each firm will be in long-run equilibrium, and price will equal long-run marginal cost.

Suppose the market for salmon is initially represented by a demand curve DD and a supply curve SS, and that the equilibrium price and equilibrium quantity are 0P and 0Q respectively, as shown in Figure 7.4. Suppose also that, at the equilibrium price of 0P, above-normal profit is earned by firms in the industry. Resources will then be attracted into the salmon industry in response to the level of above-normal profit. This can be illustrated diagrammatically by a shift of the supply curve from SS to S_1S_1. As a result, equilibrium price falls from 0P to $0P_1$, and equilibrium quantity increases from 0Q to $0Q_1$. Given that firms are price takers, as the equilibrium price falls so too does average revenue (price) and marginal revenue for each firm in the industry. As shown in Figure 7.4, equilibrium price falls from 0P to $0P_1$; and as a result the price facing the individual firm (given a horizontal demand curve where AR = MR) also falls from 0P to $0P_1$.

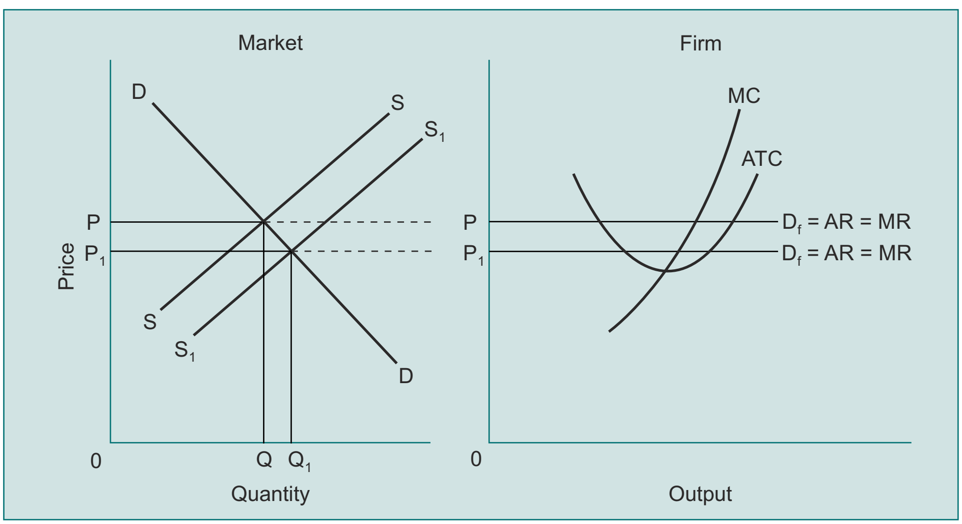


Figure 7.4 Effect on equilibrium of a good when resources move into an industry

As long as above-normal profit is being earned in the salmon industry, resources will continue to flow into the industry. This incentive to move resources into the industry will cease when, for each individual firm, price equals average cost, at which point normal returns to resources employed in the industry will exist. This situation is shown in Figure 7.5, where resources will stop flowing into the industry when price falls to $0P_2$. The firm maximises profit by equating long-run marginal cost with marginal revenue and producing output 0Y. At this output level, price (average revenue) $0P_2$ is equal to minimum long-run average cost and also to long-run marginal cost; only normal profit is earned. Hence the firm would be in long-run equilibrium by producing 0Y kilos of salmon.

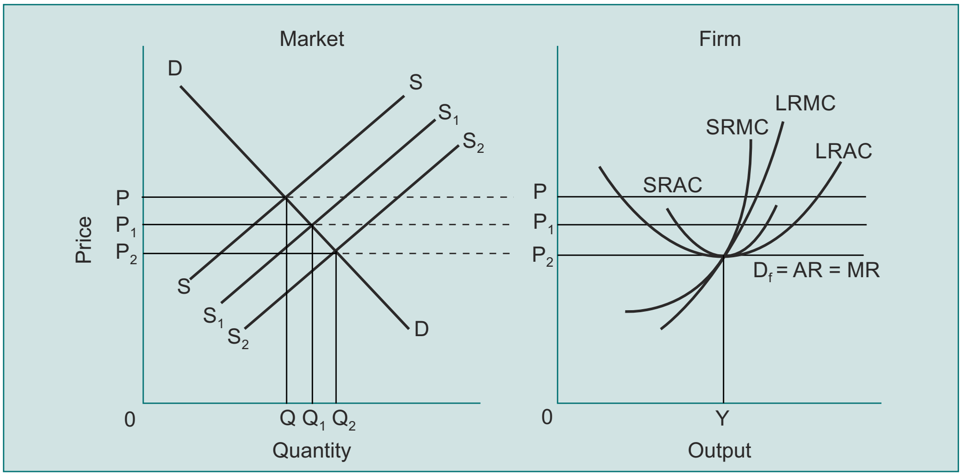


Figure 7.5 Long-run equilibrium

Given utility-maximising consumers, an economy containing perfectly competitive industries would achieve economic efficiency. Since utility-maximising behaviour requires that the ratio of $\frac{\text{Marginal Utility}}{\text{Price}}$ is equal for all goods and since profit-maximising behaviour ensures that price = marginal cost for all goods, economic efficiency would be achieved since the ratio of $\frac{\text{Marginal Utility}}{\text{Marginal Cost}}$ for all goods would be equal.

When consumers achieve equilibrium, i.e. they maximise utility from the allocation of their budget on a range of goods and services, $\frac{MU}{P}$ is the same for every good they buy. For two goods, salmon (S) and coffee (C) for instance,

$$\frac{MU_S}{P_S} = \frac{MU_C}{P_C}$$

Also, given perfectly competitive conditions in the salmon and coffee industries, in equilibrium $P_S = MC_S$ and $P_C = MC_C$. Therefore

$$\frac{MU_S}{MC_S} = \frac{MU_C}{MC_C}$$

Let us now consider the opposite of perfect competition: monopoly.

7.3 Monopoly

A *monopolist* is a producer that supplies the complete market for a good or service.

For much of the twentieth century in the USA, aluminium production from virgin ore was controlled by only one firm, the Aluminum Company of America (Alcoa), and the aircraft industry, the major user of aluminium, *had* to purchase its metal from Alcoa. In many countries the postal and telephone services are national monopolies. In local, as opposed to national markets, monopolies may also exist: in small, isolated towns the only petrol station and sole hotel are local monopolies. And in most major cities, water and sewerage services, rubbish collection, and electricity supply are monopolies.

Barriers to entry may legally prohibit new firms from entering a monopolistic market, and/or there may be economic disincentives that discourage potential entrants. Barriers may take the form of patents: for many years the Xerox Corporation had a patent of the photocopying machine, which guaranteed the firm exclusive use of the photocopying process embodied in each Xerox machine; leading pharmaceutical companies also have patents for many of the drugs they produce. A disincentive to enter a market may exist because of market conditions: for example, in a small isolated town there may be only one petrol station, which supplies petrol to all car owners in the town, and there would be a disincentive for any other individual to open another petrol station in the town unless he could quickly capture the whole local market for petrol and displace the existing firm.

Since a monopolist is the sole producer of a good or service, the monopolist's short-run marginal cost curve is the market supply curve. Also, the market demand curve for the good or service is the demand curve faced by the monopolist. Suppose a pharmaceutical company had a patent on a certain tranquilliser, guaranteeing the company sole rights of production. The market demand curve for the tranquilliser – relating the quantities that all buyers would wish to purchase at different prices – would be the demand curve faced by the company. Similarly, the company's short-run marginal cost curve – relating the change in total cost required to produce different levels of output – would be the industry supply curve.

Since the monopolist faces a downward sloping demand curve, there will be a divergence between average revenue (price) and marginal revenue.

Example

Suppose a pharmaceutical company faces a demand schedule as illustrated by columns (1) and (2) in Table 7.4.

Table 7.4 Demand and Total, Average and Marginal Revenue under Monopoly

(1) Price ($)	(2) Quantity demanded (kilos per week)	(3) Total revenue ($)	(4) Average revenue ($)	(5) Marginal revenue per kilo ($)
18	0	0	–	–
16	100	1 600	16	16
14	200	2 800	14	12
12	300	3 600	12	8
10	400	4 000	10	4
8	500	4 000	8	0
6	600	3 600	6	−4
4	700	2 800	4	−8
2	800	1 600	2	−12

Marginal revenue is less than average revenue for all levels of demand (*see* columns (4) and (5)). When average revenue decreases, marginal revenue decreases at a greater rate. When average revenue increases, marginal revenue increases at a faster rate.

Suppose the pharmaceutical firm were selling 300 kilos of a drug per week at a price of $12 per kilo. Total revenue would be $3600. In order to sell 400 kilos per week, price would have to be $10 per kilo. Total revenue from 400 kilos per week would be $4000. The marginal revenue per unit as sales increase from 300 to 400 kilos is

$$\frac{\$4000-\$3600}{100}=\$4$$

while the price per kilo of 400 kilos is $10, i.e. marginal revenue is less than average revenue. The reason for this divergence between marginal and average revenues lies in the fact that as sales increase from 300 to 400 kilos per week, the price becomes $10 per kilo not only for the extra 100 kilos but also for the remaining 300 kilos. By selling the extra 100 kilos, the monopolist receives additional revenue of $1000 (i.e. 100 × $10); however, since the price of the remaining 300 kilos falls from $12 to $10 each, i.e. by $2 per kilo, the net addition to total revenue, i.e. marginal revenue, is:

$1000 − (300 × $2) = $400, or $4 per kilo.

The average revenue (AR) and marginal revenue (MR) curves for the monopolist based on the curves in Table 7.4 are shown in Figure 7.6. The MR curve is twice as steep as the AR curve.

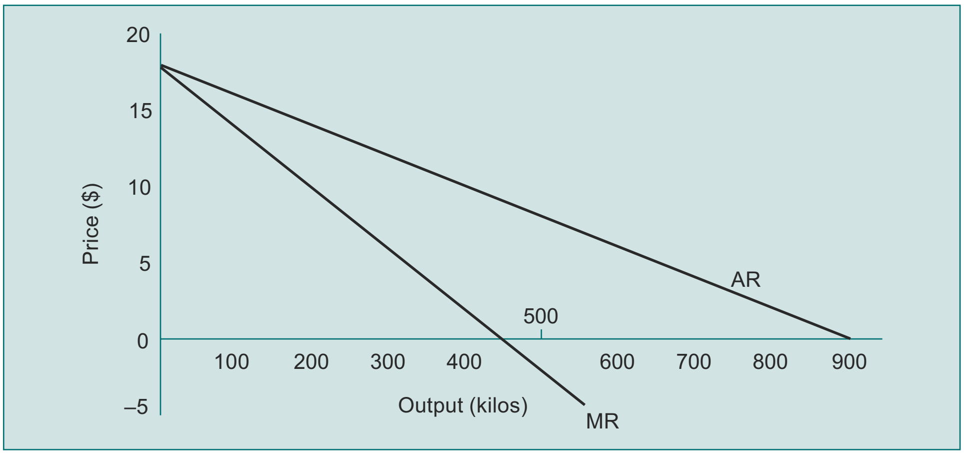


Figure 7.6 Average and marginal revenue curves under monopoly

A monopolist will adopt the same decision rule for choosing the maximum profit output level as would a perfectly competitive firm, for it will produce up to the level of output at which marginal cost equals marginal revenue. Suppose the pharmaceutical firm is faced with short-run average total cost and short-run marginal cost curves similar to a perfectly competitive firm. The short-run profit-maximisation position is shown in Figure 7.7, in which the monopolist's average revenue and marginal revenue curves are superimposed on the average total cost and marginal cost curves.

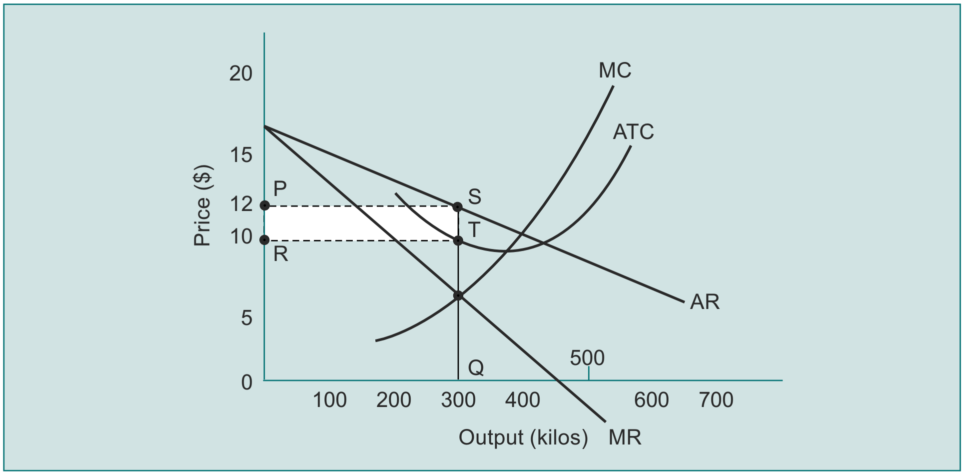


Figure 7.7 Short-run profit maximisation for a monopolist

The output level at which marginal cost equals marginal revenue is 300 kilos per week. Knowing the maximum profit output level, the monopolist will charge the price at which this output will sell in the market. From the demand curve, $12 per kilo is the price consumers would be willing to pay for 300 kilos per week. From the average cost curve, the unit cost of producing 300 kilos per week is denoted by QT, i.e. $10 per kilo. Thus in the short run, the price per unit received by the monopolist exceeds the average total cost of production by $2 per kilo, so that the average profit per unit is $2.

Total revenue from selling 300 kilos per week is given diagrammatically by 0Q (= 300 kilos) × 0P (= $12 per kilo), and is shown by the area 0QSP = $3600. Total cost incurred by producing 300 kilos per week is given by 0Q (= 300 kilos) × QT (= $10 per kilo), i.e. the area 0QTR = $3000. Thus the monopolist's profit is denoted by the area PSTR = $600 = profit per unit ($2) × output (300 kilos). This is the maximum profit the monopolist can earn, given the cost and revenue conditions under which he operates. When the monopolistic firm produces 300 kilos per week it will be in equilibrium.

Although above-normal profit is earned by a monopolist in short-run equilibrium, there will be no tendency for profit to be competed away in the long run by the entry of new firms to the market. Since the pharmaceutical firm's production of the tranquilliser is protected by a patent, no other firm may manufacture that drug and so no new firms can enter the market and compete against the existing firm. Unlike perfect competition, in the long run the monopolist's price, and hence above-normal profit, cannot be affected by the entry of new firms into the industry.

In the long run, the monopolist may expand or contract its output by adjusting its plant size. Whether it expands or contracts output depends on the relation between long-run marginal cost (LRMC) and marginal revenue (MR). If LRMC is less than MR, the monopolist will increase output; conversely, if LRMC exceeds MR, the monopolist will decrease output. In either case, above-normal profit will be higher in the long run than in the short run.

Although the pharmaceutical firm in our example is in short-run equilibrium when producing 300 kilos per week, it may not be in long-run equilibrium. If, for example, the marginal revenue gained by selling more than 300 kilos per week exceeded the long-run marginal cost of producing more than 300 kilos per week, the firm could increase profit by increasing output. This would involve adjusting its plant size to produce more than 300 kilos per week. If, however, by producing 300 kilos per week, long-run marginal cost exceeded marginal revenue, the firm again would not be in a long-run profit-maximising position. The firm could increase profit by decreasing output and adjusting its plant size to produce output of less than 300 kilos per week.

The monopolist will maximise profit in the long run, thus being in long-run equilibrium, when it produces that level of output at which long-run marginal cost equals marginal revenue. Since, under monopoly, marginal revenue is less than average revenue, price will exceed marginal cost, i.e. $P_m > MC_m$. Given that marginal revenue is less than average revenue under monopoly, as shown in Figure 7.7 equating marginal cost with marginal revenue (either in the short run or in the long run), the profit-maximising monopolist will produce an output level where price is greater than marginal cost.

When monopoly exists in an economy, economic efficiency will not prevail. The marginal equivalency conditions for economic efficiency specify that the ratios of marginal utility to marginal cost, $\frac{\mathrm{MU}}{\mathrm{MC}}$ must be equal for all goods and service. This will be achieved when the ratios of $\frac{\mathrm{MU}}{\mathrm{P}}$ are equal and when P = MC for all goods and services. Under monopoly, however, price is greater than marginal cost. Thus the marginal equivalency conditions cannot be met, since marginal cost cannot be substituted for price in the equivalency equation. The ratio of $\frac{\mathrm{MU}}{\mathrm{MC}}$ for the monopolist's product will be greater than the ratio of $\frac{\mathrm{MU}}{\mathrm{MC}}$ for competitively produced goods since, for each good produced under monopolistic conditions, P > MC. Consequently it would be possible to increase society's utility by reallocating resources so that more of the monopolist's good and fewer of other goods were produced.

Example

Suppose there are two goods, A and B, in an economy and that A is produced under monopoly conditions while B is produced under competitive conditions. Suppose also that for good A,

$$\frac{\mathrm{MU_A}}{\mathrm{MC_A}} = \frac{4}{2} = 2$$

and for good B,

$$\frac{\mathrm{MU_B}}{\mathrm{MC_B}} = \frac{3}{2} = 1.5$$

The marginal equivalency conditions are not met because the price of good A would be above its marginal cost. Consumers could become better off if resources were shifted from the production of B to the production of A. If, for example, \$2 worth of resources were shifted from the production of B to the production of A, one less unit of B would be produced since MC_B = \$2. As a result, consumers will lose three units of utility from the decrease in production of B. Because, however, MC_A = \$2, the resources transferred from B will result in one more unit of A being produced. This will increase consumer utility by four units. Hence consumers would receive a net gain of one unit of utility be reallocating resources in this way. Consumers would become no better off only when

$$\frac{\mathrm{MU_A}}{\mathrm{MC_A}} = \frac{\mathrm{MU_B}}{\mathrm{MC_B}}$$

7.3.1 Economies of Scale

Despite economic inefficiency caused by monopoly, it may be in society's interest to have only one producer of a good or service if *economies of scale* exist. A firm experiences economies of scale when, as output expands, average costs decline. Economies of scale are illustrated in Figure 7.8.

Curves ATC_1, ATC_2, ATC_3 and ATC_4 in Figure 7.8 denote the average total costs incurred at increasing scales of operation. At larger sizes, minimum average total costs are lower, i.e. $C_1 > C_2 > C_3 > C_4$. The most common sources of economies of scale are specialisation of labour, the lower per unit costs associated with the increased size of capital equipment, and the bulk buying of raw materials.

Two hundred years ago, Adam Smith pointed out the significance of the division (or specialisation) of labour, in firms as they increase their level of output. In such firms employees can concentrate on applying specialist skills to particular jobs, in comparison to firms with relatively lower levels of output where employees have to undertake a variety of jobs.

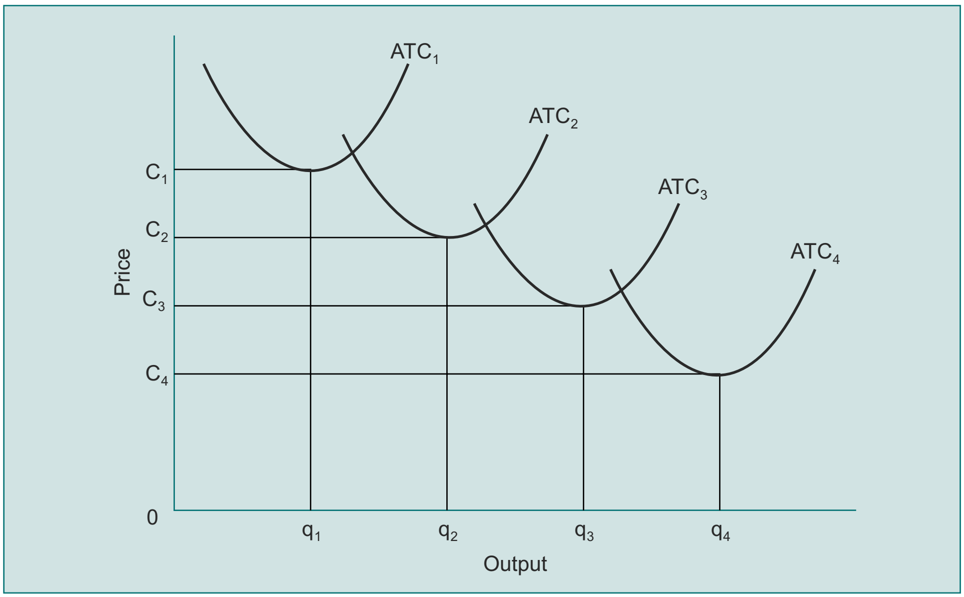


Figure 7.8 Economies of scale

An oil tanker, a piece of capital equipment in the oil transportation industry, provides an example of economies of scale through lower unit costs. The total cost of a large tanker is obviously greater than the total cost of a small tanker, and total costs will rise with tanker size in proportion to the increased surface area of the hull, measured in square metres. But the unit costs of tankers are also determined by their capacity, i.e. the volume of oil they are capable of holding, measured in cubic metres. Since capacity will increase by a greater amount than the surface area of the hull as tanker size increases, average costs will fall. In this way, if the production of a good is characterised by economies of scale, one large firm would in general be able to produce a given level of output at a lower per unit cost than many small firms, and so that fewer resources would be required for a given level of output.

Consider again the economies of scale illustrated in Figure 7.8. Four firms, each operating at ATC_1 could each produce $0q_1$ units of output, resulting in total output of $0q_4$; one large firm, operating on ATC_4, could itself produce $0q_4$. The latter would use fewer resources to produce $0q_4$, since the per unit cost to the large firm would be $0C_4$ whereas each small firm could produce its share of the total output $0q_4$ at the higher per unit cost of $0C_1$.

It is also possible that an 'enforced' competitive industry would not only have higher per unit costs but could also yield a lower total output than the monopoly. Consider Figure 7.9 and suppose that a government legislates that no firm can produce more than $0q_1$. All firms will produce $0q_1$ at minimum cost for that level of output, having plant size ATC_1. Industry output would be $0Q_1$; aggregate demand would equal aggregate supply; and the industry would be in short-run equilibrium. Each firm would be operating at the minimum point of its ATC_1; the 'competitive' price would be P_c and each firm would be in short-run equilibrium, i.e. $MC = MR = P_c$.

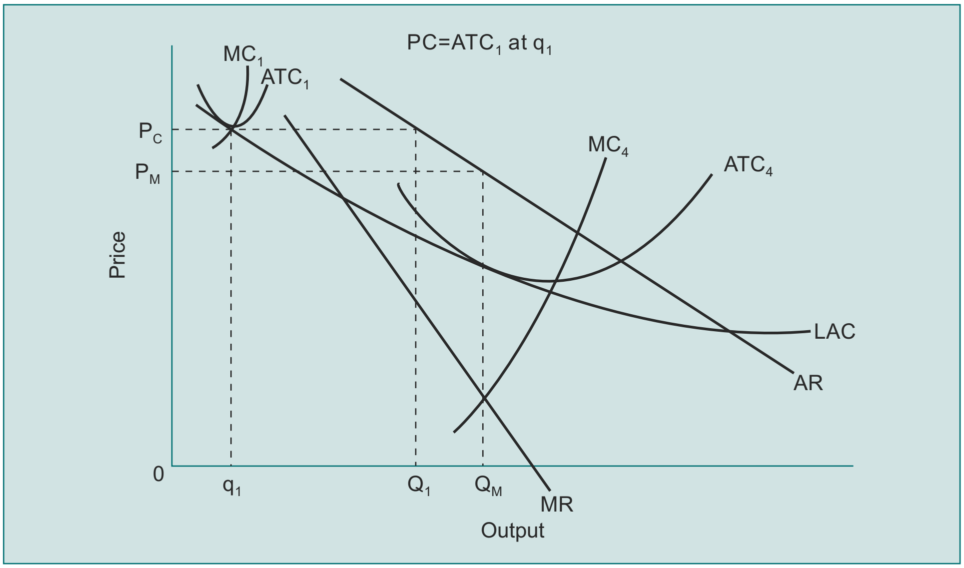


Figure 7.9 Monopoly vs competition

The monopolist, on the other hand, would produce output $0Q_M$ ($> 0Q_1$), i.e. the output where $MR = MC_4$ and charge the profit-maximising price of P_M ($<P_c$). Because $P_M > MC$, the marginal equivalency condition will not hold and we shall have an inefficient allocation of resources. Despite this fact, society is obviously better off under the monopolistic situation compared with the enforced competitive situation: there is more output of the good in question, the price is lower, and fewer resources are used in the production of each unit. Thus society wants the monopoly with the economies of scale. Later in this module we shall see how governments attempt to have the monopoly and bring about the marginal equivalency conditions.

When significant economies of scale exist in an industry then, in the absence of government interference, there will be a tendency for monopoly to arise. The first firm to move down the long-run average cost curve and command a larger share of the market will have a cost advantage over smaller firms and, given its excess profit, it can price newcomers out of the market. If a determined competitor with lots of resources managed to hang in there, however, and threatened the monopoly, the stage would become set for either a battle to the death (if the market is not large enough for two) or a 'getting together' by the companies to act together as a monopolist and 'split' the market. Such behaviour by companies is illegal in many capitalist countries.

Between perfect competition and monopoly, we have imperfect competition and oligopoly. First of these we shall analyse imperfect competition.

7.4 Imperfect Competition

An imperfectly competitive industry consists of a large number of firms, each facing a downward sloping demand curve for its goods or services. In these circumstances, the demand curve facing the firm is not completely price-elastic as it is in the perfectly competitive situation. Under *imperfect competition*, fewer firms exist than in perfect competition.

Firms have a degree of control over price because of several factors. There may be real or imagined differences between their goods or services and those of competitors, for instance. Such imagined differences may be due to advertising. In a typical city, for example, there are dozens of restaurants. Each restaurant usually incurs some expenditure on advertising its attractive qualities and the specialities it serves. The purpose of such advertising is partly to bring to the public's attention the existence of the restaurant and its food, and partly to persuade consumers about the special features or qualities offered by the restaurant. Such advertising builds up a degree of customer loyalty, which means that if a firm were to raise its price, it would not lose all its sales to competitors.

Equally there may also be elements of local monopoly that, because of consumers' valuation of time and convenience, allow a firm some control over its price. Corner grocery stores take advantage of their local monopoly position and often charge prices for homogeneous products that are higher than prices 'in town'. For example, if you were to collect data on the prices of a given can of soup, a given tube of toothpaste or a given packet of biscuits within a city's boundaries you would discover a significant variance in price for each of these products around their mean price. This does not imply that economic theory is invalid; on the contrary that is what one would expect because if a household runs out of toothpaste a visit is paid to the nearest shop to buy a tube. The value of time and the small amount of money involved does not make it worthwhile to shop carefully for this one item. If the item were highly priced, however, and took a large share of a household's budget, more careful shopping would typically result and the price variance of such items would be considerably less than those of toothpaste, soup and biscuits. If you don't believe this, try it: the variance on low-priced items in most towns and cities is very large.

Like firms in a perfectly competitive market, firms may enter or leave an imperfectly competitive industry in response to the level of profit being earned. The actions of one firm will affect the demand curve facing other firms in an industry composed of imperfectly competitive firms.

The greater the real or imagined difference between each firm's good or service, or the greater the degree of local monopoly a firm has, the more inelastic its demand curve will be. Although steaks sold in different restaurants are made to appear different in the eyes of consumers, they are nevertheless reasonably close substitutes. If one restaurant decided to increase the price of its steak, it would lose some customers to other restaurants. However, some people would still remain loyal and go to the more expensive restaurant because of real or imagined differences in the steaks supplied. The degree of price elasticity of demand would depend upon consumers' judgement about the substitutability of the steaks. If most consumers thought that the steaks sold in different restaurants were virtually identical, each restaurant would face a highly elastic demand curve, indicating that a given percentage change in price would result in a much greater percentage change in quantity demanded. In contrast, however, if one restaurant managed to convince consumers that its steaks really were unique and a large number of consumers remained loyal to the restaurant despite a price increase, then the restaurant would face a more inelastic demand curve.

Since each firm in imperfect competition faces a downward sloping demand curve, average revenue and marginal revenue will diverge as they do in monopoly situations, although the divergence will be much less. The decision rule for maximising profit for a competitive firm and for a monopolist applies also to the firm in imperfect competition: output should be expanded or contracted to the point at which marginal cost equals marginal revenue. Suppose a typical steak restaurant is faced with the cost and revenue curves illustrated in Figure 7.10. The restaurant would sell $0q_e$ steaks at a price of $0p_e$ in order to maximise profit. Because price is greater than average total costs, higher than normal profit is being earned, denoted by the unshaded area p_eXYZ.

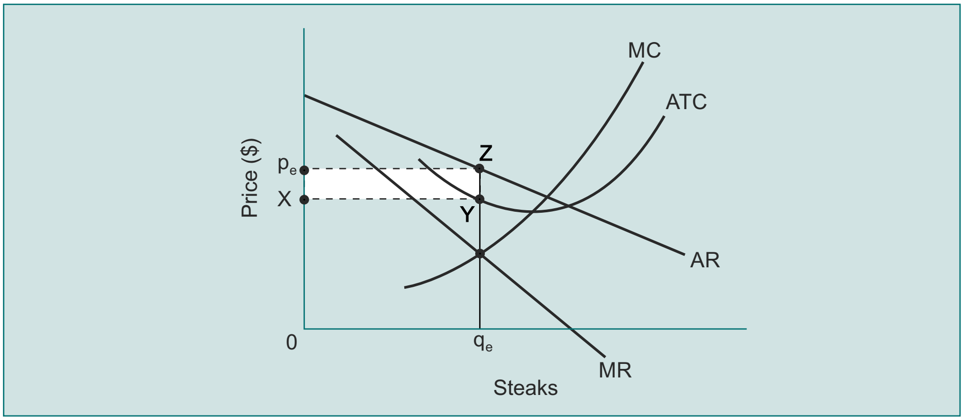


Figure 7.10 Short-run profit maximisation for a firm in imperfect competition

The existence of above-normal profit will provide an incentive for new firms to move into the industry. As new firms move in, assuming factor prices remain constant, the demand curve of existing firms will shift to the left until, in long-run equilibrium, the firm's demand curve is tangential to its average cost curve and normal profit is earned. If factor prices do in fact change as new firms enter the industry, both the firm's demand curve and its cost curve will shift until again long-run equilibrium is established when normal profit is earned.

Example

Imagine that, in response to the profit being made by existing steak restaurants, new firms decide to open up restaurants in the city. As shown in Figure 7.11, in the long run the entry of new firms will continue until our firm faces a demand curve AR_1, representing a shift to the left from AR in Figure 7.10. This is because our restaurant loses some customers to the new competitors. With a demand curve denoted by AR_1, the profit-maximising output is $0q_e$ and the price is $0p_e$, assuming factor prices remain constant as new firms enter. Graphically, the average revenue curve is a tangent to the average total cost curve at point E, and the price per steak ($0p_e$) is just equal to its average total cost of production (p_e). Thus normal profit is earned, and there is no further tendency for any more steak restaurants to be opened up.

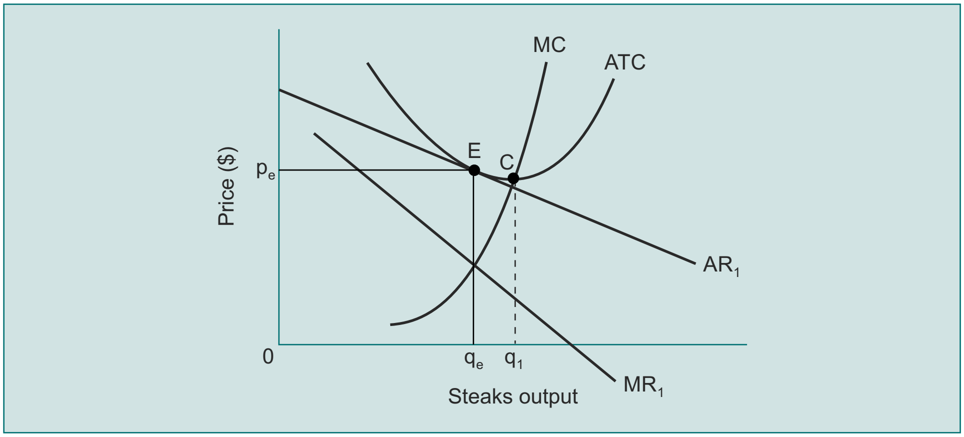


Figure 7.11 Imperfectly competitive firms in long-run equilibrium

When a firm in imperfect competition is in long-run equilibrium, the price of its product will be greater than marginal cost. Thus economic inefficiency results because the ratios of marginal utility to marginal cost in different industries will not be equal. The achievement of economic efficiency would require that price be set equal to marginal cost by imperfectly competitive firms, but since this would result in a non-profit-maximising output, given rational firm behaviour, it will not occur.

In long-run equilibrium, a firm in imperfect competition produces an output level above the minimum point on the average total cost curve. It would be, for instance, possible for our steak restaurant to produce an output at a lower unit cost than occurs in long-run equilibrium – say at point C in Figure 7.11 on the ATC curve, compared with point E. But this would result in a loss being incurred by the restaurant because, at the associated output level of $0q_1$, average total cost exceeds average revenue. Hence there is no incentive for the restaurant to expand beyond the output level $0q_e$.

Thus one implication of an economy having imperfectly competitive firms is that each has 'spare capacity', i.e. each firm is not operating at the minimum point on its average cost curve. This economic inefficiency has to be offset against the product differentiation with which such firms provide society.

7.5 Oligopoly

An oligopolistic industry is one in which a small number of firms produce the bulk of the industry's output. Each firm competes with the others, and in such an industry each firm is interdependent: one firm's sales depends upon the price it charges and the prices charged by its few competitors.

Many of the goods and services produced by oligopolistic industries are to all intents and purposes homogeneous although some of the firms involved might deny such a claim. If you wish to fly the Atlantic from London to New York, there is a limited choice of carriers. If you elect to go business class, there is not much to choose at present among American Airlines, British Airways and Virgin: the seats are remarkably similar, and so is the choice of alcohol, the meals, the first-run films and the flight time.

The problem faced by an oligopolist is one of pricing. What will happen if one firm increases its price? If none of the others follow with a matching price increase, the firm that raised its price will suffer a significant fall in sales. Conversely, if the same firm lowers its price, all the others will follow and no significant gain in sales or market share will be forthcoming. Such a situation, if it exists, produces an interesting demand curve for any one firm's output, as is seen in Figure 7.12.

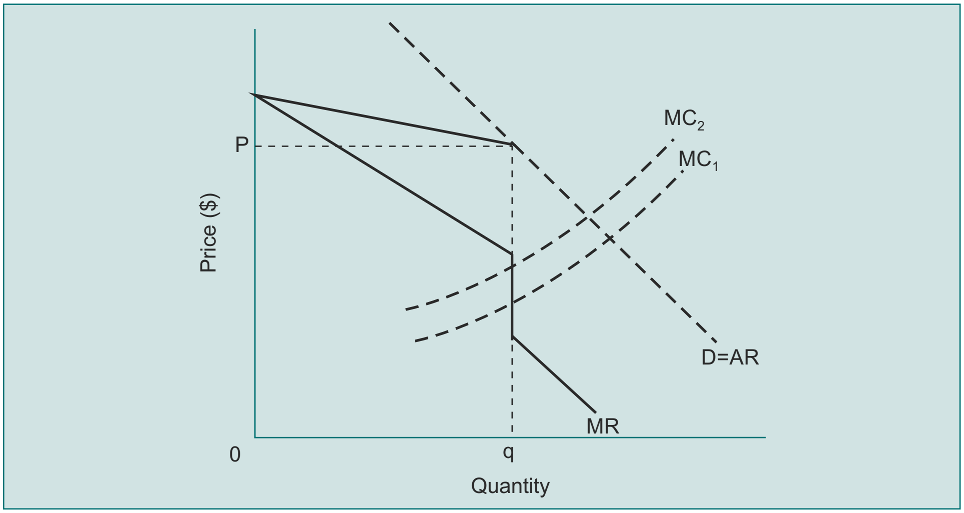


Figure 7.12 The kinked demand curve

The going price for a one-way airline ticket, London to New York, is P. The airline represented in Figure 7.12 sells 0q seats per week; this is its profit-maximising output because MR = MC_1. If the firm's costs rise – fuel costs go up, say – the marginal cost becomes MC_2. However, P still remains the profit-maximising price and 0q the profit-maximising output. But that is because of a kink in the demand curve. If the airline believes the demand curve does not kink, it will increase its price, only to discover it will be operating on the elastic portion of its demand curve – sales will fall off and losses will be incurred. Similarly, if the airline lowers its price and again does not believe in the kink, it will find other airlines following suit and, depending upon the price elasticity for Atlantic travel, may not capture substantially more passengers and suffer a loss.

What then is the optimal strategy for an oligopolistic firm? The great unknown is what competitors will do if you, an oligopolist, change your price. Suppose you are one of only four main airlines on a given route. Suppose further that the four airlines own all the landing rights negotiated with the two countries involved. Further, suppose that the demand for airline travel is price-inelastic. If you were a monopolist, there would be no problem: you would estimate the demand for air travel on that route, equate marginal revenue with marginal costs and earn monopoly profit. But being an oligopolist you have to worry about your competitors, and you are back in the kinked demand curve situation or involved in game theory, trying to guess what your competitors are going to do, matching any offers they make, and devising some of your own – remember in flying an aircraft with empty seats the marginal cost of taking on one more passenger is close to zero. You might try giving away mileage vouchers; you might try low standby fares; you might try overbooking and guess that your competitors will match you. There is, however, one simple solution: get together with the other airlines and form a *cartel*, where you all charge the profit-maximising monopoly price and divide up the market.

Fortunately for passengers (and consumers in general), most governments frown upon such practices and pass laws to make such actions illegal. It is difficult, however, for governments to make tacit collusion illegal. What oligopolists can do is to 'implicitly elect' one of the firms to be the dominant or leading firm – often the largest – and play 'follow my leader'. How often have you observed one of the major oil companies announce an increase in the price of fuel and shortly thereafter observe the other majors following suit?

Another possibility is a non-cooperative solution called a Nash equilibrium after the mathematician John Nash who devised it.

Suppose a monopoly were operating an airline between two countries and maximising profit by scheduling 30 flights per day, i.e. marginal revenue equalling marginal cost when output was 30 flights per day. Large profits invite competition. Another airline threatens to move in and operate also 30 flights per day. Both airlines get together and realise the optimal solution from their collective viewpoint is to divide the market; price and total quantity of flights remain at the initial monopoly level 30. This cartel is declared unlawful by the government and each airline is warned to act on its own. Airline A argues that if it increases the number of flights per day from 15 to 20 it will gain some of Airline B's passengers and be better off. Airline B argues that if it increases the number of flights per day from 15 to 20 it will gain some of Airlines A's passengers and be better off.

Airline A argues that if it does nothing and B increases the number of flights per day then it will lose passengers and profits and therefore should match B to minimise the loss. Airline B argues exactly the same way and thus both finish up increasing the number of flights. Both are worse off compared to the illegal cartel solution but each increases the number of flights to minimise loss in anticipation of the competitor's strategy.

This type of solution where non-cooperation makes both participants worse off is best remembered as 'the prisoner's dilemma' although as you are about to learn there really is no dilemma.

Two individuals A and B commit a robbery, are arrested and placed in separate cells so that they cannot communicate. Each is interrogated separately and each is told if he admits the crime, and his partner does not, he will be let off. If, however, he remains silent and his partner admits the crime his partner will be let off and he will be imprisoned for 10 years.

If they both admit the crime each will receive two years' imprisonment.

If neither admits the crime each will receive two months' imprisonment. The police cannot prove they committed the robbery but can imprison them for two months for 'loitering with intent'.

It is obvious that if they could communicate and trust each other, neither would admit the crime and each would receive two months' imprisonment. Now place yourself in A's shoes when he is in isolation. He argues that if he says nothing and B admits he finishes up in prison for 10 years, whereas if he admits and B says nothing he is let off. Thus he must admit to the crime. The same logic forces B to admit also. Both prisoners are worse off than if they had been able to cooperate.

In the carrying of sea cargos on recognised routes on regular schedules, major ship-owners face the same problems as airlines. Shipping cartels also became illegal, as did their practice of the 'fighting ship'. This was not a vessel designed to scare off potential entrants to the industry with conventional guns and shells; the guns and shells used were economic guns of low freight rates to deter potential entrants.

But entry into oligopolistic industries is not always easy either, because of economies of scale or the sheer size of the firms involved. To enter the mainframe computer industry, the oil industry, the shipping industry, the airline or automobile industry would require massive resources. Other barriers to entry are patents. As a result, oligopolies are here to stay, just as monopolies are. Should they be regulated to achieve economic efficiency or do attempts to regulate such industries create more problems than they solve?

In trying to answer this question, consider, for example, a government's informing a monopoly that it would be allowed to remain in business but its price would be controlled: that price would be one that covered all costs, with an additional allowance for 'normal' profit. Such a proposed solution, while sound in theory, can lead to conflicts of interest in practice. This is known as the *principal–agent problem*. An agent is an individual or group of individuals to whom a principal has designated decision-making authority. In the case of public utilities, where it is in society's interest to have only one producer to reap the benefits of economies of scale, it is necessary (if economic efficiency is also to be achieved) that price reflects marginal cost of production. However, to provide an incentive to resource owners to keep the necessary resources in the designated industry, the owners must receive a normal return on these resources. Now, if the principal, for instance the government or regulatory authority, states that prices can be set by the sole producer to cover all costs, an immediate conflict of interest arises. This conflict emerges because there is no incentive for the sole producer to minimise costs in the production process since he is guaranteed a return on whatever costs he incurs. Thus too many resources may be employed by the agent, with no concomitant increase in output. Why shouldn't the manager have a luxurious office and an expensive car plus chauffeur, and include such resources as a necessary cost of operation? His incentive to do so is personal utility; there is no disincentive because all costs are guaranteed. The principal's goal of efficiency is thus thwarted by the actions of the agent.

The principal–agent problem pervades many aspects of economics, giving rise to conflicts that are only resolved when personal interests match those of the firm or institution. For example, if the goal of a firm is increasing labour productivity and employees are paid productivity bonuses, there is no conflict. If, on the other hand, we choose a firm where all travelling salespersons are paid the same salary independent of sales achieved and compare that company with one in the same industry in which salespersons receive commission bonuses, you don't have to be an economist to predict which company is more likely to succeed in the competitive world.

Let us have one further example of the principal–agent conflict. In the former USSR, planners needed a measure of output for firms, and since prices were set by the central authority the concept of profit or return on assets was inappropriate. In a truck-making factory that produced a variety of vehicles, it was decided to measure output by total weight of vehicles produced. The truck-producing firm was adjudged to have met its quota, not by the number of vehicles produced but by their collective weight. Are you surprised to learn that in the USSR, then, you could have found some of the heaviest trucks in existence?

In these and many other situations, the principal–agent problem arises. It is essential, for the efficient operation of firms and industries, that the problem is recognised and steps taken to alleviate its impact.

7.6 Regulation and Economic Efficiency

When an unregulated profit-maximising monopolist is the supplier in a market, economic efficiency will not prevail in the economy since the marginal equivalency conditions will not be met. To achieve economic efficiency, collective action on the part of society, i.e. regulation, is required to equate the price with the marginal cost of the good in question.

In most capitalist economies monopolies are subject to some form of government regulation. In many public utilities, for example, price changes require governmental approval and, in granting or denying price changes, the marginal conditions are normally taken into account. In certain circumstances, however, economic efficiency can be achieved by replacing a monopolist with competitive firms.

Example

For example, suppose a certain drug is produced by three small firms operating under competitive conditions. As shown in Figure 7.13, each firm is in long-run equilibrium, producing $0q_c$ units at a price of $0P_c$. The price of the drug is determined by the intersection of the market demand curve D_m and the market supply curve S_m which is the summation of the firms' marginal cost curves. Market equilibrium is achieved at price $0P_c$ and output $0q_c$.

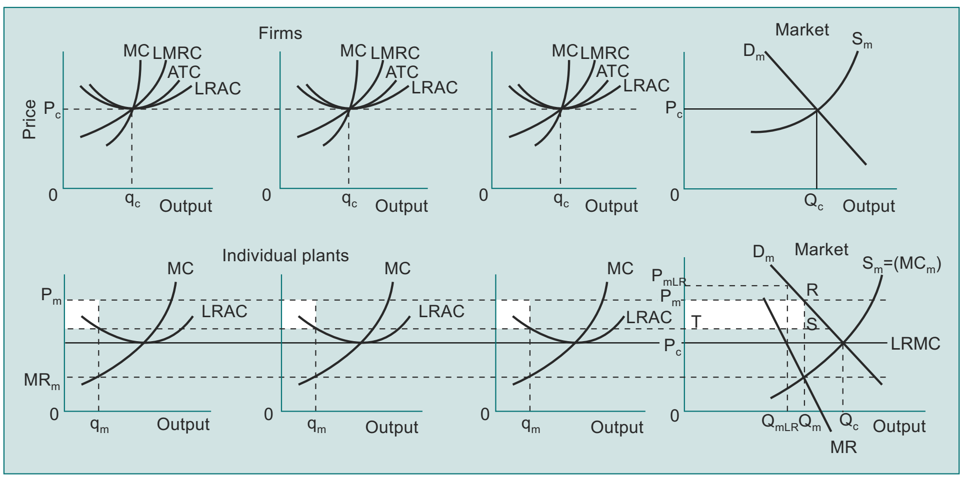


Figure 7.13 Perfect competition versus monopoly

Now suppose that one firm takes over all three drug manufacturers, obtains a patent to manufacture the drug, and thus becomes a monopolist. As shown in Figure 7.13, the monopolist's marginal revenue curve diverges from his average revenue curve, i.e. the market demand curve. Given profit-maximising behaviour, the monopolist would produce that level of output at which marginal revenue equals marginal cost, which is the industry supply curve. Thus under monopoly conditions, quantity $0Q_m$ is produced at price $0P_m$, where price is higher under monopoly than under perfect competition ($0P_m > 0P_c$) and output is lower ($0Q_m < 0Q_c$).

Under monopoly, each individual plant produces that level of output at which the monopolist's marginal revenue (MR_m) equals marginal cost ($0q_m$). Given the monopoly price of the drug, ($0P_m$), each plant thus makes an above-normal profit, shown by the unshaded areas in Figure 7.13, and the total amounts of above-normal profit earned by the monopolist is the area P_mRST.

An output level of $0Q_m$ and price $0P_m$, however, represents only a short-run equilibrium position for the monopolist. Assuming unchanged factor prices, the monopolist's long-run marginal cost curve, $LRMC_m$, will be a horizontal line tangential to the minimum points of each plant's long-run average cost curve – additional output requires additional plants. Thus in long-run equilibrium the monopolist's profit-maximising output will be $0Q_{mLR}$ and the price $0P_{mLR}$. In comparison with the short-run equilibrium position, in the long run price will be higher ($0P_{mLR} > 0P_m$) and output will be lower ($0Q_{mLR} < 0Q_m$) because the monopolist will close down one or more plants in adjusting to the appropriate long-run scale.

The government could achieve economic efficiency by splitting the monopolist into separate small firms because, as shown in Figure 7.13, under competitive conditions price would be lower and output greater than under monopoly. Given perfect competition, the price of the drug would equal its long-run marginal cost, thereby ensuring that the marginal equivalency conditions would be met. As a result, economic efficiency would be achieved.

If a monopolist's output is subject to economies of scale, however, economic efficiency will not be achieved by splitting up the monopolist into small, separate firms, because for each new firm its price will exceed long-run marginal cost.

Suppose economies of scale exist in processing a certain metal and that these economies of scale are such that one firm supplies the whole market. The monopolist is faced by a demand curve DD, as shown in Figure 7.14.

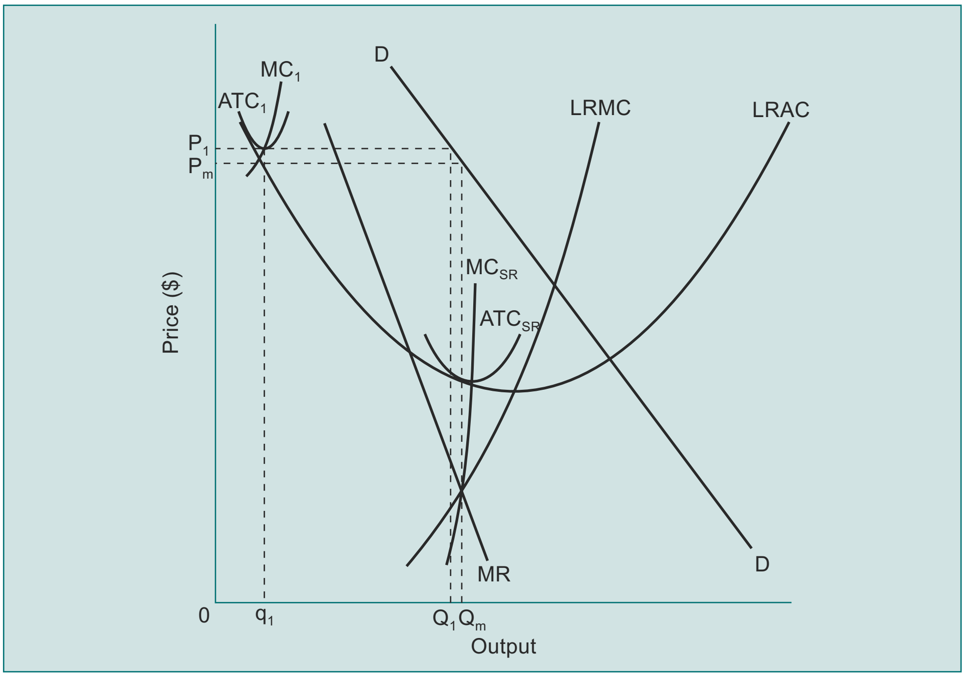


Figure 7.14 Price and output under perfect competition and monopoly with economies of scale

Further, suppose that the monopolist is in long-run equilibrium and producing an output of $0Q_m$. The conditions necessary for long-run equilibrium are shown in Figure 7.14, namely $MR = LRMC = MC_{SR}$ and $ATC_{SR} = LRAC$. The selling price would be P_m. Economic efficiency would not prevail in the economy because, in the metal-processing industry, price would exceed marginal cost.

Suppose the monopoly were broken up and replaced by a number of identical small firms each producing $0q_1$, for a total output of $0Q_1$. Equilibrium price would be $0P_1$ and each firm would be in short-run equilibrium (i.e. $P_1 = MC_1$), producing at the minimum point of its average total cost curve. However, no firm would be in long-run equilibrium because P > LRMC. Thus because of economies of scale in the industry, the price of its metal would be lower under monopoly than if produced by a large number of smaller firms ($0P_m < 0P_1$) and output would be greater ($0Q_m > 0Q_1$), so that consumers would enjoy a greater output at a lower price under monopoly than under short-run competitive conditions. In industries where economies of scale exist, regulation that replaces a monopoly by a large number of small firms will make society worse off.

When production is characterised by economies of scale, regulation to equate price with marginal cost is necessary in order to achieve economic efficiency. Suppose a monopolist steel producer were in long-run equilibrium, producing output $0Q_m$ at a price of $0P_m$, as shown in Figure 7.15. If a government wished to achieve economic efficiency and reap the benefits of economies of scale, price would be set at $0P_2$, where price equals long-run marginal cost (point Y). Thus the price of steel would be lower than under prevailing monopoly conditions ($0P_2 < 0P_m$) and output would be greater ($0Q_2 > 0Q_m$). Economic efficiency would be achieved since price equals long-run marginal cost, and hence the marginal equivalency conditions are met.

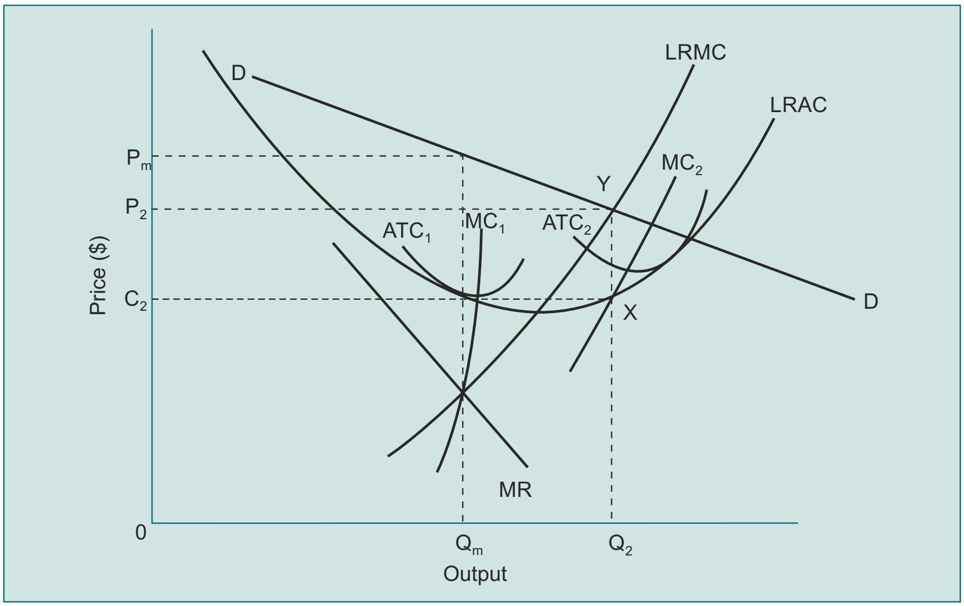


Figure 7.15 Regulating price equal to marginal cost

Note, however, that without regulation, economic forces will not lead to economic efficiency. Nevertheless, once the price has been regulated at $0P_2$ and the monopolist becomes a price taker, profit-maximising (loss-minimising) behaviour will cause the monopolist to adjust plant size in the long run to produce $0Q_2$ with the least amount of resources, i.e. be on the long-run average cost curve.

In competitive markets in long-run equilibrium, $P = MC_{SR} = LRMC = Min(ATC) = Min(LRAC)$. Under regulation, when price is equated with LRMC and the monopolist adjusts his plant size to equate price with MC_{SR}, there is no guarantee that price will also equal Min(ATC) and Min(LRAC). Under regulation, if price exceeds ATC and LRAC, above-normal profit will be earned. A profit tax will remove the above-normal profit without establishing forces that would cause price or output to change. Without such a tax, at least one factor of production will earn above-normal returns, but the existence of economies of scale will prevent factors of production being attracted into the industry seeking these returns.

If the price of metal were regulated at $0P_2$, as shown in Figure 7.15, the monopolist would still earn above-normal profit. Price per unit is $0P_2$ (= Q_2Y) whereas the unit cost is $0C_2$ (= Q_2X). Since $Q_2Y > Q_2X$, the monopolist earns above-normal profit of YX per unit. The existence of economies of scale, however, will prevent factors entering the metal-processing industry and so at least one factor employed by the monopolist would continue to earn above-normal returns. The government can remove the above-normal profit by means of a tax, thereby removing the above normal return earned by the factor(s).

Under regulation, if price is less than ATC and LRAC, insufficient returns will be earned to keep resources in the industry. A subsidy will be necessary to prevent resources from leaving the industry, or, alternatively, a regulating agency may allow a firm to set price equal to average cost, recognising that an economic efficiency-loss will occur.

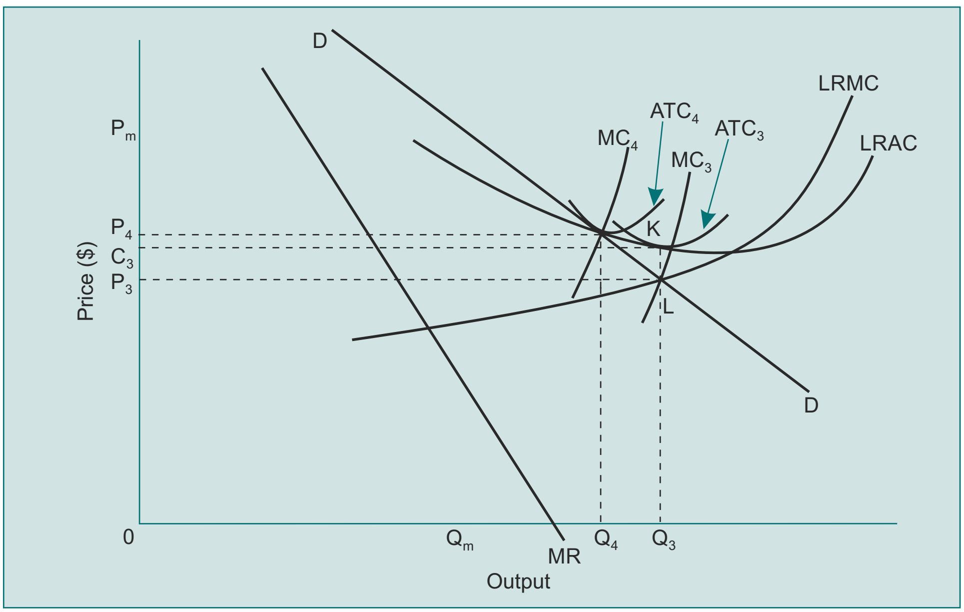


Figure 7.16 Price regulation resulting in below-normal profit

Suppose the price of a monopolistically produced good were regulated at $0P_3$ (= LRMC), as shown in Figure 7.16. The profit-maximising (loss-minimising) monopolist will produce $0Q_3$, but since at that output and price the average costs (ATC and LRAC) exceed the average revenue, the monopolist will earn below-normal profit of KL on each unit produced. At least one factor employed by the monopolist will earn less than normal returns and hence there will be a movement of factors out of the industry in the long run. Indeed, if $0P_3$ is less than average variable costs, the monopolist would cease production in the short run. To prevent resources moving out of the industry, the monopolist would require a subsidy of KL per unit, just sufficient to ensure normal profit. Alternatively, a regulating agency may allow the price to be set at the higher level of $0P_4$, where price is equal to ATC. By doing so, a lower output of $0Q_4$ is produced. In adopting this policy, the regulating agency would realise that while normal returns are earned, economic efficiency will not occur, since P > LRMC.

When monopoly exists in an economy, economic efficiency will not prevail. Since, under monopoly, price exceeds marginal cost, the ratio of $\frac{\mathrm{MU}}{\mathrm{MC}}$ for the monopolist's good is greater than the ratio of $\frac{\mathrm{MU}}{\mathrm{MC}}$ for competitively produced goods.

By regulating the monopolist's price at a level equal to marginal cost, however, more of the monopolist's good and fewer competitive goods will be produced. Society's utility will be increased by the reallocation of resources from competitive industries to the monopolistic industry, and economic efficiency will be achieved because the marginal equivalency conditions will be satisfied.

Suppose for good A, produced under monopoly conditions,

$$\frac{\mathrm{MU_A}}{\mathrm{MC_A}} = \frac{4}{2} = 2$$

and for good B, produced under competitive conditions,

$$\frac{\mathrm{MU_B}}{\mathrm{MC_B}} = \frac{3}{2} = 1.5$$

By setting the monopolist's price equal to its marginal cost, the output of the monopolist's good expands. Such expansion of the monopolist's output is possible in a fully employed economy only by transferring resources from competitive industries. Suppose that $2 worth of resources were transferred from the production of good B to good A. Since the marginal cost of both good A and good B is $2, one more unit of A and one less unit of B would be produced. Consumers would lose 3 units of utility from the decreased output of good B but would gain 4 units of utility from the increased output of good A. Hence society as a whole would enjoy a net increase in utility as a result of the reallocation of resources. When P = MC in the production of good A, there will be no further incentive for resources to be reallocated, and economic efficiency will be achieved since

$$\frac{\mathrm{MU_A}}{\mathrm{MC_A}} = \frac{\mathrm{MU_B}}{\mathrm{MC_B}}$$

However, all regulation requires resources. What a society must decide in order to make the best use of its resources is whether the formation and running of regulating agencies yields a net benefit to society. Empirical evidence suggests that, in certain major industries, regulation is worthwhile but that in many imperfectly competitive industries the costs of regulation would exceed the benefits. Professional economists are sharply divided on the issue of regulation. Many argue that the costs of regulation are both prohibitive and remove economic incentives. Others argue for more regulation on both efficiency and equity grounds.

Learning Summary

The assumption of perfect competition is relaxed and different forms of the real-world organisation of industries are introduced. You understand the differences among perfect competition, monopolistic or imperfect competition, oligopoly and monopoly, and recognise what types of goods and services are produced by the different types of firms and why.

You are able to explain why only in perfectly competitive markets do prices equal marginal costs, and you can explain the implications for economic efficiency.

Finally, you are able to understand the issues involved in regulating monopolies and oligopolies to achieve both economic efficiency and the benefits of economies of scale.

Review Questions

Multiple Choice Questions

7.1 For a monopolist, which of the following is the best definition of 'marginal revenue'?

A. The highest price it can charge without attracting competition.
B. The price at which it can sell one more unit.
C. The change in total revenue if it sells one more unit.
D. The difference in the price of the last unit sold and the price of the next unit sold.

7.2 At which operating level would a monopolist maximise profit (minimise loss)? At the output where

A. marginal cost equalled average cost.
B. marginal cost equalled marginal revenue.
C. marginal cost equalled price.
D. average cost equalled price.

7.3 Which of the following describes the most consistent difference between a pure competitor and a monopolist?

A. The monopolist has larger total assets.
B. Marginal revenue and price are the same for the competitor, but different for the monopolist.
C. The competitor is in a constant-cost industry; the monopolist is not.
D. The monopolist's price is government-controlled; the competitor's is not.

7.4 Which of the following is correct? In the long run, in monopolistic competition, high profits on the part of one firm

A. will lead to high profits for others, as they imitate the successful firm's methods.
B. will drive other firms out of the industry, and lead to pure monopoly.
C. will lead to new entry, and tend to drive profits down to normal.
D. will continue indefinitely, since the profitable firm will erect barriers to entry.

7.5 An imperfectly competitive firm discovers that at its present level of output average total cost, which is at a minimum, is $16.50 and its average revenue is $18. Marginal revenue is $12. To maximise profit what should the firm do?

A. Leave price and output unchanged.
B. Increase price and leave output unchanged.
C. Increase price and decrease output.
D. Decrease price and decrease output.

7.6 Which of the following is the key economic objection to a profit-maximising, unregulated monopoly?

A. Price will not equal average cost.
B. Price will be above marginal cost.
C. Marginal revenue will be below marginal cost.
D. Marginal cost will not equal average cost.

7.7 Suppose a certain industry is monopolistic but not subject to government regulation. Also suppose that the firm or firms in that industry is/are maximising profit. From the viewpoint of the most efficient allocation of resources in the nation as a whole, which of the following will be correct regarding the monopolistic industry?

There will be

A. no research.
B. too much advertising.
C. too many resources devoted to current production.
D. too few resources devoted to current production.

7.8 Suppose a fully employed economy had only two industries, one being monopolistic and the other being competitive. Assuming that there are no economies of large-scale production, what would be the result of government action to break up the monopoly into many competitive firms? There would be

A. an increase in output for the monopolised industry and a decrease in output for the competitive industry.
B. a decrease in output for the monopolised industry and an increase in output for the industry.
C. an increase in output for both industries.
D. a decrease in output for both industries.

Questions 7.9–7.12 are based on Figure 7.17.

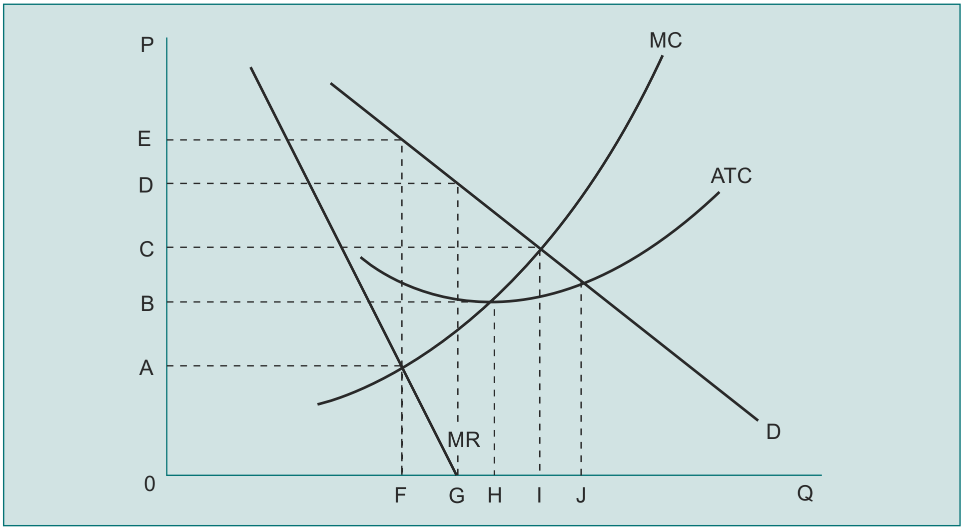


Figure 7.17

7.9 Study Figure 7.17. Which of the following is correct?

The monopolist's average revenue curve

A. is the line labelled MR.
B. is the line labelled D.
C. is the line labelled MC.
D. does not appear on the Figure.

7.10 At which price would a monopolist whose cost and revenue curves are those pictured in Figure 7.17 maximise profit?

A. B.
B. C.
C. D.
D. E.

7.11 What is the largest output the monopolist in Figure 7.17 could produce without suffering losses?

A. G.
B. H.
C. I.
D. J.

7.12 In Figure 7.17, which output would yield for the monopolist the greatest total revenue?

A. F.
B. G.
C. H.
D. J.

Case Study 7.1: The Price of Antiques

This case study applies the theory of price determination when the assumption of competition among buyers of a good does not hold. Before you tackle it you should understand:

a. supply curves
b. demand curves
c. equilibrium price determination
d. shifts of the demand curve
e. profits
f. competition
g. collusion.

The case demonstrates that

- collusion among buyers can greatly affect the operation of the market; and
- collusion among buyers can affect the distribution of income.

The Price of Antiques and Collusion by Dealers

Most major cities in the UK have auctioneering firms dealing in antiques. People usually deliver furniture, oriental carpets, clocks, etc. to be auctioned; they pay a commission – in other words a fixed percentage of the selling price – to the auctioneer when the good has been sold. The auction typically takes place in a large hall where everyone can see and hear what is going on; furthermore, the goods are on display for several days prior to the auction to allow potential buyers an opportunity to inspect them.

When the auction begins, the auctioneer describes each good, then suggests a starting price or asks for a bid. When a potential buyer bids, the auctioneer announces the bid and asks if anyone will bid higher. When the bidding stops, i.e. when no one is willing to offer a higher price than the last bid, the person who made the last bid becomes the owner of the good, after also paying a commission to the auctioneering firm. Thus competition among potential buyers determines the price of each good.

Many items at these auctions are bought by owners of antique shops. In their shops they sell the items that they have purchased at auctions. The antique dealers bid against each other at auctions just as they bid against a member of the general public. Recently, in a certain city all of the antique dealers formed a 'ring' in which they agreed not to bid against each other during auctions but only to bid against private buyers. This meant that only one antique dealer would bid against the general public. At the end of the auction the antique dealers met in private and re-auctioned among themselves the items they had purchased collectively. As each good was auctioned, the difference between the original price paid at the auction and the new price was put in a kitty; at the end of the 'ring' auction, the kitty was shared equally by all members in the 'ring'. For example, suppose Dealer A paid £200 for an oriental rug at the auction; at the 'ring' auction several other dealers indicated that they were willing to pay more than £200 for the rug and bid the price up to £300, which was eventually bid by Dealer B. Dealer B then paid Dealer A the £200 that he had paid for the rug at the official auction and put £100 into the kitty. If Dealer A had wished to retain 'his' rug, he would have had to bid, say, £320; he would then have kept the rug but would have put £120 into the kitty.

The UK's Office for Fair Trading, in a test case, took the members of the 'ring' to court in order to establish that 'rings' were illegal. The lawyer defending the antique dealers argued that the existence of a 'ring' was unimportant because:

1. the supply of goods delivered to auction houses was not affected by lower prices caused by the 'ring' – in other words, the 'ring' did not affect the efficiency of the market;
2. although the commission that the auctioneers received was reduced slightly because of lower auction prices, the substantial profits reported by the auctioneering firms were high enough anyway; and
3. no one besides the auctioneers was made any worse off by the 'ring' activities; indeed, the general public became better off because the cost of doing business by antique dealers collectively was reduced, thus enabling them to charge lower prices to their customers.

At this point, the court decided to call you in as an expert economic witness. Before appearing in court, you conducted some research and found out that the lawyers were correct in their assertion that supplies of goods coming to auction are not affected by prices; supplies of goods to be auctioned are instead determined by such things as old people moving out of larger houses into furnished flats and selling off their Victorian furniture, grandfather clocks, oriental carpets, etc.; deaths leading to property sales, with auction houses being asked to sell all furnishings in the properties; and people changing their minds about what furnishings they wanted to have in their house(s).

1 As an unbiased expert witness, please respond to the three points made by the 'ring' lawyer.

Module 8

Public Goods and Externalities

Contents

8.1 Introduction

Although a considerable amount of the economic activity taking place in our economy fits neatly into the 'perfect competition' model, there are areas of economic activity outside the arena of non-perfectly competitive firms where economic efficiency would not be achieved if we were to rely on market forces. There are two principal reasons why we as a society do not rely on market forces completely in determining the allocation of scarce resources. The first is to help achieve economic efficiency; the second is to alter the distribution of goods and services to households. The forces determining the distribution of incomes, and, concomitantly, the distribution of goods and services will be analysed in Module 9.

In our discussions to date on economic efficiency, we have not questioned the initial distribution of human and non-human capital and resulting income. Indeed, issues of economic efficiency and economic equity are separable. When we discussed using society's resources 'as efficiently as possible', we spoke about producing that basket of goods and services that would maximise society's utility using the least amount of resources. Implicitly the distribution or ownership of these resources was taken as given and not questioned. In other words, it would be possible in a market economy to have an efficient allocation of resources in the production of goods and services with a very uneven distribution of income. The existence of a population made up of 5 per cent millionaires and 95 per cent poor people would not imply economic resources were being inefficiently allocated. Conversely, an economy in which everyone had the same income would not imply an efficient allocation of resources.

In this module we shall deal with those areas of economic activity in which competitive market forces will not lead to economic efficiency and consider how society – in the form of local or national government – can interfere with, or modify, market forces so that efficiency might be achieved.

When we argued that a competitive market economy would tend to be economically efficient, there were two underlying assumptions that were never stated explicitly. We assumed, first, that whenever a firm produced a good or service it would have to pay all the costs of production; that seems to be quite a reasonable assumption. Second, we assumed that if a household wanted a good or service it would have to pay for it – remember our households voting with their dollar bills so as to indicate to firms their preferences for different goods and services. That also appeared to be quite a reasonable assumption. However, there are many goods and services produced in our economy where firms do not bear all the costs of production, and there are many goods and services produced in the economy and enjoyed by households who, in turn, do not have to pay for them.

When the costs that firms actually do pay differ from total production costs and when some households can consume goods and services without having to pay for them, a competitive market economy will tend not to allocate resources efficiently.

8.2 Private Goods and Public Goods

A private good is a good or service for which each unit is consumed by only one individual or household. The two important characteristics of private goods are excludability and rivalry. Once you buy a beer, the beer is yours to enjoy – you automatically exclude other households from consuming that particular glass of beer. The second feature, rivalry, means that your buying the glass of beer leaves less beer for other people to enjoy. Most types of good are private goods.

A public good has the exact opposite characteristic to a private good. A public good is a good or service for which each unit is consumed by everyone and from which no individual can be excluded. National defence is an example of a purely public good. If your country has a national defence industry comprising the army, navy and airforce, you consume the security provided by this national defence system without preventing others from benefiting (non-excludability). Since the amount you consume or enjoy does not reduce the amount consumed by anyone else in your country, national defence also has a non-rivalry characteristic. No one living in your country can be excluded from the national defence system – if it exists, everyone is automatically protected.

Many goods are neither private goods nor pure public goods. Consider, for example, a motorway. If relatively few cars use the motorway, it is a public good. Your using it does not stop anyone else from using it (non-rivalry) and once the motorway is built, assuming no tolls, it is available for all (non-excludability). However, when the motorway becomes congested, it takes on the characteristics of a private good: each additional car reduces the quality of services from the motorway for everyone else.

We have already seen how the efficient allocation of resources can be achieved through the market mechanism for private goods; this is not the case with public goods. The reason why is the *free rider problem*. A 'free rider' is someone who consumes a good without having to pay for it.

Example

Suppose you were one of ten wealthy families with a holiday home on the banks of a small and beautiful lake. There is only one problem in the summer – mosquitoes. You decide it is worthwhile to spray the lake with a bug killer so that you can enjoy mosquito-free evenings. Simultaneously your nine neighbours, acting independently, take a similar course of action. The result is gross overkill, i.e. the lake is sprayed ten times when once would have been quite sufficient. This is obviously a waste of resources. Incidentally the ten families enjoy their privacy and don't communicate with one other; each spends $1000 spraying the lake and each enjoys mosquito-free evenings during the summer.

Now consider a different scenario. You arrive at your lakeside home and on the first evening you observe one of your neighbours spraying the lake. You are delighted because you realise you can save $1000. The following morning your neighbour calls and suggests the cost of $1000 be shared equally among the ten households. You look your neighbour straight in the eye and tell him you enjoy mosquitoes. You are a 'free rider' – you are other things besides, but as far as enjoying mosquito-free evenings is concerned, once your neighbour has sprayed the lake there is no way he can prevent you from enjoying peaceful evenings and furthermore there is no way your neighbour can force you to pay. Similarly, if *you* spray the lake, your neighbours become free riders and you cannot withhold mosquito-free evenings from them if they do not pay. This is the free-rider problem.

Now let us expand our example. Suppose it is a large lake and to clear it of mosquitoes requires a specialised swamp-clearing firm. Suppose such a firm exists and suppose it goes ahead and clears the lake of mosquitoes. Can it then charge all residents? It can try, but there will be little incentive for residents to pay once the lake has been cleared. Recognising this problem, suppose the firm approaches households before the lake is sprayed: being a competitive market economy the firm approaches each household individually. Since we are now talking about a substantial cost, your attitude is likely to be one of it not being worth it to you to have the lake sprayed, especially when none of the cost is likely to be recouped from neighbours. In addition, if one of your neighbours *does* decide to hire the firm to spray the lake, you benefit at zero cost. As a result, no individual household is likely to hire the firm.

Recognising this further problem, the firm hits on the innovative idea of selling shares so that no individual bears the whole cost. Now what do you do? You might argue that the bulk of your neighbours will buy shares and the lake will be sprayed and you will be a free rider. Thus, since your individual contribution would be negligible, there is little point in buying a share. Alternatively, if you buy a share but too few of your neighbours do, the lake will not be sprayed and you will hold a worthless share; again you have little incentive to contribute. The solution of course, is not to have households acting independently but together as a group, each agreeing to be 'taxed' to pay part of the cost of spraying the lake. This is very different decision making from that involved by households in purchasing private goods.

Societies have long recognised the need for collective action to ensure that public goods are produced. One of the principal activities of local and national governments is to make expenditure, on behalf of households as a group, for public goods that people want but that would not be produced in a market economy. Examples of such goods are national defence, the police force and fire protection, the judicial system, pure research, roads, and public parks and recreational areas. Not all people enjoy or consume all of these equally, and because of different countries' tax structures not all people contribute to their costs, but for each of the examples noted most consumers are free to enjoy these goods whether or not they have paid for them.

In providing a range of goods and services, governments face resource constraints just as households face budget constraints in buying private goods and services. Governments, like individuals and firms, can apply the same efficiency criteria to their spending on goods and services, equating costs and benefits at the margin. The optimum provision of any public good or service occurs when the marginal social cost of production equals the marginal social benefit obtained.

In assessing whether it is providing an economically efficient level of each type of good and service, a government would have to compare the marginal social benefits of each type of programme with its marginal social costs. In the case of defence expenditure, the benefits are entirely indivisible amongst citizens. But in the case of health expenditure – if provided by the government – there may be a difference between the purely private benefits accruing to those individuals who receive certain types of medical attention and the wider social benefits that accrue to people who may not have been direct recipients of medical attention. For example, the government may provide a medical programme to immunise everyone under a certain age against a specific disease. Apart from the benefits to those who are immunised, many other sectors of the population may benefit, because of the reduced risks of contracting contagious diseases and the reduced number of working days lost due to ill health leading to higher output. Hence it is the marginal social benefit that is important in gauging the efficiency of public-sector spending.

Similarly, it is critical to measure the marginal social cost of public-sector spending programmes. When a decision is made by a government to expand the number of student places in higher education, payments have to be made for the necessary resources, such as buildings, equipment, professors, and sports and residential facilities. But while these resources have costs that are directly paid for, there are also wider social costs. By expanding the student population by, say, 10 000, the economy is sacrificing the output that they would otherwise have produced had they been members of the active labour force instead of students; hence the marginal social cost of this decision includes a valuation of the students' lost output. Only when

$$\frac{\text{MSBdefence}}{\text{MSCdefence}} = \frac{\text{MSBhealth}}{\text{MSChealth}} = \cdots = \frac{\text{MSB}_\text{Z}}{\text{MSC}_\text{Z}}$$

where MSB = marginal social benefit and MSC = marginal social cost, will the optimum level of provision be obtained.

While it may be difficult, because of the absence of consumers' valuation of benefits, to assess the optimal amount of each public good that could be produced, nevertheless some attempt must be made. Any attempt to discover the 'correct', i.e. economically efficient, number of motorways must depend on some assessment of how much citizens value the benefits of motorways. Unlike sandwiches and cars, however, we cannot directly observe the value placed on motorways because the major benefits they generate, such as reduced travelling time and lower accident rates, are not bought and sold in markets and so have no conventional price. If, however, the government attempted to derive money values for time (which is a scarce resource and has an opportunity cost) and fatal or serious injury (which has a direct economic impact on society's ability to produce through the reduction – temporarily or permanently – in the labour force), it would have a value of marginal social benefit to compare with the appropriate marginal social cost.

The issue of who pays for public goods involves considerations of equity. As a rule, governments finance public goods provision according to the ability to pay rather than according to who benefits. It might be argued that every citizen benefits equally from national security services and so each individual should pay an equal share of the cost. (Notice that this argument overrides differences in individuals' feeling towards defence; although certain individuals, e.g. extreme pacifists, may object deeply to the provision of defence at all and would not buy any defence services even if they could be allocated to markets, it is still assumed that they benefit from whatever provision is made.) If individuals did contribute equally in money terms to the provision of national security services, poorer individuals would clearly pay a higher proportion of their incomes for defence than richer individuals. If, however, individuals were taxed an equal proportion of their incomes to pay for defence, poorer individuals would pay less in money terms than individuals with higher incomes. Indeed, if the higher an individual's income the higher the proportion he pays for a public good like defence, then in money terms richer individuals would pay considerably more than poorer individuals than under the equal-proportions system.

The difference between each method of deciding who pays is one of income distribution. Under a flat-rate or equal-sum scheme, poorer families are worse off than they would be under a scheme based on an equal or higher proportion of income. Conversely, individuals with higher incomes are better off under the flat-rate system than under the other systems. Most capitalist governments adopt a system for financing expenditure where progressively higher amounts are deducted from progressively higher income levels. Although it is possible to apply economic criteria to decisions about the optimum level of public goods provision, economic criteria cannot settle the issue of who should pay. In other words, economic criteria can determine the most efficient amounts of public goods, but they cannot determine the optimal income distribution. This involves consideration of equity rather than efficiency.

Example

Suppose the government were considering building a new motorway, and three groups of citizens were debating who should pay. Group 1 argues that those who receive the greatest benefit from using the road should pay and that tolls should be charged to users. Group 2 argues that motorists in general, i.e. both users and non-users of motorways, should pay for the new road since motorists as a group receive the greatest proportion of benefits – private and social – generated by motorways. Group 3 argues that the new motorway should be financed by receipts from general taxation because all households and firms benefit indirectly from faster and safer travel.

Economic criteria, by the comparison of marginal social benefits and marginal social costs, can determine the most efficient stretch of motorway to be built. But whether only some motorists, or all motorists, or all citizens, should pay for whatever stretch of motorway is built involves a judgement of equity and cannot be decided by efficiency criteria.

8.3 Externalities: Positive and Negative

Every economic activity involves costs and benefits. In making an economic decision, the *costs* that an individual or firm takes into consideration are private costs, i.e. the costs the individual firm must bear in taking a course of action.

Example

Suppose a firm considers building a power plant to supply electricity to a city, where the firm intends to use coal as the fuel source of energy to produce the electricity. In making its decision, the firm will consider costs of resources such as the land required to build the power plant, the plant itself, the equipment and fuel needed to generate electricity, plus the labour needed to manage and operate the plant. Such costs are private costs, i.e. they are borne purely by the firm.

Similarly, in making an economic decision, the *benefits* that an individual or firm considers are private benefits, i.e. the benefits it will receive in taking a course of action.

Example

In deciding to buy a writing pad, a student will consider the benefits personally derived from the purchase, for example the facility for taking notes during lectures. Such benefits are private benefits, i.e. they accrue purely to the student who buys the writing pad. An individual or firm will consider an economic activity worthwhile if the marginal benefit derived from the course of action exceeds the marginal cost. If you consider that the marginal benefit (utility) from a sandwich exceeds the marginal cost of a sandwich in terms of the next best alternative forgone (e.g. a pint of beer), you will buy the sandwich.

For many economic activities, private costs equal social costs and private benefits equal social benefits, so that other individuals or firms are not involved in these activities, receive no benefits from those activities and do not bear any costs of those activities. For instance, since each individual student is a member of society, the total or social benefit gained from the purchase of writing pads will be the sum of the benefits gained by each individual consumer of writing pads. Similarly, the total or social cost involved in producing writing pads will be the sum of the costs of factor inputs borne by each individual supplier of writing pads. Individuals who do not buy writing pads receive no benefits from the pads produced, and only writing pad suppliers bear the costs of production.

For some economic activities, however, individuals not involved directly in the activities receive benefits from those activities. Such benefits are known as *external benefits* or *positive externalities*, for they accrue to individuals who are external to the activity.

Example

Suppose your neighbour decides to allocate resources to landscaping his garden because the marginal benefit of that activity exceeds the marginal cost. Some benefit of your neighbour's landscaped garden may accrue to you, however – a pleasant view from your window or even enhanced property values perhaps.

Similarly, for some economic activities individuals not involved directly in the activities bear some of the costs of those activities. Such costs are known as *external costs* or *negative externalities*, for they are borne by individuals who are external to the activity.

Example

In the case of our power plant, mentioned earlier, the firm may impose external costs on local households as a result of the smoke emitted from the production process polluting the air. This form of external cost is incurred by the individuals who have to pay higher cleaning bills for their clothes and their houses, and by the individuals who incur medical expenses because of ill-health contracted because of the smoke pollution in the atmosphere.

In deciding to carry out any economic activity, each individual, in order to maximise utility, will ignore benefits that may accrue to other individuals in society, i.e. they will have no incentive to consider how their actions may affect others. Your neighbour, in making the decision to landscape his garden, will consider only private benefits, i.e. the personal benefits he derives from a smart and tidy garden, and not be too concerned about whether you approve of his landscaping tastes. Similarly, each individual will ignore external costs arising from an economic activity that he undertakes, i.e. he will have no incentive to consider the costs which his actions may impose on others. Do you worry about the pollution your car causes when you drive? When you take a cold beer out of your refrigerator do you worry that someday your use of chlorofluorocarbons (CFCs) in your refrigerator's coolant will affect the atmosphere's ozone layer and increase the incidence of skin cancer in people who didn't even get a beer out of your refrigerator?

You will recall that for economic efficiency to prevail in the economy, the marginal equivalency conditions must hold. If two goods, A and B, are produced in an economy, the requirement for economic efficiency is that

$$\frac{MU_A}{MC_A} = \frac{MU_B}{MC_B}$$

In both the short run and the long run, the relevant MCs in the above equation are societal MCs, i.e. the costs to society of producing the last units of goods A and B. Similarly, the relevant MUs in the above equation are societal MUs, i.e. the benefits to society arising from the consumption of the last units of goods A and B. If the societal benefits arising from an economic activity exceed the private benefits, economic efficiency will not prevail.

If external benefits exist in the consumption of good A, then in the allocation of resources to good A the value of MU in the equation will understate the societal MU gained from the last pound's worth of resources allocated to that good. Hence the ratio of $\frac{MU_A}{MC_A}$ will be lower than the ratios of $\frac{MU}{MC}$ for all other goods, and economic efficiency will not prevail since the marginal equivalency conditions do not hold.

Similarly, if the societal costs arising from an economic activity exceed the private costs, economic efficiency will not prevail. Suppose an electricity-generating firm (E) builds a power plant on the outskirts of a city and, as a result, the amount of air pollution increases. Since the firm does not take these external costs into consideration, the marginal equivalency condition will not hold because the firm's marginal cost of producing electricity will be lower than the marginal cost to society. Hence the ratio of $\frac{MU_E}{MC_E}$ for the firm will be higher than the ratios of $\frac{MU}{MC}$ for all other goods, and again economic efficiency will not prevail.

Let us now examine what society can do in the presence of externalities to help bring about economic efficiency.

8.4 Externalities, Collective Action and Economic Efficiency

When externalities exist, market prices will not lead to an efficient allocation of resources because of the divergence between private costs and social costs and/or private benefits and social benefits. To achieve economic efficiency in the presence of externalities, there is a need for collective action on the part of individuals in a society. This is a 'legitimate' reason for a government (the people collectively) to interfere in a market economy.

Earlier it was shown that if an electricity producer did not take the costs of pollution into account in making output decisions, the marginal private cost, i.e. the extra resources cost actually incurred by the firm in producing extra output, would be less than the marginal social cost. If, therefore, some form of regulation forced the firm to pay for the marginal social costs of producing electricity, there would be a different pattern of resource allocation.

Collective action is required to equate the ratios of $\frac{MU}{MC}$ for all goods and services. One method of regulation is by a per unit tax imposed on firms that do not account for external costs in their decisions. Such a tax would add to each producer's profit-maximising price. Conversely, a subsidy paid to producers who generate external benefits by their decisions would have the effect of lowering each producer's marginal cost.

Suppose the electricity-generating industry were in competitive equilibrium, producing 0Q units of output at price 0P, as illustrated in Figure 8.1.

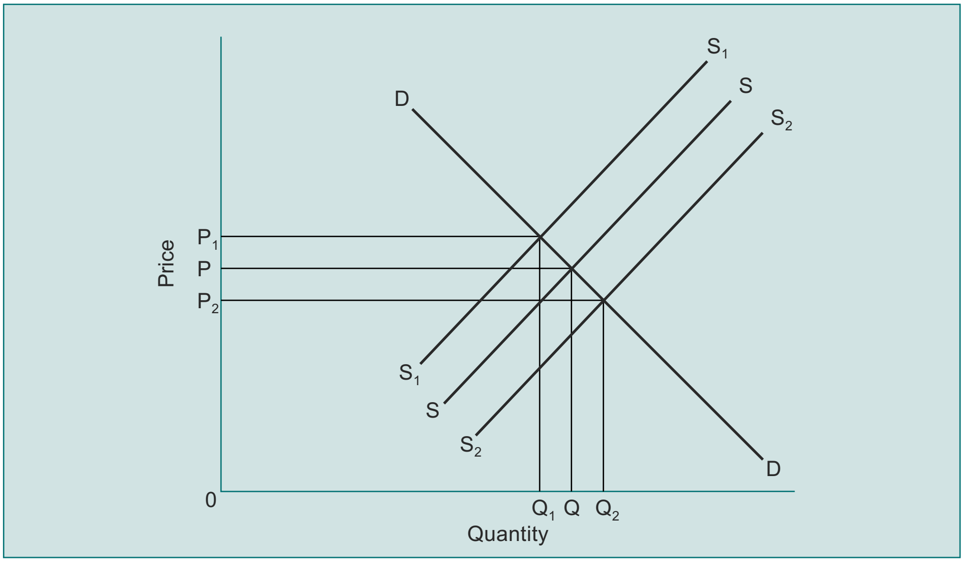


Figure 8.1 Effect of a tax/subsidy when external costs/benefits exist

If each firm were forced to pay a per unit tax that reflected the external costs of pollution that were ignored in each firm's private production decision, the industry supply curve would shift to the left, say from SS to S_1S_1, because each firm would find its marginal cost curve raised by the full amount of the tax. As a result, equilibrium output would fall and equilibrium price would rise, assuming normally sloped supply and demand curves. As illustrated in Figure 8.1, output would fall from 0Q to $0Q_1$, and price would rise from 0P to $0P_1$. The tax would affect the pattern of resource allocation in the industry because both producers (via higher marginal costs) and consumers (via higher prices) would now be taking the full social costs of generating electricity into account in their selling and buying decisions.

Conversely, if the industry represented in Figure 8.1 provided external benefits that were not accounted for, a subsidy would have the effect of lowering each firm's marginal cost curve and thereby of shifting the industry supply curve to the right – say from SS to S_2S_2. As a result, output would increase from 0Q to $0Q_2$, and price would fall from 0P to $0P_2$. The changed pattern of resource allocation following the introduction of a subsidy would reflect the external benefits derived from the good.

If an organisation that generates external costs or benefits and the bearers of those costs or benefits were to merge into a single organisation, the full costs or benefits of the organisation's activities would be taken into account. When external costs or benefits are brought within the scope of a single organisation, the externalities are internalised.

Example

Suppose a nightclub and a college bookshop were to be situated directly facing each other across a narrow street. The nightclub has a reputation for its exceptionally noisy atmosphere, which is the main attraction for most of its customers; in contrast, the bookshop requires a peaceful atmosphere in which students can concentrate on selecting books. Suppose also that, in response to the gradual increase in the noise level emanating from the nightclub over time, the owner of the bookshop is considering installing various kinds of soundproofing equipment – the more expensive, the greater the degree of soundproofing achieved. As things stand, the cost of the service provided by the nightclub does not take into account the external cost imposed on the bookshop; if, however, the nightclub were to merge with the bookshop and form a single organisation providing two different services on opposite sides of the street, the external cost associated with the noise level would be internalised. The organisation would now have to take these costs into account in making decisions about both operations. Clearly, the marginal benefits gained by increasing the noise level in the nightclub (i.e. extra revenue from customers) would then be set against the marginal costs of providing various standards of soundproofing equipment.

Example

Let us consider another example, this time with government intervention. A chemical-producing firm is located on the bank of a river. The site was not chosen randomly because the river can be used to dump its chemical wastes at zero cost to the firm. Since all chemical firms are located on the banks of rivers and use them to get rid of their wastes, our chemical firm would operate at a cost disadvantage if it were located in an urban area and had to pay to get rid of its waste products by some other means. Downstream in our river there are some famous fishing stretches, but as the firm continues to pollute the river the fishing catches start to decline. The local fishermen approach the Chief Executive Officer (CEO) of the chemical company and request he stop polluting the river. The CEO replies that the chemical industry is highly competitive and he cannot afford the cost of an alternative waste-disposal system, which would cost his company $10 000 per annum. However, he states that if the fishermen pay him $10 001 per annum he will stop polluting the river. The fishermen take their problem to their local government representative, who eventually gets a bill passed prohibiting chemical companies from dumping wastes in rivers.

What will happen? The chemical companies will each incur an additional annual cost of $10 000; thus the cost of producing chemicals to the company will rise and each firm's supply curve will shift to the left, as will the chemical industry supply curve. Given the market demand curve, the price of chemicals will rise and the real incomes of households who buy the detergents and soap powders produced by the chemical industry will fall. The externalities have been accounted for and economic efficiency prevails.

However, consider another scenario. The chemical industry exerts a powerful lobby in government circles and points out that the reduction in output of the chemical industry will cost jobs and potential political campaign funds. Why not force the fishermen to clean up the river? They argue that their industry would fail and people would start buying chemical products from foreign nations that don't worry about polluted rivers. The government is persuaded and legislates that fish caught and sold must be caught in non-polluted rivers. The fishermen get together and pay $10 000 per annum to clean up the river. However, the cost of catching fish has now risen. The price of fish will therefore rise and households who eat fish will suffer a loss in real income. In both cases we have a clean river. In the first case the chemical users ultimately pay for getting rid of the pollution; in the second, it is households who eat fish. Which solution would you, as a government official, support? The answer may depend upon whether you are a chemical consumer or a fish eater because the issue you are voting on is one of income distribution, i.e. who is going to pay.

Before you decide, or the government passes legislation, suppose the chemical company buys the fishing rights of the river. Then the problem completely disappears. Why is this so? The answer is that the chemical company, since it now owns the fishing rights, must take into account the cost of polluting the river on potential revenue from fishing catches. The company knows that the annual cost of keeping the river pollution-free is $10 000 per annum. If by so doing the net revenue from additional fish caught exceeds $10 000 per annum, it will pay the chemical company to stop polluting the river. The problem disappears because the externality – the polluted river – will now be internalised and included in the company's cost. If, on the other hand, the extra fish caught would yield less than $10 000 per annum, the river should remain polluted because, as a scarce resource, its most efficient use is getting rid of chemical waste. (It should be pointed out that in our simplified example the river has only two potential benefits for society – getting rid of waste, and fish.)

The type of problem we have just discussed may be more starkly illustrated by considering the peasant farmer located on the edge of a rain forest. He is more concerned with keeping his family alive through cultivation of three additional acres of land than the impact on global warming of burning down some of the rain forest to acquire the three acres of arable land. Or another problem for you to resolve: you are an economic adviser sent by your government to a village in an underdeveloped country; people in the village have a very low standard of living but they do have a basic electricity-generating plant, thanks to a grant from your government. The problem is the smoke pollution from the plant, which is causing some respiratory problems with the older villagers. Gas scrubbers costing $5000 can alleviate the problem but there is no more aid forthcoming from abroad; the village will have to raise the financing if the problem is to be cured. The local officials tell you that their main crop this year will generate a surplus of $5000 but unless they give the $5000 to a neighbouring village for medical supplies, 100 children are going to die. Maybe the pollution isn't all that bad.

No one likes pollution, but getting rid of pollution requires resources. The more resources society devotes to lessening pollution, the fewer resource are available for all other goods and services. If someone else is prepared to pay to get rid of pollution in your neighbourhood, you would undoubtedly welcome this action because you would experience an increase in utility at no cost to yourself. However, if you or your local community have to bear the costs, you may decide that the extra benefits from getting rid of the pollution are not worth the extra costs.

An alternative to the tax/transfer policies of government to achieve economic efficiency in the presence of pollution is to establish and enforce property rights – if this is possible. Let us consider again our chemical company–fishermen dispute. One reason the pollution problem in this example exists is because no one owns the property rights to the river and its contents, i.e. the fish. Suppose, however, that the fishermen own the stretch of river below the chemical company and the fish that swim in that part of the river. If the chemical company now pollutes the river and kills fish, it will be sued by the fishermen for the value of the fish killed by pollutants. If this value is greater than the $10 000 required to dispose of pollutants in an alternative fashion, it will be worth the chemical company's while to stop polluting the river. This example illustrates how, in situations where property rights are clearly defined and exchangeable, external costs and benefits may be taken into account through exchange without recourse to the law courts. If one individual imposes an external cost on another individual, the person adversely affected could offer the other a financial incentive not to exercise those rights. Further, if the person adversely affected is entitled by property rights to be free from such an external cost, the individual generating the externality could offer the other individual a financial incentive in compensation for the cost incurred.

An interesting hypothetical case by Professor Ronald H. Coase gave an example of a rancher and a farmer to illustrate how, in a situation of clearly defined property rights, an external cost could be accounted for by negotiation. Suppose unfenced cattle could roam and destroy the farmer's crops. Coase argued that not only was it in the joint interest of the rancher and the farmer to negotiate a solution through exchange, but also that there would be an optimal solution whether the rancher was entitled to let his cattle roam or whether the farmer was entitled to be free of crop damage caused by an external agent.

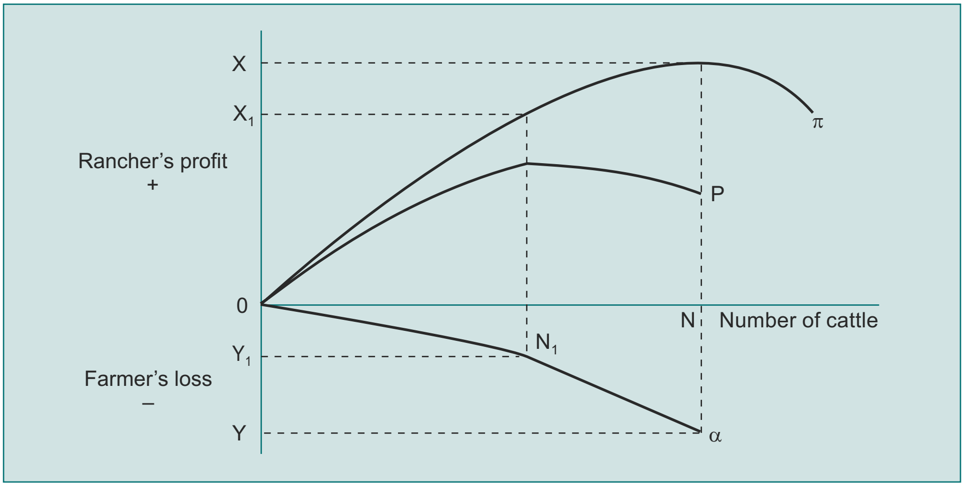


Figure 8.2 Coase's Theorem: the exchange solution to external costs

The main points of the argument can be illustrated in Figure 8.2. Curve π measures the profit earned by the rancher for different numbers of cattle reared. Curve α denotes the loss in terms of crop damage for a given crop size incurred by the farmer. The size of this loss is related to the number of cattle reared by the rancher.

Assuming profit-maximising behaviour, the rancher will rear 0N cattle if he does not have to pay for any external costs caused by the cattle. Because, however, the marginal cost to the rancher of rearing 0N cattle is less than the marginal societal cost of the value of crop damage, 0N is not an economically efficient size of herd. But by negotiating agreed terms of exchange, both parties could become better off, and the marginal societal costs of the rancher's decisions will be accounted for. From the farmer's point of view, a reduction in the number of cattle from 0N to $0N_1$ would reduce his loss from crop damage from 0Y to $0Y_1$, i.e. by YY_1. Hence the farmer would be willing to pay the rancher up to the sum of YY_1 in order to reduce his herd size by NN_1.

From the rancher's point of view, a reduction in the number of cattle from 0N to $0N_1$ would reduce profit from 0X to $0X_1$. Hence the rancher would be prepared to accept at least XX_1 for reducing his herd by that amount. Thus it would benefit both individuals if the rancher were to rear $0N_1$ cattle instead of 0N cattle and the farmer were to pay him up to an equivalent of YY_1.

Indeed, $0N_1$ would be the optimal number of cattle for the rancher to rear, meaning 'optimal' in the sense that it fully accounts for marginal social cost. This can be shown by drawing curve 0P, which represents the farmer's and rancher's joint profit from the rancher's decision. Since 0P measures the rancher's profit less the farmer's loss from different numbers of cattle, the curve reflects the external costs imposed by the rancher. The highest joint profit that the two individuals can receive from the herd-size decision is earned with $0N_1$ cattle.

Notice that if the farmer's property rights guaranteed him compensation for any crop damage, the rancher would automatically be liable for any damage caused by his cattle and would still decide on $0N_1$ cattle. In this context, the curve 0P would denote the rancher's profit minus the financial liability for damage caused by different numbers of cattle. The maximum profit for the rancher occurs with $0N_1$ cattle, and this decision explicitly covers the external costs involved.

The type of remedy adopted to deal with externalities will depend on what is practicable in the given situation. Property rights may be poorly defined and unclear; in such situations it may be impossible even to define those individuals who should come together to negotiate an efficient solution.

Example

When a factory pours smoke into the atmosphere, at least the source of the external cost can be identified. But since no one has property rights to air, it is difficult to identify those individuals on whom the external cost is imposed.

Irrespective of the definition of property rights, there may also be situations in which negotiation is extremely difficult because of the high transaction costs involved in bringing together all interested and affected parties, and the difficulty of getting information about preferences.

Example

From around 2008, when London Heathrow Airport was seeking permission to build a third runway, many groups of residents likely to be affected by the additional noise organised themselves to object to the proposals. But many other individuals, business travellers and corporations were likely to be positively affected by the decision to expand the airport. In the one case, external costs would be imposed (e.g. the possibility of increased noise level experienced by local residents); in the other case, external benefits would be received (e.g. a possible increase in demand for the goods of some companies due to the higher number of visits made by people from overseas). Some of these groups are easily identified, others much less so. In practice there would be enormous difficulty in actually arranging for all affected parties to gather together in order to agree on an efficient solution. Even if this were possible, obtaining information about individuals' exact preferences would be highly complex.

8.5 The Problems of Collective Decision Making

We have stated that collective action is required to solve efficiency problems arising out of the existence of public goods and externalities, and we have implicitly assumed that some form of voting by society will result in the 'correct' bundle of public goods and 'desired' levels of pollution. But will the voting mechanism produce such desired outcomes? (Ever since I learned that the Athenians voted to put Socrates to death, I have had severe reservations about the efficacy of the democratic process!) As you might suspect from the question, the answer is not necessarily 'yes'. The issue of 'public choice' and how to achieve it through collective action via voting is a separate subject within economics. Let us consider what has become known as the *voting paradox*.

Suppose you (A) and two friends (B and C) go to a restaurant for dinner and are told that if you all order the same meal you will be charged half price. You agree to collective action; you will all order the same meal. What is on the menu, you are told, is steak and chicken. Here is how you voted:

A. Steak is preferred to chicken (SPC)
B. Chicken is preferred to steak (CPS)
C. Steak is preferred to chicken (SPC)

By a vote of 2 to 1 you select steak. No problems here? The majority prefers steak; steak it is! Just as you place your collective order the waitress informs you there is also ham on the menu. You reconsider, and what do you choose? The answer is 'chicken'. Impossible, you might argue, in light of the above steak/chicken vote. However, consider the vote among steak, chicken and ham. Here is how you voted (given the same nomenclature as above):

A. S P C P H
B. C P H P S
C. H P S P C

Let us again look at steak and chicken, ignoring ham:

A. SPC
B. CPS
C. SPC

This is the same as before: steak is preferred to chicken. Now take steak and ham, ignoring chicken:

A. SPH
B. HPS
C. HPS

Ham is preferred to steak. Now take chicken and ham, ignoring steak:

A. CPH
B. CPH
C. HPC

Chicken is preferred to ham.

We have an interesting outcome. In the original steak/chicken choice you chose steak. However, we now see that ham is preferred to steak. Thus steak is out. However, in comparing chicken with ham, chicken is preferred to ham so ham is out. Yes, you finish up selecting chicken! That is why it is called a paradox; the pair-wise votes show:

SPC and
CPH and
HPS.

In the political arena, different voting methods can also lead to similar apparently 'undesired' outcomes. Suppose there are six political parties in a country and it is decided after each round of voting that the party with the smallest number of votes drops off the ballot until only two parties remain, A and B. Say that A wins, i.e. society prefers A to B. Now change the system and allow a final vote to take place with the last 3 parties A, B and C. C has no chance of winning but if C captures sufficient A votes, B can emerge as the ruling party even though A is preferred to B.

Furthermore, if politicians may be elected from all three parties A, B and C, representation in a parliament may finish up 48 per cent from A, 48 per cent from B and 4 per cent from C. If parliamentary voting rules involve majority voting on all issues and A and B are always opposed, the 4 per cent of votes held by party C are going to yield power completely disproportionate to numbers.

Learning Summary

You can distinguish between private goods and public goods and explain the 'free rider' problem. You understand why the price mechanism fails to provide the optimum amount of public goods and why collective action is necessary. You also understand why the decision on what public goods and how much of each public good to produce is a separate issue from who should pay for them.

You realise why, in the presence of externalities, private and societal benefits, and private and societal costs, diverge and why, therefore, economic efficiency does not prevail. You are able to explain how the establishment of property rights, as well as the use of taxes and subsidies in the presence of externalities, can bring about a more efficient allocation of resources.

Finally, you have been exposed to the difficulties inherent in reaching a social consensus through the voting mechanism – the single failure of the democratic system?

Review Questions

Multiple Choice Questions

8.1 'Public goods' are usually not produced privately and sold in the marketplace because

I. they are so expensive that only the richest buyers could afford them.
II. they are needed only by the poor, who cannot afford them.
III. if they are supplied to one buyer, they are automatically made available to non-buyers as well.

Which of the following is correct?

A. I only.
B. II only.
C. III only.
D. I, II and III.

8.2 Which of the following do economists mean by the statement that pollution is fundamentally due to 'externalities'? They mean that

A. people regard pollution as something 'external' to them.
B. individual actions have effects on other persons that are not taken into account by the decision maker.
C. pollution is fundamentally an issue that is external to economics.
D. pollution involves our external environment.

8.3 Which of the following is correct?

If a paper mill pollutes a stream so that it cannot be used for downstream city water without expensive treatment, the effect of the economic distortion involved is most likely to be

A. unemployment in the paper industry.
B. a transfer of income from paper users to the owners of the paper mill.
C. too few resources devoted to paper production.
D. a transfer of income from city taxpayers to paper users.

8.4 Which of the following is the main idea behind imposing tax on pollution? It is

A. to hurt the polluter where it really counts – in the wallet.

B. to raise money for government treatment of the effects of pollution.

C. to raise money for pollution control research.

D. to make polluters pay the social cost of their operations.

8.5 Which of the following is correct?

If a special tax were to be imposed on a polluter to 'internalise the externality' of his activities (i.e. the tax would equal the difference between private and social cost), the tax should be

A. equal to the social cost of his activity.

B. equal to the private cost of his activity.

C. equal to private plus social cost of his activity.

D. equal to the social minus private cost of his activity.

8.6 In order to establish economic efficiency in the presence of pollution, which of the following should a society do? It should

A. make pollution illegal.

B. impose a general income tax to pay for rectifying the effects of pollution.

C. impose a tax on polluters equal to the value of the damage they cause.

D. impose a tax on polluters high enough to get them to stop it.

8.7 A firm producing commodity X dumps waste in the local river, causing pollution and killing all salmon and trout. Does the pattern of river utilisation constitute an economical use of resources? Which of the following is correct?

A. No, because rivers should be preserved for fishing.

B. No, because the firm does not take into account all the costs of using the river for dumping of waste.

C. Yes, if more consumers buy X than buy fish.

D. Yes, if X can be produced more cheaply by using the river to dispose of the firm's waste.

8.8 Suppose all consumers, voting independently and honestly, indicated the maximum amount that they would be willing to pay for public good Y, and that this amount was the same for each. If each consumer contributed this amount, all of which was devoted to producing this good, then, in terms of economic efficiency which of the following would happen?

A. Too much of Y would be produced.

B. Too little of Y would be produced.

C. Just the right amount of Y would be produced.

D. Without more information, you cannot tell whether too much, too little or just the right amount would be produced.

8.9 What should we do if we really want a clean environment?

A. Increase the social awareness of our citizens.

B. Pay for it in less of other things.

C. Abandon our predominantly free-enterprise system.

D. Discipline ourselves not to be polluters.

Case Study 8.1: Pollution and the Metal Castings Industry

This case is concerned with methods of dealing with pollution of the atmosphere. Before you tackle it you should understand:

a. economic efficiency
b. externalities.

The case demonstrates that while the attainment of clean air is a laudable objective, clean air cannot be attained without incurring some costs. Furthermore, these costs may fall on certain groups, thus rendering them worse off than if no attempt had been made to clean up the air in the first place.

The Metal Castings Industry

A metal casting is the result of pouring molten metal into a mould that has been patterned to give the metal the desired shape. In its most elementary form, casting thus involves the melting of some metal, its being poured into a previously made mould, and the solidification of the metal, after which the mould is removed. Such casting, done in foundries, is by far the cheapest way to produce blocks of metal in desired shapes, and is much cheaper than machining metal into the shapes needed. Castings are made with a wide variety of metals, but iron castings are by far the largest in terms of total output.

Iron foundries are scattered all over the UK, but are generally concentrated in the Midlands region. Currently there are over 800 foundries in the Midlands, and those foundries employ 50 000 men, a high proportion of whom are unskilled. Most foundries are small, average sales being less than £2 million, on which, in a typical foundry, profit before taxes would be about £200 000 in a good year.

Nearly all iron foundries are similar in operation. They use a tall cupola with a high outside chimney. Coke is burned in these cupolas and air is forced up through the cupola at a very high velocity so as to increase the heat of the coke when it burns. The pig iron and scrap iron are melted in this heat, and poured into the mould. Under the intense heat, the air expands and rushes out of the top of the cupola into the atmosphere. This discharge over an hour from a typical small cupola will contain as much as 60 kg of fly-ash particulate.

Cast iron foundries are one of the oldest of man's inventions for handling metals, and until recently no one paid much attention to the pollution problem surrounding them. Data from government tests show that, of the dust blown upwards through the chimneys, about 80 per cent in a grey-iron foundry (the common type) is large particles that are too heavy to travel far and that usually fall very near the cupola. The remaining 20 per cent is made up of smaller particles, which are light enough to be carried further away by the wind and provide a serious nuisance in the form of soot or dirt for the surrounding area. Some of the particles (perhaps 5 per cent of the total) are very small and are taken into the atmosphere or dispersed, and some of them are eventually inhaled.

The Problem

When the Midlands authorities examined their air pollution problem, they soon found that car and lorry emissions were the largest single cause. But they also found little enthusiasm on the part of the citizens for doing anything to control car and lorry emissions. Thus, in order to reduce air pollution, they have decided to concentrate on non-car/lorry causes and have set a goal for 39 per cent reduction of Midlands air pollution. A report has suggested a reduction quota of 86.3 per cent from the foundries. Most foundry owners have protested and stated they cannot meet such standards and stay in business, and have suggested a compromise figure of 83 per cent.

The foundry owners have based their suggested compromise on three arguments. First, the 83 per cent emission control will require an investment of £2.5 million, which they are prepared to undertake, but 86.3 per cent emission control would require an investment of £7.5 million, which would mean closure for most of the foundries. Second, they pointed out that the difference between 86.3 per cent and 83 per cent foundry emission control would change the overall cleanliness of the Midlands' atmosphere by only a tiny fraction of one per cent. Third, if they went out of business, the burden of the shutdown would fall largely on unskilled workers in the area. 'What you're going to do,' bitterly protested one foundry owner, 'is to throw a lot of poor people out of work, drive us out of business, and make it impossible for many businesses to find any place to buy iron castings. What happens when a business wants to get an iron part made and there are no longer small jobbing foundries around?'

1 Many of the issues raised by the situation are both complex and difficult. Consider the following in your deliberations:

1. Should concerned students picket the Midlands foundries because the foundries are polluting the atmosphere?
2. What, if any, pollution standards should the Midlands authorities establish for iron foundries?
3. What, if any, air pollution standards should government establish for iron foundries?
4. Would a tax per pound of particles emitted into the atmosphere be more or less efficient than an absolute standard of the sort described in the case?
5. Who should pay the cost of eliminating air pollution from iron foundries?

Module 9

Income Distribution

Contents

9.1 Introduction

So far we have learned the meaning of economic efficiency and studied the conditions under which economic efficiency would be achieved in a market economy. The more efficiently a society's resources are utilised, the larger will be the total flow of goods and services and concomitantly the larger will be societal income. In determining how well off individual families are, however, we must take into account not only how large total income is but also how it is distributed.

Figure 9.1 records, for several countries, the total income of each country divided by its population: This gives us a measure of the average materialistic well-being of people in that country. This, however, is only an average measure; it takes no account of how total income is distributed within the country. To enable us to make international comparisons, we have converted the gross national income per capita datum for each country into dollars. The conversion was done by comparing the costs of a basket of representative goods and services in each country with the cost of the same basket in the USA in dollars. For example, if the representative basket of goods and services in the USA cost $100 and the same basket in the UK cost £50, the exchange rate used in converting UK gross national income per capita into dollars would buy 2:1, i.e. two dollars ($) for each pound sterling (£). Because such a conversion uses the purchasing power of each currency for the representative basket, the emerging exchange rate is known as the *purchasing power parity (PPP) exchange rate*. This PPP exchange rate can vary significantly from the 'official' exchange rate, namely the rate at which you exchange your own currency for foreign currency when you are travelling abroad; more about this in Module 10.

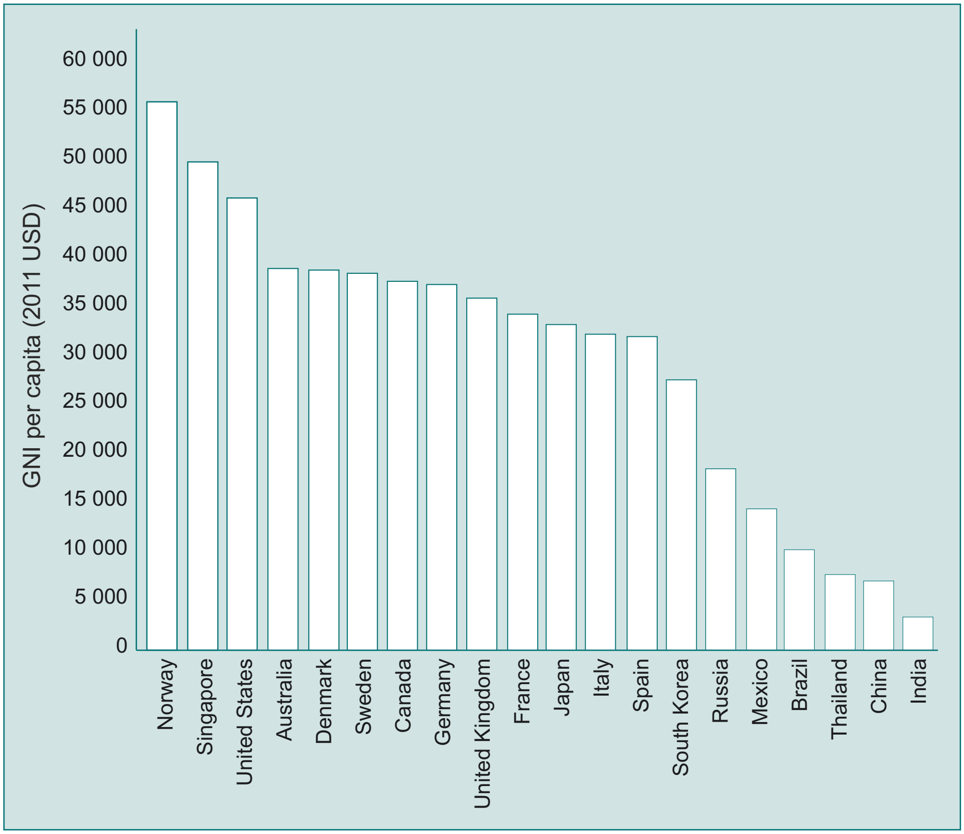


Figure 9.1 Gross national income per capita for selected countries, 2011

What determines an individual's income and, as a consequence, that individual's claim on the economy's output? In Module 5 we saw that the equilibrium price of a good or service was determined by the forces of demand and supply. In a similar fashion the equilibrium price of a factor of production is determined by the demand for that factor and the supply of that factor. Since goods and services require factor inputs in order to be produced, the greater the demand for any good so the greater the demand for the factor inputs required to produce that good. These demands, taken with the supplies of the various resources, will determine their prices. Why do the star football players, basketball players and tennis players earn so much more than economics professors? The answer is quite simple: it is all a question of demand and supply.

The income earned by factors of production is the wages and salaries paid for labour, the interest and dividends paid to the owners of capital, and the rent paid to the owners of land and mineral deposits. In all capitalist countries, labour earns by far the largest share of national income. Although each individual has the same amount of time to offer prospective employers per week, the price or wage that different individuals can command varies enormously.

9.2 Marginal Productivity

The demands for goods and services determine the demands for the factor inputs required to produce them. As the demand for cars increased steadily throughout most of the twentieth century, the demands for factories, labour, steel, rubber, glass and other resources used in producing cars grew apace.

The supplies of factors of production are determined by resource owners; they are the quantities of factor inputs that resources owners are willing to supply at different prices. Ignoring immigration, the supply of labour in any country is determined by people willing and able to work. One measure of labour supply is the number of man-days the population is prepared to supply in a year, and that is determined by wage rates and other conditions of employment.

In a resource market, the demand for a factor of production together with the supply of that factor determines the equilibrium price of the factor and equilibrium quantity exchanged. For example, in the world cargo shipping market the demand for new ships and the supplies of new ships together determine the prices of new ships and the quantity built and sold. The demands for, and supplies of, second-hand vessels determine both the equilibrium prices and the quantities exchanged. Similarly, the demands for cargo space and the supplies of cargo space determine both freight rates and the amounts of cargo transported. The demands for seamen and the supply of seamen determine the wages of seamen and the number of men at sea.

A profit-maximising firm will hire a resource up to the point at which the marginal benefit of that resource equals its marginal cost. The marginal benefit of a resource is the contribution to revenue that would result from hiring one additional unit. The marginal cost of a resource is the price of an additional unit.

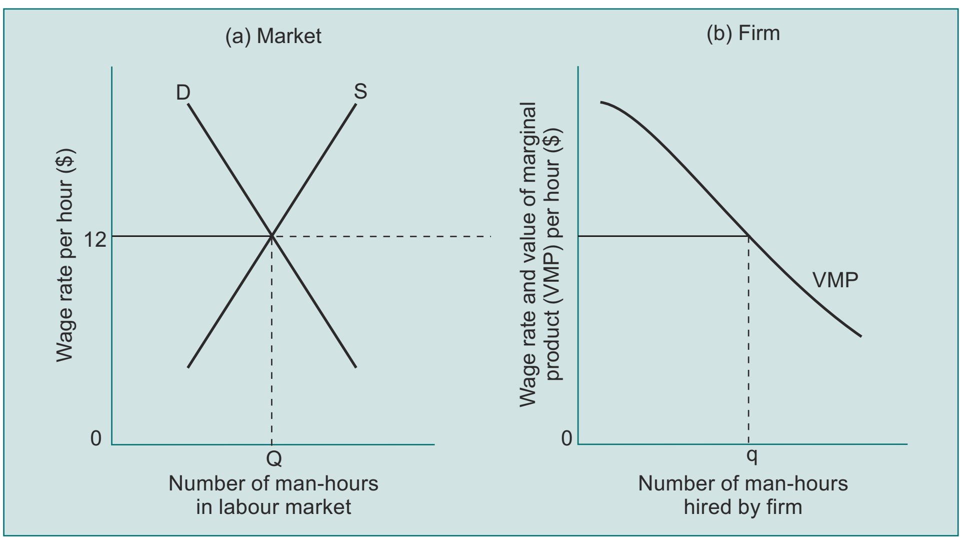


Figure 9.2 Market and firm demand for and supply of coalminers

Figure 9.2 summarises this analysis in relation to a hypothetical coalmining community where the equilibrium wage rate is $12/hour and at that price the number of coalminers hired is 0Q. Every miner who wishes to work at $12/hour can find a job and every firm that wishes to hire coalminers can do so – but remember our *ceteris paribus* assumptions when we were hiring salmon fishermen. The value of the marginal product (VMP) curve for coalminers tells us how much the value of output will increase for every additional coalminer hired by a representative firm. (The market demand curve D is the summation of the VMP curves of all the firms in the market.) Each firm is a price taker and sees that the going hourly wage rate of a coalminer is $12. Each coalminer is also a price taker and observes that the going rate is $12/hour.

The firm in Figure 9.2 will hire 0q miner-hours. If the firm were to hire fewer than 0q it would discover that each additional hour hired would add more to revenue than it would to cost. If the firm hires more than 0q miner-hours it will find that each additional hour hired adds more to cost than it does to revenue. Why? Because after 0q miner-hours are hired, the VMP (extra revenue) is less than $12, the extra cost. Thus the weekly income received by any coalminer will be $12 times the number of hours he chooses to work. Can he increase his income, given he does not wish to work more than, say, 40 hours per week? The answer is 'yes', but not in his current job unless the equilibrium wage-rate increases – and over this he has no control. What he can do is enhance his human capital, i.e. make himself a more valuable resource. He may elect to undertake a training course in the use of explosives in coalmines, and then the value of his marginal product will rise, as will his potential income. Alternatively, he may reduce his consumption expenditure out of his weekly income of $480 ($12 × 40) and with his savings buy shares in the coalmines. Now he will be a part-owner, albeit a small shareholder, of the coalmine and will share in the coalmine's profits and dividends.

The greater the amount of resources owned by an individual and the higher the marginal productivity of each resource, the higher will be the income of a resource owner. Conversely, the smaller the amount of resources owned and the lower the marginal productivity of each resource, the lower will be the income received by that resource owner. Some individuals who own land and capital and have unique talents (labour) earn high incomes. Other individuals, who own no land, little capital and can only perform low-skill tasks, earn low incomes.

Resources can be bequeathed to an individual by nature, can be bequeathed to an individual by other individuals, and can be acquired by an individual by employing resources. For example, some individuals are born with relatively scarce talents such as a fine voice, athletic ability, high intelligence and beautiful/handsome features. Some individuals have resources thrust upon them by parents or other relatives – for example, property and part-ownership of businesses in the form of stock. Most individuals acquire resources through training and/or education; for example, a student taking a medical course is hiring resources to teach him medical skills.

Although it is stated that all men are born equal, they are certainly not equal in terms of income potential. The competitive market economy rewards individuals according to their ability to produce goods and services, not according to their wants for goods and services. Not all factor markets are perfectly competitive, however, and consequently not all returns to factor inputs are determined by marginal productivity theory. We shall consider three examples, the first of which involves economic rent.

9.3 Economic Rent

The difference between the value of the marginal product of a factor of production in its most productive use and in its next best alternative is the *economic rent*. Consider our coalminer and suppose him to be working for firm A and earning $12 per hour. His next most productive employment would be working for firm B, but here he would also make $12 per hour and thus his economic rent is zero. However, suppose it is discovered, through his playing soccer in the firm's team, that he has tremendous potential as a soccer player and he is offered $50 000 per annum to sign a contract to play for a professional soccer team. The contract guarantees him $50 000 per annum whether or not he ever makes the first team but – and it is a big but – he cannot play for any other team during the term of the contract, which is for five years. He agrees; playing soccer is much more enjoyable than digging coal and his annual earnings rise from $24 000 per annum to $50 000. He very soon blossoms into a superstar and as a result attendance rises to the extent of the football club's gate receipts increasing by $400 000 per annum. Our coalminer's marginal product is thus valued at $400 000.

Given his five-year contract, his next best alternative employment is being a coalminer, earning $24 000 per annum. The economic rent in this example is the difference between the value of the coalminer/player's current marginal product ($400 000) and his opportunity cost ($24 000), i.e. $376 000. By having agreed to a salary of $50 000 from the football club, he finds the economic rent is being split 6.9 per cent to the player and 93.1 per cent to the club.

Other clubs would dearly love to have him play for them, but he is under contract. Indeed, if there were no contract, other clubs would be prepared to offer the player up to the value of his expected marginal product, namely $400 000. If, after the five-year contract has expired, the player is still making $400 000 per annum for the club and the player is a 'free agent', i.e. can sign for whatever club he chooses, his asking price will be in the region of $400 000 per annum.

This freedom of contract is comparatively new in the sports world in many countries. Previously, clubs essentially 'owned' the player and negotiated transfers without seeking a player's approval. The freedom of contract explains the enormous increase in salaries. Are these salaries outrageous? No. Although they are large compared with the return for almost all other labour inputs, your understanding of economics means you now know that the demand for, and supply of, such sporting talent intersects at an equilibrium wage rate – neither too high or too low – it should just clear the market.

9.4 Monopsony

Monopsony, the opposite of monopoly, means only one buyer in a market and occurs when only one employer of factor inputs exists in a market. In an isolated town where a company is the only major employer of labour, that company is a monopsonist. The presence of monopsony in a factor market means lack of competition on the demand side for the factors of production. For example, in our one-company town, the company does not have to compete with other firms for labour. A firm operating in a competitive factor market faces a completely elastic supply for each factor, so that it can have as many units as it wishes but at the going rate.

A monopsonist, since it is the sole employer of factors, creates the market demand schedule for each factor. The number of units of each factor hired by the monopsonist determines the prices (rates) and the levels of employment. If a profit-maximising monopsonist wishes to hire more labour, it will have to increase the wage rate to attract people into employment or attract workers from other markets. The cost to a monopsonist of hiring more units of a factor input, i.e. the marginal cost of that factor, equals the higher rate necessary to attract more of a factor, plus the increase in rate to existing employed units of that factor. Thus the marginal cost of a factor to a monopsonist exceeds the going rate. Suppose the only employer in town is in equilibrium in the labour market, with, say, a demand for labour of 50 men equal to the supply of labour of 50 men at $2.00 per hour. Further, suppose that the monopsonist wishes to expand its activities and to do so it has to hire more labour. Suppose it has to increase the wage rate to $2.50 per hour to hire one more worker. The marginal cost of labour equals the $2.50 it has to pay to employ the fifty-first man, plus the $0.50×50 it has to pay extra to existing employees. Thus, the marginal cost of labour when 50 men are employed is $27.50, which exceeds (by some way, in fact) the wage rate.

A profit-maximising monopsonist will hire a factor of production up to the point at which the benefit of hiring one additional unit of the factor (i.e. the value of the marginal product of the factor) is equal to the cost to the firm of hiring that factor (i.e. the marginal cost of the factor). The monopsonist will not hire a factor until the value of the marginal product equals the going rate for that factor since the marginal cost will exceed the going rate. This is what economists mean when they talk about *exploitation of labour*: paying a wage rate less than the value of the marginal product of the labour.

One way to counteract monopsonists' power is for labour to organise itself in the form of a trade union.

9.5 Trade Unions

The distinguishing economic feature of a *trade union* is that it represents a group of, or all, resource owners in a market. In many markets, employers negotiate rates and conditions of employment with trade union leaders. If a union were to represent all resource owners in a market, that union would have monopoly power and be the sole seller of a resource. Employers would have to deal with the union in hiring units of the unionised resource.

If unions negotiate rates for a factor of production above the level that would have prevailed in a competitive factor market, profit-maximising employers will respond by equating the value of the marginal product of the factor to the negotiated rate. Given a downward-sloping marginal product curve for a factor of production, fewer units of the factor will be hired. In the 1930s, John L. Lewis, a famous union leader of the coalminers in the United States, negotiated wage rates for coalminers that have been estimated to have been 100 per cent higher than would have existed under competitive market conditions. It has been estimated also that the effect of such rates was to reduce drastically the number of jobs offered in the coalmines by the owners.

In the presence of monopsony in factor markets, and by negotiating a forced rate for a factor that is higher than the going rate, unions can in fact reduce the marginal cost of hiring additional units of a factor and can cause an increase in employment of that factor. For example, if a union were to move into our one-company town and fix a wage rate for labour above the going rate, the sole employer would face a horizontal supply curve of labour at that rate. The company could hire labour at that rate, i.e. the marginal cost of labour would equal that rate, until no more people were willing to work at that wage. The marginal cost of labour to the firm would equal the wage rate.

The debate over the effectiveness and desirability of unionisation hinges on the extent to which they decrease employment opportunities by artificially restricting the supply of factors of production to firms, or alternatively, make factor owners better off by eliminating monopsony power. Many people argue that trade unions, by forcing up wage rates, have caused unemployment in many sectors of the economy. The workers who remain employed have higher wage rates, but such gains are at the expense of the unemployed; and total output in the economy is reduced because such factors are idle. Others argue that unions, by forcing up wages, have not caused unemployment but have caused a redistribution of income from exploiting monopsonists to workers, so that the substance of the debate is whether unions create monopoly power for workers or offset monopsony power of employers. The debate continues.

9.6 Income Distribution, Collective Action and Economic Equity

The more efficiently a society employs its resources, the greater is the total output of goods and services and, by definition, the greater is society's total income. In determining how well off an individual or family is in an economy, the *distribution of total income* must be considered as well as the size of the nation's income. In some Middle Eastern oil-producing economies, the top 10 per cent of income recipients receive approximately 90 per cent of all income, so that total income divided by the population yields a misleading figure as to the well-being of the majority of families. In the USA, where the average income level is one of the highest in the world, approximately 12 per cent of the population is defined to be *below the poverty line*, i.e. earning insufficient income to support life adequately. At the other end of the spectrum, the top 20 per cent of income recipients earn over 40 per cent of the nation's income. The lowest 20 per cent of income earners receive about 5 per cent of the nation's income.

How efficiently an economy's resources are allocated and how evenly an economy's income is distributed are not related. An economy may be operating at maximum efficiency and have a very uneven income distribution. Conversely each individual in an economy could have the same income and the economy's resources could be very inefficiently allocated. Many people argue that the socialist/planned governments' policies of attempting to make incomes equitable has a serious impact on efficient resource allocation, and they cite the low growth in total income of socialist economies as evidence of inefficiency.

All market economies elect to alter the distribution of income caused through market forces for equity reasons. Governments representing the people enact legislation to provide income for the aged, the sick and the unemployed. Tax–transfer policies constitute a common method for income reallocation: individuals with high incomes are taxed and individuals with low incomes receive transfers. In addition, all capitalist economies have a progressive income-tax structure: as income levels increase, a larger proportion of the additional income is removed in the form of income taxes. Many individuals with zero income are supported completely by the government through transfer payments. Taxes and payments in kind, namely goods and services, constitute another method of transferring income within a society. For example, in the UK, many poor families receive free school meals, subsidised housing, and family income supplements.

Any programme that redistributes income within a society causes some individuals, i.e. those taxed, to receive less than their contribution to total output, i.e. less than the value of their marginal products. The recipients of transfers in cash or kind receive incomes in excess of their contribution to total output, i.e. in excess of the values of their marginal products. When a world champion boxer earns $10 million from a championship fight, it means that society, by the fact of being prepared to pay that sum to see him fight, values his services (his marginal product) at $10 million. The government taxes the boxer a significant portion of his $10 million, some of which is received by the poor in the form of cash or kind. Out of a given national income, any redistribution scheme, by definition, can only make one group in society better off by making some other group worse off. How a society's income should be distributed is a question of value judgement and is not subject to economic analysis. Economic analysis can be used to assess the impact of any tax–transfer policy in an economy but cannot be used to determine the 'correct' income distribution.

The most conservative businessmen and the most radical students have their own opinions on how incomes should be distributed. Such opinions are value judgements and each person is as well qualified to give his opinion as is anyone else. An economist, in his professional capacity, cannot state that one opinion is superior to another, though he could offer advice on the most efficient way to achieve any given income distribution.

However, if we analyse the platform of every major political party in capitalist economies, we shall discover that income redistribution in the form of taxing higher-income households and providing transfers to lower-income households forms part of those platforms. Different political parties may disagree on the extent of these tax–transfer policies but they do exist in every platform. Since such parties represent a good proportion of the population of their economies, households collectively are indicating that the income distribution that emerges from the operation of market economies is not an acceptable distribution in terms of equity. As a consequence, a weakness of a market economy is that it produces a socially unacceptable income distribution. Precisely how equitable an income distribution should be cannot be resolved by any form of logical or economic analysis. The best way to alleviate the lot of the poor, likewise, is a subject of intense debate.

Suppose you agree with the proposition that no household in your country should have less than a certain minimum standard of living and that you vote that it is the government's responsibility to ensure each household has, at least, minimally acceptable housing, clothing and food. How, then, do we ensure in society that your, and all others', wishes are carried out? All capitalist economies have extensive welfare or income-redistribution policies. They may include minimum-wage laws, aid to dependent children, free housing, food stamps, unemployment compensation, family allowances, free school meals, income supplements, free homes for the sick and the aged, and free medical care. Running such programmes involves scarce resources and constitutes a very significant part of government budgets. Unfortunately, many such programmes do not work – poverty continues to exist amidst plenty in the most advanced capitalist nations in the world.

One solution favoured by many economists is negative income tax. Consider the simple example illustrated in Table 9.1.

Table 9.1 Negative income tax

Earned incomes ($)	Tax ($)	Income after tax ($)	Marginal tax rate (%)
0	−2 000	2 000	–
1 000	−1 500	2 500	50
2 000	−1 000	3 000	50
3 000	−500	3 500	50
4 000	0	4 000	50
5 000	500	4 500	50

Suppose it is agreed that no individual should have less than $2000 per annum because $2000 can buy enough food, clothing and shelter in a year to provide an acceptable minimum living standard. In many countries today, if any individual were receiving $2000 per annum from the government in transfers and found a job paying $2000 per annum, government assistance or transfer would cease immediately. This is a marginal tax rate of 100 per cent and consequently provides no incentive for the individual to seek gainful employment.

However, under the scheme implicit in Table 9.1 we have adopted a marginal tax rate of 50 per cent. How does the scheme work? Take the individual who cannot find a job; in this case the tax he pays is −$2000 (i.e. he receives $2000 in cash from the government); this sum is sufficient to provide him with minimally acceptable food, shelter and clothing so that he is no longer impoverished. Now he finds a part-time job for which he earns $1000 per annum. His tax liability is now −$1500 (a positive transfer received) and income after taxes is $2500. In other words he pays $500 on the $1000 he earns. This is a marginal tax rate of 50 per cent, i.e. he pays 50 per cent of the additional income he earns in taxes. He finds another job which pays $2000 per annum. Now he receives only $1000 in transfers; thus after-tax income rises to $3000; the marginal tax rate remains at 50 per cent. But he always has an incentive to find gainful employment, for he can keep one-half of what he earns (up to a certain threshold). When, then, he earns $4000 per annum he receives no transfer at all and when he earns $5000 per annum he actually pays tax of $500.

This distinguishing feature of the negative income tax is that it costs virtually nothing to run. It is administered by the existing tax authority, which can mail cheques each month to the individuals involved, who in turn file annual tax returns and receive rebates or pay tax just like regular taxpayers. All the government employees involved in existing income redistribution schemes – the social services – are available for gainful employment, i.e. more goods and services for society.

'It won't work', you might argue. The poor may take the cash and spend it on cigarettes and alcohol. Don't *you* smoke and drink? Of course you do, but you are not poor! Don't you believe the poor will use the income received from the negative tax scheme to maximise their utility? If your answer is 'no' then you really don't believe individuals know where their best interests lie and that starts you on the slippery slope of asking whether anyone knows where his or her best interests lie. The fundamental cornerstone, you may recall, is that a market economy is driven by individuals voting with dollar bills and indicating to firms which goods and services they value most highly. If you cannot accept that premise, you cannot accept the efficiency implications of the market economy and cannot accept the equity implications of negative income tax.

Learning Summary

You now understand how factor prices are determined by the forces of demand and supply, and how the demand for factors is determined by the value of the marginal product of these factors. You are able to explain the meaning of economic rent. You understand the meaning of monopsony power and trade union power in factor markets, and how the presence of either in a factor market affects returns to the factors of production and quantities hired.

Finally, you can demonstrate how market forces determine income distribution in society and the advantages and disadvantages of the tools available to alter the market-determined distribution.

Review Questions

Multiple Choice Questions

9.1 If a competitive firm operating in a competitive labour market discovered that, at its current level of output and employment, the wage rate exceeded the marginal revenue product of labour, then

I. the wage rate would fall.

II. some workers would be laid off by that firm.

Which of the following is correct?

A. I only.

B. II only.

C. Either I or II or both.

D. Neither I nor II.

9.2 Which of the following is correct?

If the market supply curve for labour is upward-sloping, a profit-maximising monopsonistic employer (a sole employer) would necessarily 'exploit' unorganised workers in the sense that it would

A. pay a wage below subsistence level.

B. pay lower wages than a competitive firm.

C. pay a wage below the value of labour's marginal product.

D. maintain substandard working conditions.

9.3 Which of the following is correct?

The differences in rates of pay between managing directors and unskilled labourers

A. are unexplained exceptions to the marginal productivity theory of wages.

B. exist because of politics and power, not economic rationality.

C. would be largely eliminated in a truly competitive, free-enterprise economy.

D. are basically the result of influences of supply and demand.

9.4 In the US in the 1930s John L. Lewis successfully negotiated wage increases for coalminers. His union, it is alleged, achieved for the workers wage rates approximately 100 per cent higher than they would have been in the absence of the union. One cost associated with Lewis's efforts was a dramatic decrease in the number of coalminers employed.

If total wage income increased because of the union efforts, then

I. employers were no longer paying employed coalminers the value of their marginal product.

II. the employed coalminers could have compensated those laid off and each would have been better off.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

9.5 In a perfectly competitive economic system in equilibrium, which of the following will be true?

A. Incomes will be equal if all people are born with equal ability.
B. Incomes will differ insofar as people differ in how long or how hard they work.
C. Average income could not be increased by labour moving from one firm to another.
D. Average income will be sufficiently high so that no working individual will starve.

9.6 The schedule below shows how many units of output a firm can produce in a day by varying only the amount of labour.

Workers per day	0	1	2	3	4	5	6	7	8	9	10	11	12
Output per day	0	10	20	35	50	60	70	78	86	91	95	98	100

If the going wage rate is $9 per day and if the selling price of the product is $1 per unit, how many people should the firm hire if it wants to maximise profit?

A. 6.
B. 7.
C. 10.
D. 12.

Questions 9.7 and 9.8 are based on the following data, which refer to a tomato farmer who hires tomato pickers at $1.20 per day.

Number of pickers	Tomato output ($)
1	3.00
2	4.80
3	6.20
4	7.20
5	8.00
6	8.70
7	9.30
8	9.60
9	9.80
10	9.80

9.7 To maximise profit, how many pickers should the farmer hire and why?

A. One tomato picker because his contribution to output is greatest.
B. Three tomato pickers because the fourth costs more than he earns.
C. Eight tomato pickers because costs will equal revenue.
D. Nine tomato pickers because the value of the marginal product of the tenth picker is zero.

9.8 The local authorities, in an attempt to make tomato pickers better off, legislate that tomato pickers must be paid a wage no less than $1.60 per day.

If our profit-maximising farmer complies with the legislation, then

I. each of the tomato pickers our farmer had previously hired will become better off.
II. the value of the average product of labour will fall.
III. the quantity of tomato output will fall.

Which of the following is correct?

A. I only.
B. II and III only.
C. III only.
D. I, II and III.

Case Study 9.1: Does the Eurozone Need a Minimum-Wage Law?

This case is concerned with the distribution of income through the market process and the effects of attempts to alter the distribution of income by controlling wages. Before you tackle it you should understand:

a. a firm's demand for labour
b. market demand for labour
c. equilibrium in the labour market
d. maximum and minimum prices.

The case demonstrates that attempting to make a particular group better off by increasing their wage by government decree can actually have the opposite effect.

Minimum-Wage Law in the Eurozone

Many people argue that a significant portion of the European labour force cannot, at going wage rates, earn sufficient income to avoid poverty. They further argue that a minimum wage should be introduced and that no firm should be allowed to hire an individual at a wage rate below the minimum. The arguments supporting some type of minimum-wage legislation to apply to all countries using the euro run along the following lines:

> No woman or man in the Eurozone should be expected to work for a wage of less than, say, €4.00 per hour. The way to help the working poor, those earning very little per hour, is to raise the minimum wage rate to €4.00 per hour; this will give them a decent living wage.

Opponents of such a scheme argue that the imposition of a minimum wage would actually harm rather than help some of the very people whom it is designed to assist – the poor. They cite evidence from the USA, which has a minimum-wage law, showing that during years when sharp increases in minimum wages have been passed by Congress, significant decreases in employment have occurred in many industries that pay low wages.

1 In this case study, consider the following points related to a minimum wage:

1. What is the effect of a minimum-wage law on a firm's demand for labour?
2. What is the effect of a minimum-wage law on the demand for labour by all firms in an industry?
3. What is the effect of a minimum-wage law on the supply of labour in an industry?
4. What is the effect of a minimum-wage law on the rate of unemployment?
5. Which group of workers is made better off by the minimum wage?
6. Which group of workers is made worse off by the minimum wage?
7. What effect is the minimum wage likely to have on training programmes?
8. What effect is the minimum wage likely to have on the total output of goods and services in the economy?
9. The minimum-wage law is intended to help those who work in low-productivity, poorly paid occupations. Is it likely to make them better off?

In your analysis, try not to confuse the objective of providing a reasonable income for everyone with the means of achieving that objective.

Module 10

International Sector

Contents

10.1 Introduction

In today's world it is almost impossible to envisage what life would be like if there were no international trade. If you think about the goods and services you currently consume, how many of them are 100 per cent home products?

We have already seen how two individuals can benefit by each exchanging a good for which marginal utility is low for a good for which marginal utility is high. It is worth noting that even if two individuals both have two goods that they like, beneficial exchange can still take place. All that is required is that the relative marginal utilities are different. For example, suppose you and your spouse are each given a box of chocolates consisting of chocolate fudge and chocolate cherries. You both like fudge and cherries but your own preference is biased towards fudge and your spouse's preference is biased towards cherries, as shown in Table 10.1. It is clear from Table 10.1 that if you trade one chocolate cherry for one chocolate fudge, both you and your spouse gain.

Table 10.1 Marginal utilities of fudge and cherries

	Fudge	Cherries
You	4	2
Spouse	3	4

Trade between individuals or groups in two different countries is no different from trade between two individuals or groups within a country. We shall now consider how nations with different capabilities to produce goods can influence trade.

10.2 The Theory of Absolute Advantage

When countries cannot produce particularly desirable goods at all, the advantages of foreign trade are obvious. For example, the British climate is too cold to grow coffee trees, and Jamaica has no petroleum. But resources of petroleum exist in the North Sea and coffee trees can easily be grown in Jamaica. Britons and Jamaicans could clearly benefit by exchanging oil for coffee.

Less obviously, even if two nations can produce a similar range of goods, they can still trade to mutual advantage. For example, Germany and France have similar resources of coal and iron, they both produce good cheese and wine, and recently France's high-technology capability in manufacturing has approached that of Germany in some industries. Despite these similarities, trade between these two countries is thriving.

Why should a country import a commodity that it could make for itself? More than 200 years ago, Adam Smith, a Scot and the father of economics after the publication in 1776 of his famous book, *An Inquiry into the Nature and Causes of the Wealth of Nations*, proposed *the theory of absolute advantage*: the same commodity can often be produced in one country using less labour and capital than in a second country trying to produce the same good. Smith noted that with intensive hothouse procedures, wine grapes could be grown in Scotland. But the resulting wine would cost about 30 times as much as wine, at least equally good, that could be bought from, say, Portugal. Because of technological progress from the industrial revolution that began in Smith's own day, machinery available for spinning and weaving woollen textiles in Scotland was more advanced than that available elsewhere – for example, in Portugal. Thus, less labour and capital were needed to be employed for each woollen shirt produced in Scotland compared with Portugal. Scotland had an absolute advantage – lower absolute labour and capital requirements – in making woollen textiles; a country such as Portugal had an absolute advantage in making wine. Smith argued that this was the basis for the mutual gains from trade.

Smith's theory of absolute advantage – or absolute productivity differences – remains useful in explaining present-day trade between similar countries such as France and Germany. The average cost of employing a worker in Germany is about the same as it is in France, and capital moves freely between the two countries; and so returns are roughly equal. Therefore, to compete successfully with a cost advantage in the French market, any particular German export must use absolutely less labour and capital when produced in Germany than if the identical item were produced in France.

Example

The Porsche – an expensive, high-powered, high-performance sports car – is cheaper to manufacture in Germany because of specialised technicians who have worked together over many years, and because of the economies of scale of concentrating production in one place. Producing an additional Porsche for export from a German plant is much cheaper than establishing a whole new production line in France. On the other hand, the smaller Citroën is an inexpensive automobile with limited speed and acceleration but with excellent fuel economy. Workers and management of the Citroën company in France have learned to produce cars cheaply – in part from relying on economies of large-scale production. Citroën perceives that the costs of producing more cars on its assembly line for export to Germany are absolutely less than setting up a new assembly line in Germany.

Trade between similar countries can occur so long as the tastes and incomes of individuals within each country differ and so long as there are absolute cost differences in producing various goods. Thus, people in France and Germany who have modest incomes and require basic transportation, and whose hearts are not set on a high-powered sports car, might purchase Citroëns made in France. Wealthy individuals in both countries with an appreciation of high-performance automobiles might purchase Porsches made in Germany. The taste and income differences among individuals within each country interact with specific absolute cost advantages across countries and provide a mutually advantageous basis for international trade. Clearly, if all Frenchmen wanted only to drive Citroëns, and if all Germans desired Porsches and had the means to buy them, the scope for trade would be restricted.

Does this theory of mutually advantageous trade hold if one country, say France, has unemployed resources? Wouldn't France be better off employing its own workers to produce Porsche-like sports cars, coupled with a ban on the importation of Porsches from Germany? The answer is 'no'. If, in this two-country example, France bans imports of German goods, then German purchases of French goods are likely to fall by an equivalent amount. In our example, French workers producing Citroëns for export to Germany will lose their jobs. The result in France is a net loss in industrial efficiency and no necessary gain in employment. Because international trade has been impeded, neither country can fully exploit its absolute industrial advantages over the other.

Cross-country differences in endowments of minerals, or in climate-affected agriculture, often provide an obvious basis for international trade. But it is not obvious why similarly situated nations such as France and Germany have differing absolute advantages in various manufactured products. The fact that one special combination of entrepreneurial and engineering skills encouraged the production of Porsches in Germany, and a second such combination led to the production of Citroëns in France, is more or less historical accident. Similarly, in Switzerland the craftsmanship required to manufacture watches and clocks could not be predicted from a general knowledge of Switzerland's capital stock and labour force.

The theory of absolute advantage not only explains the direction of trade but also explains resource movement. In the European Union (EU), at least some resources – including technical expertise – can move across national boundaries. Thus, a country's well-defined absolute advantage for manufacturing a product may attract resources not only from within that country but from neighbouring ones as well.

Example

The main assembly line for the Airbus A380 is in Toulouse, France. But parts, financial capital, craftsmen and engineers with particular skills are easily drawn from the rest of Europe. French exports of aircraft fuselages will be so much the greater, but of course they will no longer be entirely 'French'.

The determinants of trade among similar nations in a *common market* based on absolute advantage correspond to the 'accidental' geographic specialisation of manufacturing production within a country. In the USA, the Boeing Corporation has concentrated its aircraft production near Seattle, Washington State. When aircraft sales are buoyant, Boeing attracts workers, skilled engineers, and financial capital from all over the United States. Boeing aircraft are then 'exported' to other American states and to foreign countries.

From this perspective of seeking out absolute advantage, the gains from free, unimpeded international commerce are merely an extension of similar gains to be had from free domestic commerce. Indeed, Adam Smith viewed international trade as an important extension of domestic commerce, providing greater scope for the division of labour and economies of scale.

In the 1960s, the Canadian and US governments agreed to remove all barriers to trading in new automobiles, their sub-assemblies and their components. Subsequently, each large automobile company found it profitable to produce in Canada for the whole North American market one or two finished models and a few components. Other models and most components continued to be produced in the United States. Before the agreement, when restrictions on imports induced companies to produce in Canada almost all the value of each finished car sold in Canada, the same model cost a third more north of the border than in the United States. With free trade, the final prices of finished automobiles sold in Canada fell to the American level.

How do such economies of scale, associated with specialising production in one locale, come about? Many of Adam Smith's insights can be directly applied to modern industry. For example, in the case of the Canadian–American automotive agreement, the number of models assembled in Canada was reduced – but with an increase in the number of cars completed of each model. The economies were several:

(a) Any one assembly line could run with one set of tools and dies.
(b) Unskilled labour hours were not wasted in downtime associated with switching from one model to another.
(c) Fewer engineers and skilled workers were required to change the assembly line.
(d) Smaller inventories of parts and supporting materials were required.

From such economies of scale and the widening of their access to foreign resources, small countries tend to gain more – possibly much more on a per capita basis from the specialisation associated with free international trade – than do large ones.

Example

New Zealand, with only three million people, has a strong agricultural base and could feed itself in the absence of foreign trade. However, by specialising in dairy products, wool, and fruits, to be sold on world markets, New Zealand can import sophisticated dairy machinery, computers and the engineering components for automobiles. Clearly, New Zealand's average standard of living would fall sharply if it had to produce all its own manufactured goods – each on an inefficient small scale with no absolute productivity advantage in any of them.

The United States has a large population and broad manufacturing expertise. It could survive without foreign trade by producing a somewhat wider range of manufactured and agricultural goods. The scale of each industry would still be fairly large, although many would not be as efficient as their counterparts abroad. Without foreign trade, living standards in the United States would fall, but not to nearly the same extent as in New Zealand.

Is it necessary for a country to have an absolute advantage in the production of a specific good to make trade beneficial? The obvious answer is 'yes'; however, that answer is incorrect. We must now look at comparative rather than absolute advantage.

10.3 The Theory of Comparative Advantage

Like the theory of absolute advantage, *the theory of comparative advantage* assumes that resources are fully mobile within a country so that returns to equivalent labour and capital are equalised on a national basis. However, formal demonstrations of comparative advantage assume that these same resources are not mobile internationally. Thus, any randomly chosen pair of trading countries need not have – and is unlikely to have – similar standards of living or the same general level of technical proficiency. We will discuss the principle of comparative advantage, using India and the USA as an example.

Per capita income in India is very low compared with that of the United States. In every major industry, output per worker and per unit of capital is less than in the USA. Compared with the United States, therefore, India has no obvious absolute advantage in manufacturing or agriculture. Although Indian workers' wages are correspondingly lower, significant numbers of them cannot migrate to the United States because of American immigration restrictions. And, despite low yields, capital cannot flow from India to the United States because of the stringent exchange controls imposed by the Indian government. At present, owners of economic resources in India are not free to follow the model of Adam Smith and seek employment internationally in areas of their greatest absolute advantage.

Nevertheless, gains from foreign trade in commodities are still possible. The theory of comparative advantage, put forth by the English economist David Ricardo in the early nineteenth century, demonstrates that a relatively poor country (such as India) without any absolute industrial advantage, can still trade to mutual benefit with a wealthy one (such as the United States). A formal proof of Ricardo's theorem is made easier with some simplifying assumptions. We shall assume:

(a) both the United States and India produce only two goods – wheat and burlap (food and clothing);
(b) labour is the only variable factor of production, but its productivity differs in each country because of differing land availability, climate, or engineering efficiency;
(c) in each country, the productivity of labour per hour worked is constant irrespective of the scale of productivity in any particular industry;
(d) labour in each country is fully employed in producing wheat or burlap;
(e) there is no migration of labour between India and the United States; and
(f) although output per man-hour in each industry is greater in the United States than in India, the productivity gap in wheat is not proportional to the productivity gap in burlap.

Notice that assumption (c), *constant returns to scale*, prevents either country from enjoying economies of production on a large scale and thus eliminates the most important sources of gains from trade that were emphasised by Adam Smith. Yet, perhaps surprisingly, Ricardo's principle still holds, as is shown below.

Accepting our assumptions, consider in Table 10.2 the data that illustrate the productivity differences between the two countries. The United States is absolutely much more efficient in both industries. India requires ten times as much labour to produce one hectolitre of wheat but only five times as much to produce one metre of burlap. By definition, India has a *comparative advantage* in producing burlap, where its production is less inefficient; the United States has a *comparative disadvantage* in burlap and a comparative advantage in wheat.

Table 10.2 US and Indian labour requirements for production

Product	In the USA hours of labour	In India hours of labour
1 hectolitre of wheat	1	10
1 metre of burlap	2	10

Ricardo's principle suggests that more wheat and more burlap can be produced between them if the USA concentrates on producing wheat and India concentrates on producing burlap. The USA could then export wheat in exchange for Indian burlap, so that citizens of both countries could enjoy larger quantities of both goods.

Consider first how production is limited if neither country specialises, and if each produces something of both goods in a balanced fashion. Suppose, over the course of the year, 60 labour hours are available in the United States and 400 labour hours are available in India with its larger population. Suppose, further, that each country has a central planning agency that arbitrarily allocates labour to either the production of wheat or to the production of burlap. In the absence of international trade, both the Indian and American planning agencies separately decide to put half their labour hours into wheat and half into burlap. The resulting output is shown in Table 10.3. Without specialisation in production, total production over the two countries is 50 hectolitres of wheat and 35 metres of burlap.

Table 10.3 Balanced production in India/USA (without trade)

Product	India	United States	Total
Hectolitres of wheat	20	30	50
Metres of burlap	20	15	35

Now suppose that the two countries are willing to trade, and thus both need no longer to produce both goods. At the extreme, suppose the USA specialises completely in wheat with its 60 labour hours and India produces only burlap with its 400 labour hours. The resulting output is shown in Table 10.4.

Table 10.4 Specialisation in production in India/USA (with international trade)

Product	India	United States	Total
Hectolitres of wheat	0	60	60
Metres of burlap	40	0	40

The total amount of wheat now available to the two countries has risen from 50 to 60 hectolitres and amount of burlap has increased from 35 to 40 metres. The potential availability of both goods increases when each country produces according to its comparative advantage. The two countries' planning agencies, however, must bargain over how many hectolitres of wheat to trade for a metre of burlap.

Suppose there are 100 hours of labour available in the United States instead of 60; we shall retain the assumption of 400 hours available in India and also use the output labour ratios specified in Table 10.2. Now, complete specialisation of the USA in wheat and of India in burlap no longer yields more of both goods compared with having each country devote half its labour hours to burlap and half to wheat. (You can verify this by working out the arithmetic yourself to show that there will be much more wheat and slightly less burlap after specialisation.) But there are always potential mutual gains from nations specialising according to their comparative advantage. This general proposition is now demonstrated geometrically.

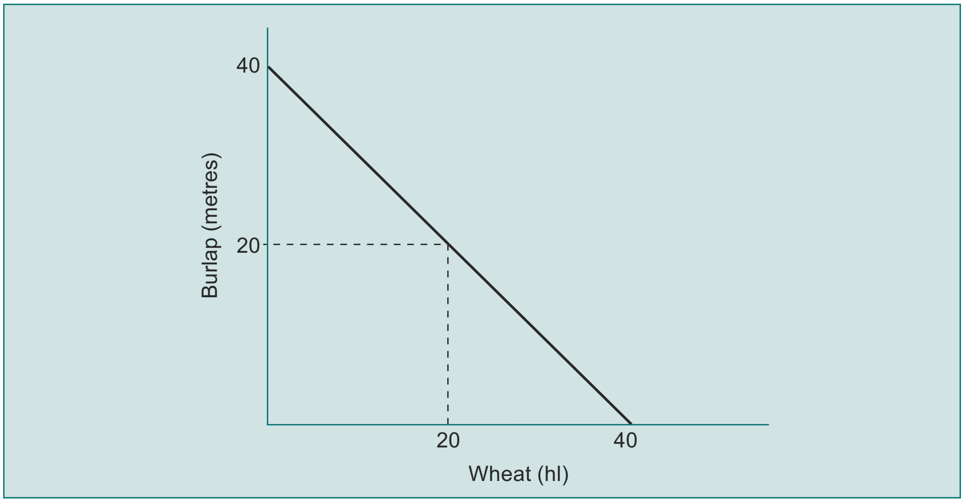


Figure 10.1 Production possibilities for India

Figure 10.1 yields the production possibilities frontier for India with 400 labour hours available, and Figure 10.2 is that for the United States with 100 labour hours. Both are linear, reflecting the assumption of constant returns to scale, and use the labour output ratios specified in Table 10.2.

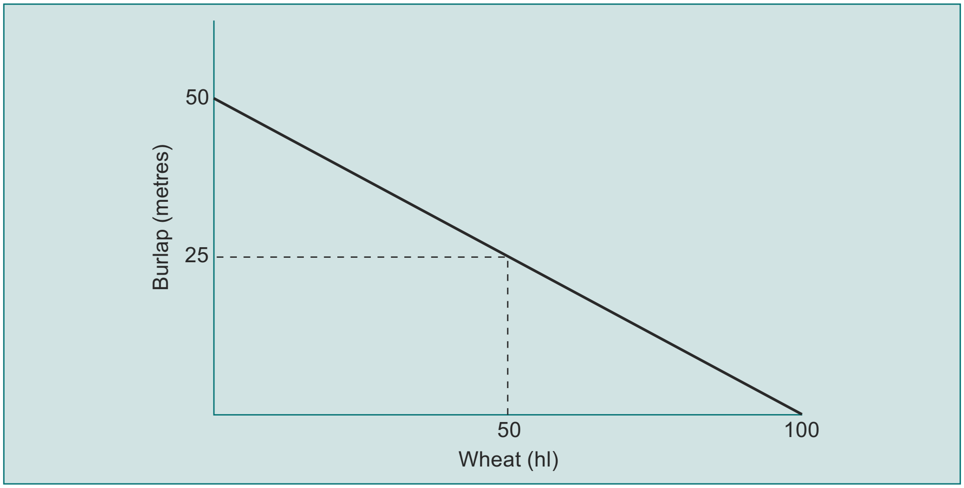


Figure 10.2 Production possibilities for the United States

To establish the world's production possibilities frontier, we sum those of India and the United States (there are only two countries in our world) while keeping the assumption of no factor mobility between the two countries. Again consider a 'socialist' solution to world resource allocation – this time by assuming a world planning agency that decides on how labour hours are to be allocated in both the United States and India. Using Ricardo's principle, our agency sees the most efficient combined production possibilities frontier to be as shown in Figure 10.3.

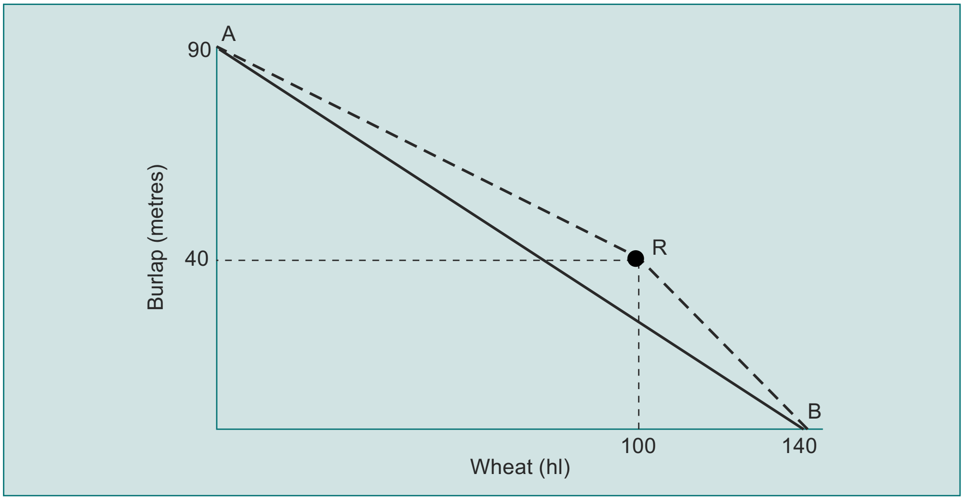


Figure 10.3 Combined production for India and the United States

Starting at point A in Figure 10.3, the planner has both countries producing only burlap: 90 metres in total. Suppose he wishes to increase wheat production by moving resources out of burlap, i.e. moving from point A downward and to the right. To minimise the opportunity cost in lost burlap output, the planner should first designate only the United States (which has the comparative advantage in producing wheat) to begin switching labour from burlap to wheat. This transformation within the United States is represented by the movement from A to R, along which only one metre of burlap need be given up to get 2 hectolitres of wheat (*see* the slope of the US production opportunity set in Figure 10.2). At point R in Figure 10.3, the 'Ricardo' point, the United States is completely specialised in wheat production and India in burlap.

Further increases in wheat production must come by instructing workers in India to switch from burlap to wheat – along the line RB at the much less favourable trade-off of one hectolitre of wheat for one metre of burlap (*see* slope in Figure 10.1) because India has a comparative disadvantage in wheat production. Finally, at point B of Figure 10.3, both countries are completely specialised in wheat, producing a total of 140 hectolitres.

The upshot is that the world production possibilities frontier ARB, constructed according to the principle of comparative advantage, is a complete menu of the most efficient trade-offs between wheat and burlap production over the two economies. Any other method of assigning production will be inefficient by comparison.

If you are not convinced, start again from point A where only burlap is being produced in each economy. Suppose now the planning agency instructs both countries simultaneously to begin moving labour out of burlap and into wheat – say 90 per cent into burlap and 10 per cent into wheat, 80 per cent into burlap and 20 per cent into wheat, and so on. Then production traced out by the economy will simply be the straight line AB in Figure 10.3. And except at the two end points where the curves touch, AB is below the efficient locus ARB. That is, as long as some of both wheat and burlap are being produced, failure to assign production across countries according to comparative advantage reduces the amount of one or both goods that would otherwise be available for consumption.

Ricardo's remarkable insight about the correct international specialisation in production is independent of whether the total economic output so produced is distributed 'fairly' or 'unfairly'. Providing that the ratios relating man-hours to output of wheat and burlap are indeed given and not sensitive to rewards to workers – a big proviso in practice – the planning agency could, for example, victimise one country in favour of the other. Almost all of the wheat and burlap produced could be commandeered for consumption in one country; the other could then be left worse off than under voluntary international exchange or even worse off than it would be if isolated from the rest of the world. Of course, this example is only likely on a significant scale if the favoured country has some military hegemony over the other. Yet the military overlord needs to understand the principle of comparative advantage in assigning production internationally if this exploitation is to be carried out efficiently.

We have considered how an omniscient central planner could allocate production efficiently according to the comparative advantage of two countries. Still it remains for us to demonstrate that voluntary international exchange for mutual gain – as happens between market economies each made up of millions of individuals – can still lead to the same production efficiency espoused by Ricardo and Smith.

10.4 The Terms of Trade and Voluntary Exchange

For assurance that voluntary international trade conducted by firms and individuals is advantageous for the economy as a whole, relative prices at which individuals may freely trade domestically should reflect accurately the opportunity costs of producing the goods in question.

Consider the production opportunity costs for the United States producing either burlap or wheat in Figure 10.2. In the absence of international trade, a metre of burlap would exchange for two hectolitres of wheat domestically so as to reflect accurately the opportunity costs. This would be true if the US economy were competitively organised, if divergences between private and social costs of production were absent, and if there were no systematic government intervention to change relative prices.

Now consider the production opportunity cost of one hectolitre of wheat for one metre of burlap for India as in Figure 10.1 without foreign trade. Suppose now the Indian government behaves more like a 'command' economy and decides to subsidise the consumption of wheat by taxing the production of burlap. Consumers in India find the price of wheat to be reduced to one-third that of burlap. This government 'wedge' in the price system falsely signals to private traders that wheat is relatively plentiful (possibly for export) and that India has a comparative disadvantage internationally in producing burlap.

Assume, therefore, that competitive pricing – without tax or other distortions – accurately reflects production opportunity costs in each country. In the absence of international trade, one metre of burlap exchanges for one hectolitre of wheat in India (Figure 10.1) and for two hectolitres of wheat in the United States (Figure 10.2). Now consider the reactions of millions of farmers, manufacturers and traders in each country to the sudden removal of all the previous restrictions on commodity trade between the two countries. We shall continue to assume that the factors of production, namely labour and capital, cannot move internationally. Thus the US farmer, who had been accustomed to buying only half a metre of burlap for every hectolitre of wheat marketed, sees an opportunity to export wheat and get as much as one metre of burlap from a private merchant in India. An enterprising Indian merchant selling burlap and accustomed to receiving only one hectolitre of Indian wheat, now sees the opportunity of receiving up to two units of wheat by exporting burlap to the United States. Without any central planner to determine the direction of trade, American wheat will start moving to India in exchange for burlap. The independent profit-seeking actions of private traders, responding to the different relative prices in each national market, ensure that each country begins to specialise according to its comparative advantage.

Once the two national markets are joined internationally, the previously separate national price ratios for burlap in terms of wheat must give way to a single world price ratio. (We are ignoring transportation costs.) The shipment of wheat to India will bid down the relative price of wheat in India, so that more than one hectolitre of wheat is then required to buy one metre of burlap. The shipment of burlap to the United States bids down its relative price, so that one hectolitre of wheat will now buy more than one half-metre of burlap. This international commodity arbitrage continues until a common international price ratio of wheat for burlap is established between the two pre-trade national price ratios.

Let P be the common international price of wheat in terms of burlap established by commodity arbitrage. The variable P is defined to be the number of metres of burlap needed to buy one hectolitre of wheat. Then $0.5 < P < 1.0$, where 0.5 is the pre-trade price (cost) ratio of wheat prevailing in the United States (Figure 10.2) and 1.0 is that which prevailed in India (Figure 10.1). Where the international price ratio P settles in the interval between the two pre-trade cost ratios depends on aggregate demand for wheat versus burlap over all consumers in both countries.

Suppose people in the two countries had a relatively strong preference for staying warm in burlap coats and were less inclined towards eating. Then P, the equilibrium price of wheat in terms of burlap, would settle at, say, 0.53, close to the minimum permitted by the underlying production technology in the United States. Suppose, alternatively, that people were inclined to eat more and wear less clothing; then P would settle at, say, 0.96, close to the maximum permitted by the underlying production technology in India. As long as the international price ratio P settles somewhere between the two technologically determined extremes of 0.5 and 1.0, the average worker in both countries will be better off. At either extreme, workers in only one country are better off, with those in the other being no worse off.

Now suppose that aggregate demand for wheat and burlap is balanced relative to the supplies available, and the international price P settles at 0.75 metres of burlap per hectolitre of wheat – halfway between the two pre-trade prices. Now Indian workers will specialise in producing burlap only, and Americans will only produce wheat. Why? An Indian worker still must spend ten hours producing a metre of burlap from the technology specified in Table 10.2. Yet now that worker no longer needs to spend ten hours directly producing a unit of wheat: instead, by working another 7.5 hours to produce 0.75 metre of burlap, he may trade that for the desired hectolitre of wheat. By specialising in producing only burlap, all Indian workers will make themselves unambiguously better off. Similarly, a US worker must still work one hour to produce a unit of wheat as specified in Table 10.2. But now instead of having to work two hours to produce a metre of burlap directly, he need only work 1.33 hours in order to produce 1.33 hectolitres of wheat to sell to India in exchange for one metre of burlap. Hence, all American workers will gravitate towards producing only wheat and will be unambiguously better off for doing so.

In both India and the United States, real wages in our example have risen: after voluntary international trade, one hour's work can buy more of one of the two goods and no less of the other. And the common *pauper labour* argument that trading with a much poorer country with low wages and low productivity will necessarily reduce real wages in a wealthier and more productive country, is simply fallacious – and this fallacy was well understood by David Ricardo 200 years ago.

Adam Smith's concern for absolute advantage and the need to improve absolute productive efficiency is highly relevant for establishing real wage levels and living standards in a country like India (or the United States). Although India may gain substantially from free foreign trade, her low absolute labour productivity in all industries consigns her to much lower real wages even when the international terms of trade are much in her favour.

Suppose the international demand for burlap is unusually strong and the international price of wheat P (measured as metres of burlap needed to purchase one unit of wheat) is 0.5, which is at its technological minimum. Although Indian workers are better off from these highly favourable terms of foreign trade, they must still work five times the number of hours to purchase one unit of wheat (by first producing burlap and selling it for wheat), and five times the number of hours to buy one metre of burlap, that an American worker must work.

Consider now just one country, say Britain, and how it might calculate its comparative advantage *vis-à-vis* the rest of the world. Would it be necessary for a British planning commission to calculate the labour output ratios across all major British industries, and then compare these ratios to those prevailing in all potential British trading partners from Japan to Peru to Monaco? Being virtually impossible, fortunately such a calculation is completely unnecessary as long as an adequate worldwide monetary system exists to support 'equilibrium' price quotations in all major commodities at which individual Britons are free to buy or sell. International trade according to British comparative advantage then simply takes care of itself, so as to increase the consumption possibilities open to the British economy. International prices contain all the relevant information needed by British traders and manufacturers.

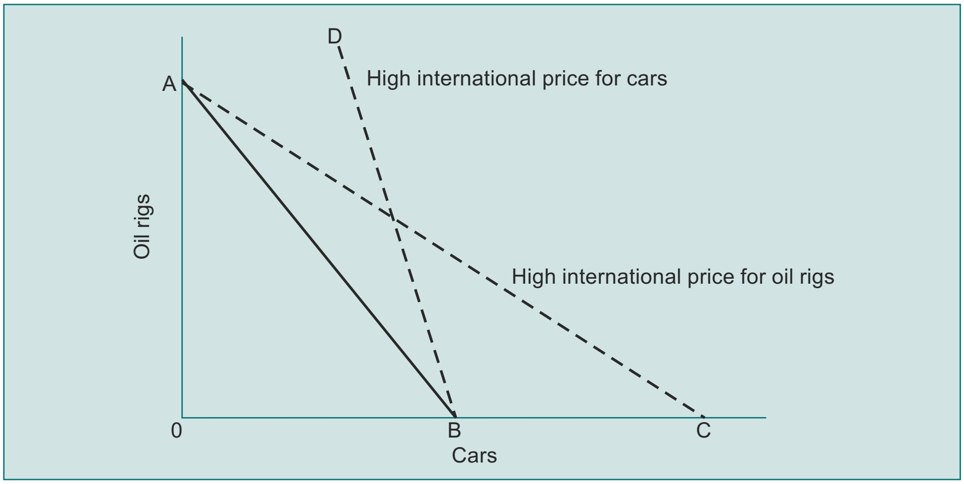


Figure 10.4 Production possibilities for Britain

Consider again only two commodities, say oil rigs and cars, either of which Britain could export or import. For simplicity, consider again fixed labour output ratios and a linear technology in Britain, so that one oil rig could be produced at the expense of one car. The British production possibilities frontier is thus the straight line AB in Figure 10.4.

The remarkable aspect of Figure 10.4 is that any international price for cars in terms of oil rigs differing from Britain's internal production opportunity costs of 1:1 will generate voluntary international trade to Britain's benefit. If cars are relatively cheap on world markets, say two cars trade for one rig, then Britain would specialise its production at point A and be an exporter of oil rigs and an importer of cars. The consumption possibility set of Britons then increases upward and to the right along the dashed line AC. If oil rigs, however, turn out to be relatively cheap in world markets, say two oil rigs trade for one car, then Britain should specialise production at point B and be an international exporter of cars in exchange for oil rigs. Again the consumption possibility set of Britons increases upwards and to the right, this time along the line BD.

In the Ricardian tradition, Britain could produce one additional oil rig at a constant opportunity cost of one less car, assuming constant returns to scale. However, different technologies abroad lead to a different international price ratio between the two goods. This difference is an important basis for the gains from trade.

In the Smithian tradition, emphasising returns to scale, there need be no such difference in foreign and domestic technological capability for foreign trade to be mutually advantageous.

Example

Britain and France may be equally capable technically of producing jet engines and aircraft fuselages; yet it still pays to concentrate engine production in Britain at Rolls-Royce and the production of fuselages at Airbus in France because of the economies of large production runs. How each country specialises may be less important than specialisation itself.

The basic principles of the potential gains from international trade were illustrated by assuming that one good could simply be bartered for another. In practice, however, thousands of goods are traded internationally and non-monetary barter is not feasible. If international trade leads to such substantial benefit among trading nations, why is free trade so vociferously opposed by some members of society, leading governments to impose trade restraints such as tariffs and quotas?

10.5 Tariffs and Quotas

The main argument put forth for restricting international trade is the protection of domestic industries. It is a logically weak but politically persuasive argument. In capitalist countries, the costs to the average household of the presence of tariffs and quotas is in the high hundreds or low thousands of dollars per annum. That is what households would save if free trade existed. Why then tariffs? Have you ever seen groups of workers parading with placards demanding free trade? Have you ever seen or heard groups of workers demanding protection against 'cheap' foreign imports of shoes, textiles, meat, wine or cars? Does that answer the question?

A *tariff* is a tax imposed by an importing country when a good or service enters that country. A *quota* is a restriction specifying the maximum amount of a good that may enter a country during a specific time period. Countries that pay lip service to free trade often resort to political pressures for trade restraint by signing agreements with exporting nations to restrain the amount of imports; these agreements are known as *voluntary export restraints*; such non-tariff barriers are now a more important impediment to free trade in many countries than tariffs. The history of tariffs since the Great Depression of the 1930s has been one of falling tariff rates, on average, but they vary for different commodities in different countries.

10.5.1 Tariffs

Figure 10.5 indicates the impact of a tariff on both price and quantity of a good bought and sold. The demand curve (D) and supply curve (S) for an imported sports car show that, in equilibrium, 6000 cars would be imported per annum at a price of $30 000 each. To protect the domestic sports car industry the government imposes a tax of $20 000 per car. The S curve shifts upwards to 'S + tariff'. The new equilibrium quantity and price are respectively 4000 and $40 000, so that imports of sports cars decrease from 6000 to 4000, with some sports car buyers now buying domestic sports cars instead. Isn't that a good thing for workers employed in the domestic sports car industry? Before you answer that question, you should ask where the resources come from to satisfy the additional demand for domestic sports cars caused by the imposition of the tariff on imported cars. The output of some other domestic industry has to be reduced.

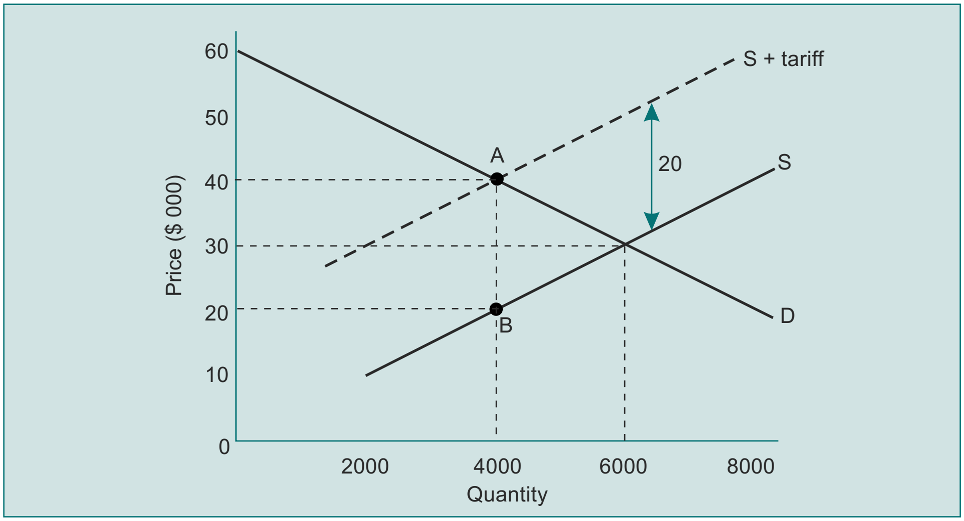


Figure 10.5 Impact of a tariff on price and output quantity

Now consider the country exporting the sports cars. Before the imposition of the tariff, that country was receiving $180 million ($30 000 × 6000) – we shall come to exchange rates shortly. If we imagine a two-country situation, that $180 million would have to be spent eventually on the exports of the tariff-imposing country, thus boosting its exports industry. After the tariff is imposed, how much will remain to buy these exports? The answer is $80 million: an output value of $40 000 × 4000 ($160 million) minus the amount paid in tariffs of $20 000 × 4000 ($80 million). The demand for the exports of the country imposing the tariff on sports cars will fall from $180 million to $80 million. Thus if we assume, in this simple example, trade has to be balanced, so that the value of exports must equal the value of imports, the imposition of a tariff will have no effect on this balance of trade but only on the volume of trade. Everyone, on average, suffers eventually by paying more for both imported and domestically produced goods and services; total world production is reduced; resources are not employed in the most efficient way globally. In the short run, however, those workers who remained employed in the 'protected' industry benefit: they don't have to go searching for new jobs and/or undertake retraining after shake-outs; the rest of society is subsidising such workers. The cost of protecting such domestic jobs is many times greater than the wages or salaries of the workers involved.

10.5.2 Quotas

A quota works in a similar fashion to a tariff but with one important difference: there is no tax advantage for the government. In the tariff example of the sports cars, given above, the government's coffers would benefit to the extent of £80 million ($20 000 × 4000), but there is no matching government benefit with a quota. The impact of a quota is shown in Figure 10.6; again we are dealing with the same sports cars.

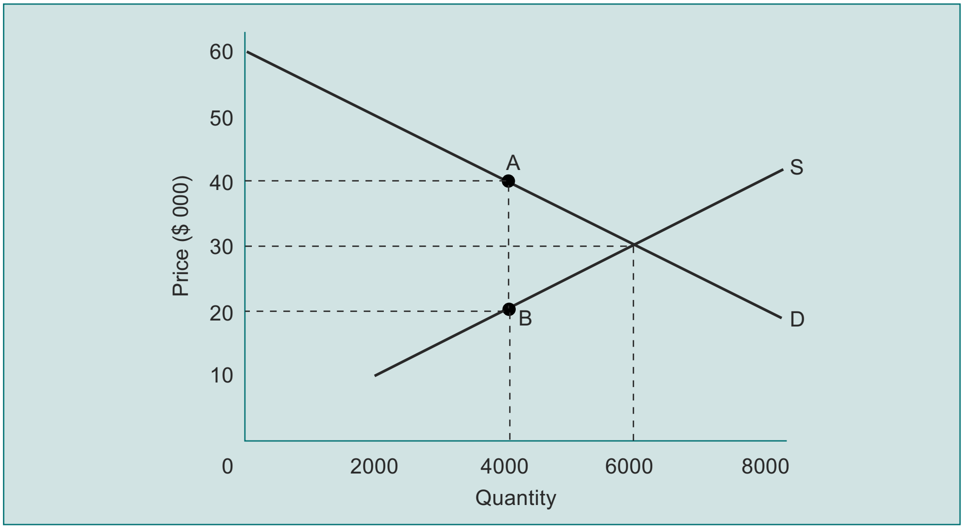


Figure 10.6 Impact of a quota on price and output quantity

Before the imposition of a quota, the equilibrium price and quantity exchanged of the imported sports cars are $30 000 and 6000 respectively. The government sets an annual quota of 4000 cars, and this will therefore be the quantity imported. The demand curve tells us, however, that if only 4000 cars are available, the equilibrium price is $40 000 (point A on the demand curve).

Importing sports cars is a highly profitable game. The car sells for $40 000 but only costs the importing firm $20 000 (point B on the supply curve). Thus at a level of sales of 4000 per annum, the sports car importer makes, or importers collectively make, $20 000 in excess profits per car ($40 000 − $20 000) for a total of $80 million ($20 000 × 4000). As you might imagine, there will be significant competition among car importers for the available quotas, and some method of allocating the quota among all the car importers will have to be found. Suppose the government were to sell the quota rights; how much would they be worth? The answer is, of course, $80 million – the amount of the above-normal profit. Remember the supply curve includes a normal return on assets in the firm.

Thus if an importer can negotiate an exclusive agreement with a foreign manufacturer to be the sole distributing agent for a certain sports car in the importing country, we have an interesting example of economic rent. The division of the rent will be the subject of negotiation, in this case the excess profit of $80 million, between the manufacturer of the sports car and the importing firm. Remember the manufacturer of the car is quite happy, given the supply curve, to supply 4000 cars at $20 000 each. But if that manufacturer can determine who the foreign distributor is going to be, he has monopoly power. He is the sole supplier – facing the known foreign demand curve.

The difference between the quota and the tariff situations in our example, therefore, is who gets the $80 million. In the tariff case it is the government; in the quota case it depends entirely on how the quota is allocated, since this determines how the economic rent is distributed. But note that in the quota case, just as in the tariff case, consumers, on average, are less well off compared with the free-trade situation. The car importers who are allocated the quota or part of it, will be strong supporters of protection for the domestic sports car industry; they might also be a promising source of political campaign contributions for any major political party!

10.6 Arguments for Trade Restrictions

Are there any economically legitimate arguments for the imposition of trade restrictions by a government? Many developing countries argue that as industries start to develop there must be a period of protection to allow them to reach a size and scale that is economically viable in the international competitive world. This is known as the *infant industry* argument. Until markets are established and scale benefits released, new industries in such countries cannot successfully compete with established firms in developed countries. Unfortunately, in practice it is politically much easier to establish tariffs and quotas than to abolish them.

A second argument concerns ‘dumping’. *Dumping* is the practice of selling a good in a foreign market at a lower price than prevails in the market of the exporting nation. The fear is that such a practice will drive domestic producers out of business and, when they no longer exist, the exporting nation will raise prices in the absence of domestic competition. Under international law, dumping is illegal and anti-dumping fiscal duties or taxes can be imposed on foreign producers if domestic firms can provide proof that dumping has affected their sales. This is a fruitful field for lawyers and accountants!

Third, tariffs may be imposed on imported goods produced by firms that receive production subsidies from their governments. Such tariffs are known as *countervailing duties* and are designed to cancel the effect of the government subsidy on the price of the imported good.

While all the above arguments can have long-term legitimacy, such that they protect consumers in the long run, they have short-term costs in that they make imported goods and services higher-priced than they would otherwise be. In practice, the most compelling reason by far for the imposition of trade restrictions lies in the fact that while *on average* everyone benefits from free trade, some people *lose*. The total benefits from free trade exceed the total costs incurred when trade restrictions are abolished but, at the margin, some people become worse off – significantly worse off if they lose their jobs and are not well qualified for alternative occupations. In contrast, those who benefit from free trade – the bulk of the population – each receives relatively small benefits, but when summed, these benefits far exceed the large losses borne by a few.

It is difficult to organise, as an effective political lobby, the large number of small gainers; it is much easier to organise the small number of large losers. Theoretically it is, of course, possible for the small number of losers to be fully compensated by the net gains from free trade – and through unemployment compensation, for example, losers (namely those whose jobs have disappeared) are partially compensated. In practice, however, it is not an easy task to determine why a particular firm has failed – because of foreign competition or, say bad management – and thus no attempt is made specifically to compensate resource owners whose business activities fail because of 'legitimate' foreign competition. As a consequence, protection remains a potent force in most capitalist countries.

As an experiment, the next time you are with a group of managers or business executives go ahead and argue the case for free world trade. Typically most of them will state that they agree with you; however, in a private conversation with any one of them you often hear the statement 'I agree with free trade in principle, but my industry is a special case because …'. There appear to be a lot of special cases around.

10.7 Exchange Rates

When an individual wishes to purchase a good made abroad, he must first buy some of the currency of that foreign country. Similarly, when a foreigner wishes to buy a good made in a home country, he must first buy some of that home country's currency. The market in which people buy and sell different national currencies is known, not surprisingly, as the *foreign exchange market*. When a UK firm wishes to buy US computers, it must first buy dollars; when a US car-importing firm wishes to buy British sports cars, it must first buy British pounds sterling. And the foreign exchange market is the mechanism for changing currencies.

Suppose, for example, the sports cars we were discussing earlier in this module sell in the UK for £10 000 each and the US importer wishes to buy 4000 of them; it thus needs £40 million to buy 4000 cars. Further, suppose that British Telecom wishes to buy 40 US-manufactured mainframe computers, each of which sells in the US for $2 million, totalling $80 million for the 40 computers. By chance, the finance directors of the US sports car dealer (SCD) and British Telecom (BT) meet. 'I'm looking for £40 million in exchange for dollars,' says SCD. 'I'm looking for $80 million in exchange for pounds,' says BT. They agree on a straight swap: £40 million for $80 million or $80 million for £40 million. Thus £40 million is the price of $80 million and vice versa. These prices are called *exchange rates*, and in our example are £1 = $2 or $1 = £0.50.

Considering only two currencies, dollars and pounds, what will determine the exchange rate? Like any good or service, the price (exchange rate) will be determined by the forces of demand and supply. What determines the demand for and supply of dollars? The number of dollars that UK households and firms want to buy for pounds is the demand for dollars, and the number of dollars US households and firms want to sell in exchange for pounds is the supply of dollars. Because Britons want dollars to buy US goods, the greater the amount of US goods that Britons want to buy, the greater will be the demand for dollars. The greater the quantity of British imports that US citizens and firms want to buy, the greater will be the supply of dollars.

How will the quantity of dollars demanded and supplied vary as the exchange rate changes? Suppose the price of dollars rises, i.e. the dollar *appreciates*. That means Britons are going to have to pay more pounds to obtain dollars and, consequently, in terms of pounds, American goods will become more expensive. Because a higher price for US goods exists, fewer will be bought; a higher price of dollars will mean a smaller demand for dollars.

However, a higher price for dollars also means Americans can obtain more pounds for one dollar. This means that British goods will be less expensive in terms of dollars and Americans will want to import more British goods. As long as the elasticity of demand for British goods is greater than one, i.e. as long as the dollar price times the quantity bought increases when the dollar price falls, the supply of dollars will increase.

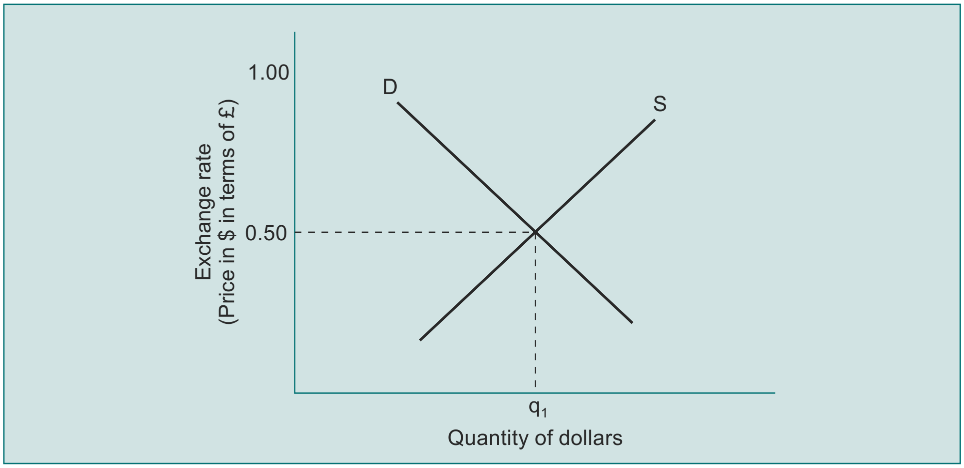


Figure 10.7 Exchange rate determination

Possible supply (S) and demand (D) curves for dollars are shown in Figure 10.7. As the value of the dollars in terms of pounds increases, the quantity of dollars supplied increases and the quantity of dollars demanded decreases. Where the two curves intersect, the quantity of dollars supplied will just equal the quantity of dollars demanded and the price of dollars in terms of pounds will be determined. Thus, as long as there is a free market in which dollars can be exchanged for pounds, the exchange rate will adjust to the level where the quantities demanded and supplied are equal. Such exchange rates are known as *flexible exchange rates*.

While many countries today have adopted a flexible exchange rate policy, this has not always been the case with the leading capitalist countries. The opposite of a flexible exchange rate system is a *fixed exchange rate* system, to which some countries adhere. Under a fixed exchange rate system, the governments involved agree to buy or sell sufficient quantities of their currencies to keep the exchange rate at an agreed value. Fixed exchange rates are demonstrated in Figure 10.8.

We shall continue our dollar/pound example because although the USA and UK today have a flexible exchange rate system, a fixed rate system existed until the early 1970s.

In Figure 10.8 the free-market equilibrium exchange rate is ER_2. Suppose, however, the US and UK governments agree to fix the rate at ER_1. At ER_1, the quantity of dollars demanded exceeds the quantity supplied by the amount cd. In order to keep the exchange rate from moving from ER_1 to ER_2, the US government would have to sell cd amount of dollars in exchange for pounds. Alternatively if the rate had been fixed at ER_3, the US government would have to buy ab dollars and pay for them with pounds to prevent the exchange rate from falling to the free-market rate of ER_2.

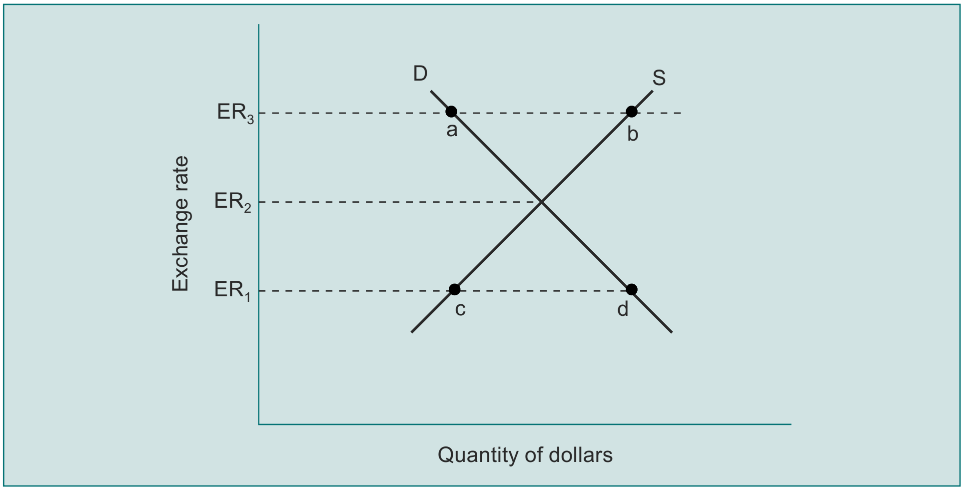


Figure 10.8 Fixed exchange rate

Now, if under a fixed exchange rate system there is an excess demand for dollars, the US government can always obtain enough dollars to sell on the foreign exchange market by borrowing internally and/or by raising taxes. However, when there is an excess supply of dollars at the fixed exchange rate and the government has to buy dollars, it will only be able to do so if it has sufficient reserves of pounds sterling. Suppose the US government runs out of sterling; what can it do? It has two options: it can borrow pounds from the UK government and use these pounds to buy the excess supply of dollars, or it can try to persuade the UK government to buy dollars to eliminate the excess supply. As you might imagine, limits exist to both of the above solutions as it is highly unlikely that the UK government would be prepared to loan pounds to the US government and/or continue to buy its currency *ad infinitum*.

In the world today there are many more currencies than dollars and pounds and consequently many more exchange rates than the dollar/pound rate. But the market for each currency operates just like the dollar market we have analysed in Figure 10.7 and Figure 10.8.

The move to flexible exchange rates amongst the world's major currencies in the early 1970s was designed to stabilise exchange rates and avoid large devaluation and revaluation, arising in the fixed exchanges by government edict. Such stability has not evolved, and the past four decades have witnessed greater swings in exchange rates than experienced during the days of fixed exchange rates. Why is this so? To answer the question we must drop the implicit assumption we have been making, that the demand for a nation's currency depends upon the demand for that nation's exports.

10.8 The Balance of Payments

A country's *balance of payments* accounts record its international trade, its international borrowing and its international lending. It must be understood that when we buy anything whatsoever from a foreign country, we pay for it in that country's currency.

Buying a good like a foreign sports car is straightforward, but what about a service? Well, if we buy a Volvo Turbo, it has to be transported to our shores by ship. It may be shipped by Wallenius Lines (a Swedish carrier), and Lloyds of London may be the insurer of that cargo; each is supplying a service for which the importer must pay in the relevant currency. What else can we buy abroad? When you are on vacation in Paris (assuming you are not a French resident), your hotel bill, your restaurant bill and your ticket for the Louvre are classified as the import of services to your own country – you pay for them in euro, which you acquired with your own currency. You might also buy a foreign firm, a holiday home in Spain, or some other real estate. If a Japanese car firm (again assuming we are not in Japan) builds a car plant in your country, the return on that investment to the Japanese owner is treated as an import of services from Japan.

The flows of foreign currencies required to carry out all these transactions are recorded in each country's balance of payments. Not all countries use identical accounting procedure, although they are broadly similar. Our exposition will be a simplified version of these complicated accounts. There are three major accounts in a country's balance of payments account; they are:

(a) the current account or trade account;
(b) the capital account; and
(c) the official settlements account.

The principal items in the current account are imports of goods and services, and exports of goods and services. Imports and exports, and any other items that do not result in the addition to, or subtraction from, a country's claims on foreign resources, are included in the current account. Thus if your country purchases foreign ski boots, gives donations to Save the Children Fund, or sells machinery or wheat to foreign nations, all these items would appear in the current account.

All transactions that affect the amount of claims that your country has abroad and that foreign countries have in your country enter the capital account, which records all international borrowing and lending transactions. Thus if the Japanese purchase a US government bond or the US builds a factory in Japan, these transactions would appear in the Japanese and American capital accounts.

The third account is the official settlements account, which shows the change in a country's holdings of foreign currency.

Table 10.5 shows a balance of payments account for a hypothetical country X, which is running a deficit on its current account, where the value of imports exceeds the value of exports by $100 billion.

Table 10.5 Country X's balance of payment account

	$bn
Current account	
Import of goods and services	500
Exports of goods and services	400
Current account balance	−100
Capital account:	
Foreign investments in X	300
X investments abroad	190
Capital account balance	110
Official settlements account	
Change in official X reserves	10
Net effect	0

This deficit has to be offset by foreign borrowing, which occurs in the capital account as part of foreign investment in X. The difference between the deficit in the current account and surplus on the capital account is accounted for by a change in holdings of foreign currency. Thus, by definition, the balance of payments must always be in balance.

Now, undoubtedly you have heard of *balance of trade* deficits or surpluses – that's simple, that's the current account. But what is the meaning of a balance of payments deficit or surplus? A balance of payments deficit can be thought of as the excess supply of a country's currency that results from international transactions and that the government must purchase if the exchange rate is to be preserved. Conversely, a balance of payments surplus is the excess demand for a nation's currency that the government must sell if the currency is to remain at its existing exchange rate. What happens if the government does not intervene, as many don't under a flexible exchange rate system? A balance of payments deficit in a country will cause the country's currency to depreciate, i.e. its price to fall. How far will it fall? It will fall, like any other price, until demand for and supply of the currency intersect at a new equilibrium price, the new exchange rate. Conversely a balance of payments surplus will cause a country's currency to appreciate against foreign currencies.

In the recent past, some major currencies, such as the US dollar, the Japanese yen and the British pound sterling, have fluctuated quite violently against each other, so that there has been significant appreciation and depreciation of these currencies. For example, in the early 1980s the dollar/yen exchange rate was 250 yen/dollar; a mere five years later, it was 170 yen/dollar and in the later 1990s, it was 130 yen/dollar. In a similar period of time, the British pound changed from $2.00/pound to $1.00/pound and by 2008 was back to $2.00/pound. As you can imagine, such changes cause havoc with the plans and prospects of multinational firms, firms highly dependent on imported resources, and major exporting firms. The reasons for such fluctuations in exchange rates cannot be blamed on fluctuations in the import or export of goods and services; their role is minor. The value of world trade today is a very small percentage of the value of world financial transactions. These flows are extremely volatile and, as a consequence, so are many exchange rates. Such fluctuations cause severe problems for firms involved in international trade.

A detailed study of the complicated world of international trade and finance is beyond the scope of this course. However, to give you a basic understanding of the complex world, we shall construct a simple model of a hypothetical country's international sector.

10.9 The Operation of the International Sector (Simplified Model)

We shall assume, in our simplified model, only two subsectors in our international sector, a trade sector (current account) and a capital sector (capital account). The balance of payments account will consist of the balance of trade plus net capital flows, which must sum to zero. Thus it is quite possible for an economy to have a large surplus (deficit) in the trade account offset by a large deficit (surplus) in the capital account.

Let's start off with the current account. What factors will affect the demand for a country's exports? We shall assume three major factors:

(a) foreign demand (Y_F), which we shall assume is a function of foreign national incomes;
(b) domestic (P_D) versus foreign (P_F) prices – the lower are our domestic prices, the higher will be exports;

(c) an exchange rate (ER) – if our domestic currency depreciates, foreigners will receive more of it for one unit of their own currency and consequently can buy more imports for some given amount of their own currency.

Thus we have in general terms the following expressions for X, a measure of exports:

$$X = f(Y_F, P_D/P_F, ER)$$

and

$$\frac{\Delta X}{X} = f(\Delta Y_F/Y_F, INF_D/INF_F, App/Dep)$$

where INF is the relevant inflation rate (both domestic and foreign), App is the currency appreciation factor, and Dep is the currency depreciation factor. The growth rate of exports will depend upon the growth rate of foreign income, relative inflation rates and the degree of appreciation or depreciation of the currency against a basket of foreign currencies.

What about imports (Z)? The factors affecting exports will be the same as the factors affecting imports except they act in the opposite direction and will depend upon domestic national income (Y_D) rather than (Y_F). Thus:

$$\frac{\Delta Z}{Z} = f(\Delta Y_D/Y_D, INF_D/INF_F, App/Dep)$$

Now let us suppose the foreign income growth rate increased from, say, 2 per cent to 6 per cent. This will boost our exports but have no impact on our imports, which are dependent on the growth rate in our national income. Further, suppose we have low inflation compared with foreign inflation. This will boost our exports and make imports less attractive. If, at the same time, our currency has depreciated, there will be a further stimulus to exports and a diminution in imports. Thus the trade balance could be very difficult to predict and, as we shall see in our consideration of macroeconomics (Module 11 onwards), could provide problems for policy makers.

Conversely the impact of relatively low inflation domestically could be offset by appreciation of our currency, which would make exports less attractive and boost our imports. A critical question is what determines the exchange rate. Well, we know the answer to that; it is the demand for and supply of our currency. We have already dealt with demand for and supply of a currency in the current account, but what about the capital account? Here is where the problems arise.

We shall assume that two factors affect the demand for our currency in the capital account. They are the domestic (R_D) versus foreign (R_F) interest rate and expectations about whether our currency will appreciate or depreciate. Suppose our interest rate (R_D) is 10 per cent and foreign interest rates are all 5 per cent. Managers of a pension fund in Japan with several billion yen to invest can double their rate of return by buying our currency and investing their funds in our money markets. Is there any risk in doing so? The answer is 'yes': our currency may depreciate after the investment is made, thus effectively diminishing the interest-rate differential. Of course, our currency may appreciate; if this happens, the Japanese pension fund managers are laughing all the way to their bank, because they gain on both the interest rate differential and the appreciation of our currency. As illustrated, expectations about potential appreciation or depreciation of our currency will play a critical role in the demand for our currency. The supply of our currency may affect these expectations and as a consequence *demand for and supply of our currency will not be independent of each other.*

10.9.1 An Example of International Action and Reaction

Hopefully the example we are about to set out will be more easily understood after you have studied the macroeconomics modules (Module 11 onwards). Accept the following in blind faith. A country is experiencing high inflation and the government states that it is going to restrict the money supply to cure the inflation. The restriction of the money supply means interest rates are going to be 'high' for the next year or so. Suppose foreign investors believe our government. They decide to invest their funds in our financial markets. That means an increase in demand for our currency and a potential appreciation of our currency on foreign exchange markets. Now suppose other foreign speculators believe foreign fund managers are going to invest in our country because of the interest rate differential and predict our currency is going to appreciate. They may transfer funds into our financial markets, not because of the interest rate differential *per se* but because of the expected appreciation of our currency. And all of this has been caused by the announcement by the government of its intention to cure inflation by restricting the money supply. Of course, if they don't believe our government and/or they think the policy will not work, they may divert their funds to other financial markets.

Now suppose it is rumoured that, in light of the coming election, the government is going to adopt policies to lower interest rates and make business and mortgage holders happy. What would you do if you were a foreign pension fund manager with lots of money to invest to make the highest possible return? With interest rates about to come down in our hypothetical country, it is not such an attractive haven for foreign funds, so you look to other countries. At the same time, you might think that other fund managers will be thinking along similar lines and so the quicker you withdraw funds from our country the better. Why? Because if every fund manager is thinking along the same lines, the withdrawal of these funds, i.e. a shift to the left of the demand curve for our country's currency, is going to cause the currency to depreciate. Thus the expected decrease in interest rate caused by an increase in money supply will cause the foreign demand for our country's currency to shift to the left and could result in a dramatic depreciation of the currency.

Thus an unfortunate exporting firm ordering foreign raw materials or component parts for future delivery – say two years from now – has a problem (or several problems). If it has little idea of what the exchange rates are going to be in the next two years, what prices for imported component parts designated in foreign currencies, and what prices for finished goods for sale abroad designated in its own currency, is the exporting firm prepared to accept? If it guesses wrong and its competitors guess right, it could go bankrupt; conversely it could make a fortune. One way to minimise risk is to buy foreign currencies in the *futures market*[2] and have speculators take the risk if the exchange rate fluctuates, but that involves additional costs. So what is the answer?

Economists and politicians are sharply divided on the solution. Those who believe in the efficacy of the market mechanism argue that a flexible exchange-rate mechanism is the only answer and that if governments were to embrace 'sensible' economic policies, flexible rates would enhance world trade and lead to a globally efficient allocation of resources. Their opponents argue that a fixed exchange-rate system is necessary to produce worldwide economic efficiency and that the leading world economies – the USA, China and the Eurozone – should stabilise international exchange rates. They further argue that the smaller

[2] Futures markets are markets in which buyers and sellers of currencies contract today on exchange rates and quantities of currencies to be exchanged at some future date.

trading nations, essentially the rest of the world, should tie their currencies to one of those major partners and enact policies to ensure exchange-rate stability.

The international trade and international finance world straddles the micro- and macroeconomics world, and so we had best now turn our attention to macroeconomics.

Learning Summary

You are now able to distinguish between the theory of absolute advantage and the theory of comparative advantage, and understand why all countries can gain from international trade. You understand why trade restrictions lower the volume of exports and imports, and consequently the world's consumption possibilities but may benefit a few protected workers at the expense of many.

You can explain how international trade is financed, what constitutes the balance of payments, what determines the amount of international lending and borrowing, and what determines exchange rates.

Review Questions

Multiple Choice Questions

10.1 Two countries, X and Y, can produce only wheat and cloth according to the following schedule:

	Marginal costs of production	
Commodity	**Country X**	**Country Y**
A unit of wheat	1 man-day	2 man-days
A unit of cloth	3 man-days	4 man-days

Assuming no other production costs, transportation costs, or trading restrictions, which of the following is correct?

A. Country X will export wheat and import cloth.
B. Country X will export both wheat and cloth.
C. Country X will neither export nor import wheat or cloth.
D. The pattern of trade cannot be determined from the above information.

10.2 Higher living standards throughout the world would be made possible by

I. an increase in the skills of the labour force in each country.
II. an increase in the stock of capital goods in each country.
III. an increase in protective tariffs in each country.

Which of the following is correct?

A. I only.
B. I and II only.
C. I and III only.
D. I, II and III.

10.3 Suppose that diplomatic initiatives by the United States led to the opening of extensive trade with Russia. Which of the following best describes the effects this would have on the United States?

A. Incomes in export industries would rise, but not by as much as income in import-competing industries would fall.
B. Incomes in import-competing industries might fall, but not by as much as income in export industries would rise.
C. Per capita income would rise as long as income in both export- and import-competing industries would rise.
D. Incomes in both export- and import-competing industries could either rise or fall.

10.4 If all restrictions to international trade were removed, there would be

I. higher average living standards.
II. lower living standards for some persons.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

10.5 'The possible gains from trade are greatly exaggerated. If a country's exports and imports are equal, it means that it is exchanging with the rest of the world goods of equal market value. A country can only gain if it imports more than it exports and, for the world as a whole, imports and exports must be equal.'

Which of the following makes this statement incorrect?

A. It is possible that traders would be willing to give up more than they have to for what they get.
B. An economy can gain from trade only if export earnings exceed payments for imports.
C. Equality of imports and exports does not necessarily imply exchange of goods of equal market value.
D. Imports add to and exports subtract from the amount of useable output available to an economy.

10.6 If, because of continued UK balance of payments surpluses, the pound sterling appreciated, the effect on UK residents would be that

I. a holiday in Europe would cost less.
II. imported goods would become cheaper.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

10.7 Suppose that the pound sterling on foreign exchange markets was being traded for 2.5 Swiss francs, and further suppose that the following day the exchange rate was one pound sterling for 2.8 Swiss francs. Which of the following would be the immediate effect of the change in the exchange rate between pounds sterling and the Swiss franc?

A. Swiss goods would become more expensive for British consumers.
B. Swiss goods would become less expensive for British consumers.
C. Swiss goods would become neither more nor less expensive for British consumers.
D. British goods would become less expensive for Swiss consumers.

10.8 If Europe had fixed exchange rates with all other trading nations, and if its imports exceeded its exports by €4 billion, and if capital investment abroad by its residents exceeded foreign capital investment in Europe by €5 billion, there would be a net inflow of foreign exchange into Europe in the absence of any other foreign transaction.

Which of the following is correct? The analysis above is

A. correct because the inflow from foreign capital investment would exceed the outflow from the trade deficit.

B. incorrect because whenever imports of goods and services exceed exports of goods and services the balance of payments must be in deficit.

C. correct because the net foreign capital investment would increase Europe's assets abroad by €5 billion, which would exceed the trade deficit.

D. incorrect because the deficit in the balance of trade plus the net outflow in the capital account would cause an outflow of foreign currency holdings.

10.9 It is possible for a country to run a persistent and sizeable trade deficit over several years only if

I. its currency is constantly appreciating.

II. it is a consistent international borrower.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

Case Study 10.1: A Tale of Cigarettes and Whisky

This case is concerned with the benefits accruing from exchange and how they are distributed. Before you tackle it you should understand:

a. the benefits from exchange

b. the rate of exchange.

The case demonstrates that when exchange takes place, all parties can be made better off. It shows that all parties to a deal can conclude that they have struck a highly profitable bargain.

A Tale of Cigarettes and Whisky

A US Strategic Air Command base was established on an island off the north-west coast of Scotland. Each US airman on the base was given 400 cigarettes a week as part of his living allowance. Spirits, however, were not available on the base.

The principal occupation of the island's residents was the distillation of malt whisky. The 200 employees in the distillery were each given 10 miniature bottles of whisky per week as part of their wage. Because of general shortages, cigarettes were almost unavailable to the distillery workers.

US security regulations prohibited contact between the Americans and the locals. This led to the situation where the Americans had plenty of cigarettes but no whisky, and the islanders had plenty of whisky but no cigarettes.

Joe Bloggs, the local MP, was the only Scot who had free access to the airforce base. Being a born politician, he asked a group of US airmen how he could make their stay on the island more pleasant. They replied that they would be most appreciative if they could obtain some of the island's excellent whisky and stated that they would be willing to trade two packets of cigarettes for every miniature bottle of whisky delivered.

However, what they did not tell Joe was that at the limit they would in fact have given up three packets of cigarettes for every miniature bottle of whisky.

Joe left the base with 400 packets of cigarettes one night. At a local political rally the following day, he asked a group of distillery workers at what rate they would be prepared to swap whisky for cigarettes. Sensing that a deal was in the wind and that they were dealing with a wily politician, they conferred and decided that one packet of cigarettes was worth three bottles of whisky; however, they told Joe that they were willing to give up one miniature bottle for one packet of cigarettes. Joe immediately produced 300 packets of cigarettes and returned to the airforce base with 300 miniature bottles of whisky. He turned over to the airmen 200 of the bottles, since this was at the agreed exchange rate, two packets of cigarettes for each bottle. The airmen were quite delighted and thanked Joe most profusely for his services. The fact that he had not attempted to haggle over the deal was taken as a token of his concern for their welfare and was viewed by the Americans as a sign of his being a true Scottish gentleman. Back at the distillery the canny Scots were smoking and drinking in celebration at having pulled the wool over Joe's eyes. They were amazed that Joe had not attempted to haggle and reckoned that they had struck a remarkably good bargain. Joe, meanwhile, was resting very comfortably in his hotel room surrounded by 100 packets of cigarettes and 100 bottles of whisky.

1 Consider, then, the following:

1. Must this mean that either the Scots or the Americans or both were worse off? That did not seem to be the case. How can this be?
2. When regulations were relaxed on the island relating to the fraternisation of Americans and the islanders, on what terms would you expect to observe Americans and Scots exchanging cigarettes and whisky?
3. What happened to Joe's standard of living at the point described in consideration 2?

Module 11

Macroeconomics Overview

Contents

This module provides the same kind of overview for macroeconomics that Module 2 provided for the whole course. As in Module 2, there are no review questions.

11.1 Introduction

Macroeconomics is about the 'big picture'; it describes and explains the working of the whole economy. *Microeconomics*, as we have seen, is about individual parts that make up the big picture; microeconomics describes and explains the working of the units that make up the whole economy – the household, the business firm and the government.

Macroeconomics tries to explain the behaviour of the level of prices of all goods and services taken together; the change in this price level from one period to another is known as the *inflation rate*. *Deflation* occurs when the price level falls – a rather infrequent event. The price of one good compared with another good is called the *relative price*; microeconomics attempts to explain relative prices.

Furthermore, macroeconomics is concerned with the behaviour of total income and total output for the whole economy; microeconomics is concerned with income and output of the individual units that comprise the total economy. Macroeconomics studies the employment and unemployment rates for the economy as a whole; microeconomics is concerned with the hiring and firing practices of individual firms.

11.2 Potential and Actual Output

In Figure 11.1 the nation's resources are represented by the two rectangles. The first is the capital stock consisting of natural and man-made resources. Natural resources comprise agricultural land, forests, mineral deposits, rivers and fish stocks, in other words the resources nature has bestowed on a nation. Man-made resources consist of all that we have inherited from past generations, such as roads, harbours, houses, factories and machine tools. The second is the labour force, consisting of that proportion of the population able and willing to work at going wage rates and working conditions. Part of the labour force is made up of entrepreneurs and managers, whose task is to combine capital and labour to produce goods and services.

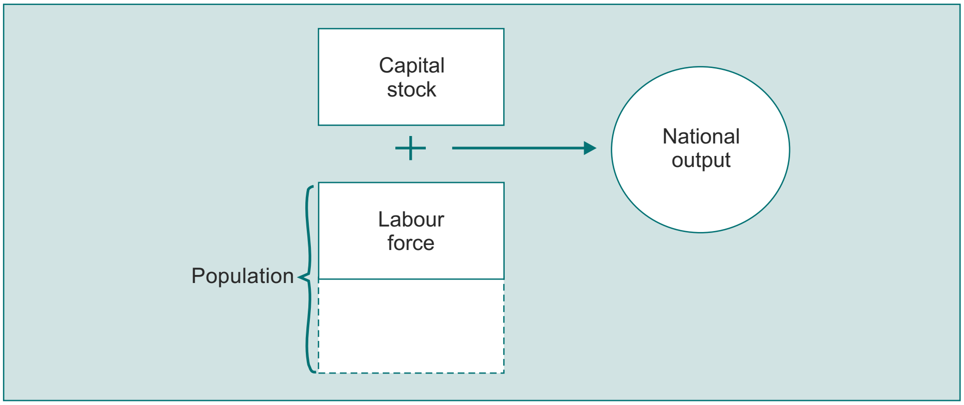


Figure 11.1 Potential output

How large total national output *can* be, therefore, depends upon the quality and quantity of the capital stock, the quality and quantity of the labour force, and the existing technology. The total output of goods and services is represented by the circle, national output, in Figure 11.1.

The world today is a very ill-divided place in terms of allocation of capital and labour, and consequently in terms of world output. Figure 11.2 shows the distribution of world output and world population. As can be seen, a relatively small proportion of the world's population (19 per cent), consisting of the people of North America, Europe and Japan, produces and consumes over 80 per cent of the world's output. It follows that average material living standards vary enormously by country throughout the world.

Average living standards for a nation can be found by dividing a nation's total output by its population. Remember we are talking about *averages*, which do not take into account the distributions within nations. There can be, and are, some very low incomes and some very high incomes within any nation, no matter how relatively high is the average national output per capita. The USA today has many millionaires; it also has 35 million people who are classified as living below the US poverty line. As we shall see, there are problems in accurately measuring a nation's output and, consequently, even more problems in making accurate international comparisons.

One method of comparing across countries makes use of a technique called *purchasing power parity* (*PPP*). By this technique the cost of a similar basket of goods in different countries is compared and that cost is translated into a common currency such as the dollar.

The stark realities of the plight of nations at the lower end of the output league have become all too evident in the past decades through the power of television reporting. But these are not new problems; a considerable portion of the world's population continues to live in abject poverty – we are just more aware of such poverty today. The long-term solution to such poverty is not the transfer of food and clothing from the rich to the poor nations but rather assistance in establishing a capital stock and a labour force over time to permit the poor nations to become self-sufficient. In the short run there is virtually nothing a nation can do internally to acquire resources, but can it avoid wasting existing resources?

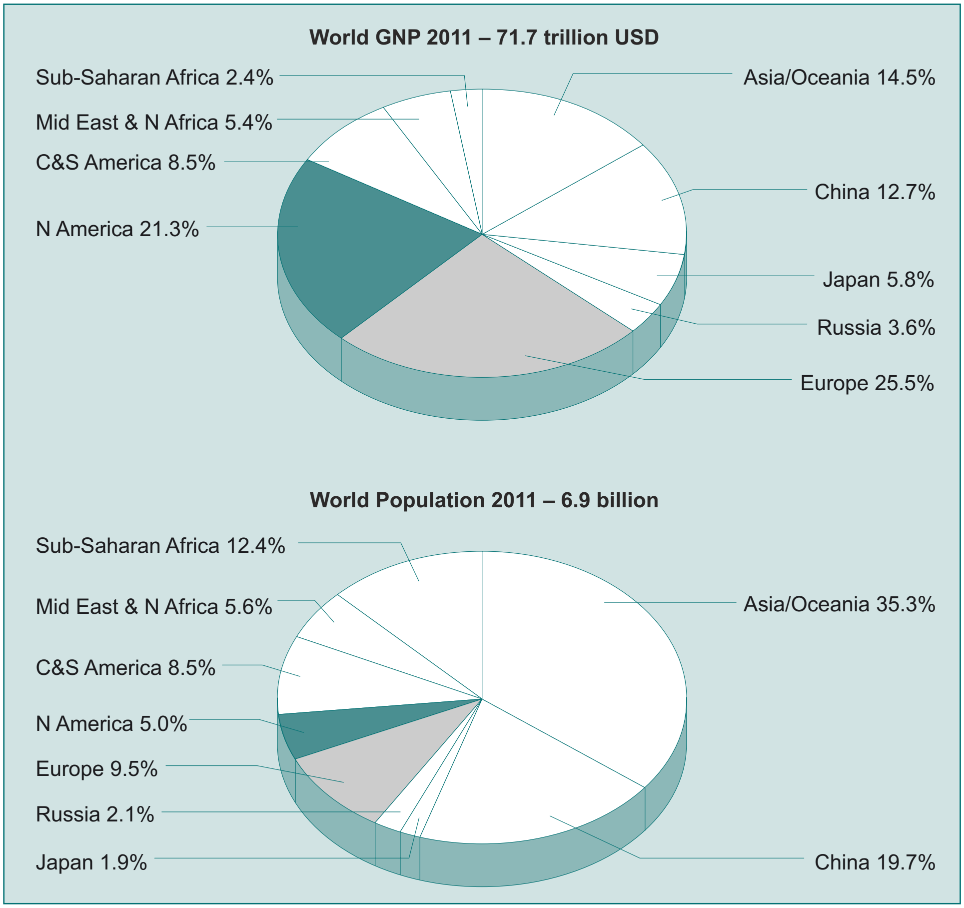


Figure 11.2 World output and population in 2011

Figure 11.3 shows something quite obvious. If part of the capital stock (factories and machine tools) is idle and part of the labour force is unemployed, national output will be less than its potential. In other words, some goods and services that could have been produced with the idle resources will be lost and gone for ever. As a result, people will be less well off than they otherwise could have been. Thus having idle resources does not appear to make any economic sense. Yet the history of market economies shows many periods when actual output was less than potential output. As we shall see, matching actual and potential output in a nation over time is not an easy task. Let us consider both potential and actual output and the factors determining them.

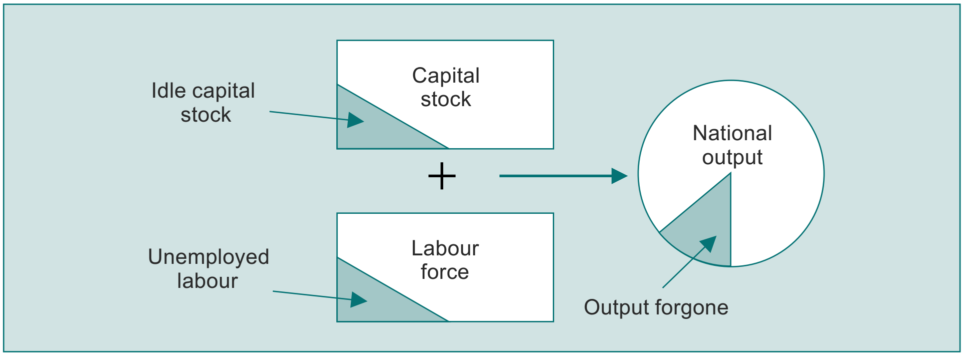


Figure 11.3 Unemployed resources

11.2.1 Potential Output

Potential output is the output that the economy would produce if both the labour force and the capital stock were fully employed. As we shall see, the idea of *full employment* does not mean that every adult works for 24 hours a day, 7 days a week, 52 weeks of the year. In order to be realistic, full employment is defined as a percentage of people in the labour force being employed for an expected number of hours per week (a normal work week) for an accepted number of weeks per year, an allowance being made for recognised holiday periods. An additional allowance is made for 'frictions' in the labour market, i.e. the time required to match people with jobs. As a result, each nation has a *full-employment rate of unemployment* (also known as the *natural rate of unemployment*), and this rate varies according to the socio-economic characteristics of each nation's labour force. Each nation also has a concomitant *full-employment rate of down-time* (or unused capacity) for factories and machine tools.

The potential output of an economy depends upon the quality and quantity of the capital stock and the quality and quantity of the labour force and existing technology. Fortunately for society, potential output is not static; normally, it grows over time. This expansion is caused by:

(a) growth in the quality and quantity of labour;
(b) growth in the quality and quantity of the capital stock; and
(c) technological advances.

Figure 11.4 shows a simple model of potential output growth. Part of the national output in year t consists of new factories and machine tools, i.e. new capital goods. They are added to the capital stock of year t, producing a larger capital stock CS_t the following year. In addition, these new machine tools can embody the latest technological advances, e.g. faster computers, more advanced engines etc. Thus a nation that allocates a large part of current output to the production of capital goods scores on two counts in terms of economic potential: it will possess more capital goods, and it will have more technologically advanced capital goods. For most nations the labour force (human capital) grows each year, albeit slowly; therefore the more current resources are devoted to educating and training labour, part of CG_t, the higher will be the quality and potential of the labour force and, consequently, the higher will be the nation's potential output in the future.

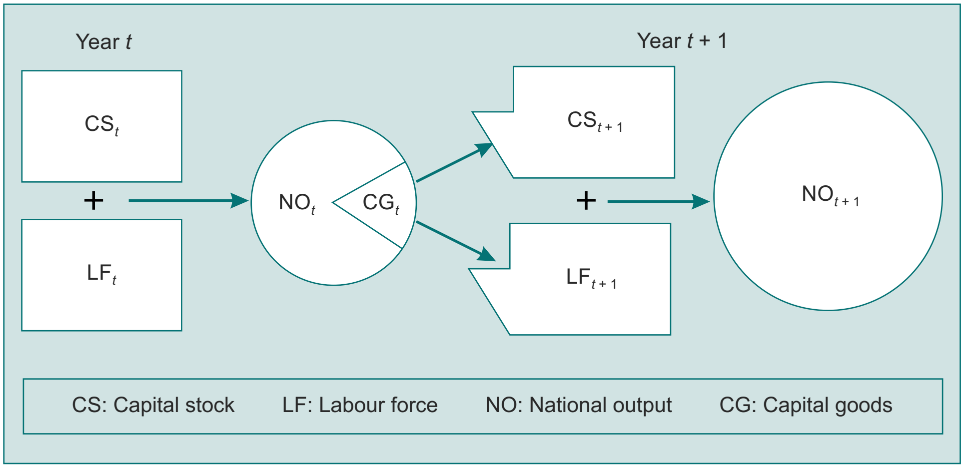


Figure 11.4 Growth in potential output

The proportion of current output devoted to new capital goods varies widely among the leading capitalist nations in the world. For several decades in the late 1900s, Japan led the way, allocating as much as 35 per cent of output in some years to new capital goods. The USA and UK have lagged behind most of the developed countries, averaging respectively 15 per cent and 18 per cent in the same period. In many of the poorest nations in the world, the proportion of total output allocated to capital goods is negligible. Similar wide disparities exist in the proportion of resources devoted to training and educating labour forces.

The cruel dilemma facing poor nations is that resources allocated to capital goods, education and training today means even greater deprivation for the current generation. In the short run, there is little a government can do to influence a nation's potential output; it is dominated by the existing quality and quantity of the capital stock and labour force. Even if a nation were to double the amount of resources allocated to capital goods production, the impact this would make on the capital stock the following year would be negligible. It is only if such a policy were pursued over a number of years that the impact on potential output would become significant. Similarly with the labour force: unless there is a dramatic change in immigration or emigration policies, the quality and quantity of the labour force is determined primarily by past birth rates, participation rates, retirement practices, schooling and training. Thus in countries like the USA and the UK, potential output today grows at a rate of 2.5–3.0 per cent per annum; in emerging capitalist nations such as India and China, current growth rates in potential output have reached double digits similar to the rates achieved by Japan in the 1960s.

If such growth rates are sustained the combined GNPs of China, India, Latin America, including the Caribbean (the 'industrialising countries') will be approximately equal to the combined GNPs of the USA, Western Europe and Japan (current 'industrialised nations') by the middle of the century.

11.2.2 Actual Output

The actual output of an economy is known as the *gross national product* (*GNP*); it is the value of all final goods and services produced in the economy in a year. It is calculated by multiplying each good and service produced by its price and adding them all together, so that

$$\text{GNP} = \sum(ap_a + bp_b + \cdots + zp_z)$$

In words, GNP equals all the tons of apples produced times the price of a ton of apples plus all the bottles of beer produced times the price of a bottle of beer plus all the kilos of zucchini produced times the price of a kilo of zucchini, plus the sum of all other national goods and services produced multiplied by their individual unit prices.

In the calculation of GNP each good and service is thus weighted by the number of physical units produced and by its price. For example, if in a nation's GNP there was one car that sold for \$40 000 and 40 000 hamburgers that each sold for \$1, the car and the 40 000 hamburgers would have equal weight. Why? The answer is that if you bought the car for \$40 000 you would be able to exchange it for 40 000 hamburgers. In this case \$40 000 and \$1 are the respective equilibrium prices of cars and hamburgers, reflecting the resources required to produce them.

For the economy as a whole, there is a corresponding flow of income to resource owners matching the value of all final goods and services produced. This is shown in Figure 11.5. To simplify the national output and income flow figure, we shall ignore temporarily government action and the international sector. In the simple model shown, all resources are owned by households, some of which form entities called firms. Those firms hire resources owned by households and produce goods and services, which in turn flow back to the households. The firms pay resource owners for the use of their resources in the form of wages and salaries for labour services, rent for the use of land, and interest and dividends for the use of capital. The households, in turn, pay firms for the goods and services produced and bought.

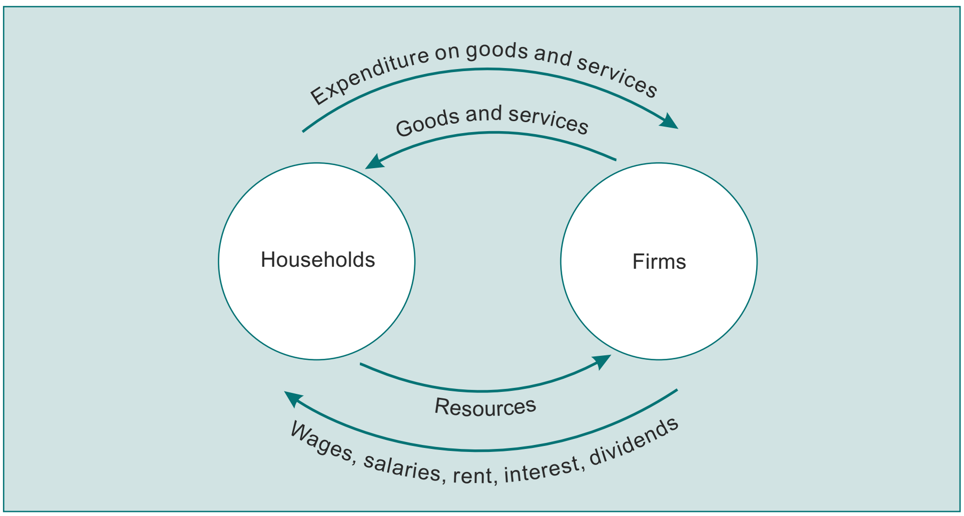


Figure 11.5 National output and income

Gross national product (GNP) is the value of the final goods and services produced. *Gross national expenditure* (*GNE*) is the value of total spending by the households. *Gross national income* (*GNI*) is the value of the services of the factors of production hired by the firms. These are shown in Figure 11.6.

Thus there are three ways to measure economic activity in a nation each year. We can look at the total output of goods and services produced (GNP), the total expenditure made by households for these goods and services (GNE) or the total income flowing to resource owners (GNI). Since they are all measuring the same thing, economic activity – albeit each from a different perspective – they must be identical.

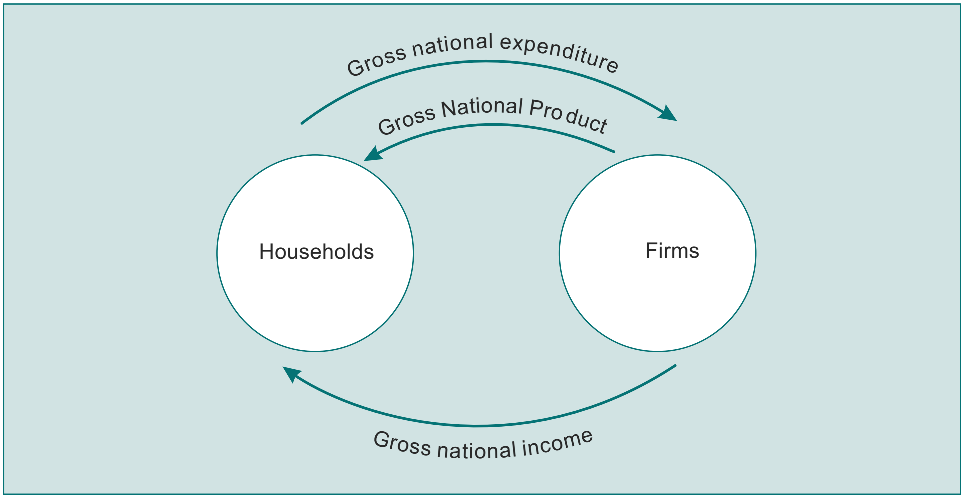


Figure 11.6 National output, expenditure and income

Remaining with our simple model, we can look at the economy from the viewpoint of firms and from the viewpoint of households. There are only two types of goods that firms can produce: consumer goods (and services) and capital goods. Thus

$$GNP = C + I$$

where C = consumer goods produced and I = investment (capital) goods.

Similarly, there are only two things households can do with their incomes: spend them on consumer goods and services and/or save them. Thus

$$GNI = C + S$$

where C = consumer goods purchased and S = savings.

Now, we know GNP must equal GNI. However, the decision on the production of consumer and capital goods is made by firms, whereas the decision on how much to spend on consumer goods and how much to save is made by the households. It would be quite remarkable if the amount of consumer goods produced by firms always matched the amount of consumer goods that households planned to purchase. If, however, this were to happen, it tumbles out of the above equations that the amount of capital goods a nation produces each year to add to its existing capital stock will depend upon the proportion of national income that households elect to save. Thus high-saving nations like Japan will be able to allocate substantial resources to the production of investment goods, while high-consumption nations like the USA will have low investment rates. Such propensities will have long-term implications for economic growth.

Let us return to our two equations. What will happen if the amount of consumer goods produced by firms exceeds the amount of consumer goods households wish to buy? The answer is that firms will discover that they have unsold goods, their inventories will be higher than they would wish, and when they come to reorder they will order fewer goods than previously. This effect will ripple down from shops and stores to wholesalers and eventually to factories, which will then produce fewer goods. But producing fewer goods means a decrease in output, and taking all firms together means a fall in GNP. Concomitantly, firms will not need as many resources to produce the smaller output and will therefore hire fewer resources – there will be unemployment of labour, unused capacity in factories and GNI will fall. With lower incomes in total, households will buy fewer goods and services and our economy will be heading for a recession. The gap between potential and actual output will widen.

When will GNI and GNP stop falling? When inventories fall below desired levels, the whole process will be reversed: orders will increase, output will rise and unemployment will fall. However, we shall have lived through a period of waste, a period when actual output fell below potential, a period when unemployment rose and a period when society was less well off than it otherwise could have been.

Back to our equations again. What happens if the amount of consumer goods and services produced by firms is less than that which households wish to buy? The first thing that firms will experience is sales from inventories to meet the unanticipated demand. They will step up their orders, which will ultimately be reflected in firms hiring more resources to produce more goods and services. This will mean an increase in GNP and GNI and a further round of increased demand from the higher GNI. Now the economy is headed for a 'boom' and the expansion will only be constrained by the economy's potential, i.e. by the capital stock and the labour force. When will the boom end? What we have learned in microeconomics can now be applied to the economy as a whole: if total demand rises faster than supply, prices will rise; at higher prices households will buy fewer goods and services, inventories will eventually exceed desired levels, orders will decrease and the cycle will be repeated.

These processes are the essence of the so called *business cycle*. Fortunately, the government has tools at its disposal to help alleviate this cyclical tendency: this is the subject of macroeconomic policy making. Before we reach policy making, however, we must know much more about the principal players in the macroeconomic world and the manner in which the tools available to government impact on the behaviour of these players.

11.2.3 Gross National Product versus Gross Domestic Product

While this text uses gross national product as the measure of an economy's output throughout the macroeconomics portion of this course, some readers may also have come across the measure known as *gross domestic product* (*GDP*).

We have already defined GNP as the value of all final goods and services produced in the economy in a given period, and GDP measures the same thing. The difference is that GDP defines the economy as existing within the country's borders, while GNP defines the economy according to the ownership of factors of production. While GDP measures the output of an economy based on location, GNP measures the output of an economy based on ownership. In a closed economy there is no difference between GDP and GNP, because all factors of production are owned domestically, but for an open economy (one in which international trade takes place and multinational companies operate) the two are different.

In the case of an economy with production interests abroad, the output from these production interests is counted as part of GNP but not as part of GDP. Similarly, when a country has foreign-owned production assets located within its borders, the output from these facilities counts as part of GDP but not part of GNP. So, for an open economy

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

where Net Income from Abroad consists of all production abroad by domestic companies minus all domestic production owned by foreign companies.

Table 11.1 GDP versus GNP for the G7 economies, 2010, USD trillions

	GDP	GNP
Japan	5.03	5.17
Germany	3.33	3.38
France	2.62	2.67
United Kingdom	2.17	2.20
Canada	1.34	1.32
Italy	2.11	2.08
United States	14.04	14.01

Table 11.1 shows GNP and GDP data for the G7 economies in 2010. It can be seen that the numbers are in fact very close to one another. It can also be observed that Japan, Germany, France and the UK all have GNP greater than GDP; their net income from abroad is positive. Canada, Italy and the USA are in a situation where their GDP is greater than their GNP; their net income from abroad is negative.

Newspapers and politicians will quote GDP or GNP data depending on which better suits their purposes at the time. You may come across GNP data in the press referred to simply as 'national income'. For the purposes of teaching the fundamentals of macroeconomics, the distinction is purely academic; while some textbooks favour the use of GNP, others use GDP. In this text we use GNP exclusively as the primary measure of production.

11.3 The Demand for Gross National Product

Gross national product is purchased by four separate groups. The sum of the expenditure of the four groups is known as *aggregate demand*. The four groups are:

(a) consumers (households)
(b) firms
(c) government
(d) international (households, firms and governments).

Each of these groups' expenditures can be affected by governmental policies; consequently the policies enacted by government can influence aggregate demand.

If we assume potential output in the short run is fixed, the ability of a government to affect aggregate demand means that government policies can affect the unemployment rate and the inflation rate. Unfortunately, many of the policy tools at a government's disposal act with a lag and as a result it is extremely difficult to keep aggregate demand constantly at the desired level. In addition, as we have seen, potential output normally increases continuously; thus the government is constantly pursuing a moving target.

The whole process is complicated by two additional factors. The first is that households and firms have minds of their own. Household expenditure patterns are certainly more predictable than those of firms, but often governments under- or overestimate household expenditure out of additional income. And firms' investment expenditure is the most volatile component of aggregate demand. A significant component of firms' investments is not geared to the current state of the economy but depends on managers' long-term expectations; such expectations vary by industry and are well-nigh impossible to predict.

The second factor that is completely unpredictable takes the form of exogenous shocks. These are outside-world events over which policy makers have no control but that can have a significant impact on an economy. For example, an effective cartel of the oil-producing countries can create havoc (and indeed has created havoc in the past) in oil-importing countries. What happens in the short run when such a cartel decides to fix the price of world oil at, say, twice the current level? The price of oil products will rise dramatically, and given that the demands for oil and oil-related products are price-inelastic, countries that have large oil imports will have less income to spend on all other goods and services. Demand for these goods and services will fall and unemployment will rise, while at the same time the high oil prices cause price rises in oil-using and oil-related products.

Thus there are some areas of economic activity difficult to predict and other areas impossible to predict. This makes the task of government policy makers both challenging and politically hazardous, and such policy makers are ultimately governments appealing to electorates for political survival. Not surprisingly, one of the dominant factors affecting the outcomes of elections is the state of the economy.

Before we review the tools available to governments, let us restate the macroeconomic relationships. As we have discussed, GNP must equal GNE must equal GNI. Let the symbol Y stand for Gross National Product. Thus

$$\mathrm{GNP} \equiv \mathrm{GNE} \equiv \mathrm{GNI} \equiv \mathrm{Y}$$

(NOTE: ≡ indicates an identity, i.e. something that is necessarily true by definition.)

For a domestic economy, aggregate demand equals consumption expenditure (C) plus investment expenditure (I) plus government expenditure (G). Thus domestically

$$\mathrm{GNP} \equiv \mathrm{Y} \equiv \mathrm{C} + \mathrm{I} + \mathrm{G}$$

However, in an international setting, part of a nation's resources are allocated to the production of export goods (X) and simultaneously households, firms and governments may buy goods from foreign nations, in other words imports (Z).[3] Thus internationally

$$\mathrm{Y} \equiv \mathrm{C} + \mathrm{I} + \mathrm{G} + \mathrm{X} - \mathrm{Z}$$

This last equation is known as the *national income identity*, and policy makers must never lose sight of it. Any policy action they take will have an impact on at least one of the components of Y, which in turn will have an impact on Y itself, which in turn will have a further impact on each of the component parts. These 'ripples' will have further impacts on Y and its components.

11.4 Policy Tools

The policy tools that governments control fall into two categories: *fiscal policy* and *monetary policy*. *Fiscal policy* involves control of government expenditure and tax rates. Tax rates take many forms, for instance income taxes, sales taxes, customs and excise taxes and corporation taxes. *Monetary policy* involves control over the supply of money, which directly affects interest rates. Let us examine briefly how each of those policy tools works.

Consider the simple representation of an economy in Figure 11.7. If the level of aggregate demand were D_1, GNP would be Y_1 and the rate of employment E_1. If the level of aggregate

[3] Why, you might ask, am I using 'Z' for imports? Well, I have used 'I' for Investment, I am keeping 'M' for Money and 'Y' for GNP, and that leaves 'Z'.

demand were D_2, GNP would be Y_2 and the rate of employment E_2. Finally if aggregate demand were D_3, GNP would be Y_3 and the rate of employment E_3. Why cannot this process continue to infinity? The answer, of course, is that the system is constrained by potential output. Let us call potential output Q. Remember Q will be determined by the quantity and quality of the capital stock and the quantity and quality of the labour force. Let us suppose $Q = Y_3$ as in Figure 11.8. If the level of aggregate demand were D_3, then the economy would be at full employment, with $Y = Q$ and the employment rate E_3. In this case the unemployment rate would be the natural rate of unemployment.

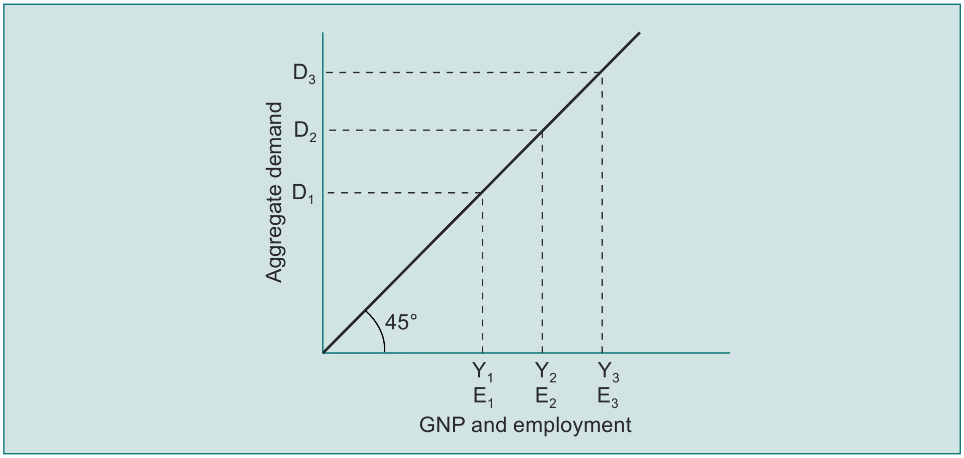


Figure 11.7 Simple model of an economy

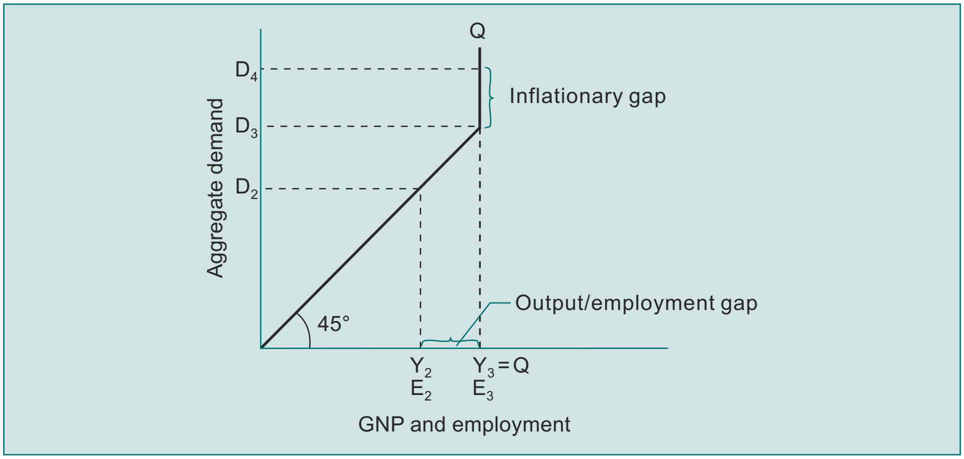


Figure 11.8 Simple model of an economy, including potential output

Suppose, however, that the level of aggregate demand were D_2. Then the level of GNP would be Y_2 and the employment rate E_2. Unemployment would exist over and above the natural rate of unemployment. Thus we would have what is called an *output gap* (or *employment gap*), with less output than could be produced, idle labour and unused capacity in the capital stock – in other words, a waste of resources.

But what happens if households, firms and governments combined (the aggregate demand) attempt to buy more than the economy can produce; for instance, what happens if the level of aggregate demand is D_4? The answer is that prices will rise and we will have what is known as an *inflationary gap*, the extent of the gap and therefore the inflationary forces being dependent upon by how much aggregate demand exceeds full-employment output. Prices will rise, marginal buyers will drop out of the various markets, and the economy will return to full-employment equilibrium but at a higher price level. Why, you might ask, cannot this excess aggregate demand be satisfied with imports? It can be, as long as your country is creditworthy abroad, i.e. as long as other nations are willing to loan you the use of some of their resources – the resources that produce the imported goods and services. The debt, of course, will have to be serviced regularly and be repaid in the future.

Let us return to Figure 11.8. If the government is to avoid the inflationary gap and the output gap, it must arrange for aggregate demand to be D_3. At no other level of aggregate demand does $Q = Y$; in addition the simple model shows no inflationary pressures at this full-employment GNP level. Suppose for a moment that the government is successful and somehow manages through fiscal and monetary policy to arrange for aggregate demand to be D_3. Now what happens? Figure 11.9 has the answer.

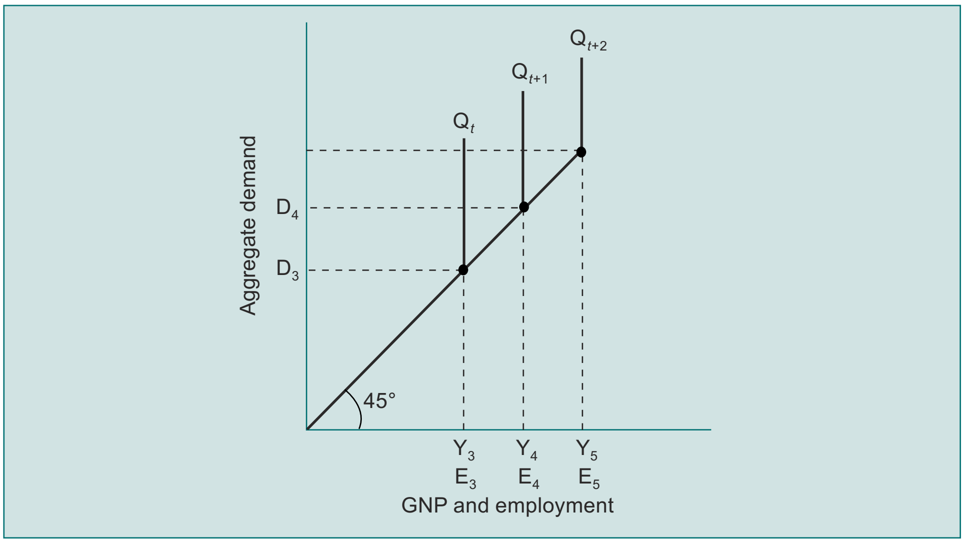


Figure 11.9 Simple model of an economy: potential output growing

Potential output grows continuously and in Year *t* will equal Q_t. Aggregate demand, remember, was perfect for potential output Q_t, but if aggregate demand remains at D_3 we shall have an output gap of $Y_4 - Y_3$ and an unemployment gap of $E_4 - E_3$. Thus to achieve full employment, the government will have to use fiscal and/or monetary policy to expand aggregate demand from D_3 to D_4. Again, if it is successful it cannot rest on its laurels because the following year potential output will be Q_{t+1}. The value of Q will keep on growing no matter how well or badly the government manages the economy because of *supply-side factors*, which we are assuming are outside the control of the government. Such factors include growth in the quality and quantity of the capital stock (net investment), growth in the quality and quantity of the labour force (education and training), and technological change. So how would the government expand aggregate from D_3 to D_4? There are three ways:

(a) increase government expenditure;
(b) reduce tax rates;
(c) increase the money supply.

Suppose the government increases expenditure on road building. This will cause Y to rise as soon as the first piece of road is built; road builders will receive additional income, some of which they will spend on goods and services. However, for the firms producing these goods and services, increased sales mean increased incomes, some of which will be spent on still other goods and services. This process will continue as long as additional expenditure and additional income are generated. The final result is a multiple increase in Y resulting from a given increase in G. The ratio of the total change in Y to the initial change in G is known as the *government expenditure multiplier*. Thus, all the government has to do to expand the economy from Y_3 to Y_4 is create additional aggregate demand of $D_4 - D_3$. If the value of the government expenditure multiplier is 2 then an increase in G of $(D_4 - D_3)/2$ on road building will, with the help of the successive rounds of spending and income generated, result in an increase in total aggregate demand of $D_4 - D_3$, an increase in GNP of $Y_4 - Y_3$ and an increase in employment of $E_4 - E_3$, which will result in the economy being at full employment.

Suppose the government decides to close the output gap by cutting taxes. How will this work? The process is very similar to the increased government expenditure case above, except that in the 'first' round there is no increase in output corresponding to the new road. Suppose every individual is paying $5000 in income taxes initially, but then the government reduces the marginal tax rate or raises the exemption threshold. The result is that each individual now pays only $4000 in taxes. Note that this cut in tax is not associated with any increase in Y; it is merely a government taking less in tax than previously. What will happen is that individuals' disposable incomes (income earned less taxes) will rise, and some of the increase will be spent on additional goods and services; this expenditure is the first round of the multiplier process we encountered in the government expenditure example.

The monetary policy solution to the output/employment gap involves increasing the money supply faster than growth in the demand for money. Such a policy will cause the price of money to fall; the price of money is the rate of interest (R). Borrowers, be they households or firms (or sometimes even governments), normally take the cost of borrowing into account when seeking a loan. Thus a decrease in R is normally associated with an increase in borrowing and a subsequent increase in expenditure; for consumers we tend to think of them borrowing to buy a house, a car, or other consumer durables; for firms, much borrowing is concerned with building new factories, buying new machine tools and computer systems and undertaking investment projects. Once again, like government expenditure, the initial round of purchases will generate additional spending and further rounds of income and spending. Thus all the policy maker has to do is calculate the size of the multiplier and figure out how sensitive consumer and business firm expenditures are to interest rate changes; then calculate how fast the demand for money is growing and increase the money supply by the amount required to obtain that value of R that will yield the required initial increase in consumer spending and business investment – and hope nothing significant changes until the multiplier process is well under way!

We see now how we can achieve full employment by any one of three methods. If we think about it, there must be an infinite number of ways we can mix these three policy tools to achieve an objective we can reach using one tool on its own. For instance, if we return to Figure 11.9 we could close the output gap $Y_4 - Y_3$ by a mixture of increased government expenditure, a tax cut and an increase in the money supply. There is no single unique solution in terms of, say, how large the increase in government expenditure should be. That will depend on the size of the tax cut coupled with the amount of the increase in the money supply.

So which tool or combination of tools should be used? If the only goal is closing the output gap $Y_4 - Y_3$, then it doesn't matter. Most governments or societies have more than just one macroeconomic goal, however, and the weights given to each of the goals may well determine the most appropriate mix of policy tools.

Macroeconomic goals generally accepted as desirable are:

(a) low inflation rate;
(b) low unemployment rate;
(c) balanced government budget;
(d) a positive balance of trade;
(e) stable currency in international exchange markets.

There is less agreement on the ideal mix of Y among C, I and G. People on the political left favour a higher G than people on the political right. There is also disagreement in many nations about the ideal levels of C and I, in other words how much of C should be given up today in order to allow more I and consequently a higher Q in the future. The wide variance in income tax rates by nations in the world today shows no generally accepted idea of the ideal tax structure.

Our simple model did not portray the real dilemma facing most policy makers. In the simple model, it appeared possible to achieve full employment without inflation. This does not occur in the real world; at full employment the inflation rate is positive; this is known as the *inflationary bias of market economics*. To achieve a zero inflation rate might require a very high unemployment rate. Thus, policy makers must choose between two evils, given a trade-off between inflation and unemployment. To add to the policy makers' woe, the terms of the inflation/unemployment trade-off are not constant over time.

This brief overview module has hopefully given you some idea of macroeconomics and macroeconomic policy making. It *may* also have shown politicians and policy makers in a more sympathetic light.

Module 12

Potential Output

Contents

12.1 Introduction

If we could 'stop' our economy at some given point in time, it would be possible – theoretically – to identify the stock of resources and the maximum output that the economy could produce. In our simple model we classified resources as capital goods and labour, although a more detailed classification would subdivide both these categories as

(a) Capital Goods:
   - *Land*: all natural resources including agricultural land, rivers, forests, climate and mineral deposits;
   - *Capital*: all factories, offices, warehouses, shops, houses, machinery and equipment, stocks of materials, roads, railways, harbours, airports, vehicles, ships, schools, colleges, universities and similar assets produced in previous time periods;

(b) Labour Force:
   - *Labour*, which is that proportion of the population able and willing to work, the quality being dependent upon variables such as age, effort, hours of work, education, skills, aptitudes and attitudes;
   - *Enterprise*, which consists of those individuals in the labour force with organisational and managerial ability or with financial skills, such individuals taking risks in starting new companies and launching innovative products – sometimes with their own and sometimes with borrowed resources.

If sufficient time, energy and resources were allocated to the task of increasing a nation's stock of human and non-human resources, virtually all of the above factor inputs could be increased and enhanced. In the short run, however, say a year, a not unreasonable assumption in tackling macroeconomics is that the supply of factors of production and the state of technical knowledge are given or fixed. Thus it follows there must be a fixed upper limit to the total output that can be produced. This upper limit is the *productive potential* of the economy. This still leaves open the question as to whether potential output (Q) will in fact be equal to actual Gross National Product. It also still leaves open the question of how the composition of output is to be determined. What is important to note is that choosing one combination of outputs produced with a fixed quantity of inputs means forgoing alternative combinations. This is the well-known dilemma that all societies encounter: how to allocate scarce resources among competing uses.

Consider a two-good economy where only guns and butter are produced, guns symbolising defence purposes and butter peaceful purposes. The production possibilities frontier sets the limit to the various combinations of the two goods that can be produced, if all resources are fully and efficiently employed (*see* Figure 12.1).

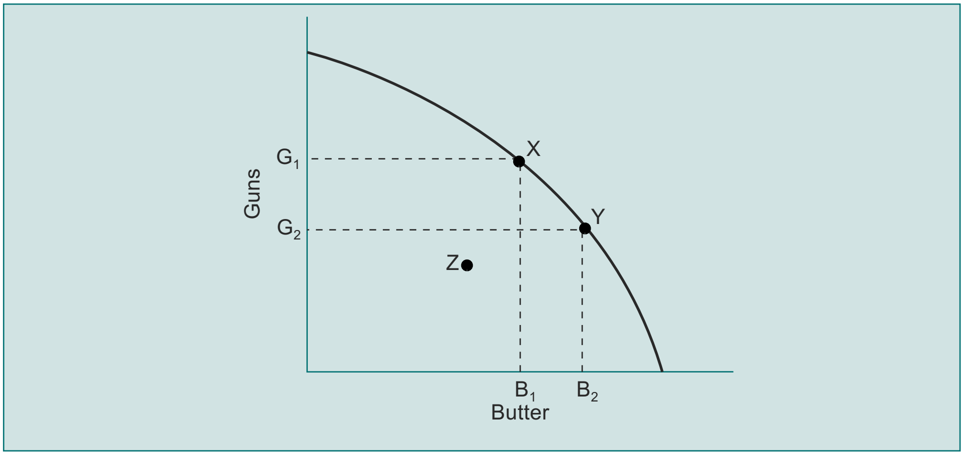


Figure 12.1 The production possibilities frontier in a two-good economy

The production possibilities frontier can be used to illustrate the microeconomic question of *what* should be produced. At a point *on* the frontier, such as X, the production of more butter, such as would be implied at Y, involves a sacrifice of some gun production. In other words, the additional output of butter at Y, given by B_1B_2, has an opportunity cost. That opportunity cost is the gun output sacrificed by moving from X to Y, namely G_1G_2.

The example illustrates that microeconomics is concerned with explaining how choices are made, given that resources are limited relative to human wants. Clearly, if a society is on its production possibilities frontier, then the additional production of one good or service requires a reduction in the production of another good or service; in other words, when all resources are fully employed, additional output of one commodity requires the sacrifice of some other output. The opportunity cost involved is determined by the *slope* of the production possibilities frontier. Microeconomics is concerned with explaining how a society determines the allocation of resources between competing uses, given the fact of scarcity.

The problem of resource allocation is particularly acute on the production possibilities frontier, for once the frontier is reached additional production of one good necessarily involves the sacrifice of production of another good. Suppose, however, that society is initially at a point such as Z. Both positions X and Y are superior to Z, in that X and Y offer a greater output of *both* goods than that possible at Z. For this reason, we can see that X and Y will be preferred to Z so long as society is rational in the economic sense, so that it wants more of preferred commodities rather than less of those same preferred commodities. Indeed, as long as economic rationality exists, then for any point *within* the production possibilities frontier there must be a preferred point *on* the production possibilities frontier.

In short, there may be circumstances when a society can produce more of one good without involving the loss of output of some other good and, indeed, it may be possible to increase the output of both (all) goods, by moving from a point such as Z to a point such as X or Y. The first basic macroeconomic issue to resolve is thus: what determines whether a society is on its production possibilities frontier, employing all factors of production and making maximum use of its productive potential, or whether it is within its production possibilities frontier, with unemployed factors of production and unused productive potential? In the former case, a society will need to sacrifice production of one good if it decides to produce more of another good. In the latter case, it will be possible to increase output of one good without sacrificing the output of another good, or even to increase the output of both goods.

A point like Z in Figure 12.1 represents the situation experienced by all capitalist countries within the past two decades and a situation that was extremely pronounced in the Great Depression of the early 1930s. There is a clear presumption that economies would prefer to operate on the production possibilities frontier, with more goods and services, than within the frontier. There must be compelling reasons if a government enacts policies to put their economy at a point like Z rather than on the frontier. Before we explain such reasons, let us examine potential output in the long run.

12.2 Potential Output in the Long Run

We began this module by assuming as fixed: the supply of the factors of production, and technical knowledge. Over short periods the assumptions of a fixed supply of factor inputs and given technical knowledge are not unreasonable. The supply of land is inelastic; land can be reclaimed from the sea or the productivity of the existing land can be improved, but changes in supply will be marginal compared with the existing stock. As Mark Twain put it, 'they ain't making any more', although he had not visited Singapore where land reclamation has been significant. The supply of the other factors of production is more elastic but the supply of labour is heavily conditioned by the population and its age structure, by the conventional retirement age, the extent to which people participate in the labour force, and similar factors conditioned by social and demographic factors that change only slowly. Again, new skills and industrial experience are acquired only slowly. New investment in factories and machine tools and other forms of physical capital will add to the capital stock, but these additions will be small relative to the size of the existing capital stock. The supply of entrepreneurship appears least amenable to rapid change, depending, essentially, on sociological factors and on the values of the existing society. Finally, although some technical change occurs even in the short run, the addition to the existing stock of knowledge usually takes place incrementally, although there are exceptional periods of technological change, such as the Industrial Revolution and (today) the IT Revolution.

Given that the supply of factors of production changes only slowly, the production possibilities frontier may be taken as fixed for short periods, determined by the existing supply of the factors of production and the stock of technical knowledge. Hence, in the short run the most important macroeconomic question is whether a society attains its given production possibilities frontier, thus making full use of the existing supply of the factors of production and the given stock of technical knowledge.

Over longer periods of time, however, the production possibilities frontier is not fixed – as is evident from the fact that the last two centuries have witnessed a massive rise in the living standards of the developed industrial economies. Indeed, in the last half-century, the developed industrial economies have experienced growth rates that are very high by historical standards, and it has not been unusual to see living standards double within a 20-year timespan – the production possibilities frontier has shifted to the right very rapidly over time. Economic growth has occurred in response to an increase in the supply of the factors of production, increased factor productivity and increased technical knowledge. The second major macroeconomic issue is then: what determines the rate of growth of potential output through time? Why do we observe differences in growth rates across economies in a given period, or in the same economy at different periods of time?

Economic growth can be depicted graphically by a rightward shift of the production possibilities frontier, and society may be conceived as making choices in the current period between current consumption and capital investment. The ultimate objective of all economic activity is consumption, by which we mean the present enjoyment of material goods and services, but it can be seen that no society devotes all its resources to present consumption. If a society uses all its resources to satisfy present consumption, this would entail the penalty of lower consumption in the future, as no resources would be available to make good the diminution in the productivity of the existing capital stock that occurs with the passage of time through wear and tear in use (called *depreciation*). The productive capacity of an economy, as represented by its stock of factories, shops, commercial premises, transport systems, machine tools, industrial commercial and professional skills, etc., will diminish unless resources are set aside to offset depreciation.

To maintain the productive capacity of the capital stock, resources must be devoted to replacement investment. Likewise, if the labour force is to remain at the same quantity and with the same skill level, not only do new entrants have to enter the labour force to replace those who retire or die, but they have to have been trained to the same skill levels. Therefore, to maintain potential output at a constant level requires that part of the output of the economy must be diverted from consumption goods towards investment to replace physical and human capital. In Figure 12.2 the economy is stationary through time, in the sense that the level of investment, both physical and human, is equal to replacement investment. The shaded area in both the capital stock and the labour force represents the depreciation of physical and human capital, and the shaded area in investment represents the replacement investment necessary to maintain the capital stock and the labour force.

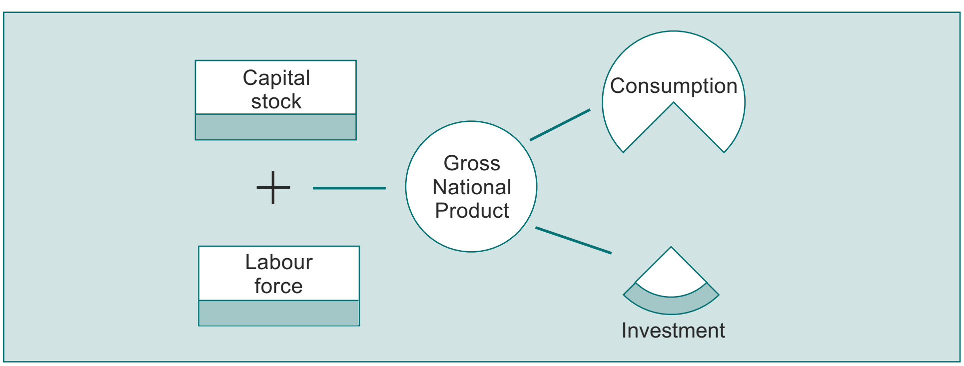


Figure 12.2 Factor inputs and national output; replacement investment

Society, however, may make decisions that result in an increase in (rather than just a maintenance of) the stock of human and physical capital. For this to happen, investment must exceed that necessary to replace the human and physical capital stock, i.e. there must be positive net investment. Net investment is equal to gross investment minus replacement investment. If net investment is positive, then the stock of inputs increases, and so does productive potential of the economy.

Figure 12.3 illustrates the growth due to net investment. National output in the upper portion of Figure 12.3 (GNP_t) consists of consumer goods and services, replacement investment (RI) and net investment (NI). In the lower portion of the figure, RI allows the capital stock and labour force to remain at Capital stock$_t$ and Labour force$_t$, and the net investment adds to the Capital stock$_t$ – making it Capital stock$_{t+1}$. The net investment plus more workers from a growing population add to the Labour force$_t$ – becoming Labour force$_{t+1}$. Thus, potential GNP grows, this growth being attributable to the net investment.

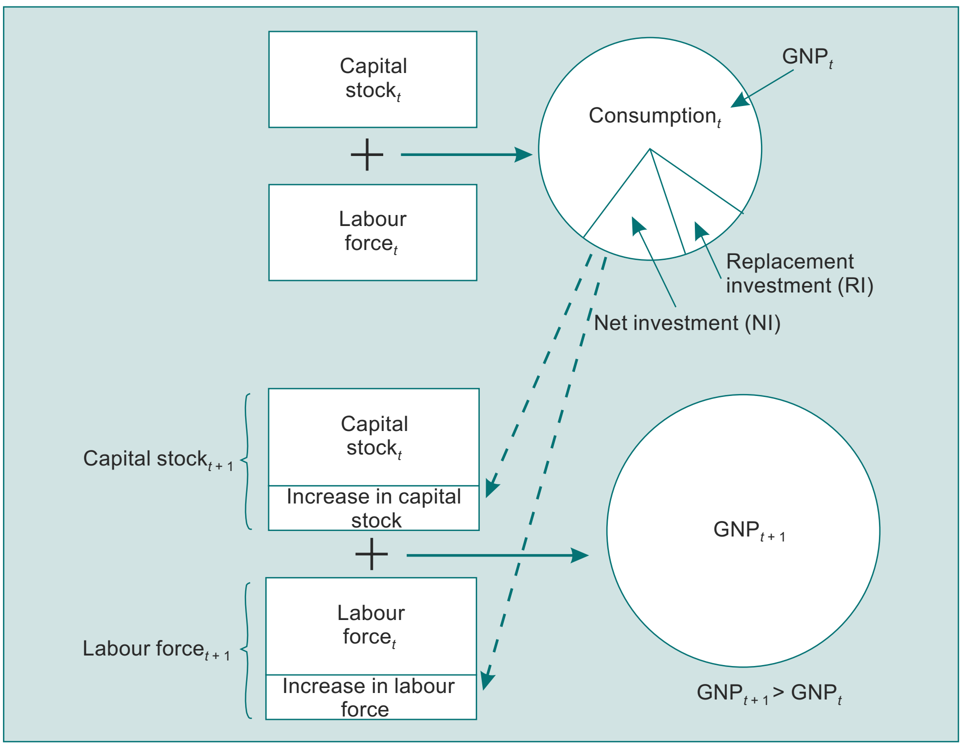


Figure 12.3 Economic growth

If society chooses, by sacrificing current consumption, to *add* to its stock of resources sufficient extra resources to make good any depreciation and allow some *net addition* to the stock of capital goods or the labour force, then the productive capacity to produce goods and services would increase and the production possibilities frontier would shift upwards/outwards.

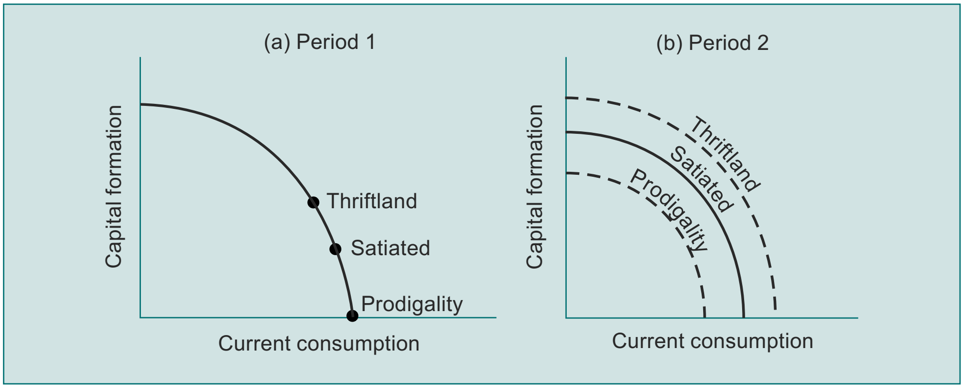


Figure 12.4 Shifts in the production possibilities frontier

Figure 12.4 shows the outcome for three different countries, 'Prodigality', 'Satiated' and 'Thriftland'. In period 1 (Figure 12.4a), all three countries have the same productive capacity and all are operating on the production possibilities frontier. However, the distribution of resources between current consumption and capital formation is different. 'Prodigality' devotes all of its resources to current consumption, 'Satiated' provides sufficient resources to offset capital depreciation alone, and only 'Thriftland' undertakes sufficient capital investment to offset depreciation and make some net addition to the existing capital stock. In consequence, in period 2 (Figure 12.4b), 'Prodigality' experiences a downward shift in its production possibilities frontier, indicating that the productive capacity of the economy has declined as the capital stock has depreciated. In 'Satiated', the production possibilities frontier does not shift in either direction, showing no change in the productive capacity of the economy. In 'Thriftland', the production possibilities frontier shifts upwards, indicating an increase in productive potential compared with period 1 as a consequence of the net addition to the capital stock.

Historically, the production possibilities frontier has shifted upwards at different rates for different economies and different periods. This process of economic growth, and the factors that determine the growth rate, are of crucial importance because small differences in growth rates will have substantial implications for material living standards even over a period of a few years. For example, if the output of goods and services is increasing by 2 per cent annually, then the total output doubles every 35 years, compared with every 18 years for a 4 per cent compound rate of growth. The rule of thumb for making this calculation is to divide 72 by the growth rate. This will give you, approximately, the number of years it takes for output to double. In the 1960s, when the Japanese economy was growing at a rate of approximately 11 per cent per annum, its GNP was doubling every 6–7 years. The UK growth rate in the same period was around 2.5 per cent per annum and the US around 4 per cent per annum.

Most economies experience some growth over time. Occasional situations arise where there is a leftward shift of the production frontier; fortunately, these situations are rare and are due usually to war, flood or earthquake, i.e. where some disaster destroys a significant portion of an economy's production capacity. Very poor nations, where the goal of current survival means the bulk of resources are allocated to the production of consumer goods and services, lie close to the 'satiated' production frontier. For many such countries, foreign aid in the form of capital goods and skilled labour has helped move the frontier outwards.

12.3 Measuring Potential Output

Potential national output is a supply concept, a measure of productive capacity that may be represented diagrammatically by the production possibilities frontier. It is achieved only when the economy reaches full employment of all the factors of production. Should there be underemployment of any of the factors of production, then actual output will fall short of potential output and the economy will be within its production possibilities frontier. Clearly, the greater the extent of underemployment, the greater the difference between actual and potential output.

Potential national income and output cannot be directly observed but may be inferred indirectly. Likewise, we can only estimate the difference between actual and potential output. To estimate potential output, some measure of the utilisation of the factors of production is required and the two most widely used in market economics are the *unemployment rate* (the number unemployed as a percentage of the labour force in work or seeking work) and the extent of the *capacity utilisation* as measured in industry surveys.

There is a clear relationship between capacity utilisation and unemployment. As the degree of capacity utilisation increases, the rate of unemployment falls. For the economy as a whole, however, the unemployment rate never falls to zero and capacity utilisation never reaches 100 per cent. The question immediately arises as to what rate of unemployment and capacity utilisation should be taken to be consistent with the full utilisation of productive capacity. Evidently, if the unemployment rate that defines the productive capacity were zero, then the concept of productive capacity might not be operationally meaningful. Thus, the definition of productive capacity should allow the possibility that it might be realised in some periods.

Full employment is an important goal of economic policy and the achievement of full employment is desirable for economic, social and political reasons. Partly because of this, but also because more accurate data on employment, unemployment and the labour force are more easily collected than data on the other factors of production, full employment and potential output are usually defined in terms of labour utilisation. Full employment is not defined in terms of a zero rate of unemployment, but is usually taken to be some target rate of unemployment. We have to consider the logical basis of such a definition of potential output.

12.3.1 Types of Unemployment

All economies undergoing change must experience some unemployment. There are three types of unemployment: frictional, structural and seasonal. These three types will always be present when change is occurring and if the labour market operates with any degree of inefficiency.

12.3.2 Frictional Unemployment

In any labour market there will always be some people changing jobs, and the process of searching for a new job takes time. This is because information is imperfect. Movement in the labour market is impeded by such imperfections, just as the movement of one surface across another is impeded by friction. Hence the term 'frictional unemployment'. It takes time and effort to seek out a new job opportunity, even when a person is simply changing employer without changing either occupation or area of employment.

Example

Suppose a secretary leaves an employer in a given city and looks for another job as a secretary in the same city. Some time will be spent in looking at job advertisements in newspapers, visiting job centres, and discussing job prospects with friends and relatives. More time will be used attending interviews and deciding among any job offers received. The more difficult it is to collect information about job opportunities, the longer will be the period of job search and of intervening unemployment.

The amount of frictional unemployment depends upon the number of persons changing jobs, and the average length of job search. Frictional unemployment does not result from a lack of jobs, or even from a lack of suitable jobs. Frictional unemployment arises when the unemployment is matched by unfilled job vacancies in the same occupations and same locations. For example, in any period in any city there will be some persons seeking work as secretaries and some employers seeking secretaries to fill existing job vacancies. Some unemployment will exist even though there are job opportunities available in the labour market – and even if the number of opportunities matches the number of jobseekers. Unemployment occurs despite the unfilled vacancies because it takes job searchers some time to seek out new jobs and employers some time to seek out and select persons who are looking for jobs. Such unemployment arises from imperfect information. The labour market does not work perfectly. There are frictions. By its nature, frictional unemployment is generally short-term unemployment.

12.3.3 Structural Unemployment

Structural unemployment can be thought of as more stubborn frictional unemployment. Again, it is unemployment that occurs despite unfilled job vacancies, but in this case there is a 'mismatch' between the characteristics of the unemployed and the characteristics of the job vacancies.

Example

The unemployed may be steelworkers or carworkers and the job vacancies may be for deep-sea divers or electronic engineers. Or, there may be unemployed school teachers in one part of the country and job vacancies for school teachers in other parts of the country. A mismatch of occupation or location is the most common cause of structural unemployment.

While frictional unemployment arises from imperfect information and will eventually be rectified by a search for work in the same occupation within the same location, structural unemployment cannot be cured so easily. As there is a mismatch between the characteristics of the unemployed and job vacancies, the unemployed can only be matched to the jobs available if they undertake retraining, or change their location of work, or both. For instance, if unemployment exists amongst steelworkers while there are job vacancies for electronic engineers, then the unemployed can only fill the job vacancies through retraining. If the unemployment exists among school teachers in one location while there are job vacancies in distant counties, then the unemployed can only fill the job vacancies through moving home. If the unemployment exists among coalminers while there are job vacancies for deep-sea divers, then the unemployed can only fill the job vacancies by both retraining and moving home. Retraining and moving home take time and impose heavy costs on those having to change job and/or location. For this reason, structural unemployment can persist for long periods.

12.3.4 Seasonal Unemployment

This is the easiest type of unemployment to understand; indeed, the term is self-explanatory. Seasonal unemployment will occur in any activity in which the level of production is dictated by weather or by the calendar, for example farming, fishing and forestry, food processing, construction, and tourism. Such activities are characterised by periods of high employment, combined with overtime working; these alternate, in a regular seasonal pattern, with lower unemployment, short-time working, and unemployment.

12.3.5 Factors Determining Unemployment

The most important factors determining the amount of frictional, structural and seasonal unemployment are:

(a) the level of economic activity;
(b) the transmission of information;
(c) the rate of structural change;
(d) the ease of changing occupation and home;
(e) institutional restrictions and barriers; and
(f) the dependence on seasonal industries.

Unemployment is affected by the general level of economic activity. When actual output is close to potential output, employment is high and there will be strong competition among employers to hire labour. Employers will be more active in advertising job opportunities and the improved flow of information will reduce frictional unemployment. Employers will also offer better retraining allowances and financial assistance to those moving home in order to attract people from other occupations and areas. These measures will reduce structural unemployment. When actual output is well below potential output there will be less competition among employers to hire labour and frictional and structural unemployment will increase.

Unemployment is also affected by the quality of transmission of job information. Jobs cannot be filled unless employers and employees can be brought into touch with each other. Employers advertise jobs through the Internet, local and national newspapers, radio and TV, private agencies, shop windows, billboards at the factory or office gate, and 'word of mouth'. Employees use some or all of these methods of looking for work. The better the methods of communicating information, the lower will be unemployment.

There is always some structural change in an economy. Tastes change, new products emerge, goods produced domestically are displaced by imports, and technical change occurs. For example, the increased demand for holidays overseas has reduced the demand for domestic holidays; the development of man-made fibres has reduced the demand for natural fibres; Japanese cars have taken an increasing share of many domestic car markets; the railways displaced the stagecoach and canals, and the motor car has largely displaced the railways; the manual typewriter displaced the quill, the electric typewriter displaced the manual typewriter, and the PC displaced the electric typewriter. New economic activities expand and old ones contract. These changes require a redistribution in the skills and location of the workforce. The higher the rate of structural change, the higher will be the level of structural unemployment.

Structural unemployment will be minimised where it is comparatively easy to acquire new skills and move from one home to another. This will depend on factors such as the availability and cost of retraining programmes, the cost of travel and moving home, suitable schools for children, and suitable hospitals for the infirm. The higher are these costs, the higher the level of structural unemployment.

Institutional restrictions and barriers also affect unemployment. These include any action by government, or by organisations of employers and employees that restrict the efficiency of the labour market. For example, unemployment will increase if the government imposes minimum-wage levels above the level that would be established in the marketplace; if organisations place undue restrictions on entry to certain jobs; if unions impose restrictive practices, restrict employment to union members, or impose unduly long periods of training; if local authorities discourage geographical mobility by favouring local residents in allocating houses; if pension arrangements impose heavy costs on those who change jobs. Anything that limits the ease of changing or taking employment will increase unemployment.

Finally, unemployment is affected by seasonal industries. Economies that are dependent on farming, fishing, forestry and tourism will experience seasonal fluctuations in unemployment. Some economies have clear natural advantages in these industries. Hence, they continue to use many of their resources in them in the full understanding that unemployment will result at certain predictable periods of the yearly cycle.

12.4 The Relation Between the Unemployment Rate (U), Potential Output (Q) and Actual Output (Y)

Is there a uniform and persistent relationship between the gap between potential and actual output and the proportion of the workforce that is unemployed? If there were, this would be of major interest and concern to policy makers anxious to reduce unemployment to a minimum because, as will be demonstrated, policy makers are more confident about eradicating demand-deficient unemployment than they are about frictional, structural and seasonal unemployment.

Here it is necessary to introduce the fourth main category of unemployment: *demand-deficient unemployment*. As the name implies, demand-deficient unemployment arises because of insufficient demand for labour. Thus we can argue that full employment is realised when the number of unfilled job vacancies (V) is equal to or greater than the number unemployed (U). In this case, there would be no demand-deficient unemployment, as there would be sufficient job vacancies to absorb all those who are unemployed – although unemployment would still exist for all the reasons already discussed.

If accurate measures of job vacancies and the numbers seeking work were available, full employment would occur where $V \geq U$. Unfortunately, even where job vacancy data are available, they are subject to severe deficiencies (for example, in the UK the vacancies notified to the local employment offices probably account for only about one-third of all vacancies) and most countries have no job vacancies data at all. For this reason, full employment is usually taken as some target rate of unemployment.

The natural rate of unemployment varies over time for any one country and varies significantly among countries. In the last three decades the rate has varied in the USA and UK from 2 per cent to 6 per cent, whereas in countries like Switzerland and Japan the rate has varied from less than 1 per cent to 2 per cent. The changing socio-economic characteristics of the labour force cause variation within a country over time. The characteristics of the labour force, employment practices, and the operation of labour markets cause the international difference.

If we know that the economy is operating at the full-employment rate of unemployment, in other words the only unemployment is frictional, structural and seasonal, then actual output (Y) and potential output (Q) will be equal. Then we shall at least be able to measure the potential output of the economy. It also follows that the larger the gap between the values of Q and Y, the higher will be the rate of unemployment; and the smaller the gap, the closer the unemployment rate will be to the full-employment unemployment rate.

The output gap can be expressed as the percentage difference between potential and actual output, namely

$$\%\ \text{output gap} = \frac{Q-Y}{Y} \times 100$$

A US economist, Arthur Okun, conducted research on the empirical relationship between the unemployment rate and the output gap and found stability in such a relationship over a 25-year period in the US economy commencing in the late 1940s. The equation that summarises this relationship is now known, not surprisingly, as *Okun's Law*. The equation is:

$$U = \left[\frac{1}{3} \times \frac{(Q-Y)}{Q}\right] + U_F$$

where U_F = full-employment rate of unemployment. In words Okun's Law states that for each one per cent by which actual output falls short of potential output, the unemployment rate will exceed its full-employment rate of unemployment by one-third of one per cent. Thus if Y is 6 per cent below Q, then U will be 2 per cent above U_F. Expressed alternatively, Okun's Law suggests that each percentage point increase in the unemployment rate above U_F is associated with a 3 per cent reduction in real output.

Since Okun's Law was first formulated, further statistical testing has suggested that the parameter value of one-third is not immutable. Be that as it may, Okun's Law provides a reasonable measurement of the loss in real output attributable to demand-deficient unemployment.

Figure 12.5 records the unemployment rates of the world's seven leading market economies for the past four decades. These data may differ from 'official statistics' because the official data often do not make allowances for part-time working and individuals who are no longer looking for jobs because those jobs do not exist. If you take the lowest unemployment rate for each country in the figure and apply the Okun formula, you will get a reasonable estimate of the massive amount of real output lost and gone for ever during these decades for all nations.

If government policy makers are aware of Okun's Law and have the policy tools to affect national output, why have they been willing to tolerate such high rates of unemployment? To answer that question we must turn to the relationship between unemployment and inflation.

Figure 12.5 Unemployment rates for the G7 economies, 1960–2015

12.5 Output and Inflation

The demand for output by consumers, firms and government (aggregate demand) will determine the actual output of goods and services. Given potential output Q for any short time period under consideration, the level of aggregate demand will thus determine the unemployment rate. This principle can be shown in Figure 12.6.

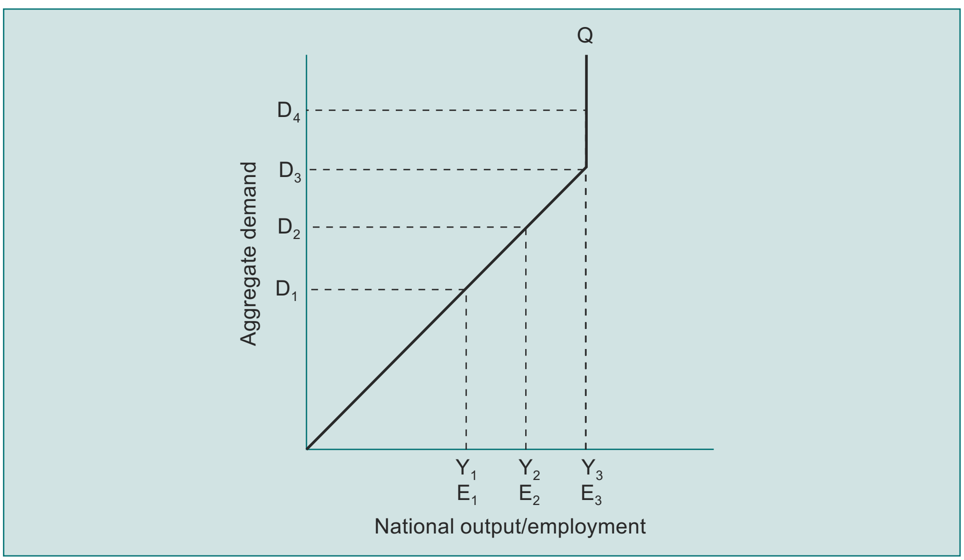


Figure 12.6 Aggregate demand, national output and employment

If aggregate demand is D_2, national output and employment will be Y_2 and E_2 respectively. If aggregate demand is D_3 the economy will be at full employment because Y_3 will equal Q. Aggregate demand of D_4 will create an inflationary gap: prices will rise, and the economy will remain at full employment at a higher price level. Unfortunately, our simple model no longer suffices in the real world. At full employment, or even before we reach full employment, the economy normally experiences rising prices, equivalent to inflation. Why does this happen? Before we answer that question we must define what we mean by inflation.

12.5.1 The Inflation Rate

The *inflation rate* is the percentage increase per year in the average price level from one time period to another. Thus if P_t is the average price level in period t and if P_{t+1} is the average price level in period $t + 1$, then the inflation rate in time period t is defined by

$$\mathrm{INF}_t = \frac{P_{t+1} - P_t}{P_t}$$

Expressed alternatively, the price level in period $t + 1$ is:

$$P_{t+1} = P_t(1 + \mathrm{INF}_t)$$

But what do we mean by the price level? The most commonly cited measure of the average level of prices is the c*onsumer price index*. This index measures the average level of prices of the goods and services consumed by the typical household. In most market economies this index is calculated monthly. Several hundred goods and services regularly purchased each month by average households are selected, initially to establish a base period, and the quantities purchased and the prices paid are noted; an index is then constructed.

To simplify the calculation, assume only three goods and/or services, as in Table 12.1.

$$\text{Price Index for Base Period} = \frac{\$30.00}{\$30.00} \times 100 = 100$$

The price index for the base period is always 100. Assume the base period is year t.

Table 12.1 Consumer price index in base period t

Good/service	Quantity	Price ($)	Total expenditure ($)
Loaf of bread	10 loaves	1.00/loaf	10.00
Glass of wine	5 glasses	1.20/glass	6.00
Haircut	2 haircuts	7.00 each	14.00
			30.00

Now let us consider the next time period, period $t + 1$. In time period $t + 1$ we observe the prices of the same goods and services and multiply them by the base period quantities. This has been done in Table 12.2.

Table 12.2 Consumer price index in period t + 1

Good/service	Base period Quantity	Period t + 1 Price ($)	Expenditure in period t + 1 (on base period quantities) ($)
Loaf of bread	10 loaves	1.10/loaf	11.00
Glass of wine	5 glasses	1.20/glass	6.00
Haircut	2 haircuts	8.00 each	16.00
			33.00

$$\text{Price Index for Time Period } (t+1) = \frac{\$33.00}{\$30.00} \times 100 = 110$$

The increase in the price index is $\frac{110-100}{100} \times 100\% = 10\%$

Thus the annual inflation rate as measured by the index is 10 per cent.

The typical household, of course, does not purchase an average selection of *all* goods and services that make up total output or Gross National Product; for a start, investment and government goods make up between 35 per cent and 45 per cent of the total output of the major capitalist countries and do not feature as consumer goods. The index that includes all goods and services produced in the economy is known as the *GNP deflator*. Its calculation is similar to the consumer price index calculation.

Now you understand how the inflation rate is determined, it is necessary to investigate how firms set prices in some given time period. We discussed in the microeconomics modules how the forces of demand and supply determine equilibrium prices, but we also saw how, in markets where perfect information is absent and where many retail outlets implicitly possess a degree of local monopoly, some variation in prices of homogeneous goods and services at some given time is consistent with market forces in action. In other words, very little of our ordinary day's shopping takes place in auctions. When we visit the local supermarket, most goods on the shelves carry a price tag; normally there is no opportunity to 'haggle' the price downwards, though if you are buying in bulk and paying cash there is no harm in asking for a discount or 'best price'. Prices of new cars are also advertised, but bargaining normally occurs when you purchase a new car. Haggling or bargaining is also a distinguishing feature of many bazaars or open markets in countries throughout the world.

For the non-bargaining markets, i.e. the markets for clothes, food, entertainment, restaurants, taxis, airline tickets etc., how are prices set by the firms involved in any given time period? Since most of the firms involved are private businesses, we can assume that they set prices to maximise the profits they can expect during the short-run period for which prices will be fixed. Many variables are likely to have an effect on what price the firm would want to charge. These include the *expected* costs of production and the *expected* demand for the goods and services to be produced. The higher the cost of production is expected to be in the coming time period, the higher the price the firm will have to set at the beginning of the time period to make a profit during the time period. The higher the demand is expected to be for the output in the coming period, the higher will be the price the firm will set to maximise profit. Thus a crucial question to be answered in understanding how firms set prices is how firms form expectations about costs and demand for the coming period. Two major influences on expectations are likely to be firms' most recent experiences and their interpretation of what is happening in the whole economy. Of particular importance for both is the most recent level of aggregate demand.

It is clear that the higher has been recent aggregate demand, the higher will have been the average firm's own demand, and the higher will be its expectations of its own demand in the immediate future. Furthermore, the higher was aggregate demand, the higher will have been the demand for factors of production and the higher will be firms' expectations about what they will have to pay for labour, raw materials and other factor inputs in the coming period. Thus, the previous period's aggregate demand will influence firms' expectations of both demand for their own products and the prices they will have to pay for inputs in the coming period.

The foregoing analysis suggests that the average price level that will be set at the beginning of the next period will depend on the level of aggregate demand in the current period. This implies that, given the price level set at the beginning of the current period, the rate of increase in the price level from the beginning of the current period to the beginning of the next period will be positively related to the level of aggregate demand in the current period. That is, the higher the level of aggregate demand in any short-run period, the higher will be the rate of inflation during that period.

Since, in any short-run period, aggregate demand will influence both the unemployment rate and the inflation rate, there will be an implied relationship between unemployment and inflation. For each possible level of aggregate demand, there will be a corresponding combination of unemployment and inflation rates. When aggregate demand is high, there will be a low rate of unemployment and a high rate of inflation. When aggregate demand is low, there will be a high rate of unemployment and a low rate of inflation.

The combinations of unemployment and inflation rates that would result from various levels of aggregate demand are summarised by the curve in Figure 12.7. This curve is called the Phillips Curve, named after the New Zealand economist A.W. Phillips, who first identified a systematic relationship between inflation and unemployment. Each point on the Phillips Curve corresponds to a particular level of aggregate demand taken from Figure 12.6, which is largely reproduced as Figure 12.7. For example, point A in Figure 12.7a is the combination of inflation and unemployment rates that results when aggregate demand is D_2. Point F on the Phillips Curve corresponds to full-employment output, an employment rate of E_3, and an unemployment rate of U_F which is, of course, our full-employment rate of unemployment.

Point B is the 'tricky' point that did not exist in our simple model of the economy in Figure 12.6 where Q represented the potential of the economy. We can now relax the constraining influence of Q because of the way we have defined the full-employment rate of unemployment. In the face of aggregate demand exceeding Q, some factories and a proportion of the labour force now work overtime and the average frictional and structural rates of unemployment fall because there are so many unfilled vacancies in the labour market. Thus the economy in the short run squeezes some extra output from the economy, represented by $Y_4 – Y_3$. However, there is a cost to be borne; and that cost is the higher prices of goods and services and the higher costs that firms must pay for factor inputs in the economic 'boom'. The steepening of the slope of the Phillips Curve between points F and B indicates an *increase* in the *rate* of inflation.

It is important to realise that the Phillips Curve is a short-run relationship. It shows, for a particular short-run period, the combinations of U and INF that would result from different levels of Y. As will be seen in later modules, it is possible that this short-run relationship may shift from one short-run period to the next in an unpredictable fashion.

The Phillips Curve has a number of characteristics that need to be noted. The first of these has already been seen, namely that it slopes downwards from left to right. This means that, for any short-run period, lower unemployment will be associated with higher inflation. Which combination will actually result will depend on the level of aggregate demand. This inverse relationship between unemployment and inflation is sometimes referred to as the *trade-off* between unemployment and inflation. Given the position of the Phillips Curve, lower unemployment can be achieved only at the expense of higher inflation and vice versa. Thus, in any particular short-run period, had aggregate demand been higher, the result would have been lower unemployment, but also higher inflation, than actually occurred.

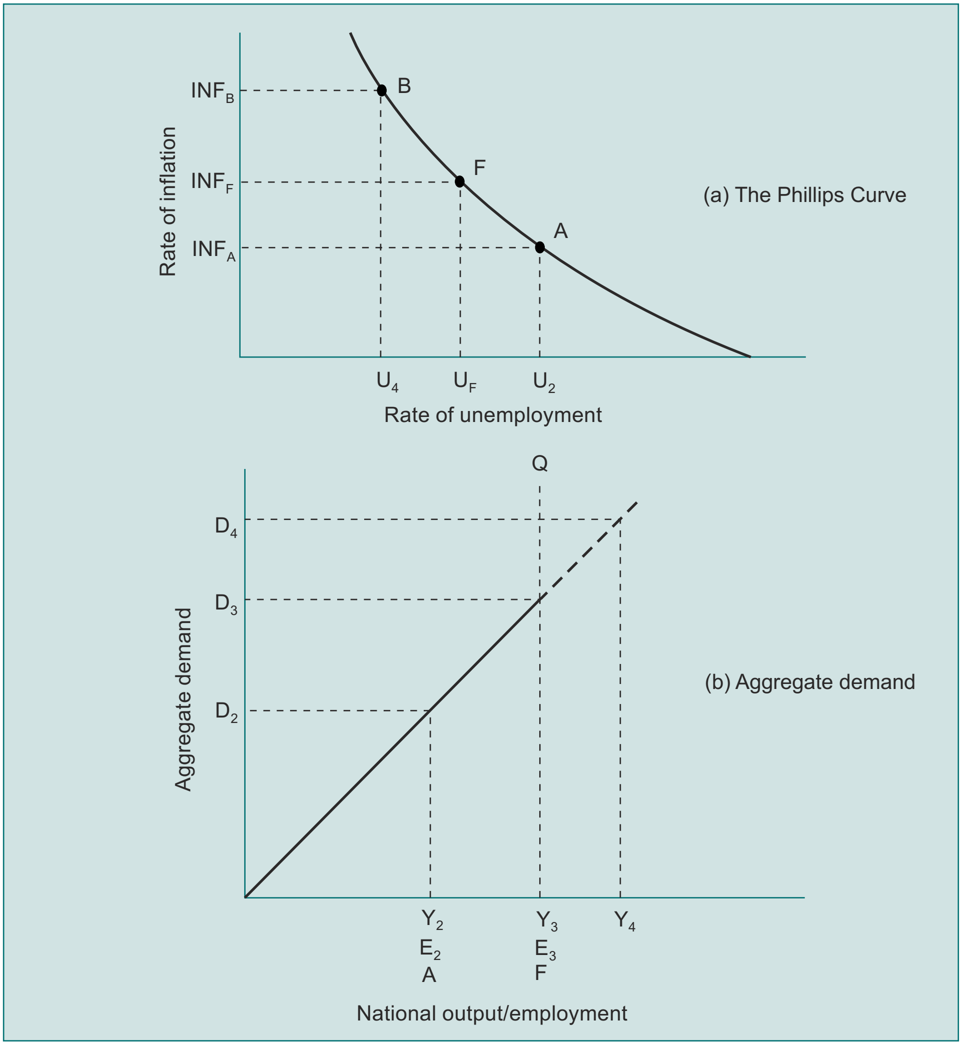


Figure 12.7 The Phillips Curve and aggregate demand

A second important feature of the Phillips Curve is that historically it has usually been in a position such that the inflation rate that occurs at full employment is positive. This is indicated in Figure 12.7a by the point F on the Phillips Curve. If U_F is taken to be the full-employment unemployment rate, then INF_F is the accompanying inflation rate, which is greater than zero. This typical characteristic of the Phillips Curve reflects what is often called the *inflationary bias* of the economy.

One might expect that if markets for individual goods and services, including the markets for the services of factors of production, worked perfectly, the inflation rate at full employment would be zero. The reason for this follows the ensuing line of argument. At overfull employment, when firms are employing more resources than are normally supplied, the forces of supply and demand put upward pressure on wages, rents, and other factor costs. As a result, firms expect production costs to rise and set higher prices at the beginning of the next short-run period. Thus, at overfull employment the result would be inflation. At excess unemployment, when firms are employing fewer resources than are normally supplied, firms might well expect the excess supply of labour and other resources to drive wages and other costs down and, therefore, set prices at a lower level at the beginning of the next period. If this occurred, there would be deflation when there is excess unemployment. At full employment, where there is neither excess demand nor excess supply in the markets for resources, there would be neither inflationary nor deflationary pressure and a stable average price level would seem a plausible outcome, as shown in our earlier simple model of the economy.

The reason why this description does not accurately fit the real world can best be understood in terms of the market for labour, which is quantitatively the most important factor of production in that it accounts for roughly two-thirds of all income payments by firms. First, it should be understood that the market for labour is not just one market but many markets. There are markets for skilled and unskilled types of labour in many different industries and the situation in these markets can differ radically from one to the other. In some years, for example, the market for mechanical engineers has been characterised by substantial excess supply at the same time as a shortage of electronic engineers. Second, these individual markets operate imperfectly in the sense that wages do not always adjust quickly to equate supply and demand. In many labour markets when there is excess supply, wage rates do not respond to the downward pressure. As a result, wage rates appear to be sticky in the face of high unemployment.

Many reasons are given for the downward rigidity of wage rates. One is that in many labour markets wage rates are determined by collective bargaining between labour unions and management representatives for an entire industry. The political nature of union decision-making is such that it is very difficult for elected union leaders to recommend to the union members a proposed settlement that calls for a wage decrease, regardless of the economic circumstances. Furthermore, a majority of union members will have sufficient seniority that they would not be laid off even when there is high unemployment. Since they do not fear unemployment, they have no reason to vote in support of a contract that involves a wage cut to avoid unemployment. In addition, given the opportunity to collect unemployment compensation in the event of being laid off, many workers would find it advantageous to work for a higher wage and accept periodic spells of unemployment rather than work at a lower wage with a lower probability of being laid off work. Thus, while a perfectly competitive market for labour would result in wages falling during periods of excess supply, in the real world there appears to be downward wage rigidity in many countries even in the face of high unemployment.

Consider now in a full-employment situation the implication of there being many distinct labour markets, each of which is characterised by downward wage rigidity. As discussed earlier, full employment does not mean zero unemployment; it means that there are enough jobs available for all those who are in the labour force, but there is still an imperfect match between jobs and workers. In other words, even though there is full employment, there is frictional and structural unemployment. The existence of such unemployment at full employment can be thought of in terms of some labour markets having an excess demand for labour while others have an excess supply. In those markets with excess demand, wage rates can be expected to rise. Because of downward wage rigidity, however, wage rates will not fall in those labour markets with excess supply. As a consequence, firms that hire workers in markets where there are job vacancies will expect wage rates to rise and will set their prices at a higher level at the beginning of the next period. Firms that hire workers in markets with high unemployment will not expect wage rates to fall as they would in perfect markets and consequently will not be able to reduce prices at the beginning of the next period. Thus, in the full-employment situation, the fact that wage-rate increases in tight labour markets are not offset by wage rate decreases in slack labour markets means that price increases in some industries will not be offset by price decreases in other industries. The conclusion, therefore, is that at full employment the average price level will tend to rise at least somewhat, which was defined earlier as an inflationary bias.

In terms of the Phillips Curve, the inflationary bias is reflected in the curve crossing the horizontal axis at an unemployment rate to the right of the natural rate of unemployment. At an unemployment rate that is high enough, the downward pressure on wages will be sufficient to overcome the downward wage rigidity and inflation rates will become negative, although there have been very few years when this has occurred in a major industrialised nation. The most striking example was during the Great Depression in the 1930s, when in the presence of extremely high unemployment rates the inflation rate was negative for several years in many countries. Just where the Phillips Curve crosses the horizontal axis may vary from short-run period to short-run period, depending on other factors to be discussed later.

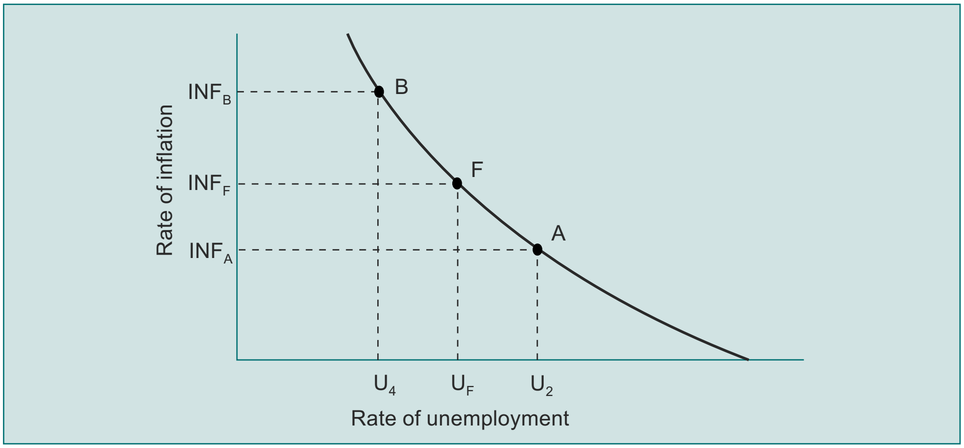


Figure 12.8 The Phillips Curve

The Phillips Curve has still another important feature – its curvature. As shown in Figure 12.7a (reproduced as Figure 12.8), the Phillips Curve is a curved line bowed toward the origin. This means that what the trade-off is between unemployment and inflation depends on where the economy is on the curve. At high rates of unemployment, the curve is relatively flat. In that region, a given change in the unemployment rate would be accompanied by a smaller change in the inflation rate. In Figure 12.8, for example, the change in the unemployment rate from U_2 to U_F is the same as from U_F to U_4. The corresponding changes in the inflation rate, however, are quite different. The increase in the inflation rate from INF_A to INF_F is less than the increase from INF_F to INF_B.

The reason for this curvature can be understood by considering the effect on inflation of an increase in unemployment in two extreme situations: one in which there is initially high unemployment, and the other in which there is initially low unemployment. At very high rates of unemployment, very few labour markets would be characterised by excess demand. An increase in the overall unemployment rate would tend to reduce the upward pressure on wages in those few markets. As a result, those firms that hire in those markets would not expect as large an increase in costs in the next time period and, therefore, would raise prices at the beginning of the next period by a relatively small amount. Most labour markets would be characterised by excess supply and the unemployment increase would increase this excess supply. This would not have much effect on wages, however, because of downward wage rigidity, and would not, therefore, have much effect on the pricing decision of enterprises hiring labour in those markets. Thus, at high unemployment, there would be some decrease in the rate of inflation as a consequence of an increase in the rate of unemployment.

In contrast, however, at very low overall unemployment rates, most labour markets would be characterised by excess demand. In those markets, an increase in the overall unemployment rate would reduce the upward pressure on wages and, therefore, prices. As before, wages in those markets with initial excess supply of labour would not be significantly affected by the increase in the unemployment rate, but there would be relatively few such markets in this case. Thus, at low initial unemployment, the increase in unemployment would have a much more widespread effect on wages and, as a consequence, a much more pronounced effect on the rate of inflation.

While the Phillips Curve in any short-run period will typically be as it appears in Figure 12.8, it does not necessarily remain in the same position from one period to the next. For a number of reasons, which will be analysed in a later module, it is possible for the Phillips Curve to shift up or down over time. If it shifts up, then the full-employment inflation rate will be higher and the unemployment rate where inflation is zero will be higher. The rate of inflation that results in any period depends not only on where the economy is along the Philips curve, but depends also on the position of the Phillips Curve itself.

The fact that the Phillips Curve can shift over time causes some difficulty in interpreting historical data. Suppose the Phillips Curve were to shift upward from one period to the next and further suppose that aggregate demand declined at the same time. This is shown in Figure 12.9. Suppose, too, that the economy is at full employment initially, U_F on the Phillips Curve, with an accompanying inflation rate of INF_F. The corresponding levels of aggregate demand and national output are D_3 and Y_3 respectively and the corresponding employment rate (full-employment unemployment rate) is E_3. Now, the Phillips Curve shifts to the right and for some unspecified reason aggregate demand falls to D_2. The resulting unemployment rate on the Phillips Curve becomes U_2, associated with D_2, Y_2 and E_2 in Figure 12.9b. The inflation rate increases to INF_c on the new Phillips Curve. Driving the economy back to full employment, U_F, would result in an inflation rate of INF_F on the shifted Phillips Curve. In this case an increase in unemployment would be associated with an *increase* in inflation, and it would appear as though there were no trade-off between unemployment and inflation. In real life in national economies, because the Phillips Curve has shifted over time in the past two decades, there have been years in which the unemployment and inflation rates have moved in the same direction. This has led some people to conclude that there is no Phillips Curve.

To avoid confusion, it is important to remember that the Phillips Curve is a short-run relationship. The trade-off it describes pertains to a given short-run period. It indicates what inflation rate would result for different unemployment rates that might occur within the short-run period. It does not necessarily reflect what would happen if the unemployment rate changed from one period to the next. To determine that, it would be necessary to know how the Phillips Curve shifted, if at all, from one period to the next. The Phillips Curve simply shows what the consequences for inflation and unemployment would be in any given short-run period for various levels of aggregate demand.

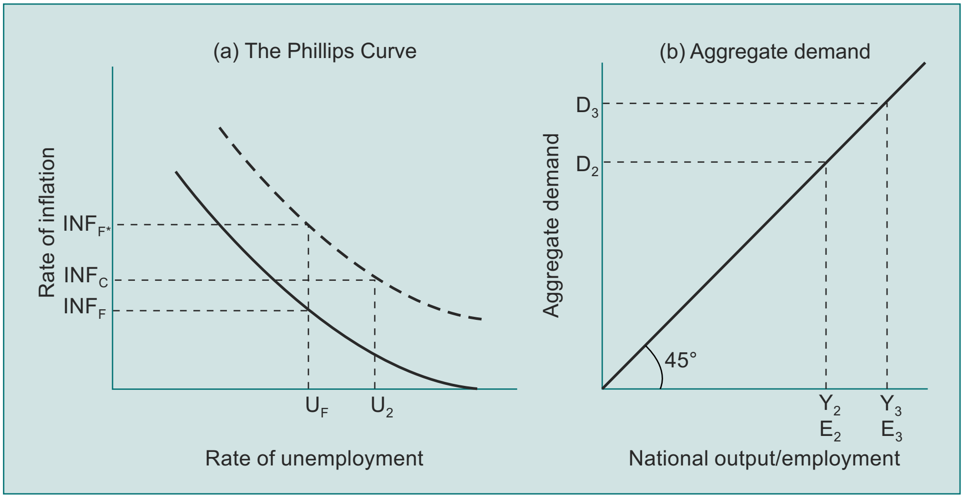


Figure 12.9 A shifting Phillips Curve and aggregate demand

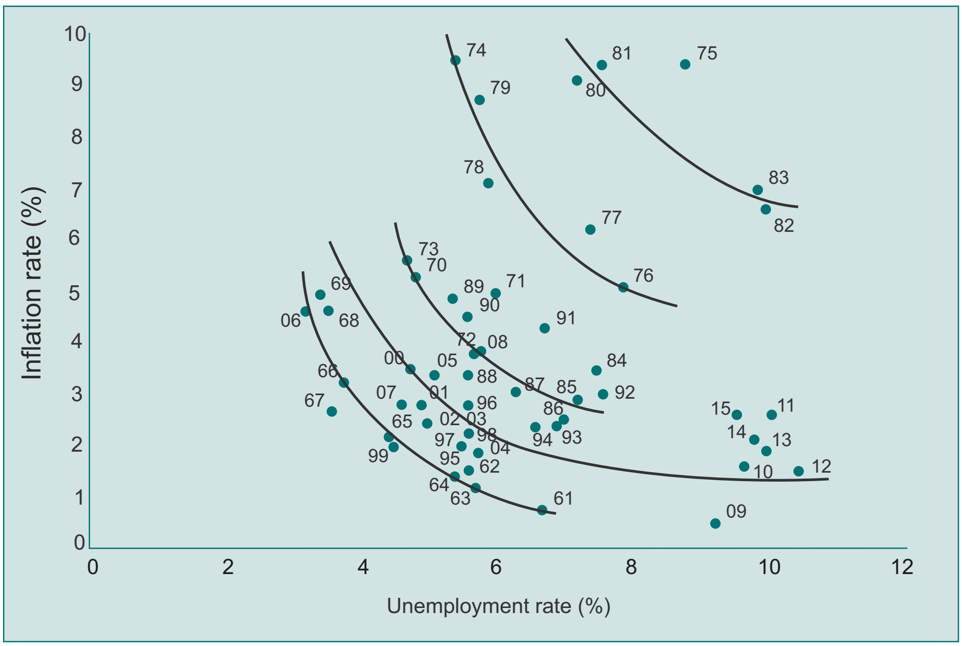


Figure 12.10 Shifting Phillips Curve (USA), 1960–2015

The existence of a short-run Phillips Curve poses a serious dilemma for government policy makers. To the extent that it can influence the level of aggregate demand and, thereby, where on the Phillips Curve the economy will be, the government must make a choice as to what combination of unemployment and inflation it will set. Its aggregate demand policies cannot be used, in the short run at least, to fight both unemployment and inflation at the same time. If they are used to get to full employment, some inflation will usually result; if they are used to eliminate inflation, high unemployment will usually result. Faced with this dilemma, policy makers must choose between two evils. They must decide where along the inflation–unemployment trade-off the economy should be. How a government decides this depends in part on the position of the Phillips Curve: when it is relatively low and to the left, policy makers tend to choose targets close to full employment. When the Phillips Curve is high and to the right, the short-term goal is likely to emphasise low inflation, even though this will mean higher unemployment.

It is important to understand that no policy decision about aggregate demand can be made solely on the basis of economic analysis. The appropriate policy also depends on the preferences of the people it affects. Whether or not the government should emphasise the battle against inflation in its policy goals will depend on which evil, unemployment or inflation, society considers to be more burdensome. Economic analysis can identify the trade-off between inflation and unemployment and it can indicate how the economy can be moved along the trade-off. The question of what point on the trade-off makes a society best off, however, depends entirely on the subjective opinion or judgement of the members of that society.

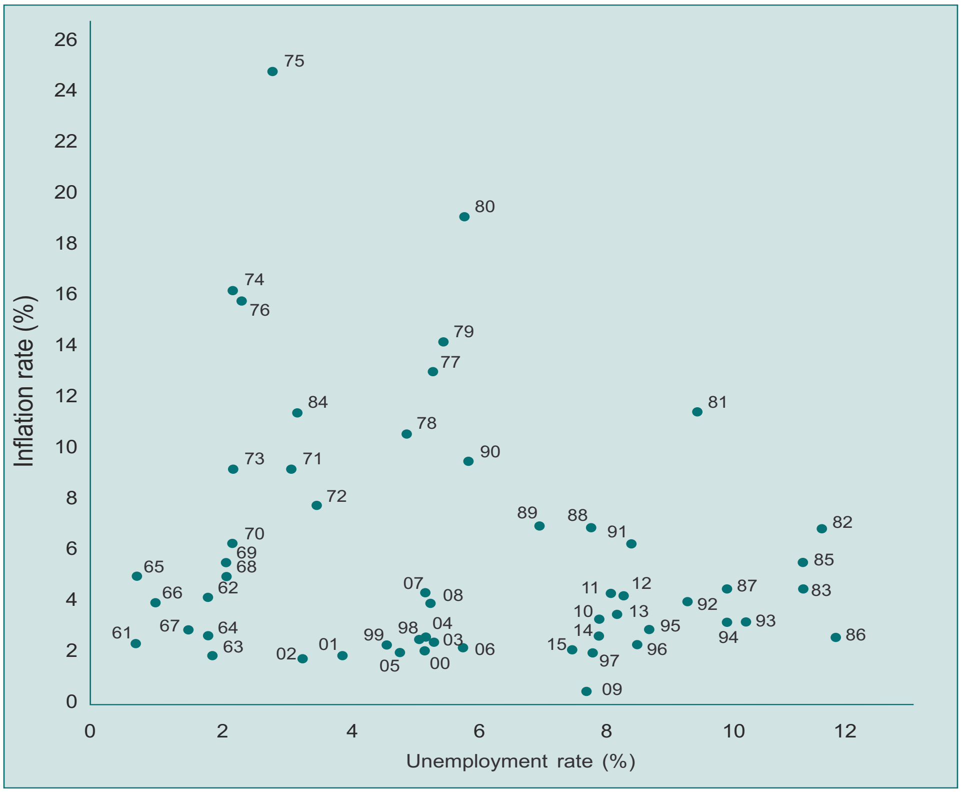


Figure 12.11 Inflation and unemployment data for the UK, 1960–2015

The problem of choosing the appropriate policy is compounded by the fact that society is composed of many members, who may differ as to whether inflation or unemployment is a worse problem. It is obvious that people such as construction workers, who are likely to become unemployed if aggregate demand is low, would regard a low-unemployment/high-inflation combination as relatively painless. Others, such as retired persons relying on their pensions for income, who would lose real income because of inflation if aggregate demand is high, would consider a high-unemployment/low-inflation combination to be relatively easy to bear. This is one reason why there is so much disagreement as to what constitutes the 'right' policy for a government to undertake. Because the goals of policy are essentially subjective in nature, people can agree about how the economy works and how policy affects the economy, yet disagree sharply about what should be done. Thus, policy makers will usually be in the position of being unable to please everyone, and decision making remains a political as well as an economic process.

Short-term policy making is also complicated by another factor: the short-run policy goal should not be decided without taking into account the long-run consequences of that policy. It might be possible, in normal situations, that high unemployment and low inflation today will permit certain preferred combinations of unemployment and inflation tomorrow that would be impossible otherwise. In such situations, the appropriate short-term policy goal might be to choose an aggregate demand target that is below potential GNP. Before the long-term implications of any policy can be analysed, however, it is desirable to understand fully the nature and short-run consequences of policies designed to control aggregate demand. This is the next topic to be tackled.

Learning Summary

You can now distinguish potential output from actual output and are thoroughly familiar with the representation of both of these concepts on the production possibilities frontier. You understand the factors that cause potential output to grow over time, i.e. growth in the quantity and quality of the labour force, and growth in the quality and quantity of the capital stock, recognising the important role of technological change. You also understand the concepts of gross investment, net investment and depreciation, and their impact on the capital stock and, simultaneously, the causes and characteristics of different types of unemployment in the labour force.

While later modules will analyse the causes of unemployment and inflation in detail, you currently understand why, when actual output is less than potential output, unemployment must result. You also have a basic grasp of the mechanisms that cause prices of goods and services and factor inputs to rise and fall, and why some prices are 'sticky' downwards, leading to the inflationary bias that exists in market economies. Understanding what inflation is, requires a grasp of price indexes, from which various measures of inflation are derived.

You are familiar with Okun's Law, the relationship between changes in the unemployment rate and changes in the output gap.

You recognise the Phillips Curve and understand how changes in aggregate demand influence both the unemployment and inflation rates in the short run. You are also aware that a shifting Phillips Curve over time can have inflation rates and unemployment rates moving in the same direction, whereas a movement along a Phillips Curve shows higher inflation rates associated with lower unemployment rates. The curvature of the Phillips Curve suggests the short-run inflation–unemployment trade-off is not constant but varies depending upon where an economy actually is on the curve.

Finally, you understand that the existence of the Phillips Curve can pose a dilemma for policy makers desiring low rates of both inflation and unemployment.

Review Questions

Multiple Choice Questions

12.1 The average well-being of people in a nation is determined by

I. the quantity and quality of resources under the nation's command.

II. how fully the nation utilises these resources in producing goods and services.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

12.2 Which of the following is correct? A production possibilities frontier illustrates that:

A. an economy's capacity to produce increases in proportion with population.

B. if all the resources of an economy are employed efficiently, more of one good can be produced only if less of another good is produced.

C. an economy automatically adjusts to that level of output at which all of its resources are employed efficiently.

D. various combinations of output can maximise the welfare of society.

Questions 12.3–12.6 are based on Figure 12.12.

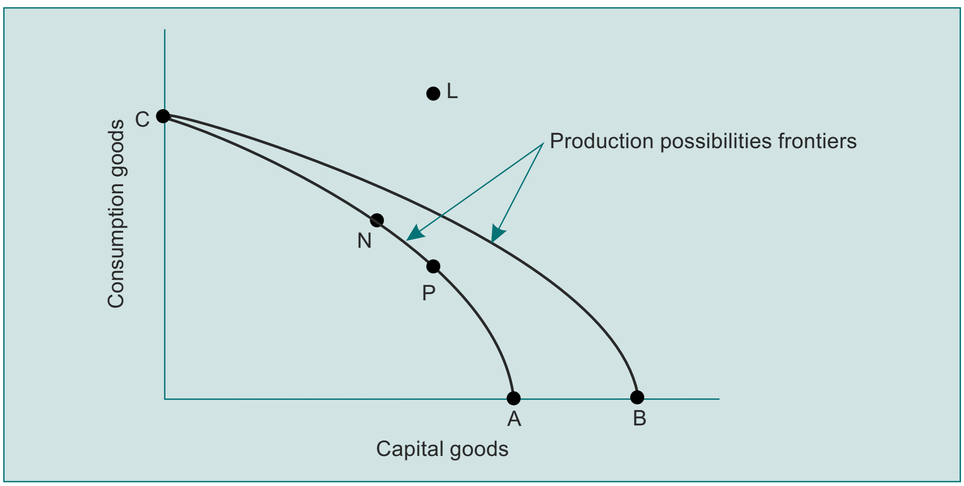


Figure 12.12

12.3 For an economy operating on the production possibilities frontier CA in Figure 12.12, which of the following correctly describes the combination of capital and consumer goods indicated by point L? It

A. would entail substantial unemployment.

B. would entail an inefficient use of society's resources.

C. is beyond the productive capacity of the economy.

D. suggests the productive capacity of the economy is declining.

12.4 Which of the following is correct? The choice of point P on production possibilities frontier CA in Figure 12.12 will

A. result in a faster rate of economic growth than would the choice of point N.
B. result in a slower rate of economic growth than would the choice of point N.
C. result in the same rate of growth as would the choice of point N.
D. be unattainable because P exceeds productive capacity.

12.5 Which of the following is correct? A shift of the production possibilities frontier in Figure 12.12 from CA to CB indicates

A. a movement from unemployment to full employment.
B. an improvement in capital goods technology but not in consumption goods technology.
C. an improvement in consumption goods technology but not in capital goods technology.
D. a decline in the total output of the economy.

12.6 After the production possibilities frontier in Figure 12.12 has shifted from CA to CB, the economy has moved from N to P. The result is

I. an increase in unemployment.
II. an increase in output.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

12.7 Assuming all resources in an economy are fully employed, which of the following would make potential GNP in that economy lower in the future than it otherwise would have been?

A. An increase in output of consumption goods and a decrease in output of capital goods.
B. An increase in output of capital goods and a decrease in output of consumption goods.
C. Technological progress that reduces capital inputs per unit of output.
D. Technological progress that reduces labour inputs per unit of output.

12.8 In advising the government on economic policy, the economist's role is

I. to outline the likely consequences of alternative policies.
II. to determine society's goals and the policies for achieving these goals.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

12.9 Which of the following would increase the potential GNP of an economy?

A. An increase in government expenditure on government employees' pay.
B. An increase in the proportion of the population that is available for employment.
C. A decrease in the rate of unemployment.
D. A decrease in output per worker.

12.10 Which of the following is correct? In practice, the potential output of the economy is calculated as the level of output

A. that would be achieved when demand-deficient unemployment was zero.
B. at which all unemployment would be demand-deficient unemployment.
C. at which the number of people seeking work exceeds job vacancies.
D. achieved in the current period.

12.11 Which of the following is correct? Actual output of an economy is the output

A. that would be produced if all resources were fully employed.
B. produced by currently employed labour and capital.
C. produced in the consumption goods sector.
D. produced in the capital goods sector.

12.12 Which of the following is correct? The higher the actual level of output in the economy in a given year,

A. the higher will be potential output.
B. the lower will be potential output.
C. the higher will be the unemployment rate.
D. the lower will be the unemployment rate.

12.13 If potential output in an economy exceeds actual output,

I. some labour must be unemployed.
II. some capital must be underutilised.

Which of the following is correct?

A. I only.
B. II only.
C. I or II or both I and II.
D. Neither I nor II.

12.14 If, in an economy, actual output equals potential output, the following types of unemployment cannot exist:

I. demand-deficient unemployment;
II. frictional unemployment;
III. structural unemployment.

Which of the following is correct?

A. I only.
B. II only.
C. II and III only.
D. III only.

12.15 In an economy, potential output has increased and this is accompanied by an increase in the unemployment rate. This could be due to

I. actual output increasing more slowly than potential output.

II. an increase in the number of married women wishing to work.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

12.16 Which of the following is an example of structural unemployment?

A. A building worker leaving one company and seeking employment from other builders in the local community.

B. A credit squeeze reducing the demand for all goods and hence employment.

C. Cheap oil from the North Sea reducing the demand for coalminers.

D. A bricklayer retiring because he can no longer face the demands of his job.

12.17 The level of frictional unemployment in an economy

A. is determined solely by macroeconomic policy.

B. depends on the nature and extent of job search.

C. results from obsolete industries closing down.

D. can be eliminated by lowering real wages.

12.18 Assume an economy moved from a state of high unemployment at the beginning of the year to full employment by the end of the year. Assume also no change in the capital stock or labour force (employed + unemployed). It can be concluded that

I. the flow of goods and services at the beginning of the year was below the economy's potential output.

II. average output per worker employed was higher at the end than at the beginning of the year.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

12.19 The following data refer to the output of guns and butter (both standard quality) in three economies in different years:

	Output of guns (no.)		**Output of butter (tons)**	
Economy	Year *t*	Year (*t*+1)	Year *t*	Year (*t*+1)
X	100	80	90	130
Y	100	100	100	90
Z	100	120	100	130

In which of the economies did economic growth occur?

A. Z only.

B. X and Z only.

C. X, Y and Z.

D. Not X, not Y, and not Z.

12.20 The following data in index form refer to a hypothetical economy:

Year	Actual GNP	Potential GNP	Population
t	100	150	100
t + 1	125	150	100
t + 2	150	150	100

From the data it can be concluded that

I. per capita GNP did not increase between Year *t* and Year *t* + 2.

II. the labour force did not increase between Year *t* and Year *t* + 2.

III. technological progress did not occur between Year *t* and Year *t* + 2.

Which of the following is correct?

A. I and II only.

B. I and III only.

C. II and III only.

D. Not I nor II nor III.

Questions 12.21 and 12.22 are based on the Phillips Curve in Figure 12.13.

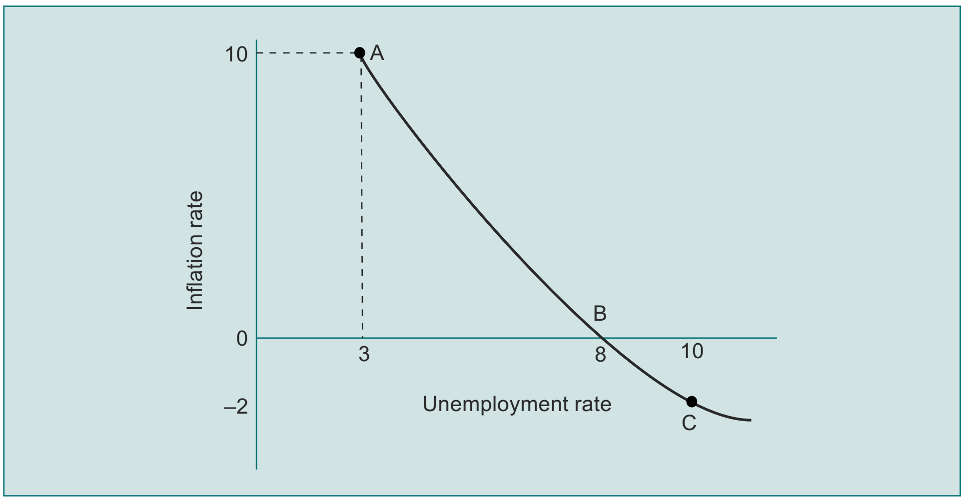


Figure 12.13 Sample Phillips Curve

12.21 The government is considering which policies to adopt and is using the Phillips Curve in Figure 12.13 in its deliberations. It has decided that

I. Policy A will result in a 10 per cent inflation rate and a 3 per cent unemployment rate.

II. Policy B will result in a zero inflation rate and an 8 per cent unemployment rate.

III. Policy C will result in a −2 per cent inflation rate (i.e. a 2 per cent deflation rate) and a 10 per cent unemployment rate.

Which of the following is correct?

A. The best policy is A because it yields the lowest unemployment rate.

B. The best policy is B because it yields the lowest inflation rate.

C. The best policy is C because it will lower inflationary expectations.

D. The best policy is dependent entirely on the subjective opinions of the population of the economy.

12.22 If the Phillips Curve in Figure 12.13 were to shift to the left, the following would occur in the economy:

I. the unemployment rate would decrease.
II. the inflation rate would decrease.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

Module 13

The Circular Flow of Income

Contents

13.1 Introduction

This module begins the process of developing a short-run theory of income determination. It is short-run because it is concerned with explaining the relationship between actual and potential output at a given point of time, and to emphasise the short-run nature of the model it is assumed that technical knowledge is unchanging and therefore that the production possibilities frontier is fixed. Such an assumption is unrealistic in the long run, because historically the production possibilities frontier has shifted outwards with time, and the productive capacity of the economy has increased. Nonetheless, over short periods of time the assumption is a reasonable approximation of reality because, in normal circumstances, the supply of the factors of production increases only slowly, as does technical change. The assumption that the production possibilities frontier is fixed does not do undue violence to the facts *provided* that the short-run nature of the theory is recognised. The corollary to this, however, is that the short-run theory that has been developed, while it may be useful in explaining whether a society falls short of, reaches, or exceeds its current production possibilities frontier, is unlikely to be of much assistance in explaining the process of economic growth.

From the assumption that the stock of technical knowledge is fixed, it follows that there will be a fixed relationship between output (income) and employment. Given that technical knowledge cannot be changed, the productivity of labour cannot be altered, so that output and employment must move in the same direction: it will not be possible to increase output without increasing employment, or to increase employment without increasing output. Again, this assumption is not entirely consistent with actual experience, but it is an observed fact that, in the short run, output and employment do tend to move in the same direction, illustrating the need to employ more labour if output is to be increased and also that reduced output is accompanied by a fall in employment. Hence, while the assumption of a fixed relationship between employment and output is an oversimplification, it is useful as a pedagogical device, as it allows us to treat the short-run theory of income and output determination as synonymous with the short-run theory of employment determination.

The essence of the short-run theory of income determination is that the level of actual income realised is determined by aggregate demand. The relationship between aggregate demand and aggregate supply determines whether a society experiences unemployment of factors of production and a failure to realise productive potential, full employment at capacity output, or overfull employment and rising prices. The remainder of this module develops the theory of the circular flow of income, which underlies all short-run income–expenditure models.

13.2 The Structure of the Economy: A Two-Sector Model

To understand how aggregate demand is determined, it is necessary to study the structure of the economy or, to put it another way, to study the transactions that take place among the principal groups performing economic activities. We shall assume, initially, a very simple economy consisting of only two groups of decision makers: households and firms. Hence the title of 'a two-sector model'.

The population is composed of households, some of which consist of only one individual, others of families. All households perform one economic activity and most perform two. The activity in which all households participate is as consumers: they buy goods and services from firms to satisfy their wants. The second activity arises from the fact that all resources in the economy are owned by households; all the land, labour, capital goods and enterprises are owned by households. These resources are hired by firms to produce goods and services. For example, households are known to buy food, clothing, heating, light, shelter, consumer durables, entertainment, cigarettes, alcohol, switchblade knives, horror comics, and pornographic literature, to name but a few items, from the firms that produce them. Most households also sell labour services to these firms to produce the goods and services.

Households, either singly or operating in groups, also form entities, known as firms, that hire resources from other households and produce goods and services. They also buy goods and services from other firms. IBM, for example, is owned by hundreds of thousands of households, though some households own a larger share than others. IBM has large numbers of employees working in its facilities throughout the world; it buys raw materials, goods and services from thousands of other firms and produces a broad range of electronic goods. The corner grocery store, on the other hand, may be owned by a single family but it also deals with a large number of other firms to enable it to produce a variety of consumer goods for its customers.

In order to build up a simplified version of how an economy operates, it is necessary to classify markets into two groups:

(a) the markets for goods and services that producers buy from households; and
(b) the markets for goods and services that producers sell to households.

In this simplified economy there is a circular flow of income and expenditure among households and producers, as shown in Figure 13.1. Households (H) sell services to firms (F), which are needed to produce goods and services; in turn households use the income they receive to purchase the goods and services that are sold by producers.

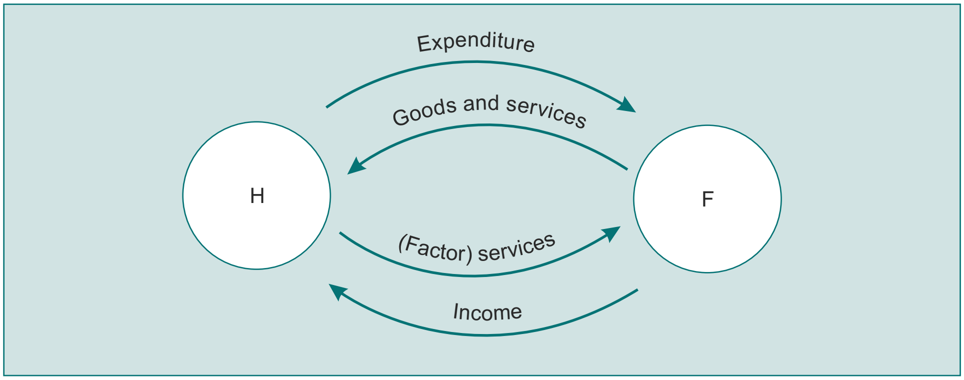


Figure 13.1 The circular flow of goods and services, and income and expenditure

13.2.1 National Income Accounting

The national output is the flow of all final goods and services in an economy within a given period (usually a year). It is useful to distinguish between Gross National Product (GNP) and *Net National Product* (NNP). Net National Product sets the limit to consumption that does not entail the penalty of a lower level of output and consumption in the future. If consumption were to exceed net national output, society would not provide sufficient resources for capital investments to make good the depreciation of the stock of capital assets. This is 'prodigality' in our production possibilities frontiers in Figure 12.4. The relationship between GNP and NNP is given by:

$$\text{GNP} - \text{Depreciation} = \text{NNP}$$

In calculating GNP (NNP), it is necessary to place values on the output of the economy. The values are taken to be the prices paid for goods and services. However, it is important that only *final* goods and services are included and that *intermediate* goods and services are excluded from the calculation; intermediate goods and services are used up in the production of the final goods and services, and if their values were included in the calculation of GNP, double counting would occur.

Example

Consider the output of two goods, cars and steel. Suppose that steel is used only to produce cars, and the values of the two goods produced this year was $30 million for cars and $10 million for steel. Because the steel is used up in the production of the cars, part of the value of the cars comes from the steel used to produce them. Consequently, if the value of cars and the value of steel produced are added together, the value of the steel will be counted twice, providing a value for the output of both industries of $40 million instead of the $30 million which is the value of the final good, cars, produced.

GNP can be calculated in three different ways:

(a) by finding the expenditure on final goods and services;
(b) by finding the value added by each producer; or
(c) by finding the total income earned by each factor of production.

Gross National Expenditure (GNE) is the total amount spent by households (and by government and/or firms acting on their behalf) on final goods and services. In computing GNE, it is often difficult to determine in practice the distinction between final and intermediate goods. *Final goods and services* can be defined as all goods and services not used up in the production of other goods and services during the time period under consideration. *Intermediate goods and services*, by contrast, are any goods or services used up in the production of other goods or services. Continuing the cars/steel example, suppose that $5 million worth of cars was bought by firms producing chauffeur-drive hire services; for their services, households paid a total of $10 million. Table 13.1 shows the sale and purchase of intermediate and final goods and services, and the identification of GNE, i.e. the expenditure by households on final goods and services.

Table 13.1 GNE: The elimination of intermediate sales ($m)

Buyers	Steel	Cars	Hirers	Households	Total purchases
Steel	–	–	–	–	–
Cars	10	–	–	–	10
Hirers	–	5	–	–	5
Households	–	25	10	–	35*
Total sales	10	30	10	–	50

* Gross National Expenditure

The value added by a producer is the value of the output of the firm in a given period less the value of the inputs used in production; hence, addition of the values added by producers automatically eliminates the problem of intermediate goods. The sum of *values added* by all producers in the economy is another expression of the concept of GNP. Consider again the production of steel, cars and hires. Suppose that steel is the only intermediate good used in producing cars and that no intermediate goods are used in producing steel. The value added by the three industries is shown in Table 13.2.

Table 13.2 GNP as value added by producers ($m)

	Value of input	Value of output	Value Added
Steel	0	10	10
Cars	10	30	20
Hirers	5	10	5
Total Value Added			35

The figure for value added is the same as that obtained for GNE, and does not involve the problem of determining which are final and which are intermediate outputs; for example, it is not necessary to differentiate between the cars bought by households and the cars bought by car hirers.

The income that accrues to those who provide the factor services that producers use in the production of goods and services is derived from the value added at each stage of production. All of value added must accrue as a factor income to some household. Hence, the sum of all factor incomes is also equal to the value of GNP. In the steel/cars/hire example, the value added at each stage of production will be distributed to households in the form of wages, salaries, rent, interest, profits or dividends. If all these factor incomes are added together, they will amount to $35 million, so that gross national income = gross national product = gross national expenditure.

In addition to intermediate goods, there are other inputs to production. These are primary factors of production and include labour, land, capital and enterprise. Primary factors of production are all inputs used to produce output in the current period. Labour services are primary inputs because they are used to produce current output, but are not themselves produced in the current period, as skills and experience are acquired as a consequence of past decisions. Similarly, buildings and machinery produced prior to the current period are primary factors since they are used to produce current output, but are not themselves produced in the current period.

The income paid to the owners of primary factors employed by a producer must be financed from the firm's sales receipts and are payments to the owners of the factors of production in return for their use in the production of output. Income is normally classified as:

(a) *wages and salaries* – paid in return for the use of labour services;
(b) *rent* – paid in return for the use of land and capital goods not owned by the producer;
(c) *interest* – paid to the householders who have loaned money to purchase land and capital; or
(d) *gross profits* – residual accruing to the firm after all payments of wages, salaries, rent and interest have been made.

Example

Consider a firm with sales receipts of $1000 per week and payments that must be met to ensure that production takes place:

Payments	**$**
Intermediate goods	350
Wages and salaries	300
Rent and interest	150
Total	800
Residual: gross profit	200

The sum of payments by a producer for intermediate goods and for primary factors of production equals total receipts from sales of output; therefore a producer's value added is equal to the sum of payments to primary factors.

We can then arrange the firm's receipts in slightly different manner:

Receipt	**$**	**Payments**	**$**
Sales	1 000	Intermediate goods	350
		Wages and salaries	300
		Rent and interest	150
		Gross profit	200
Total receipts	1 000	Total payments	1 000

Total payments equal total receipts because one factor income (gross profits) is, by definition, what is left of receipts after all other payments have been made. By subtracting purchases of intermediate goods from both sides, the firm's account can be reconstructed as follows:

		$		$
Sales		1 000	Wages and salaries	300
less	Purchases of		Rent and interest	150
	intermediate goods	350	Gross profit	200
Value added		650	Payments to primary factors of production	650

Receipts add up to equal the firm's value added. The payments side adds up to the firm's payments of income to owners of primary factors of production.

Since GNP equals the sum of producers' values added, GNP is also equal to the sum of producers' payments to primary factors of production, i.e. Gross National Product equals Gross National Income (GNI). Since GNP also equals Gross National Expenditure, it follows that GNE equals GNI.

These three methods of measuring national output or income are summarised in the national income accounts of every country. The accounts provide a method of assessing past economic activity, allow the measurement of living standards over time, and give an assessment of differences in living standards among different economies. Table 13.3, Table 13.4 and Table 13.5 show the national income accounts for a hypothetical economy, national income being measured by the expenditure, income and output methods, with the circular flow of output, income and expenditure.

Table 13.3 GNP (expenditure approach)

Category	Symbol	$bn
Consumption expenditure	C	59
Gross private domestic investment	I	19
Government purchase of goods and services	G	20
Exports minus imports	X −Z	2
Gross National Product	Y	100

Table 13.4 GNP (factor incomes approach)

Category	$bn
Compensation of employees	61
Rental income	1
Interest income	9
Self-employment income	7
Firms' profits	6
Indirect taxes − subsidies	6
Depreciation	10
Gross National Product (Y)	100

Table 13.5 GNP (output approach)

Industry	$bn
Agriculture, forestry and fishing	3
Mining and quarrying	3
Manufacturing	26
Construction	8
Transport	5
Gas, electricity and water	3
Distribution trades	12
Insurance, banking, finance and business services	14
Other services	14
Public administration, defence, health, education	15
Statistical discrepancy	−4
Gross domestic product (GDP)	99
Net income from abroad	1
Gross National Product (Y)	100

The categories in the hypothetical accounts in the three tables do not accord precisely with the categories in the national income accounts of any of the leading market economies in the world today. They are a sort of pastiche of the world's leading economy. Similarly, the relative magnitudes of the data in each category will not accord with those for any given country.

The importance of these accounts, nevertheless, is to reveal the anatomy of a typical developed market economy and to assist in the understanding of the circular flow of output, income and expenditure. We shall now explore how the actual level of national output is determined.

13.3 The Equilibrium Level of National Income

As stated in Section 13.2, national income can be measured in three ways: as a sum of expenditures, as a sum of outputs (value added) and as a sum of factor incomes. All three measures, in theory, will yield the same estimate of national income. In practice, of course, they seldom do. (To 'ensure equality', we introduce (as we did in Table 13.5) an item known as 'statistical discrepancy' or 'residual error'.) It is this equivalence between national output and national income that is at the heart of short-run macroeconomic analysis.

The equivalence between national output and national income is a consequence of definition of profits as a residual, obtained after deducting from the value of output all other factor incomes: wages and salaries, rent, interest and dividends. Given that profits are the residual, it follows that the sum of all factor incomes – wages and salaries, rent, interest, dividends, and profits – must equal the value of national output. In other words, the whole of the value of output must accrue as a factor income to some household; or, to put it another way, the income enjoyed by a community is determined by the flow of final goods and services it produces.

This equivalence between output, income and expenditure underlies the 'circular flow' model; income created by the process of production is equivalent to the value of that production. At each stage of the productive process the income that accrues to households owning, directly or indirectly, the factors of production is exactly equal to value added. In short, the very process of production will create factor incomes equal to the value of that output.

It follows that the income received by the households must be sufficient to purchase all the output produced by business firms. Now, consider why any firm sets up in business. It does so in the expectation of finding a market for the product or service produced. Provided that the business produces a good or service people want, it would seem that the increase in supply (output) must generate an equivalent increase in demand (income). That is, the very process of production creates a market, in that it creates an income accruing to households that is equivalent in value to the value of output produced. However, the business will only be able to sell its additional output if the increase in income is translated into *effective demand*, i.e. if the additional income is translated into actual *expenditure*. Clearly, if the additional income received by households is not translated into expenditure, then the business will not be able to find a market for the additional goods produced. The business will incur losses, or even go into liquidation, and as a consequence output will fall.

The level of output that can be sustained is dependent, therefore, on the level of expenditure or effective demand. Potential output sets the limit to the output (income) of a community, but while capacity output sets the limit, actual output may fall short of that limit. In the short run, actual output can and does diverge from capacity output, and we have to devise a short-run theory of income determination that is consistent with this observation. This short-run theory sees the level of actual income and output as dependent upon the level of aggregate demand, with capacity output setting the limit to real income and output. The remainder of this module develops the basic theory of the circular flow of income and shows how the equilibrium level of national income depends upon the level of aggregate demand.

We shall begin with a simple model of the circular flow, which is subsequently elaborated to reflect more accurately the structure of the economy. We shall assume initially that:

(a) the stock of technical knowledge is fixed, as is the supply of factors of production, so that potential output of the economy is given and does not change;
(b) there is a fixed relationship between output (income) and employment, so that output and employment move in the same direction, and once a short-run theory of output (income) determination has been developed, the same theory serves to explain the level of employment;
(c) there is no international trade – the economy is a 'closed' economy;
(d) there is no government sector, i.e. there is no taxation and all goods and services are provided by private businesses;
(e) all saving is carried out by private households, so that there are no 'retained profits' by business firms;
(f) all investment is carried out by business firms; and
(g) all prices are constant, so that any change in national income (Y) is a consequence of a change in *real national income*, i.e. in the *output* of final goods and services.

We shall also define some terms more precisely to avoid confusion, as follows:

- *Consumption* (C) consists of expenditure on goods and services to satisfy current needs. Complications are introduced by expenditure on consumer durables that yield a flow of services over time, e.g. a car; we shall ignore consumer durables, therefore, for the sake of clarity.
- *Savings* (S) is income not spent on consumption, i.e. it is the residual obtained by subtracting consumption from household income.
- *Income* (Y): it follows that income has two components, consumption (C) and saving (S) so that:

$$Y \equiv C + S$$

- *Investment* (I) is the production of goods that are not used for consumption purposes; these goods are known as *investment goods*. There are two main categories of investment goods: inventories and capital goods.
  (a) *Inventories*: most firms hold stocks of inputs required in the production process and stocks of their output. These stocks reduce delays in meeting production targets and in meeting customers' orders. Changes in inventories are considered as part of investment, because they represent changes in the volume of goods produced that are *not* used for current consumption. An increase in inventories represents additional investment, while a reduction in inventories represents *disinvestment*. Changes in inventories can be intended or unintended. If a firm decides to increase (decrease) output, it is likely to increase (decrease) its inventories of inputs and outputs. In this case, the change in inventories is *intended* or *planned*. Alternatively, the change in inventories may result in an error from forecasting sales – for example, if sales are less than expected, stocks of outputs will rise, and such investment is *unintended* or *unplanned*.
  (b) *Capital goods*: the productive capacity of an economy is partly dependent upon its capital stock, consisting of factories, offices, machine tools, airports, harbours, roads, railways, etc. Investment includes the production of all new capital goods. As has been previously stated, investment may be to make good depreciation, or it may involve net addition to the stock of capital goods.

  Total (gross) Investment (I) thus includes investment in inventories and investment in capital goods:

$$\text{Gross Investment (I)} \equiv \text{replacement investment} + \text{net investment.}$$

  Remember from our previous discussion that net investment can be positive or negative. As the concept of investment (I) used in the following model is *gross investment* (including replacement investment), Y will be taken to represent *Gross National Product*. In short-run macroeconomic analysis, this is usually taken as a more appropriate measure of income than *Net National Product* because expenditure on replacement investment, which is included in GNP, is part of effective demand as it creates factor incomes and employment.

13.4 The Savings–Investment Schedule

Gross national product (Y) is a flow of goods and services available for satisfying wants. When these goods and services are used to satisfy *immediate wants*, this is called consumption (C). The remainder of output represents goods that are not consumed immediately, but are added to inventories or the stock of capital goods and together constitute investment (I). Hence

$Y \equiv C + I$

As we have seen, national income (Y), can be used to purchase consumption goods, and what is not consumed is called saving (S). So

$Y \equiv C + S$

but, given that also $Y \equiv C + I$, we must have that $I \equiv S$. That is, investment (I) and saving (S) are so defined that they are necessarily equal in the national income accounting sense. This does not mean that planned saving always equals planned investment; on the contrary, planned saving and planned investment are only equal at equilibrium national income. From the point of view of national income accounting, actual saving by definition must always equal actual investment. From the point of view of economic theory and analysis, it is the relationship between planned saving and planned investment that is important. To this topic we now turn.

13.5 The Concept of Equilibrium

Given the assumptions adopted above, it is possible to develop a simple model of the circular flow of income in which there are only two principal players, private households (H) and private business firms (F). The model demonstrates that while actual savings and actual investment, as defined for national income accounting purposes, must always be equal, planned savings and planned investment are not always equal and, indeed, are only equal when equilibrium national income is established.

If we suppose initially that all income is consumed, then:

(a) all value added (output) would accrue to private households and factor incomes; and
(b) all factor incomes would be used by households to purchase consumption goods and services provided by business firms (producers).

If output (income) were \$100 billion, then the circular flow of income could be represented by Figure 13.2.

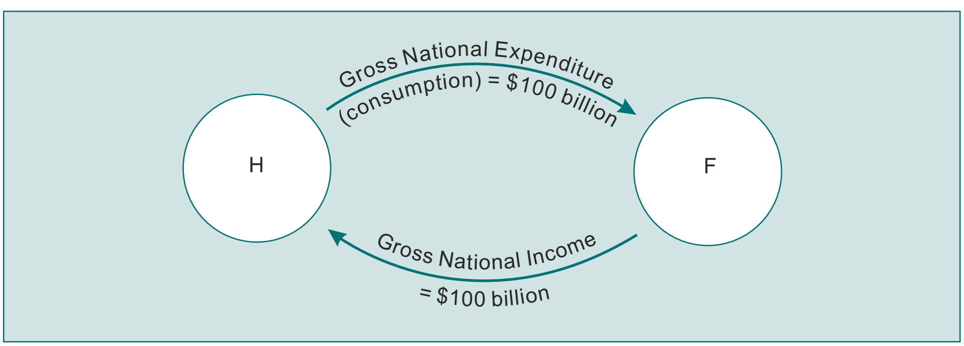


Figure 13.2 Circular flow of income: all income consumed

In such an economy, supply would create its own demand, as all income would be consumed; i.e. whatever the households received in the form of factor incomes from the business firms, the households would return as consumption expenditure to the business firms. There would be no withdrawals from the circular flow of income and no injections, so that any flow of national income established would continue indefinitely. The system would be in equilibrium, and there would be no tendency for the level of national income to change.

While this is strictly true given the assumption of a fixed capacity output, it is clear that any such assumption will be unrealistic in the real world. All the goods produced by an economy would be goods used to satisfy immediate consumption needs. There would be no investment to make good the depreciation of the stock of capital goods and, if this situation were realised in practice, then our first assumption of a fixed capital stock could not hold. The productive capacity of the economy would diminish as the productive capacity of the existing capital stock declined in the absence of any replacement investment. In other words, the economy would be in the situation described for 'Prodigality' in Module 12's Figure 12.4.

In practice, most economies save and invest a proportion of national income (output). Savings are not passed on in the circular flow of income, but are a *withdrawal* from that circular flow. Thus, savings do not constitute a component of aggregate demand (expenditure) and the act of saving does not create a demand for output, thereby generating income and employment. On the other hand, investment is an *injection* into the circular flow of income. It is part of aggregate demand (expenditure), as the act of investment creates a demand for output and results in income and employment creation.

A *withdrawal* is any part of the income of private households that is not passed on in the circular flow of income. An *injection* is any addition to the income of domestic firms that does not accrue from the expenditure of private domestic households. Any withdrawal has a *contractionary* effect on the level of national income. The act of saving does not constitute part of aggregate demand; it does not create a demand for output which gives rise to income and employment. A rise in savings will, other things being equal, tend to reduce national income, and vice versa. Any injection has an *expansionary* effect on the level of national income. The act of investment does constitute part of aggregate demand, as it creates a demand for investment goods and results in income and employment. A rise in investment will, other things being equal, tend to increase national income, and vice versa.

It follows from this that equilibrium national income can only be achieved when there is consistency between the plans of savers and the plans of investors. If planned saving (assumed to be undertaken by households) and planned investment (assumed to be undertaken by business firms) are equal, then national income will be in equilibrium and will show no tendency to change. If planned savings and planned investment are *not* equal, national income cannot be in equilibrium and, instead, *must* change – and must continue to change until the equality of planned saving and planned investment is achieved.

This can be illustrated by reference to Figure 13.3, which shows the circular flow of income modified to allow for savings from private households and investment by private firms. Initially it is assumed that planned savings are one-tenth of income of \$100 billion and that investment is fixed at \$10 billion, so that planned saving equals planned investment (S = one-tenth × Y; Y = \$100 billion; therefore S = \$10 billion = I).

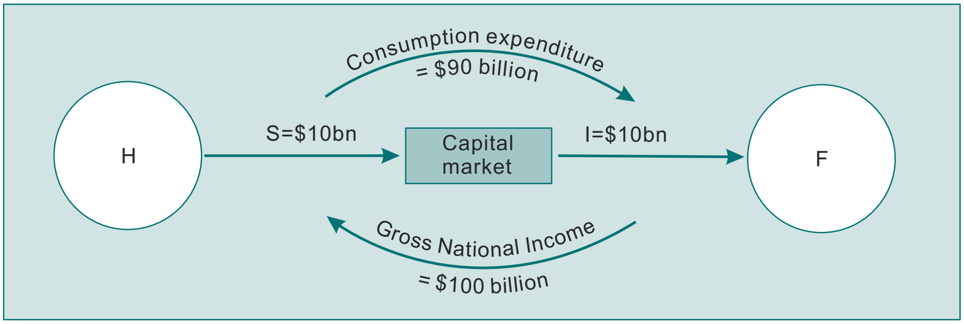


Figure 13.3 Circular flow of income: planned savings equals planned investment

In Figure 13.3 the system is in equilibrium. Notice that the composition of output is different from that shown in Figure 13.2 because output comprises both consumption goods ($90 billion) and investment goods ($10 billion). Because the withdrawal from the circular flow of income of planned savings is equal to the injection to the circular flow of income of planned investment, the system remains in equilibrium with national income equal to $100 billion. Aggregate demand, equal to consumption plus investment expenditures (C + I), is equivalent to the aggregate supply of goods and services produced by the business firms. The situation will be sustained so long as there is no change in the plans of savers and/or investors.

A change in planned savings and/or planned investment will produce a change in national income as follows:

(a) an increase in planned saving will produce a reduction in national income, unless it is offset by an increase in planned investment;
(b) a reduction in planned saving will produce an increase in national income, unless it is accompanied by a reduction in planned investment;
(c) an increase in planned investment will produce an increase in national income unless it is offset by an increase in planned saving;
(d) a reduction in planned investment will produce a reduction in national income, unless it is accompanied by a reduction in planned saving.

Thus, changes in national income will be in the same direction as changes in planned investment, and will be opposite in direction to changes in planned savings, unless offset by other factors.

Now let us assume that households plan to save one-fifth of income instead of one-tenth of income. This is shown in Figure 13.4. Given no changes in the plans of investors, national income will fall. Why?

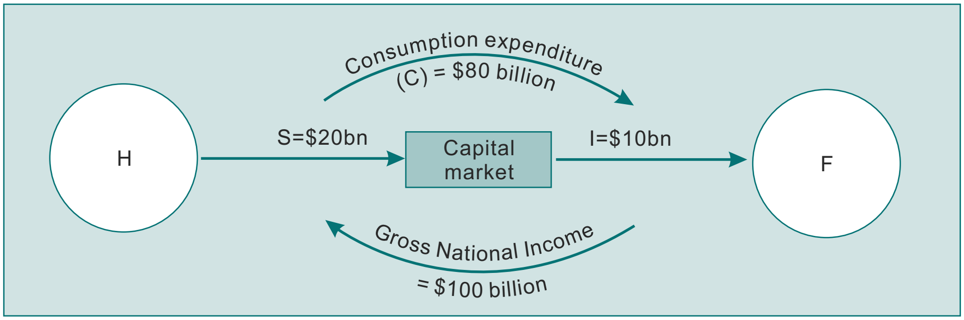


Figure 13.4 Circular flow of income: inequality of planned saving and planned investment

If planned savings equal one-fifth of Y, then planned saving will rise from $10 billion to $20 billion, so that planned savings (withdrawals) will be greater than planned investments (injections). In consequence, national income will fall. Initially, aggregate demand was equal to $100 billion with consumption (C) equal to $90 billion, and investment (I) equal to $10 billion; now, consumption expenditure has fallen to $80 billion and investment will be unchanged at $10 billion, so that aggregate demand equals $90 billion. The receipts of business firms will have fallen accordingly. In consequence, business firms will not be able to sell all the output produced and there will be an *unplanned increase* in inventories, which must continue so long as output is sustained at the initial level.

Evidently, this cannot represent the equilibrium position and, in time, business firms will react to the fall in aggregate demand by reducing output, as in Figure 13.5. Notice that the new equilibrium level of national income will be lower than $90 billion; the fall in national income will be some multiple of the initial increase in planned saving of $10 billion. This is so because if national output (income) were $90 billion, then planned saving (S) would be $(1/5)Y = 1/5 \times \$90$ billion = $18 billion, and would still be greater than planned investment. In our example the initial decrease in consumption expenditure of $10 billion produces a decrease in national income of $50 billion. The size of this multiple contraction depends, as we shall see later, on the size of something called the *multiplier*.

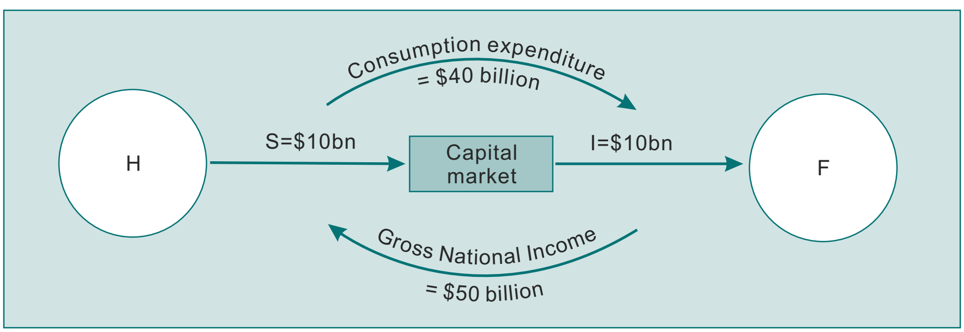


Figure 13.5 Circular flow of income: new equilibrium position

In a modern economy, the ownership of the means of production is dispersed. The greater part of economic activity is controlled by large corporations, usually owned by shareholders scattered across thousands of families. Retained business profits are still an important source of finance for business investment, but in most modern economies the larger part of business investment is financed by sources external to the firm, rather than by retained profits. Typically, the greater part of saving in a modern economy is carried out not by businesses in the form of retained profits but by private households.

Most important of all, the motive for saving is often different from the motive for investment. In addition, a firm may retain profits not to invest in the current period but to finance investment at an unspecified future date, or as a precaution against some anticipated or unforeseen adverse development in trading conditions, or to improve the liquidity position of the firm. Individual households may save to buy some goods in the future; to protect against a 'rainy day' (including becoming unemployed); because they have contracted into some insurance or savings programme; to leave wealth for their children; or through sheer miserliness. The desire to save is not determined by the availability of real investment opportunities, namely the likely profitability of investment in inventories or capital goods. But, in contrast, it is precisely the prospect of such profitability that determines investment. Investment takes place in the expectation of profitability and the volume of investment is clearly a function of the availability of investment opportunities. Even when investment opportunities are considered unfavourable, individual households will still wish to save and will hold their savings either in cash or in financial assets, in the absence of alternative outlets. The motives of savers and investors being different, there is no guarantee that the plans of savers and investors will be consistent in the short run at all levels of income. For this reason, the level of national income realised can and does depart from full-employment income.

Learning Summary

You now understand, given the assumptions of the simple model devised, that output and employment move together in a fixed relationship.

There are two vital lessons to be learned from this module. The first is that economic activity is like a coin – it has two faces; it can be recognised and measured by looking at the output side or the income side – it doesn't really matter which, for you get an identical measure of economic activity. If Robinson Crusoe and Man Friday shake down 20 coconuts, the output of their island economy is 20 coconuts; their income is also 20 coconuts. The second important lesson is that while income recipients earn sufficient income to purchase the output they produce, they may not elect to spend that income on the currently produced output of goods and services. This decision determines whether output in the next time period increases, remains constant or decreases.

You now understand the difference between intermediate goods and services, and final goods and services; only final goods and services are considered in the GNP accounts. You can also recognise that GNP can also be calculated by summing the value added of all firms in the production process.

If households plan to spend less on goods and services than firms anticipate, household planned savings will exceed firms' planned investments and there will be a 'leakage' from the circular flow of income, causing total output to contract. Conversely, if planned savings are less than planned investment, with people buying more goods and services than firms anticipated, output will expand – provided always that sufficient unemployed resources exist.

Review Questions

Multiple Choice Questions

13.1 In year 1 an economy is operating on its production possibilities frontier. In year 2 the level of technology, the capital stock and the labour force are the same as in year 1 (i.e. the production possibilities frontier has not moved) but output is lower than in year 1. This means that

I. at least one resource must not be fully utilised.

II. the output of capital goods must have decreased.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

13.2 In the circular flow of income, households

I. purchase goods and services.

II. sell labour and capital services.

III. invest in plant and equipment.

Which of the following is correct?

A. I only.

B. I and II only.

C. III only.

D. I, II and III.

13.3 In the theory of income determination

I. the excess of expenditures on goods and services by consumers over the value of the goods and services consumed equals the returns to factors of production.

II. the returns to factors of production minus the value of goods and services produced provides investment resources.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

13.4 If, in a certain economy, the flow of goods and services equalled £200 billion, it follows that

A. the value of the capital stock was £200 billion.

B. the value of labour services was £200 billion.

C. national income was £200 billion minus depreciation.

D. expenditure on final goods and services was £200 billion.

13.5 In a simple two-sector model of an economy, the higher the level of

I. national income, the higher the level of national output.

II. national output, the higher the level of national expenditure.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

13.6 For which of the following is the simple two-sector model of the economy useful? To show

A. the equality between expenditure on goods and expenditure on services.

B. that households and firms have conflicting interests.

C. the equality between national income and national expenditure.

D. that wages + salaries equals expenditure on final goods and services.

13.7 Calculation of Gross National Product includes

I. purchase of 100 shares in IBM by a company pension fund.

II. purchase of $1 million of machine tools by car companies.

III. the accumulation of tractors by the manufacturers in their warehouse.

Which of the following is correct?

A. I only.

B. II only.

C. II and III only.

D. I and III only.

13.8 Last year all the bakers in an economy produced bread which sold for $55 million. The value of the flour required for the bread was $35 million and the value of the wheat required for the flour was $20 million. The contribution to GNP of the bread industry last year was which of the following?

A. $5 million.

B. $50 million.

C. $55 million.

D. $105 million.

13.9 For any particular receipt to be counted as part of the national income of an economy, which of the following must be true? It must

A. be in return for a currently produced good or service.

B. be in return for the use of resources in the production of current output.

C. represent payment on borrowings.

D. represent an increase in the resources of the recipient.

13.10 Which of the following is correct?

The sum of all expenditures on final goods and services is equal to

A. all income paid out as wages and salaries.

B. the sum of money in circulation.

C. the total of all 'values added' at each stage of production.

D. disposable income plus investment expenditures.

13.11 In theory, Gross National Income (GNI) equals Gross National Expenditure (GNE). In constructing the National Income Accounts, such equality is attained normally by including a ‘residual error’ term. This procedure is necessary to achieve balance because of which of the following?

A. It is necessary to subtract savings from GNI to arrive at total spending.

B. GNE includes spending on intermediate goods, whereas GNI is derived only from ‘values added’.

C. Although GNI and GNE each measure economic activity, statistical errors result in inaccuracies.

D. The wishes of savers may not match the investment plans of firms.

13.12 If all sales in an economy were to be added up for a particular year, which of the following would be true with respect to the grand total? It would

A. exceed GNP.

B. equal GNP.

C. be less than GNP.

D. equal GNP plus sales tax.

13.13 Assuming inventories are constant, if replacement investment is less than gross investment in a given year, then

I. net investment is positive.

II. net investment is negative.

III. the productive capacity of the capital stock is increasing.

Which of the following is correct?

A. II only.

B. III only.

C. Both I and III.

D. Both II and III.

13.14 National income equals national output only when

I. planned saving equals planned investment.

II. the amount of consumer goods produced equals the amount of consumer goods bought.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

13.15 Which of the following describes a ‘closed economy’? Exports and imports are

A. equal.

B. fixed.

C. both zero.

D. dependent on government policy.

13.16 A survey carried out in Year t gathered information about the intentions of firms to invest and individuals to save. It was found that planned saving was greater than planned investment. However, in the national accounts for Year $t+1$ it was found that saving and investment were equal. Which of the following is the reason for this?

A. Businessmen must have lied about their intentions.

B. Households must have lied about their intentions.

C. Unintended stocks accumulated in Year $t+1$.

D. It was purely a coincidence.

13.17 National Income ≡ C + S

National Output ≡ C + I

Given the above identities, when the amount of consumer goods produced exceeds the amount of consumer goods purchased

I. unintended inventories of consumer goods will occur in firms.

II. planned savings will exceed planned investment.

III. actual investment will exceed actual savings.

Which of the following is correct?

A. II only.

B. I and II only.

C. II and III only.

D. I, II and III.

13.18 A change in the equilibrium level of national income in the short run will occur if there is an increase in

I. private sales of second-hand cars.

II. the rate of saving.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

13.19 Saving is a leakage from the flow of income in the economy because

A. it is typically a lower proportion of income than consumption spending.

B. it is never channelled into useful spending.

C. it is that part of income that is not currently spent.

D. it is usually greater than investment spending.

13.20 Which of the following is the best description of macroeconomic equilibrium?

A. Full employment.

B. Full capacity utilisation.

C. Any level of national income at which there are no forces acting to bring about a change.

D. That level of national income at which net investment is zero.

Module 14

A Simple Model of Income Determination

Contents

14.1 Introduction

In Module 13 we demonstrated the circular flow of income and expenditure. The purpose was to show how the economic actions of two principal players, households and firms, operating through markets, add up to form aggregate output of goods and services. No attempt was made to explain, however, what determined aggregate economic activity. In order to say why GNP was a certain amount in a given time period or how GNP might have been made larger or smaller, it is necessary to develop a theory about how GNP is determined.

A theory is simply a possible explanation of observable phenomena. Every theory begins with assumptions and proceeds through logical reasoning to conclusions. Consider, for example, the following simple theory: 'If the earth is spherical, then a person travelling straight ahead along the surface of the earth will eventually return to the starting point.' Is this conclusion valid? There are two ways in which a theory's conclusion can be tested. First, check to see whether the theory's assumptions are true and if its logical reasoning is correct. In the example, observe the earth from space to see if it looks spherical and try to prove mathematically if the conclusion is implied by the assumption. Second, test the conclusion directly. In the example, send people travelling straight ahead and see if they return to their starting point. Clearly, in this example, the theory's conclusions are largely correct – which is, of course, why Columbus did not fall off the edge of the earth.

The usefulness of theorising for finding explanations of natural phenomena is widely accepted. This is less so for economic theories, which are received with far greater caution than their physical, chemical or biological counterparts, largely because of the difficulty involved in verifying the validity of a theory's conclusions (let alone its assumptions) due to the limited amount of data available. Whereas natural scientists are usually able to test and retest their hypotheses many times over in a laboratory, economists have only the data provided by history, which are often insufficient to decide which of competing theories is correct. This is not to imply that physical scientists resolve all problems in the laboratory. Were they able to do so, test pilots in experimental aircraft would not command the 'high' salaries they do, for part of these salaries constitutes a risk premium, the risk arising partly because of an imperfect theory and untested assumption and partly because of exogenous (external) factors, some of which cannot be anticipated.

Despite the difficulties involved, and even though some ambiguities may never be eliminated, it is only through the persistent application of the scientific method that an increased understanding of economic behaviour can be achieved. In the context of macroeconomics, this means continued formulation and refinement of macroeconomic theories and rigorous testing of their validity with data generated in the real world – which is, after all, the economist's only laboratory.

14.2 The Development of Macroeconomic Models

Recall once again the circular flow of income and expenditure. Clearly, this involved numerous variables – various components of expenditure and income as well as their totals. Necessarily, a theory that explains the behaviour of these variables will be complicated. For this reason, a macroeconomic theory often takes the form of a *macroeconomic model*. At an advanced level, a model is usually written as a system of mathematical equations. It may also be expressed diagrammatically or verbally as, for the most part, will be done here.

Sometimes, in order to explain an important point, it is helpful to use models that are extremely simple compared with the real world. Such models are unrealistic in the sense that they do not explain everything that occurs in the real world. At the same time, however, they can be helpful in understanding some aspect of the real world because they contain an essential feature that also exists in the real world and have other complications stripped away. The strategy of this and the following modules is to start with the simplest macroeconomic model and, once that is understood, proceed to formulate increasingly realistic models by adding new, complicating features to the simpler ones. Even the simplest model offers lessons to be learned about the way the real world works; for that reason, it is important to be sure that the simpler model is well understood before going on to the more complicated ones.

Since no model, not even the more complex ones, can explain everything, it is necessary to regard some variables in the model to be determined externally. Such variables are considered to be *exogenous variables*. The variables whose behaviour the model purports to explain are called *endogenous variables*, and they can be thought of as being determined within the model. Both types of variable play a part in the model. Clearly, the endogenous variables are essential to the model since their behaviour is what the model is all about. But exogenous variables are also important in that, even though their behaviour is not explained by the model, they influence the endogenous variables and their effects must be taken into account. The distinction between these two types of variables can be illustrated by a couple of examples. In physics, there is a theory that says the volume and pressure of a given amount of gas will depend on the temperature of the gas. Volume and pressure are the endogenous variables and the temperature is the exogenous variable. The model is concerned with explaining the volume and pressure of the gas, and the temperature is treated as being determined by forces outside the model. In microeconomics, the position of the demand curve for ice cream may be thought of as being influenced by weather conditions. In this case the price and quantity of ice cream would be endogenous variables, while temperature and precipitation would be exogenous variables.

A model is a complete model when it can be used to predict what the values of the endogenous variables will be when the value of the exogenous variables are known. In the previous example of the model for the ice cream market, the model would be complete if it contained both supply and demand curves. If it included only the demand curve, then it would be impossible to predict the price and quantity of ice cream even if you knew the weather conditions, since every point on the demand curve would be a possible combination of price and quantity. Only when the supply curve is also known would the model provide a prediction of a unique price–quantity combination. In the following sections a complete, but simple, macroeconomic model will be developed and used to reach conclusions about the determination of actual GNP in the short run.

One of the most important aspects of constructing a model is specifying the endogenous and exogenous variables. We are not completely free to decide whether a variable should be endogenous or exogenous. A variable can properly be treated as exogenous only if that variable is not influenced by the endogenous variables in the model. In a model of the market for ice cream, it is correct to treat the weather as exogenous since the weather clearly is not influenced by the price or quantity of ice cream. It would not be appropriate, however, to specify the quantity of ice cream as exogenous, since both the quantity demanded and the quantity supplied are likely to be influenced by price.

The most crucial assumption to be made here in constructing a simple macroeconomic model is that the price level is exogenous. At first, this might seem rather strange. Is it reasonable to expect that the price level will not respond to the endogenous variables, particularly GNP, in the model? When it is clearly understood that the model being constructed is a model of GNP determination *in the short run*, then this assumption is reasonable. The purpose of the model is to explain how, *for a given period*, various factors affect actual GNP. How would actual GNP in that period be different if, for example, tax rates were higher or lower? The way in which many people believe the real world works is that in the short run the price level has been predetermined by past events. Actual GNP and any other endogenous variables that might be included in the model will in fact influence the price level, but not in the short run. They will affect how the price level changes from its present to its future level, but they will not affect its level in the present.

Our inflation equation is helpful in clarifying the distinction between the present price level and the rate of change in the price level, which must be understood if this assumption is to seem plausible:

$$\mathrm{P}_{t+1} = (1 + \mathrm{INF})\mathrm{P}_t$$

This equation says that next period's price level is equal to this period's price level times (one plus the rate of inflation). It is possible for some variable, such as this period's actual GNP, to affect the rate of inflation and thereby next period's price level, but to have no effect on *this* period's price level. It is in this sense that the model to be constructed here will treat the price level as exogenous: the current period's price level will be regarded as having already been determined by events that took place in the previous period. The inflation rate may be regarded as endogenous, but any change in the inflation rate will be reflected in next period's (but not this period's) price level.

The assumption that the price level is exogenous in the short run lies at the heart of what is called *the Keynesian Revolution*. John Maynard Keynes was a British economist who published in 1936 *The General Theory of Employment, Interest and Money*. This book revolutionised the way economists thought about macroeconomics. Prior to Keynes' *General Theory*, economists typically treated the price level as endogenous in the short run. Keynes showed that if the price level is assumed to be exogenous in the short run, the model behaves in a radically different manner and certain real-world events are more easily explained. The validity of this assumption is still debated by economists. Because it is both important and controversial, it will be discussed again in a later module; at this point, however, the assumption will be made and the process of constructing the simple macroeconomic model will proceed so that its implications can be deduced.

14.3 The Consumption Function

The assumptions outlined in Module 12 are still held to apply, namely that the economy is a private-sector closed economy, with no government sector and no international trade, and with only private households and private business firms. All savings are undertaken by households and all investment by business firms. Potential output is taken as fixed.

We have shown that the level of national income achieved will depend upon the level of aggregate demand that, given the stated assumptions, in turn depends upon consumption demand (C) and investment demand (I). We shall now consider the factors that determine the level of consumption demand and assume that the level of investment demand is fixed at all levels of income. This enables us to develop a simple model of the circular flow of income and the determination of national income. Subsequently, we shall investigate the factors that underlie investment demand, the other component of aggregate demand in this simple model.

Consumption expenditure may be influenced by many factors: the level of income; the distribution of income and wealth; tastes, habits and social arrangements; advertising and hire purchase regulations; and these are quite apart from other factors currently excluded by our assumptions, such as the level and structure of taxation and the provisions of the welfare state. From all these potential influences, economists usually single out one factor as having primary importance: income. Given our assumption, total income and disposable income are the same but, as soon as allowance is made for taxation and transfers, then the appropriate income concept is *disposable income* (Y_d), income minus taxes plus transfers, rather than total income. Most people find this plausible enough from introspection and from casual empiricism. Introspection regarding their own behaviour suggests that satisfaction is derived from consuming goods and services; the more goods consumed, the greater consumer satisfaction; the greater a household's disposable income, the more consumption the household will be willing and able to undertake. Casual empiricism lends support in that there appears to be – and this appearance is confirmed by the data available – a direct relationship between consumption and disposable income across households. Typically, high-income households have a higher level of consumption expenditure than low-income households, and, for society as a whole, as real national income has increased through time, the level of aggregate consumption expenditure has also increased.

Although economists have long recognised the link between consumption and income, the concept was not given formal expression until the publication of J.M. Keynes's book in 1936. His important departure from received wisdom was a suggestion that the relationship between consumption and income might be stable and systematic. This being so, the form of the relationship could be established by empirical evidence. Once established, the relationship between consumption and income, called the *consumption function*, could be used to predict how consumption would behave as income changed, thus providing a basis for estimating and forecasting the behaviour of the single most important element of aggregate demand.

Keynes provided few clues as to the precise relationship between consumption and income, confining himself to the following general statement:

> The fundamental psychological law, upon which we are entitled to depend with great confidence, both *a priori* from our knowledge of human behaviour and from the detailed facts of experience, is that men are disposed, as a rule and on the average, to increase their consumption as income increases, but not by as much as the increase in their income.

This 'fundamental psychological law' is extremely important, because (as we shall see) if the *marginal propensity to consume* (that is, the change in consumption accompanying a change in income) were unity or greater than unity, then the economic system would prove to be highly unstable. In other words, the fundamental psychological law is an important stability condition for the economic model we are developing. For our purposes here, however, it tells us very little about the precise relationship between consumption and income. Should income increase by $1 million, the fundamental psychological law simply states that the increase in consumption will lie somewhere between $1 and $999 999. Any value within this range would satisfy the condition that individuals would spend part, but not all, of an increase in income.

Empirical studies following on from Keynes's *General Theory* have provided a mass of statistical material, and while much of it is subject to differing interpretations, there is fairly general agreement on the following conclusions:

(a) At any point in time in a given community, households with relatively low incomes typically consume a high proportion of household income, or even dissave, while households with relatively high incomes typically save a higher proportion of household income.
(b) Over long periods of time, over which living standards have increased appreciably for society as a whole, the relationship between consumption and income appears roughly proportional. Even in economies where real national per capita income has grown substantially over a number of decades, the relationship between consumption and income has changed very little. In consequence, the proportion of disposable income saved by the community has remained fairly stable over long periods of time.
(c) While the long-run relationship between consumption and income has been stable historically, the relationship between changes in income and changes in consumption is much less predictable. The evidence suggests that there is a time lag before consumption responds to changes in incomes, this time lag resulting from habit, custom, existing institutional arrangements and the desire to ensure that any change in income is permanent rather than temporary. For these reasons, changes in consumption tend to lag behind changes in income, and after some period of time, consumption adjusts to new income levels. In other words, the short-run relationship between consumption and income is different from the long-run relationship.

These short-run and long-run relationships between consumption and income are presented diagrammatically in Figure 14.1 and Figure 14.2. Figure 14.1 depicts a series of successive short-run relationships between consumption and income, while Figure 14.2 depicts the long-run relationship. This type of figure is widely used in macroeconomic analysis and it is important to understand its construction and implications.

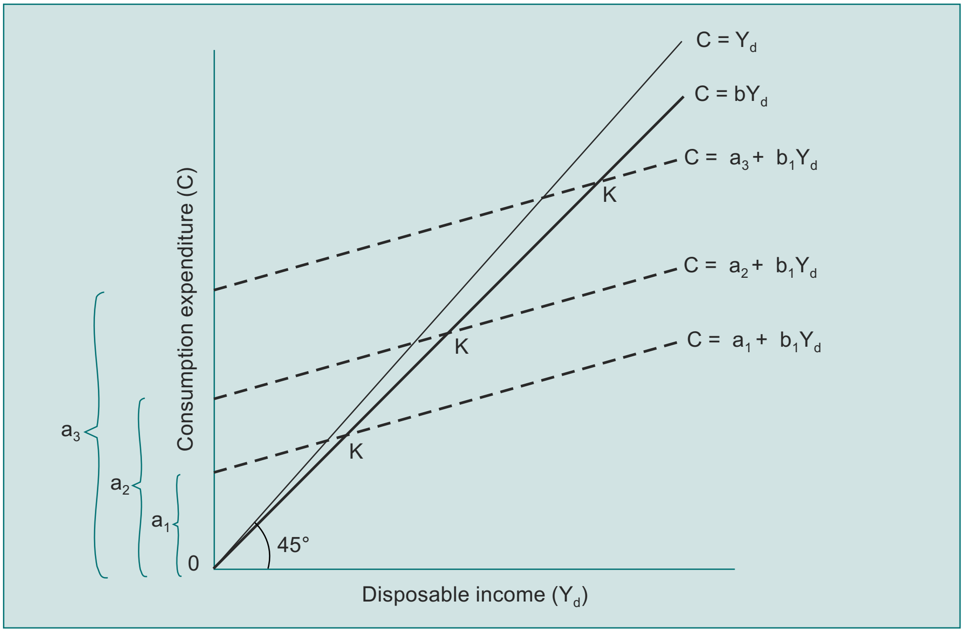


Figure 14.1 Short-run consumption function

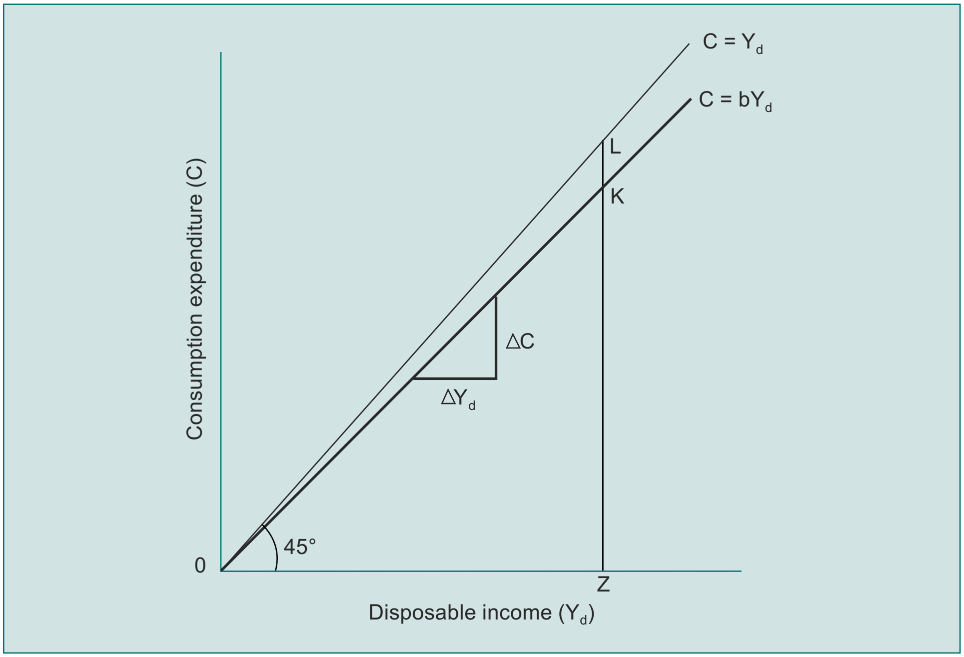


Figure 14.2 Long-run consumption function

The following points in construction of the consumption functions in both diagrams should be noted. Consumption is plotted against disposable income (Y_d), reflecting the fact that consumption is influenced by income net of taxes and transfers. In the model we are developing here, total income (Y) and disposable income can be treated as identical, for there is no government sector. More realistically, when a full model is developed, we have to recognise that 'death and taxes are always with us', and treat consumption as a function of disposable income, which differs from total income ($Y > Y_d$).

The 45° construction line reflects the locus of points where consumption equals income. If the consumption function lies along the 45° line, then all income would be consumed and there would be neither saving nor dissaving. It follows from this that if, at some levels of income, the consumption function lies above the 45° line, then, at those levels of income, consumption would exceed income and there would be dissaving. Similarly, for those levels of income where the consumption function lies below the 45° line, consumption would be less than income and there would be saving. The amount of dissaving or saving is given by the vertical distance between the consumption function and the 45° line. For example, in Figure 14.2 where Y_d equals 0Z, the amount of saving is given by KL (this representing the vertical distance between the height of the consumption function and the 45° line), and the amount of consumption is given by KZ.

Figure 14.1 shows a series of short-run consumption functions that are relatively flat. This implies that as income changes, consumption changes will, at least initially, tend to lag behind income changes. The short-run changes in consumption accompanying a change in income, the marginal propensity to consume, will be relatively small and the proportion of income consumed will decline in the short run. After some time period, however, the relationship between consumption and income will shift upwards, reflecting the adjustment of consumption to new and higher levels of income. Hence the series of short-run consumption functions as shown.

Given a series of short-run consumption functions, such as depicted in Figure 14.1, the long-run relationship between consumption and income may emerge as a series of points marked by K in Figure 14.1. Taken together, the series of points marked by K produce a long-run relationship, such as that shown in Figure 14.2. Notice that, while the long-run relationship between consumption and income might be stable and predictable, the flat shape of the short-run consumption function, and the fact that the timing of upward shifts in the short-run function may be difficult to predict, may produce a relatively unstable relationship between consumption and income in the short run.

The linear consumption function demonstrated in Figure 14.1 can be represented algebraically as follows:

$$C = a + bY_d$$

where C is consumption; a is the amount of consumption that is *independent* of income (i.e. the amount of consumption that occurs when Y_d equals 0, given by the intercept of the consumption function with the vertical axis); b is how consumption changes as income changes, known as the marginal propensity to consume and given by the slope of the consumption function, i.e. $b = \Delta C/\Delta Y$; and Y_d is the disposable real income, i.e. income less taxes plus transfers. In this case, disposable income equals gross income because of the assumption of no government sector, and so no taxes and transfers.

It should be emphasised that such a formula, while applicable to the short-run relationship between consumption and income, does not appear to be consistent with any observed long-run relationship. The long-run linear function shown in Figure 14.2 can be represented by the equation:

$$C = bY$$

There is no intercept a, and consequently at all levels of income the same proportion of income is consumed. Empirical evidence suggests that the long-run relationship between consumption and income is roughly a linear one.

It is important to distinguish between the proportion of income consumed at any level of income (the average propensity to consume, APC) and the change in consumption that accompanies any change in income (the marginal propensity to consume, MPC).

Average Propensity to Consume (APC) is the proportion of income consumed and is calculated by dividing consumption expenditure by the level of disposable income (Y_d). Thus:

$$APC = \frac{C}{Y_d}$$

In Figure 14.2 when Y_d = 0Z then C = KZ. Therefore, at disposable income level 0Z, the average propensity to consume is given by

$$APC = \frac{KZ}{0Z}$$

Marginal Propensity to Consume (MPC) is the change in consumption (ΔC) that accompanies a change in income (ΔY_d). Thus:

$$MPC = \frac{\Delta C}{\Delta Y_d}$$

The MPC is given by the slope of the consumption function and is represented in the formulas $C = a + bY_d$ (short run) and $C = bY_d$ (long run).

Given a consumption function such as shown in Figure 14.2, APC = MPC and this is consistent with the long-run evidence. Given a consumption function, such as shown in Figure 14.1, MPC may be substantially less than APC in the short run, because of the lag before consumption adjusts to changes in income.

The MPC is of particular importance in determining the manner in which the economy may react to an initial change in a component of aggregate demand. If the MPC is high, then a large part of any initial increase in demand is 'passed on' in the circular flow of income in the form of higher consumption. This is known as 'the multiplier process' and it played a crucial role in the Keynesian revolution for policy making.

It is extremely important that you should be able to distinguish between movements along a given consumption function curve and whole shifts of the curve itself. The notion of the consumption function is a deliberate simplification of the real world. It does not deny that variables other than income influence consumption, but it attempts to see how consumption will change with changes in income, *ceteris paribus*. Of course, other factors clearly do influence consumption – the distribution of income and wealth, tastes, habits, social conditions, advertising, liquidity, hire purchase regulations, etc. – and, should these parameters change, the consumption function is likely to shift. For example, if, through taxes and transfers, the distribution of disposable income shifts in favour of poorer households and against high-income households, then the proportion of aggregate national income that is consumed is likely to rise as, typically, poor households consume a higher proportion of income than rich households. This will be shown diagrammatically by an upward shift in the aggregate consumption function for the economy.

A schedule showing the relationship between consumption and income is drawn up on the assumption that changes in these other factors influencing consumption are relatively rare and/or occur only slowly through time, so that the most important factor influencing consumption is income. Clearly, the validity of this assumption depends upon prevailing circumstances. For many periods the assumption appears a reasonable approximation to reality, but from time to time changes in these other factors do occur and cause shifts in the relationship between consumption and income, shown diagrammatically by a bodily shift in the consumption function.

The ability to distinguish between a movement along a given schedule and a shift in the schedule is an important part of the apparatus of economic analysis. A movement along the consumption function reflects changes in consumption that are associated with changes in disposable income. A shift in the position of the consumption function itself occurs because of changes in the other factors, leading to changes in the relationship between consumption and income. From the microeconomics modules, the appropriate analogy is a movement along a given demand curve caused by changes in price, and shifts in that curve that are the consequence of changes in:

(a) incomes,
(b) the price of complementary or substitute goods,
(c) changes in tastes.

A *shift* in the demand curve changes the quantity demanded for any given price. A *movement along* the demand curve shows the changes in the amount demanded, which are caused by changes in price. Similarly, a *shift* in the consumption function curve indicates that the amount consumed is different at each level of income, the direction of the change depending upon whether the consumption function schedule shifts upwards or downwards. A *movement along* a given consumption function curve indicates the changes in consumption that are a consequence of changes in income.

14.4 The Solution to the Simple Model

Given the consumption function relationship depicted in Figure 14.2 above, a simple graphical illustration of national income determination can now be presented. The consumption function relates consumption expenditure to income. If investment is taken as fixed at all levels of income, then the aggregate demand function can be determined since aggregate demand in the model consists only of consumption and investment expenditure.

The model is based on three equations:

$Y = C + I$	where national income (Y) is determined by the sum of consumption expenditures and investment expenditures and C + I = aggregate demand or aggregate expenditure (E).
$C = bY_d$	where consumption is a function of income. The relationship is proportional at all levels of income, so that APC = MPC = b, as in Figure 14.2; Y is equivalent to Y_d because there is no government sector.
$I = I_0$	where the level of investment is fixed exogenously or autonomously, i.e. its value is determined outside the economic system.

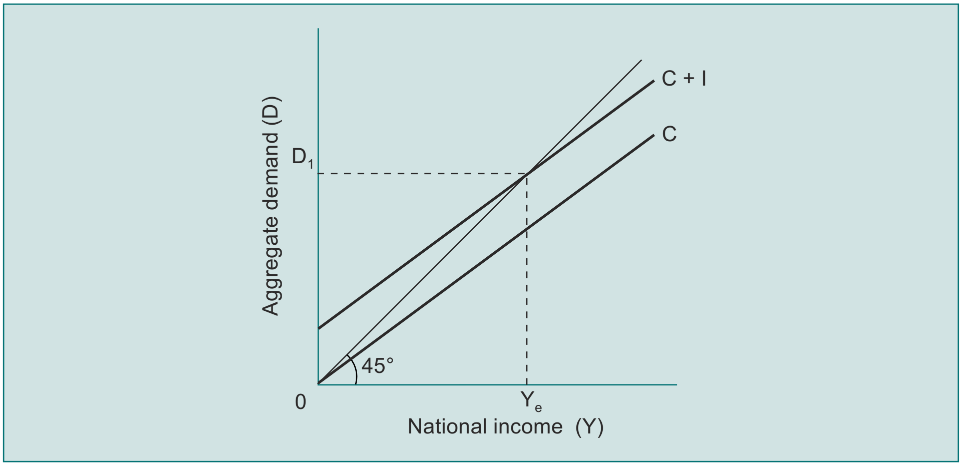


Figure 14.3 Simple model of income determination

This model is demonstrated in Figure 14.3. Variable C is the consumption function. The line C + I represents aggregate demand or expenditure and runs parallel to the consumption function, indicating that at all levels of income investment expenditures are fixed. The aggregate expenditure function is composed of consumption and investment expenditures. Where the aggregate demand function lies above the 45° line, aggregate demand is greater than national income Y; where the aggregate demand lies below the 45° construction line, aggregate demand is less than national income. At the point of intersection of the aggregate demand function and the 45° line, aggregate demand = national income. This gives the equilibrium level of national income, Y_e.

In terms of planned saving and planned investment, why does the point of intersection determine national income? At the point of intersection at income Y_e, planned saving is given by the vertical distance between the consumption function and the 45° line, and is equal to planned investment. The plans of savers and investors are consistent and there will be no tendency for the level of national income to change.

14.5 The Multiplier

Suppose our economy is in equilibrium and a series of new inventions promises to make investment in capital goods more profitable to undertake. This increase in profitability on investment means that, for any national income Y, firms will want to make greater investment expenditure than before. As a result, the investment function will shift upwards, and this is shown in Figure 14.4.

The initial investment function I, being a horizontal line, indicates the exogenous nature of investment, so that we are assuming businesses invest an amount I independent of the level of national income Y. The increase in investment expenditure ΔI gives a new investment function I + ΔI. The shift of function will cause aggregate demand to increase; but by how much? By ΔI? No, the total increase in aggregate demand will be a multiple of ΔI; hence the name 'multiplier'.

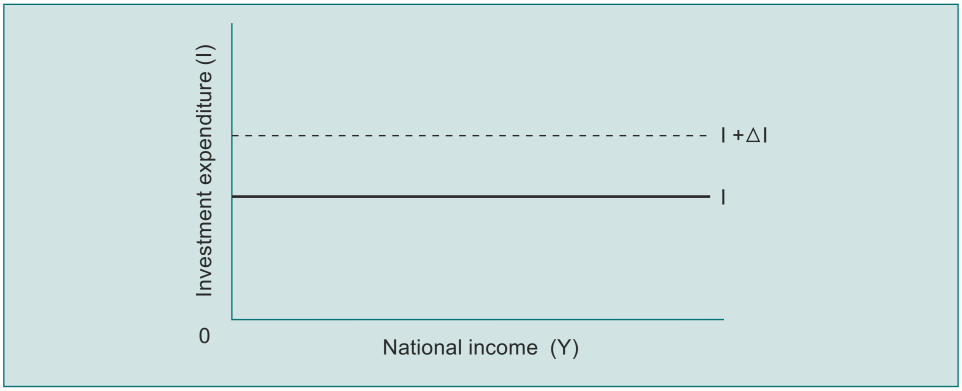


Figure 14.4 Shift in the investment function

We can superimpose the increase in investment expenditure, ΔI, on the aggregate demand curve C + I from Figure 14.3, and this is seen in Figure 14.5.

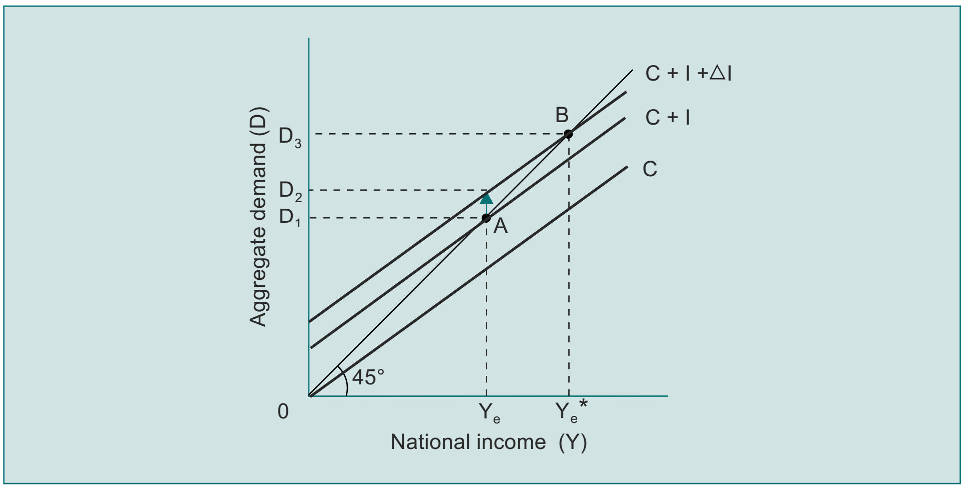


Figure 14.5 Autonomous and induced changes in aggregate demand

The consumption function remains unchanged as C but the autonomous increase in investment expenditure I shifts the aggregate demand curve C + I upward from C + I to C + I + ΔI. The amount of the shift, as shown in the vertical axis, is D_1D_2. The effect of this upward shift is to move its point of intersection with the 45° line from A to B. This indicates that, as a result of the shift of the aggregate demand function from C + I to C + I + ΔI, aggregate demand increases from D_1 to D_3. Thus, the autonomous increase in aggregate demand of D_1D_2 leads to a total increase from aggregate demand of D_1D_3. It is clear that the total increase is greater than the autonomous increase. Simultaneously, the increase in national income from Y_e to Y_e^* is greater than the autonomous increase in aggregate demand.

How can this be? For the whole process to occur, there must be unemployment in the economy; given our assumptions of a fixed capital stock and given the state of technology, any increase in national income must be associated with an increase in employment and consequently a decrease in the rate of unemployment. Thus, the autonomous increase in aggregate demand due to the increase in investment expenditure, I, leads initially to an equal increase in national output/national income. But that increase in national income induces a further increase in aggregate demand. This induced increase in aggregate demand will lead to an increase in national income which in turn will lead to a further increase in aggregate demand. This is the multiplier process at work, as shown in Figure 14.6.

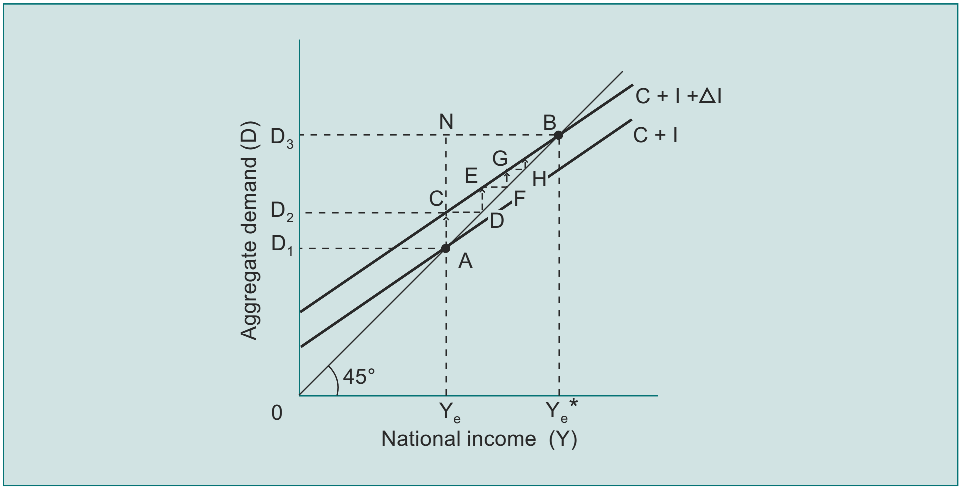


Figure 14.6 The multiplier process at work

The autonomous increase in aggregate demand (I) represented by AC leads to an equal increase in national income of CD. This, in turn, leads to an induced increase in aggregate demand of DE, which leads to a further equal increase in national income of EF, a further increase in aggregate demand of FG and a further induced increase in national income of GH. This process continues until point B is reached. The total increase in aggregate demand is therefore AN, which equals the autonomous increase of AC plus the induced increases DE, FG and so on. Similarly, the increase in national income from Y_e to Y_e^* will be the sum of CD, EF, GH and so on.

In summary, the autonomous increase in aggregate demand has a multiplier effect on national output and income. This multiplier effect shows how many times greater the total increase in national output will be than the autonomous increase in aggregate demand. For example, if an autonomous increase in aggregate demand of $10 million causes an increase in total national output of $40 million, then the value of the multiplier would be 4.

Why doesn't the multiplier process go on indefinitely? Why, as is illustrated in Figure 14.6, is each successive round of increase in national output smaller than the previous round? The reason is that an increase in national output will not lead to an equal increase in consumption expenditure; rather it will induce an increase in aggregate demand that is less than the increase in output – in other words, not all of the increase in national income will be spent. Some of it will be saved, representing a withdrawal from the expenditure stream. How much of the additional increase in income will be spent? To answer this question, we must return to the marginal propensity to consume (MPC). Remember what the MPC was? It was the slope of the consumption function measured by $\Delta C/\Delta Y$, i.e. the increase in consumption attributable to the increase in disposable income. As we have seen from our consumption functions, households are likely to spend part of an increase in disposable income and save part of it. When part is saved, the value of MPC must be less than 1. Thus,

$\text{MPC} = \frac{\Delta C}{\Delta Y} < 1$

In Table 14.1 we trace out the multiplier effects from an autonomous increase in investment expenditure I. We shall assume I = 10, MPC = 0.75 and sufficient unemployed resources exist to allow the multiplier process to work itself out.

Table 14.1 Multiplier effects of an autonomous increase in investment expenditure

Round	ΔC	ΔI	ΔAggregate demand	ΔY
1	0	10	10	10
2	7.5	0	7.5	7.5
3	5.625	0	5.625	5.625
4	4.219	0	4.219	4.219
5	3.164	0	3.164	3.164
6	2.373	0	2.373	2.373
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Total change	30	10	40	40

In the first round, ΔI = 10; this is the autonomous increase in investment expenditure that starts the multiplier process. Because there is no autonomous increase in consumption expenditure in the first round, ΔC = 0. The increase in aggregate demand in the first round is therefore 10, as is the increase in national output and national income ΔY. In the second round, because the value of MPC is 0.75, the increase in consumption, stemming from ΔY of 10 of round 1, is 7.5 (10 × 0.75). Because there is no induced increase in investment expenditure, ΔI = 0. Thus, the increase in aggregate demand in round 2 is 7.5, which causes an equal increase in Y, so that ΔY in round 2 is 7.5. In the third round, ΔC is 5.625 (7.5 × 0.75), ΔI is 0, Δaggregate demand is 5.625, and ΔY is likewise 5.625. This process continues round after round in a convergent manner.

As can be seen from Table 14.1, the sum of all the autonomous plus induced changes in aggregate demand is 40. As a result, the total increase in national output is also 40. Because the autonomous increase in aggregate demand is 10 and the total increase in national output is 40, the value of the multiplier in this example is 4.

It is no accident that the value of the multiplier in the example is 4. It results directly from the fact that the value of the MPC is 0.75. Suppose the value of the MPC had been 0.5. In that case, in the second round the increase in aggregate demand and national output would only have been 5.0 instead of 7.5. In successive rounds, Δaggregate demand and ΔY would have been 2.5, 1.25, 0.625, 0.3125 etc. and the sum of the autonomous plus induced increases in aggregate demand would have totalled 20, as would have the increase in national output and income. Thus, the larger the value of MPC, the larger the value of the multiplier. An MPC of 0.75 produces a multiplier of 4; an MPC of 0.5 produces a multiplier of 2. The formula for calculating the value of the multiplier is:

$\text{Multiplier} = \frac{1}{1 - \text{MPC}}$

Now, we have already calculated from our example that the autonomous increase in aggregate demand times the multiplier equals the increase in national output, i.e.

$$\text{autonomous } \Delta\text{aggregate demand} \times \text{multiplier} = \Delta Y$$

We can now restate this equation to read:

$$\Delta Y = \text{autonomous } \Delta\text{aggregate demand} \times \frac{1}{1-\mathrm{MPC}}$$

In the example in Table 14.1, the autonomous change in aggregate demand was an autonomous increase in I. Thus we have

$$\Delta Y = \frac{\Delta I}{1-\mathrm{MPC}}$$

giving us the *investment multiplier*.

As we shall see, when we expand the simple model to include other sectors, such as the government and international sectors, the multiplier process is not confined to an autonomous change in investment expenditure. We shall also see that if we relax our assumption about induced effects applying only to households (e.g. induced investment expenditure), the multiplier formula becomes much more complex.

Before we move on, however, two additional points are worth mentioning. For the multiplier process to work its way to completion, there must be sufficient unemployed resources. If the economy were operating at full employment, any autonomous increase in aggregate demand would result in inflation, instead of an increase in real output. If the economy were close to full employment, an autonomous increase in aggregate demand could result in an increase in national output *and* inflation, as shown in Figure 14.7.

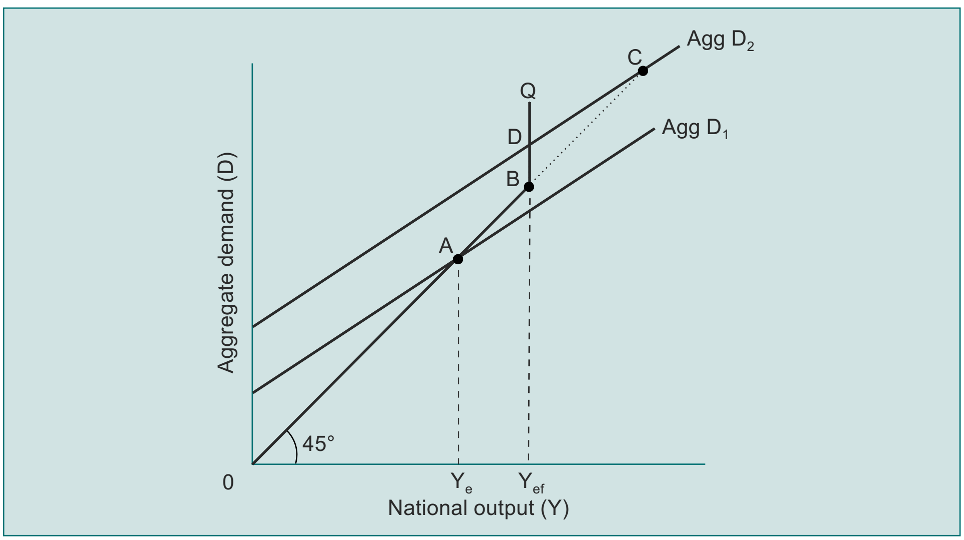


Figure 14.7 Absence of a full multiplier effect

Initially, the economy is in equilibrium; aggregate demand is D_1 and the equilibrium level of national output is Y_e. Potential output is Q. Because of an autonomous increase in investment expenditure, aggregate demand shifts to D_2. For the full multiplier effect to be realised, the economy must reach point C – but this is impossible because actual output cannot exceed potential output Q. Thus, the economy reaches full-employment national output Y_{ef} and the shift in the aggregate demand curve also creates an inflationary gap of DB.

The second point to note concerns the MPC. If the value of MPC were zero, the multiplier would be 1. If the value of MPC were 1, the value of the multiplier would be infinity and the economy would be unstable, swinging from zero output to full-employment output. The economy does not behave in this fashion; the multiplier in the real world is not 1, but calculating the multiplier in a modern economy is not simple, as policy makers have discovered. As a consequence, trying to guide the economy to full employment is no easy task. For example, suppose that, in order to stimulate the economy in the light of a high unemployment ratio, the government cuts taxes, thereby increasing disposable income. Further suppose that all households spend the increase in disposable income on imported goods and services. Such an action will cause multiplier effects in foreign nations and will not contribute to domestic employment. In addition, we shall see that if increased aggregate demand causes the demand for money to increase and interest rates to rise, then the resulting decrease in investment and consumption expenditure can have negative multiplier effects. Such exciting complications will be the focus of the later modules as we move from our oversimplified models of the economy to greater reality.

Learning Summary

You are strongly advised to accept the reasoning, however apparently divorced from reality, of economists. Such reasoning attempts to isolate the ‘important’ variables in the macroeconomic world and estimate how changes in certain factors affect incomes, output and employment.

You now understand the similarities and differences between economic models and hard science models. These differences in the real world are small. While the physical sciences have laboratories, economics has as its laboratory the real world. The constantly changing nature of the world poses problems for the economics researcher. Exogenous shocks and the imperfect specification of models provide challenges for physical scientists and economists alike.

You have a firm grasp of the short-run and long-run consumption functions and understand how the multiplier is derived from the slope of the consumption function. You are able to solve the simple model and understand the meaning of equilibrium. You understand how autonomous changes in aggregate demand lead to induced changes in demand, and how these changes affect output and employment. You are also aware of the factors that limit the full multiplier effect.

Review Questions

Multiple Choice Questions

Questions 14.1 and 14.2 are based on the table below, which shows levels of disposable income and corresponding levels of consumption expenditure.

Disposable income ($)	Consumption expenditure ($)
200	205
225	225
250	245
275	265
300	285

14.1 Which of the following is correct? The marginal propensity to consume out of disposable income

A. is sometimes positive and sometimes negative.

B. varies between 0.20 and 1.25.

C. cannot be calculated from the data provided.

D. is 0.8.

14.2 Which of the following is correct? The average propensity to consume

A. is constant.

B. increases as disposable income increases.

C. decreases as disposable income increases.

D. has a value less than 1.

14.3 Which of the following is correct? A straight-line consumption function with a positive intercept (cuts the vertical axis above zero) and a slope of less than 1 (slopes upward to the right at less than 45° angle) would be one for which

A. the average propensity to consume increases with the level of income.

B. the average propensity to consume is constant for all levels of income.

C. the marginal propensity to save is positive at all levels of income.

D. the average propensity to save is positive at all levels of income.

14.4 For a certain economy, the consumption function is $C = 15 + 0.75\ Y_d$ (where Y_d = disposable income). In addition,

I. saving = 10 when disposable income = 100.

II. consumption = 240 when disposable income = 300.

III. a level of disposable income exists for which consumption = saving.

Which of the following is correct?

A. I only.

B. I and II only.

C. II and III only.

D. III only.

14.5 Which of the following explains why an increase in investment can have a 'multiplier' effect on national income in an economy which is operating well below full-employment output? Investment spending

A. raises potential GNP.
B. yields income to resource owners, who in turn spend part of that income on consumption and hence income increases again.
C. increases the marginal propensity to consume, which increases income flows.
D. enters the flow of economic activity more quickly than any other type of spending.

14.6 If the marginal propensity to consume were zero, which of the following could result from a decline in investment demand?

A. A reduction of equal value in the equilibrium level of national income.
B. No change in the equilibrium level of national income.
C. An unending downward spiral of national income.
D. A compensating increase in the level of consumption expenditure.

14.7 Which of the following is correct? The consumption function shows

A. that consumption depends primarily on the level of business investment.
B. that households consume more when interest rates are low.
C. that the marginal propensity to consume increases with national income.
D. the amounts households plan to consume at various possible levels of income.

14.8 In constructing economic models

I. the value of an exogenous variable is determined by the endogenous variables.
II. exogenous variables affect the values of the endogenous variables.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

Questions 14.9 to 14.14 are based on the following model:

$$C = 50 + 0.5Y$$
$$I = 100$$
$$Y \equiv C + I$$

14.9 Which of the following is correct?

A. Only I is an exogenous variable.
B. Only I is an endogenous variable.
C. Only C and I are exogenous variables.
D. Only Y is an endogenous variable.

14.10 What is the equilibrium value of Y?

A. 50.
B. 100.
C. 250.
D. 300.

14.11 Which of the following is correct? Independent of the level of Y, the average propensity to consume is

A. 0.5Y
B. 200
C. $\frac{50}{Y} + 0.5$
D. $\frac{150+0.5Y}{Y}$

14.12 What would be the equilibrium value of C if I were to increase from 100 to 200?

A. 200.
B. 300.
C. 400.
D. 500.

14.13 What is the change in the value of the investment multiplier when I changes from 100 to 200?

A. It remains constant at 2.
B. It increases from 2 to 3.
C. It doubles from 2 to 4.
D. It remains constant at 0.5.

14.14 If I became zero, what would the equilibrium value of C be?

A. 0.
B. 0.5Y.
C. 50.
D. 100.

14.15 The consumption function in Figure 14.8 relates household consumption expenditure (C) to disposable income (Y_d). It shows

I. a constant marginal propensity to consume (MPC).
II. a constant average propensity to consume (APC).
III. dissavings for Y_d < \$4000.

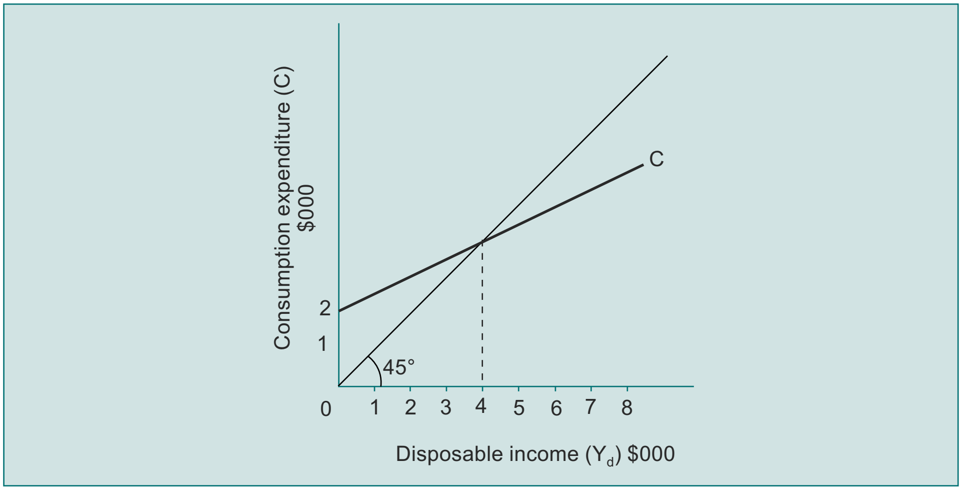


Figure 14.8 Consumption function

Which of the following is correct?

A. I only.
B. I and III only.
C. II only.
D. I, II and III.

Module 15

Expanded Model of Income Determination

Contents

15.1 Introduction

In this module we shall expand our simple model of income/output determination to include the major features of a modern market economy. We shall begin by relaxing the assumption that all investment expenditure is exogenous; we shall add a government sector and an international sector; and we shall have firms retain part of their incomes as undistributed profits.

While this expanded model is much more realistic, it still retains simplifying assumptions. For example, we shall assume the government is our single decision maker rather than the myriad of decision makers in both local and national governments that pervade the real world. We shall retain the assumptions of a given potential output (Q), a fixed relationship between output and employment, and all investment being carried out by private firms.

15.2 The Investment Function

Figure 15.1 shows investment expenditure (I) as a proportion of GNP for the world's leading capitalist countries for 2011. As can be seen from the figure, Japan and India are 'high' investment economies and the UK and the USA are 'low' investment economies. While the levels of investment in China, Brazil and Russia appear to be low, there is a significant amount of government investment in these economies that is not included as part of I in our model.

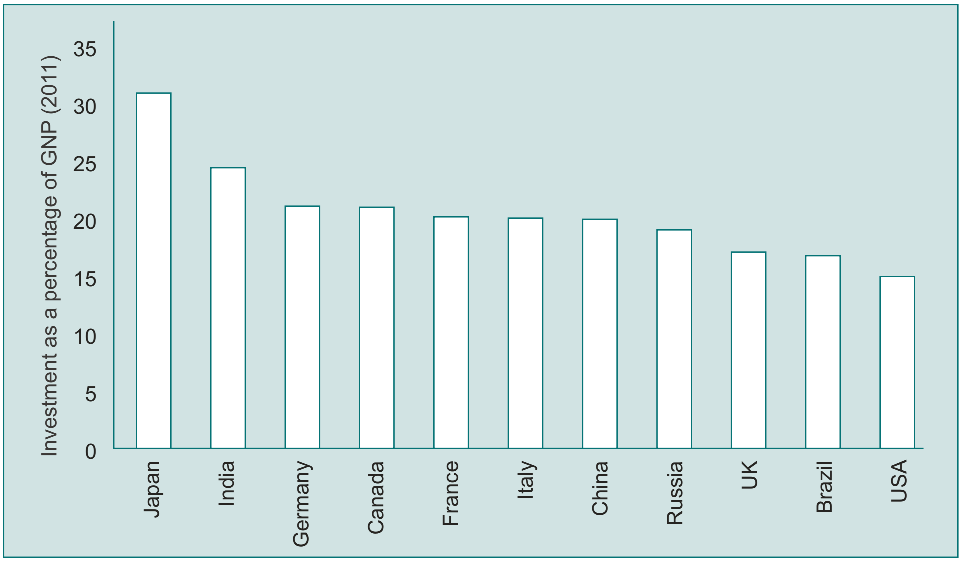


Figure 15.1 Private investment expenditure as a percentage of GNP for G7 and BRIC economies, 2011

In the model developed in Module 14, there are only two components of aggregate demand: private consumption and private investment. The consumption function suggests that the relationship between consumption and income is reasonably stable and predictable. Although there are difficulties in accurately forecasting consumption, private consumption expenditure is much less volatile than private investment expenditure through time. For this reason the analysis seen in Module 14, while helping to show how consumption and investment expenditures interact to determine income in a simple model, is unrealistic and misleading. To take investment as fixed at all levels of income denies one of the most important features of investment behaviour, namely its pronounced *instability over time*.

It is this instability of investment expenditure that explains its central role in the theory of income determination and in business cycle theory. Investment expenditures are *quantitatively* less important than consumption expenditures, which average about 60 per cent of GNP for the 11 nations noted in Figure 15.1; but because investment expenditures fluctuate more violently over time, they are important in explaining the short-run behaviour of income and output.

Any explanation of investment expenditures must begin from this pronounced instability, an instability that might be explained by reference to 'subjective' and 'objective' facts. The 'subjective' fact is that investment decisions are grounded in *expectations about the future*, and all expectations are subject to uncertainty. The 'objective' fact is that the technology of modern industry, which requires for efficient production relatively large amounts of capital, is likely to produce greater instability in investment expenditure than in consumption expenditure. From this it can be shown that changes in income are likely to produce proportionately larger changes in investment expenditure.

Investment is undertaken in the hope and expectation that it will be profitable. To assess whether any proposed investment scheme will be profitable, a business has to arrive at an explicit or intuitive judgement on three factors:

(a) the cost of the investment;
(b) the expected returns from the investment in the form of increased income;
(c) the cost of financing the investment.

The first two factors give us the rate of return over cost, or what is known as the *marginal efficiency of investment* (*MEI*). The third factor, the cost of financing the investment, is the rate of interest (R).

For some investments, such as large construction projects and industrial projects with long gestation periods (e.g. aircraft development), the cost of the investment can only be estimated at the time of the initial decision. For other investments, for example the purchase of a machine tool, the cost can be determined accurately. In all investments, the second factor involves uncertainty, because investments are durable and usually yield income over a period of time stretching some years into the future. Thus a firm has to *estimate* the expected returns from all investment projects.

Let us take an example of how a firm might relate the expected *return* from an investment to the *cost* of that investment. Suppose that the investment considered is the purchase of automatic welding equipment; further suppose that the cost of this equipment is $4 million and that the equipment will have a productive life of five years, during which period it will give a return of $1 million in each year, i.e. a total return over five years of $5 million. Is this a sound investment? In attempting to answer this question, the first problem that arises is that the expected returns accrue over a long period of time. The expected returns can only be summed directly if $1 million received in the first year has the same value as the $1 million received in any subsequent year. Intuitively, most individuals will recognise that a sum of money received today is more valuable than the same sum of money received tomorrow, and that the sum of money received tomorrow is more valuable than the same sum of money received two days from now. Is there a rational justification for this intuitive view?

To answer this question, let us begin with a formula with which you are undoubtedly familiar – the *compound interest formula*. If you lent a *principal amount* (P_0) to the value of $5 million at a rate of interest of 5 per cent, then the principal received back at the end of one year (P_1) is given by the formula:

$P_1 = P_0(1 + R)$ where R is the rate of interest.

Solving this equation with the values from our example, we have:

$P_1 = \$5 \text{ million } (1 + 0.05) = \$5.25 \text{ million}$

If this sum of money were lent for two years, then the principal received back at the end of the second year would be:

$P_2 = P_0(1 + R)(1 + R) = P_0(1 + R)^2$

By extrapolation, if this principal is lent for *n* years, then the principal after *n* years (P_n) is:

$P_n = P_0(1 + R)^n$

In our example, where *n* is 5 years, then:

$P_5 = P_0(1 + R)^5 = \$5 \text{ million}(1 + 0.05)^5 = \$6.381 \text{ million}$

In the case of our firm considering its welding equipment investment, which is estimated to cost $4 million and is expected to yield an income of $1 million a year over five years, the above process has to be reversed. The firm expects to receive $5 million in total over the five-year period, but would have preferred to have received the $5 million in a lump sum in the first year because the future dollars are not as valuable as the dollars received in the first year. The $1 million expected in the second year is worth less than the $1 million expected in the first year, the $1 million received in the third year is worth less than the $1 million received in the second year, and so on. The firm has to reduce this future stream of income to a common denominator, which reflects the fact that a given sum of money becomes less valuable the greater the number of years it is removed from the present. It is necessary to calculate what is called the *present value* of this stream of future dollars. This is done by *discounting* each of the future dollars, the amount of *discount* increasing the further the receipt of the income is removed from the present.

The discounting process is simply the reverse of the compound interest procedure. Consider the question, What is the present value of $1 million received one year from today? The compound interest formula is:

$$P_1 = P_0(1 + R)$$

Therefore, to get P_0, which is the present value of the principal of $1 million received one year from today, the above equation is simply transposed to solve for P_0, and to illustrate that we are *discounting* a future stream of income we substitute r for R. Thus:

$$P_0 = \frac{P_1}{1+r}$$

If the rate of discount, r, is 5 per cent, then:

$$P_0 = \frac{\$1\text{ million}}{(1+0.05)} = \$952\,000$$

Similarly, if we want to know the present value of $1 million received two years from today:

$$P_0 = \frac{P_2}{(1+r)^2} = \frac{\$1\text{ million}}{(1+0.05)^2} = \$907\,000$$

Extending this, the present value of a stream of future income of $1 million per year over five years is:

$$\begin{aligned} P_0 &= \frac{\$1\text{ million}}{(1+r)} + \frac{\$1\text{ million}}{(1+r)^2} + \frac{\$1\text{ million}}{(1+r)^3} + \frac{\$1\text{ million}}{(1+r)^4} + \frac{\$1\text{ million}}{(1+r)^5} \\ &= \$4.329\text{ million} \end{aligned}$$

As far as the investment decision is concerned, P_0, the present value, is the 'value' or purchase price of the investment, which we can represent by PV. The stream of future income can be represented by $FI_1, FI_2, \ldots, FI_n$ and r represents the rate of discount. Hence

$$PV = \frac{FI_1}{(1+r)} + \frac{FI_2}{(1+r)^2} + \cdots + \frac{FI_n}{(1+r)^n}$$

What this formula tells us is that a firm, in considering an investment decision, will consider the cost (or value) of that investment and the expected future stream of income; and, given that the expected income accrues over the future, the firm will discount that future stream of income to reflect the fact that future dollars are less valuable than present dollars. The above formula is the forerunner of *discounted cash flow*, a concept widely used in the accountancy profession.

The curve showing the relationship between the marginal efficiency of investment, the rate of return over cost, and the volume of investment, is drawn sloping downwards from left to right, as shown in Figure 15.2.

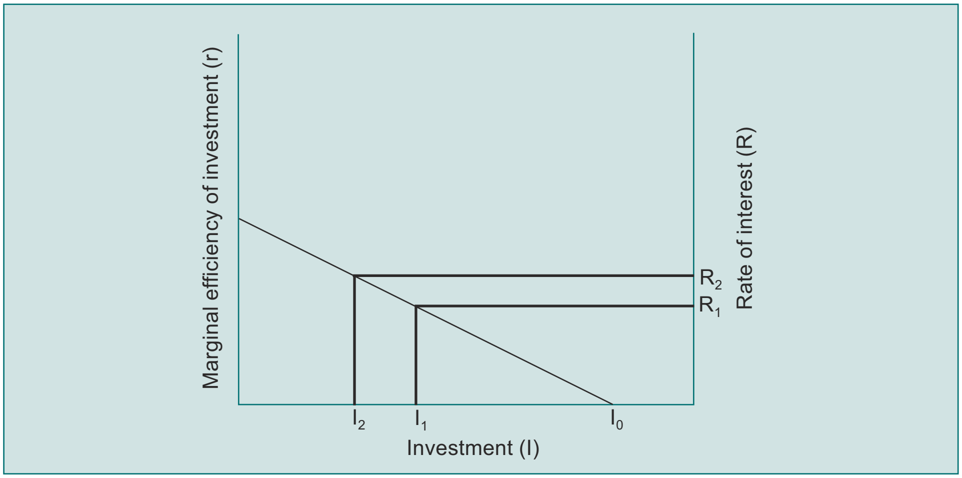


Figure 15.2 Factors determining the volume of investment

The shape of the curve assumes that, given technical knowledge, the marginal efficiency of investment will decline as the volume of investment increases. Thus, each firm is considered to have a list of possible investment projects ranging from projects with high expected profitability to projects with low expected profitability. As the volume of investment undertaken by the firm increases, the rate of return over cost declines. Again, as the volume of investment increases, it will put greater pressure on the productive capacity of the capital goods industries and, as costs in these industries rise, the increasing cost of capital equipment will lower the marginal efficiency of investment. However, there is no convincing historical evidence to suggest that the rate of return over cost has declined through time primarily because technical knowledge improves through time and an improvement in technical knowledge shifts the marginal efficiency of investment curve to the right.

If the only factors influencing investment decisions were the expected returns from investment and the cost of that investment, then, in terms of Figure 15.2, the volume of investment would be I_0, where all investment projects with a positive rate of return over cost, however low, would be undertaken. Yet this would not be rational, because the funds used to finance investment projects have an opportunity cost – resources used for one purpose cannot be used for another. The opportunity cost of investment funds is the return those funds could have earned if employed in their next best alternative use. This opportunity cost may be measured by the rate of interest (R), i.e. what those funds could earn if loaned out. If the marginal efficiency of investment (r) is greater than the market rate of interest (R), then it is rational to proceed with the investment project. If the marginal efficiency of investment (r) is less than the market rate of interest (R), then the business would maximise profits by abandoning the investment project and lending, at the market rate of interest, the funds that would have been used to finance the investment. As shown in Figure 15.2, the volume of investment will be determined where the marginal efficiency of investment equals the market rate of interest.

As the marginal efficiency of investment curve slopes downwards to the right, then the higher the market rate of interest the lower will be the volume of investment and vice versa. Thus it follows that:

$$I = f(R)$$

where the volume of investment is a diminishing function of the rate of interest. Investment will be low when the rate of interest is high and high when the rate of interest is low.

The above analysis is a formalised approach to those factors that impinge on investment decisions. While it is useful in illustrating the essential elements involved, it does tend to imply that investment decisions are mechanistic, the consequence of detailed calculation and painstaking accuracy. In the real world, many investment decisions seem to be unrelated to such detailed calculations. To understand the fluctuations in investment expenditure we must investigate the underlying assumptions of our model.

The curve in Figure 15.2 showing the relationship between the marginal efficiency of investment (r) and the volume of planned investment (I) is drawn on the assumption of a given set of conditions. As these conditions change, the curve will shift. The two most important conditions which are subject to sudden changes are business expectations and the degree of uncertainty.

The marginal efficiency of investment is based on an estimate of future returns, and expectations and uncertainty are fundamental in determining the view that firms take of future returns. One characteristic of investment goods is that they are durable; that is, in estimating returns, firms have to consider a future that usually stretches over a number of years and they have to calculate the likely return on the investments in a situation in which there is no precise and reliable information. A firm must make judgements about future prices and costs, about technological advance, political disturbance, the reaction of competitors, and the behaviour of governments, to name but a few of the factors involved.

The importance of expectations is a key factor in explaining why investment is one of the more volatile components of national income. If the managers who make investment decisions in firms become pessimistic about the future because they fear an international recession or a rise in raw material prices, the consequence will be a reduction in the volume of investment. As shown in Figure 15.3, the marginal efficiency of investment curve will then shift to the left (from AA to BB), giving a lower volume of investment at any given rate of interest. Conversely, if managers become more optimistic about the future, if they believe that the world economy is going to expand rapidly, that a new government will introduce measures more favourable to profits, or that raw material prices are going to fall, then the marginal efficiency of investment curve will shift to the right (from AA to CC), giving a higher volume of investment at any given rate of interest. At a given rate of interest (R_0) the volume of investment varies between I_0, and I_2 depending on the position of the marginal efficiency of investment curve, the position being a function of expectations about the future.

Closely related to expectations is the degree of uncertainty under which investment decisions are made. Investment decisions are based on some view of the future and the future cannot be known with certainty. So, when managers make judgements about the marginal efficiency of investment, i.e. the rate of return over cost, the return is not anticipated with complete confidence – it is an *expected* return. There is a risk that expectations will not be fulfilled, so that an investment that was expected to be profitable may become unprofitable. If the degree of uncertainty increases, the marginal efficiency of investment curve will shift to the left; if uncertainty decreases, the marginal efficiency of investment curve will shift to the right.

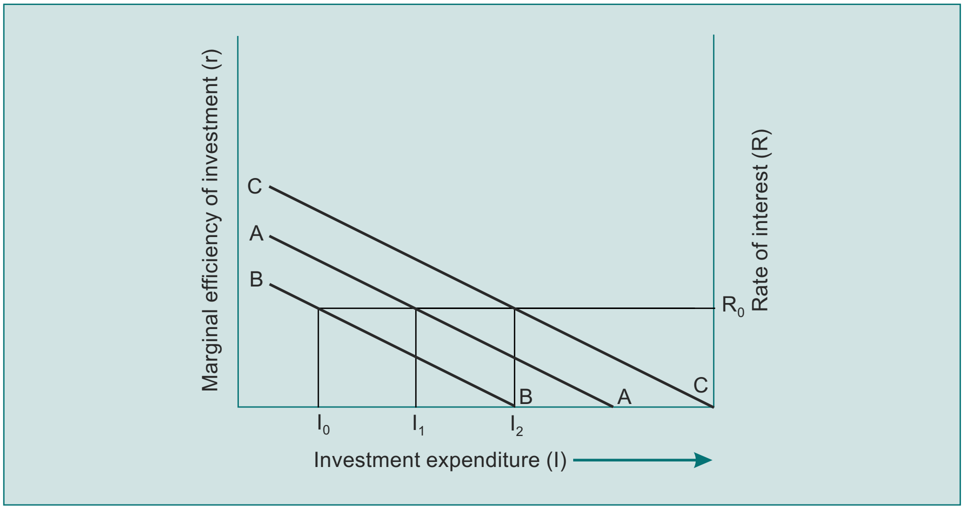


Figure 15.3 Shifts in the marginal efficiency of investment curve

This analysis implies that the marginal efficiency of investment curve is likely to be subject to sudden, sharp shifts depending on expectations and uncertainty. Thus in discussing the marginal efficiency of investment, shifts in the position of the curve are likely to be more important than movements along a given curve. It is misleading to suppose that planned investment is subject to smooth, continuous adjustment in response only to changes in the rate of interest. Instead, many economists believe that investment decisions are relatively little affected by changes in interest rates, and that changes in the volume of investment usually reflect changes in underlying business expectations and the degree of uncertainty. Certainly, this is consistent with the observed behaviour of investment, which is subject to substantial fluctuations over time.

A situation where the volume of investment is influenced by shifts in the curve rather than by movements along the curve is represented as in Figure 15.4. The marginal efficiency of investment curve is drawn almost vertical and subject to large and sudden shifts, suggesting that the volume of investment may not be interest-elastic, i.e. that changes in interest rates may not have much influence on the volume of planned investment. Hence, if the marginal efficiency of investment curve is originally at AA with interest rate R_0, then a change in the rate of interest will have relatively little effect on the volume of investment. For example, for interest rates between R_1 and R_2 the volume of investment varies only between I_1 and I_2 as long as the marginal efficiency of investment curve is AA. However, if the curve shifts, then at any given rate of interest there will be a substantial change in the volume of investment. If the rate of interest is R_0, then a shift in the curve to BB will give investment of I_3, while a shift to CC will give investment of I_4.

Given such a situation, it may be difficult to influence the volume of investment by variations in interest rates induced by monetary policy. Investment expenditures would depend on the state of business expectations and on the degree of uncertainty, and these factors may not be subject to close control through the conventional instruments of economic policy. If so, investment expenditures would introduce a substantial element of uncertainty into the economic system. Another factor contributing to fluctuation in investment expenditure is based on the idea that some investments are determined by the rate of change of national income/output. This is known as the *accelerator principle*, i.e. I =

f (Y). Many economists claim that the *accelerator*, combined with the *multiplier*, provides an explanation for the observed fluctuations in the level of economic activity. To understand the accelerator principle consider the following simple example.

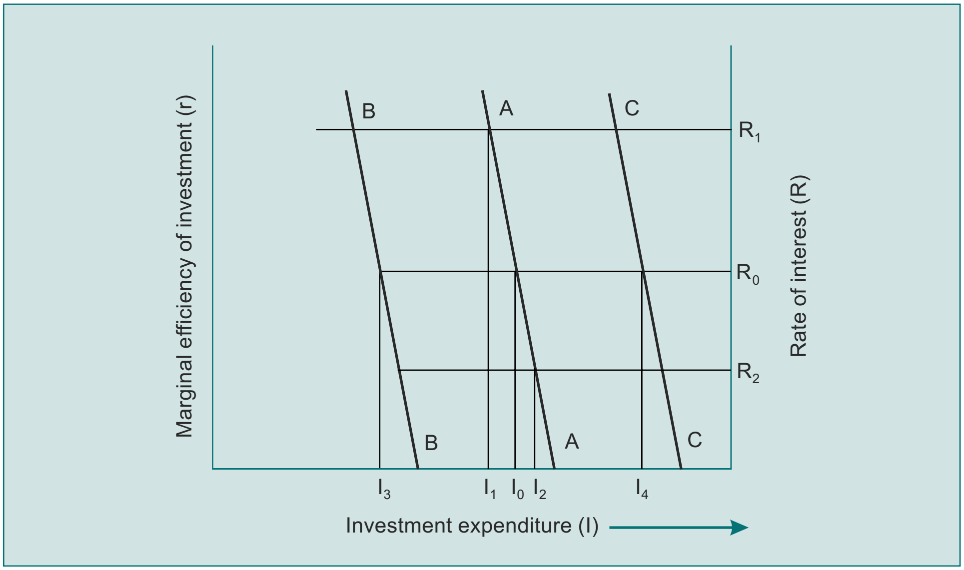


Figure 15.4 Shifts in and movements along the marginal efficiency of investment curve

Example

A firm producing shoes requires $2 million of capital goods to produce a $1 million output of shoes and this capital–output ratio is fixed at all levels of output. Capital goods are assumed not to depreciate, and the new capital investment is deemed instantaneous (these assumptions simplify the subsequent arithmetic rather than influence any point of principle). The level of sales of shoes is shown in Table 15.1.

Table 15.1 Net investment and sales of shoes

Year	Sales	Required capital stock	Net investment
1	$1.0 million	$2.0 million	–
2	$1.0 million	$2.0 million	nil
3	$1.2 million	$2.4 million	$0.4 million
4	$1.5 million	$3.0 million	$0.6 million
5	$1.6 million	$3.2 million	$0.2 million

As long as sales remain at $1 million, no new net investment is required. As sales rise, the firm requires additional capital goods to allow greater output. Given the capital–output ratio of 2:1, each extra $1 of sales requires $2 of net investment. Hence, the volume of net investment, shown in the right-hand column, is simply twice the increase of sales recorded in the left-hand column. As long as the capital–output ratio is greater than unity, the increase in net investment will exceed the increase in sales. Increases in sales are, presumably, a function of changes in the level of income rather than a function of the level of income itself. Thus, in the above example, the volume of net investment in period 5 is lower than the volume of net investment in period 4, although in period 5 sales, and presumably income, are higher. The lower volume of investment in the last time period is due to the slower rate of increase of sales (income) in that period. The relationship can be expressed algebraically as

$$I = \alpha Y$$

where α is the capital–output ratio, sometimes called the *accelerator coefficient*.

The accelerator principle suggests that the change in net investment is a function of the *rate of change* of income.

The accelerator principle does not yield a comprehensive explanation of investment, for it does not deal with replacement investment, nor with investment associated with new innovations and new goods, nor with 'autonomous' investment (investment that is *not* closely associated with current output – for instance investment in atomic energy, or hydro-electric developments). Even where investment decisions are influenced by changes in income, the causal connection is often less immediate and simple than the accelerator principle suggests. Nonetheless, the empirical evidence does indicate that a significant part of observed fluctuations in investment may be associated with changes in income levels, and the accelerator principle does provide a clear explanation for the observed cyclical instability of the output of the capital goods industries. In combination with the multiplier, the accelerator helps to explain the observed fluctuations of economic activity that are such a feature of economic history.

The multiplier demonstrates that an initial change in aggregate demand has secondary effects on the level of income, causing the final change in income to be some multiple of the initial change in aggregate demand. The eventual change in income depends on the size of the multiplier and the size of the multiplier is determined by the marginal propensity to consume. The higher the marginal propensity to consume, the lower the leakages from the circular flow of income, the greater the multiplier, and the greater the change in income that will be associated with any initial change in aggregate demand.

If we take the initial change in aggregate demand as resulting from a change in investment, then:

$$\Delta Y = k \times \Delta I$$

where k is the multiplier.

Changes in investment, through the multiplier process, have secondary effects on the level of income. The causal sequence runs *from changes in investment to changes in income*. The accelerator principle suggests that the secondary effects may be complicated, and are likely to be greater, because changes in income resulting from the multiplier process are likely to have consequential effects on the volume of investment. The causal sequence runs *from changes in income to changes in investment*, where

$$\Delta I = \alpha \Delta Y$$

where α is the capital–output ratio or accelerator coefficient.

Suppose there is an initial autonomous increase in the volume of investment expenditure. The higher investment expenditure creates additional incomes of factors of production previously unemployed (note, again, that the multiplier and accelerator process can only create an increase in real income if the economy starts from a position where some factors of production are unemployed) and this increase in income results in an increase in consumption expenditure. That is, the multiplier process comes into effect. As income and consumption rise, firms find that the productive capacity of their existing capital stock is inadequate to meet the new and higher level of demand, and additional net investment will take place. Thus the accelerator process comes into effect. Two forces are at work: increased investment, through the multiplier, generates higher income; and higher income, through the accelerator, generates higher investment. Of course, the reverse applies when the initial disturbance is a *fall* in investment.

The behaviour of the economy, in the reaction of income, output and employment to any initial change in aggregate demand, is therefore influenced by the size of the multiplier (given by the marginal propensity to consume) and the size of the accelerator (given by the capital–output ratio). The higher the multiplier and the greater the accelerator, the more explosive the reaction of the economy to any initial change in aggregate demand and the greater the duration and strength of the cumulative movements in income, output and employment that result. The lower the multiplier and the smaller the accelerator, the more stable is the economy likely to be, in that fluctuations in income, output and employment are likely to be relatively small and the economy is unlikely to be characterised by substantial and sustained cumulative movements.

15.3 The Government Sector

Government taxation and government expenditure are analogous, in their effects on the circular flow of national income, to household saving and business investment. Taxation is a withdrawal from the circular flow of income, as the act of taxation is not a claim on output that creates factor income and employment. On the other hand, government expenditure is an injection into the circular flow of income, constituting a claim for output and creating factor income and employment. If the government purchases goods and services in order to provide medical services, build roads and airports, and provide law, order and defence, then the expenditure creates both income and employment. Government expenditure is therefore a component of aggregate demand or expenditure.

The effect of government taxation and expenditure on the circular flow of national income is shown in Figure 15.5.

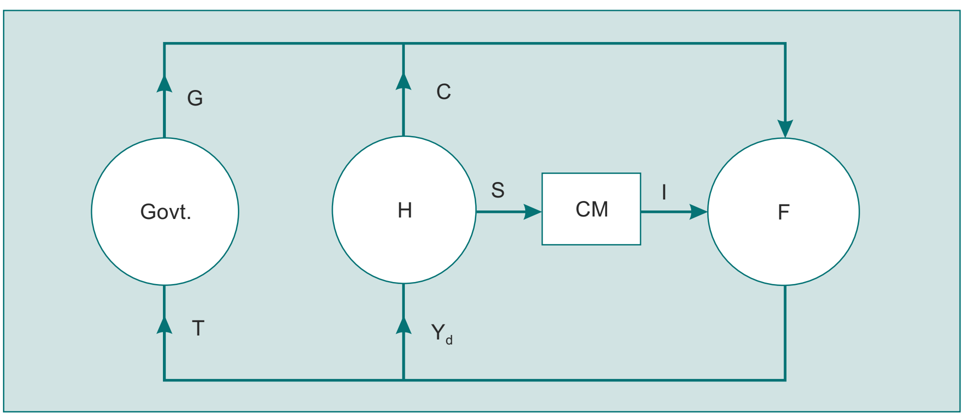


Figure 15.5 The circular flow of income and expenditure: impact of the government sector

Notation: Y_d=disposable personal income (i.e. gross income less taxes); T=Taxes; G=Government expenditure, H=Households, F=Firms, S=Savings, I=Investment, CM=Capital Market, C=Consumption.

In Figure 15.5, withdrawals consist of household savings and government taxation, and injections of business investment and government expenditure. Equilibrium national income is achieved when

$$S + T = I + G$$

where S is household savings, T is taxation, I is business investment, and G is government expenditure.

For equilibrium, the sum of planned savings and taxation must equal the sum of planned investment and government expenditure. It follows from this that planned savings can diverge from planned investment without necessarily causing a change in equilibrium income. This can happen provided the inequality between planned savings and planned investment is offset by an inequality, of opposite sign but the same magnitude, between taxes (T) and government expenditure (G).

In the instance where $S > I$ and $T < G$, there is a budget deficit – that is, government expenditure is greater than the revenue raised in taxes; in the second instance, where $S < I$ and $T > G$, there is a budget surplus – that is, government expenditure is less than the revenue raised in taxes. The effect of a budget deficit is expansionary, leading to a rise in national income and employment as long as there are unemployed resources in the economy: injections into the circular flow of income in the form of government expenditure are greater than withdrawals from the circular flow of income in the form of taxation. The effect of a budget surplus is deflationary, leading to a fall in national income and employment: withdrawals from the circular flow of income in the form of taxation are greater than injections into the circular flow of income in the form of government expenditure.

Consider again equilibrium national income being reached when

$$S + T = I + G$$

and the implications of a budget deficit, a common feature of many market economies. A budget deficit means $G > T$ and as a consequence $I < S$.

Rewriting the equation as

$$S = I + (G - T)$$

indicates that when an economy is in equilibrium for any given S, the greater is the budget deficit $(G - T)$, the smaller must be private investment I, i.e. private savings (S) must now accommodate private investment expenditure I and also the budget deficit, $G - T$ (where $G > T$). In our expanded model we can now rewrite our national income identity so as to include the government sector; we have:

$$Y\ (\text{output}) \equiv C_p + I + G$$

$$Y\ (\text{income}) \equiv C_b + S + T$$

On the output side, the economy produces three types of goods: consumer goods (C_p), investment goods (I), and government goods (G). On the income side, there are three things households can do with this income: spend them (C_b), save them (S) or pay taxes (T). Since national output, by definition, must equal national income, we have:

$$C_p + I + G = C_b + S + T$$

Thus, where the amount of consumer goods produced just equals the amount purchased, i.e. when the $C_p = C_b$ are equal, the economy will be in equilibrium. If, in addition, the budget is balanced (i.e. $G = T$), then planned saving (S) will equal planned investment I. Remember, however, that the decisions on how much consumer goods and services to produce are made by firms, whereas the decisions on how many to buy are made by households. What will happen if the amount produced does not equal the amount purchased? As in our simple model, national income/output will not be in equilibrium and the level of national income and employment will change until equilibrium is established.

Let us restate our equations

$$Y_o \equiv C_p + I + G$$
$$Y_i \equiv C_b + S + T$$

Suppose $C_p > C_b$. Firms will discover initially unsold consumer goods, which will accumulate as inventories. These unintended inventories can be represented by

$$Inv_u = C_p - C_b$$

which can be reformulated as

$$C_p = Inv_u + C_b$$

But, remember, inventories in the national income accounts are part of investment. Thus we can rewrite our national output identity as

$$\begin{aligned} Y_o &\equiv C_p + I + G \\ &\equiv C_b + Inv_u + I + G \end{aligned}$$

and with national income remaining as

$$Y_i = C_b + S + T$$

we can see that, with a balanced government budget of G = T,

$$S = I + Inv_u$$

In these circumstances, savings (S) must always equal investment ($I + Inv_u$)but clearly planned savings (S) no longer equal planned investment I. Thus the economy is in disequilibrium and firms, wanting to get rid of their unintended inventories that are unsold consumer goods, will decrease their output of consumer goods in the next time period. This will cause Y_o to fall and unemployment to increase. The equilibrium level of national output will be reached only when planned savings and planned investment are again equal.

This analysis has assumed a balanced budget, i.e. G and T equal. What happens if the government decides to increase its level of expenditure in light of the output gap (rising unemployment)? Similar to an increase in autonomous investment expenditure in our simple model of Module 14, the government expenditure will also cause multiplier effects in the economy and will increase national output and employment. Before we turn to the analysis of how this occurs, it is important to define what we mean by *government expenditure*.

There are two principal measurements of government in the economy, government expenditure on current goods and services (the G in the national income accounts) and total government outlays, which equal G + transfers. These transfers do not enter the national income identity because they do not represent expenditures on goods and services produced in the circular flow of income and output – they can be thought of as negative taxes and income in the form of transfers to the old, the infirm and the unemployed. How important is the government sector in the modern market economy?

Figure 15.6 shows the two measures of government expenditure, as a proportion of GNP, for the leading seven capitalist nations in 2011.

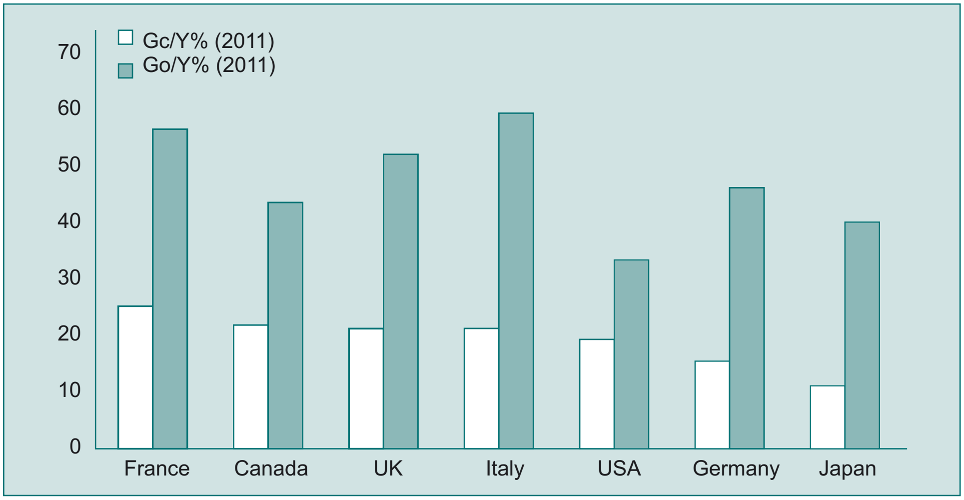


Figure 15.6 Government consumption and total outlays for the G7 economies, 2011

As can be seen from Figure 15.6, with the exception of Japan and Germany there is a striking similarity among the other nations, i.e. approximately one-fifth of GNP is directly controlled by governments, through government consumptions. This influence rises to one-third to almost two-thirds if we include transfers. The latter proportion has been rising over time, indicating the growing importance of social expenditure on items such as state pensions, unemployment benefit and sickness benefits. Thus, with the exceptions of Japan and Germany, G as a percentage of GNP is quantitatively similar to I for the five major capitalist countries shown, with all around 20 per cent of GNP. As a consequence for those five, C is around 60 per cent of GNP. Germany has high private consumption, low government and high investment, and Japan has low private consumption, low government and high investment.

We can now analyse the effects of changes in government expenditure (G) and/or taxes (T) in the level of national income/output and employment. We shall assume that government expenditure on goods and services (G) is completely autonomous, i.e. the government has total power in determining G, although in practice this is not true: the data presented in Figure 15.6 include central, state (regional) and local government expenditures and, as a consequence, central government, the new principal player along with households and firms in our model, has only indirect control over both the amount and composition of expenditures of lower tiers of government. In addition, in practice, there are specific categories of expenditure that no government can be expected to reduce below existing levels.

The government purchases goods and services from the private sector (households plus firms) to produce government services such as defence, law and order, health, and education. The government typically does not sell these services and so finances them from taxes levied on households and firms. If taxes are insufficient to meet government outlays, then the government has to borrow – and from whom does it borrow? The answer is from households and firms.

In expanding our model to include government we shall retain our assumptions affecting households and firms, namely consumption expenditure (C) is a function of disposable income (Y_d) and investment expenditure (I) is autonomous. For government we shall assume government expenditure (G) is autonomous; as far as taxes are concerned, we shall have two versions of this model: in the first, taxes (T) will be autonomous, i.e. independent of the level of national income (Y). (An example of such a tax would be a tax per head or 'poll' tax, where each person pays a specified tax independent of income level.) In the second version we shall assume the government determines a tax rate that is proportional to income (Y). In this case the government no longer controls the tax yield since it does not control Y but only the rate applied to Y.

In version 1,

$T = T$

and in version 2,

$T = t(Y)$, where t is a constant.

We can now specify the algebraic interpretation of our expanded model to include the government sector.

15.3.1 Version 1 of the Expanded Model

In this version we have the following set of equations:

$$
\begin{aligned}
Y &\equiv C + I + G \\
C &= bY_d \\
I &= I \\
G &= G \\
T &= T
\end{aligned}
$$

$$
\begin{aligned}
Y_d &= Y - T \\
C &= b(Y - T) \\
Y &= b(Y - T) + I + G \\
Y - bY &= -bT + I + G \\
Y(1 - b) &= -bT + I + G \\
Y &= \frac{-bT + I + G}{1-b} \\
Y &= \frac{-bT}{1-b} + \frac{I}{1-b} + \frac{G}{1-b}
\end{aligned}
$$

Recall the similar equations we encountered in Module 14:

$$
\begin{aligned}
Y &= C + I \\
C &= bY \\
Y &= bY + I \\
Y - bY &= I \\
Y(1 - b) &= I \\
Y &= \frac{I}{1-b}
\end{aligned}
$$

and, therefore,

$$\Delta Y = \frac{1}{1-b} \cdot \Delta I$$

where the multiplier was

$$\frac{1}{1-b}\left(\text{or } \frac{1}{1-MPC}\right)$$

In our expanded model

$$Y = \frac{-bT}{1-b} + \frac{I}{1-b} + \frac{G}{1-b}$$

Let us again increase investment expenditure by ΔI. Because T and G are both autonomous, their values do not change when Y changes. Thus a change in I, which causes multiplier effects on Y (and C) will not affect taxes (T) or government expenditure (G). We thus have

$$\frac{\Delta Y}{\Delta I} = 0 + \frac{1}{1-b} + 0$$

or

$$\Delta Y = \frac{1}{1-b} \cdot \Delta I$$

What happens if we leave I alone – remember I is autonomous – and change G? Our starting equation is

$$Y = \frac{-bT}{1-b} + \frac{I}{1-b} + \frac{G}{1-b}$$

and therefore

$$\frac{\Delta Y}{\Delta G} = 0 + 0 + \frac{1}{1-b}$$

or

$$\Delta Y = \frac{1}{1-b} \cdot \Delta G$$

This shows that our investment and government expenditure multipliers are identical. This is because both G and I are assumed to be autonomous.

Suppose, now, that I is not completely autonomous; suppose an element of investment expenditure is related to the *level* of income Y (introducing part of I as dependent on Y or last year's income Y_t complicates the mathematics horribly!). Thus the investment function would be represented by the sloping function in Figure 15.7, i.e. $I = I_0 + kY$.

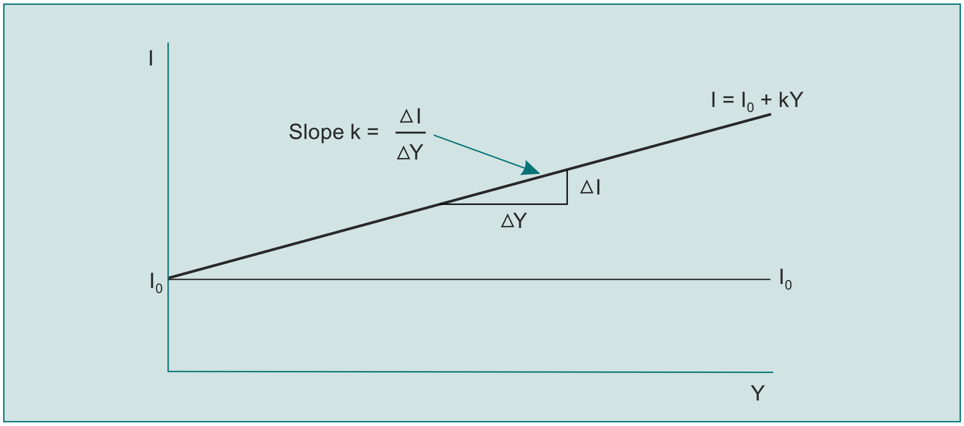


Figure 15.7 Autonomous and endogenous investment variables

Investment expenditure in Figure 15.7 now consists of two elements, I_0 (which is independent of the level of Y) and kY (which is endogenous). Restating our set of equations, we have

$$Y = C + I$$

$$C = bY$$

$$I = I_0 + kY$$

$$Y = bY + I_0 + kY$$
$$Y - bY - kY = I_0$$
$$Y(1 - b - k) = I_0$$

Therefore

$$Y = \frac{I_0}{1-(b+k)}$$
$$\frac{\Delta Y}{\Delta I_0} = \frac{1}{1-(b+k)}$$
$$\Delta Y = \Delta I_0 \cdot \frac{1}{1-(b+k)}$$

As expected, the multiplier is larger because any increase in autonomous investment I_0 causes multiplier effects through increased Y, leading to increased C, leading to increased Y, etc., but also because increased autonomous investment expenditure ΔI_0 leads to induced investment expenditure because of increased Y. This induced I further increases Y and induces even more I until the multiplier process works its way to completion.

15.3.2 Version 2 of the Expanded Model

Let us now look at version 2 of our expanded model as it attempts to take account of the government sector. Here we no longer have a poll tax; the government sets an income tax rate. For simplicity's sake we are assuming the tax rate is constant. In most capitalist countries we have a progressive tax structure out of income, where each higher band of income is taxed at a higher rate. Such a sample tax structure is shown in Table 15.2.

Table 15.2 Progressive income tax structure

Income	Income range $	Marginal tax rate %	Marginal tax	Total tax	Average tax rate %
5 000	0 – 5 000	0	0	0	0
10 000	5 000 – 10 000	20	1 000	1 000	10.0
15 000	10 000 – 15 000	50	2 500	3 500	23.3
20 000	15 000 – 20 000	80	4 000	7 500	37.5

An individual earning up to \$5000 per annum pays no tax. An individual earning \$10 000 per annum pays no tax on the first \$5000 but pays 20 per cent on the next \$5000, i.e. \$1000 (representing a marginal tax rate of 20 per cent). Thus the tax paid on \$10 000 = 0 + \$1000, yielding an average tax rate of 10 per cent (\$1000/\$10 000).

An individual earning \$15 000 pays 0 on the first \$5000, \$1000 on the next \$5000 and \$2500 on the third \$5000 (at the marginal tax rate of 50 per cent). This sums to \$3500 (0 + \$1000 + \$2500) and equates to an average tax rate of 23.3 per cent.

If the individual's income now increases by a further \$5000, that individual will pay \$4000 in tax on the additional \$5000, for a marginal tax rate of 80 per cent applies – although the average tax rate rises to only 37.5 per cent.

Returning to our version 2 of the government sector, the tax rate is going to be constant, not progressive, i.e.

$$T = tY$$

where t is a constant (for a linear relationship). Therefore

$$Y_d = Y - tY$$
$$= Y(1 - t)$$

Respecifying our equations, we have

$$Y = C + I + G$$
$$C = bY_d$$
$$Y_d = Y(1 - t)$$
$$I = I$$
$$G = G$$

Therefore

$$Y = bY_d + I + G$$
$$= bY(1 - t) + I + G$$

Therefore

$$Y - bY(1 - t) = I + G$$
$$\therefore Y[1 - b(1 - t)] = I + G$$
$$\therefore Y = \frac{I}{[1-b(1-t)]} + \frac{G}{[1-b(1-t)]}$$

Because G is autonomous, we have

$$\frac{\Delta Y}{\Delta I} = \frac{1}{[1-b(1-t)]} + 0$$

and because I is also autonomous, we have

$$\frac{\Delta Y}{\Delta G} = 0 + \frac{1}{[1-b(1-t)]}$$

15.3.3 The Differences

The major difference between the two versions is that the poll tax of version 1 does *not* affect the value of the multiplier whereas the imposition of a marginal tax rate in version 2 reduces the value of the multiplier. Why? Because part of the increase in income resulting from an autonomous increase in G or I 'leaks' from the circular flow in the form of tax, and then the proportion of income that households pay in tax is no longer available to purchase goods and services, and create additional income and further rounds of consumption expenditure and income.

Let us return to version 1, defining a per capita tax. Suppose there is unemployment in the economy and the government wishes to reach full employment; should it increase government expenditure or cut tax? Does it make any difference if it increases G by a given amount or cuts T by an equal amount? The answer is 'yes'. Why? Suppose the gap between Q and Y equals 100 units. Further, suppose the value of MPC = 0.5, yielding a multiplier of 2 (multiplier = 1/1 − MPC, remember). Thus the required increase in G is 50 units. The multiplier process is spelled out in Table 15.3; Table 15.4 spells out the multiplier process for a tax cut of 50 units.

Table 15.3 Government expenditure multiplier

Round	ΔC	ΔG	ΔY
1	–	50	50.0
2	25.0	0	25.0
3	12.5	0	12.5
4	6.25	0	6.25
.	.	.	.
.	.	.	.
.	.	.	.
Total	50	50	100

Table 15.4 Tax-cut multiplier

Round	ΔC	ΔG	ΔY
1	–	−50	–
2	25.0	0	25.0
3	12.5	0	12.5
4	6.25	0	6.25
.	.	.	.
.	.	.	.
.	.	.	.
Total	50	0	50

What is missing in Table 15.4 compared with Table 15.3 is the first round of income/output generation. A cut in tax of 50 units does not generate any output, whereas an increase in G of 50 does. In the tax-cut case, households discover collectively that disposable income has increased by 50 units and, given MPC of 0.5, 25 units of the tax cut will be spent and 25 will be saved. The 25 spent on goods and services means an increase in Y of 25, which in time leads to a further increase in C of 25 and resultant multiplier rounds. Thus if the government in this example wishes to reach full employment, the required tax cut is 100.

The above example leads to an interesting corollary. Suppose the government wishes to increase G but, for political reasons, must balance its budget, i.e. ΔG must equal ΔT. What will happen to Y? Will ΔY be positive, negative or zero in value? As can be seen from Table 15.3, an increase in G of 50 will increase Y by 100. In Table 15.4 we could trace out the impact of an *increase* in tax compared with the existing decrease, and we would discover that ΔY would be −50. Thus an increase in G of 50 accompanied by a tax increase of 50 would increase Y by 50. Thus the *balanced budget multiplier* = 1; the increase in Y is exactly equal to the initial increase in G.

The geometric interpretation of version 1 of the expanded model is shown in Figure 15.8. Equilibrium is achieved where aggregate demand crosses the 45° line (aggregate supply). Thus D_1 produces an output level of Y_1 and employment level of E_1; at this equilibrium level, consumption expenditures C_1, I and G are autonomous. Potential output (Q) would be reached only if national income/output equalled Y_F and E_F. To achieve full employment, aggregate demand has to be increased. This is done in Figure 15.9 by increasing G by ΔG. The increase in G of ΔG causes an increase ΔY in Y of $Y_F - Y_1$.

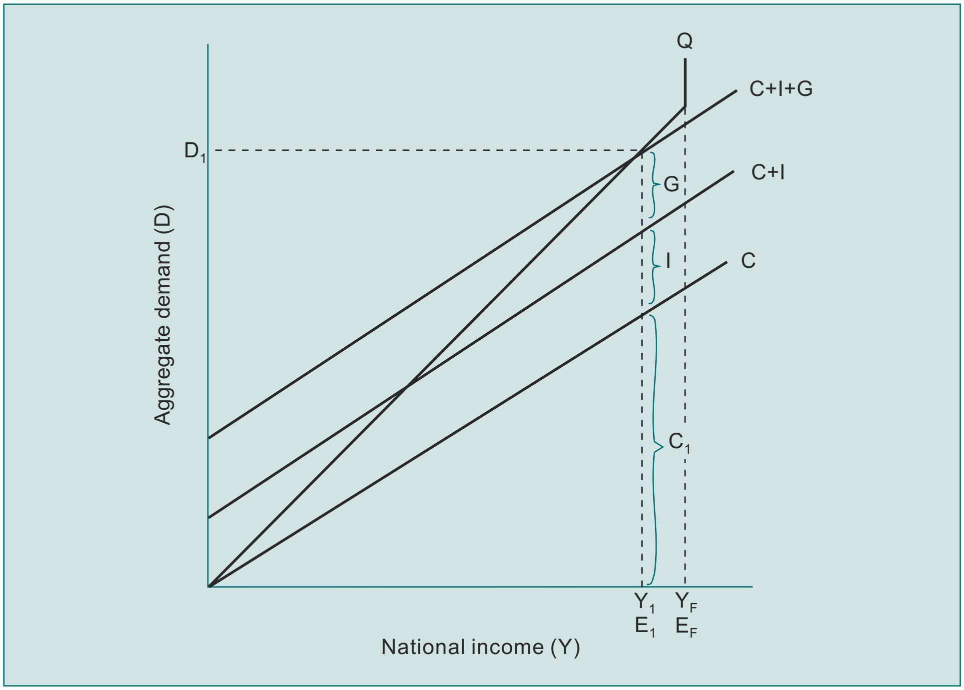


Figure 15.8 Expanded model of income determination (version 1)

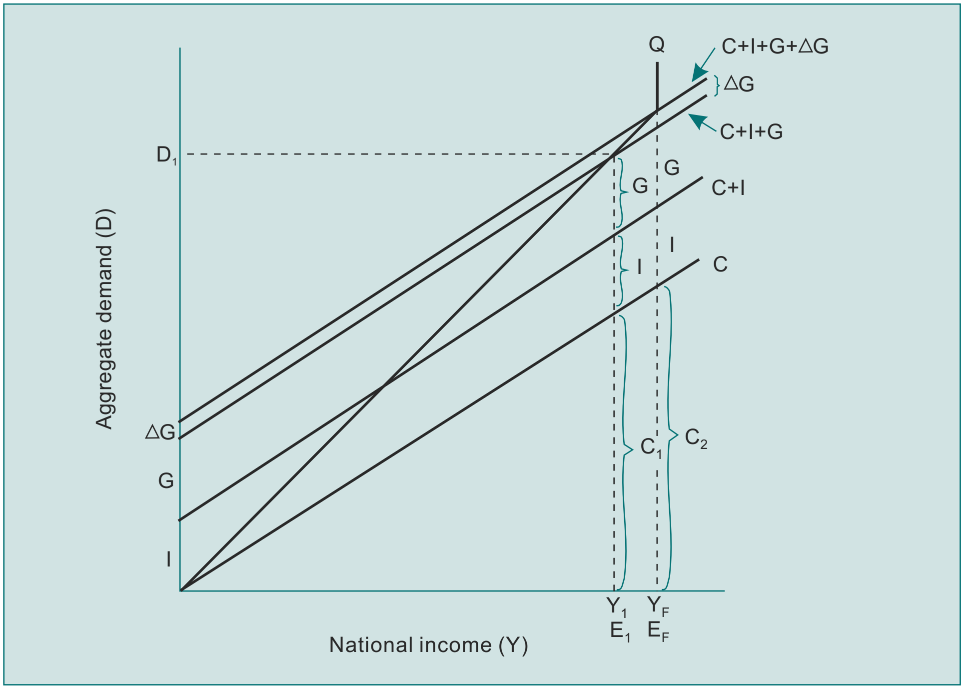


Figure 15.9 Impact of an increase in variable G

Using our multiplier formula we know that

$$\Delta Y = \Delta G \times \text{Multiplier}$$
$$= \Delta G \times \frac{1}{1-\text{MPC}} = \Delta G \times \frac{1}{1-b}$$

MPC being determined by the slope of the consumption function (b). Thus the increase in Y, $(Y_F - Y_1)$, is made up of two components, ΔG and the induced consumption expenditure $C_2 - C_1$.

Had the investment function not been autonomous, i.e. had there also been induced investment expenditure, the amount of ΔG required to achieve full employment would have been less.

To accommodate version 2 of our expanded model geometrically, we have to shift the consumption function – and consequently the aggregate demand function – if the tax is of the poll tax variety. This must be done such that the slope of the consumption function will not change. If the tax is a tax on income, then we have to change the slope of the consumption function.

15.4 In-Built Stabilisers

The discussion in Section 15.3 might be taken to imply that explicit action must be taken by the fiscal authorities to influence the level of economic activity through changes in the balance between government expenditure and taxation. This is not so. Although the analysis does suggest a need for a *discretionary* fiscal policy that manipulates expenditure and taxation in a deliberate attempt to influence aggregate demand, the nature of government expenditure and taxation regimes creates a degree of *automatic stabilisation*, so that some stabilisation will take effect even if the fiscal authorities remain passive.

The nature of government expenditure and the structure of taxes is such that a rise in employment and incomes will tend to raise government taxation relative to expenditure, acting as a brake on the economy. Conversely, when employment and incomes fall, this will tend to reduce government taxation relative to expenditure, thus tending to stimulate the level of economic activity. This tendency will come into play without any explicit action on the part of the government, and such an in-built stabiliser may be important in reducing fluctuations in the level of economic activity.

An *in-built stabiliser* can be described as any aspect of government taxation and expenditure policies that automatically reduces government expenditure and/or increases government tax revenue when income and output are increasing, or that automatically increases government expenditure and/or reduces government tax revenue when income and output are falling. The most important in-built stabilisers are taxes, transfers and price supports.

Almost all taxes vary directly with the level of income, increasing when incomes are increasing and falling when income is falling. This applies to sales taxes, excise duties, taxes on profits, payroll taxes and income taxes. The effect is most marked in the latter case, particularly if the income tax regime is highly progressive, as shown in Table 15.4. Where the income tax regime is progressive, tax revenue will rise more quickly than household income as income increases. This will occur because, with rising incomes, more households will be drawn into ranges of income where higher rates of tax apply. These results are reversed if income falls. Hence, tax revenues rise more quickly than household incomes when incomes rise, and they fall more quickly than household incomes as incomes fall.

Most capitalist economies have a comprehensive system of unemployment compensation financed by employer and employee contributions, which are related to current income. In a recession, when unemployment rises and employment incomes fall, the amount paid out in unemployment benefit will increase and the income obtained from contributions will fall. The effect of unemployment compensation is to reduce the decline in income experienced by the unemployed, which allows them to maintain higher consumption than would otherwise be possible. The adverse multiplier effect of a fall in employment is therefore reduced, and it will be smaller the higher is unemployment income relative to employment income. Conversely, in the recovery stage, expenditure on unemployment benefits will fall and revenue from contributions will increase. Other welfare programmes, for example statutory redundancy payments and social security payments, will exhibit the same pattern of rising expenditure when employment income is falling and falling expenditure when employment income is rising.

The best example of price support systems are those applied to agriculture. If the demand for agricultural products declines, price support systems come into effect, thus reducing the decline in agricultural incomes and maintaining consumption above the level that would otherwise have been possible. As with unemployment compensation, this reduces the adverse multiplier effects of a fall in demand for agricultural products. Conversely, if the demand for agricultural products rises, then, as market prices recover, the amount of price support will decline and agriculture incomes will grow more slowly than agricultural output. This will act as a brake on consumption by the agricultural sector.

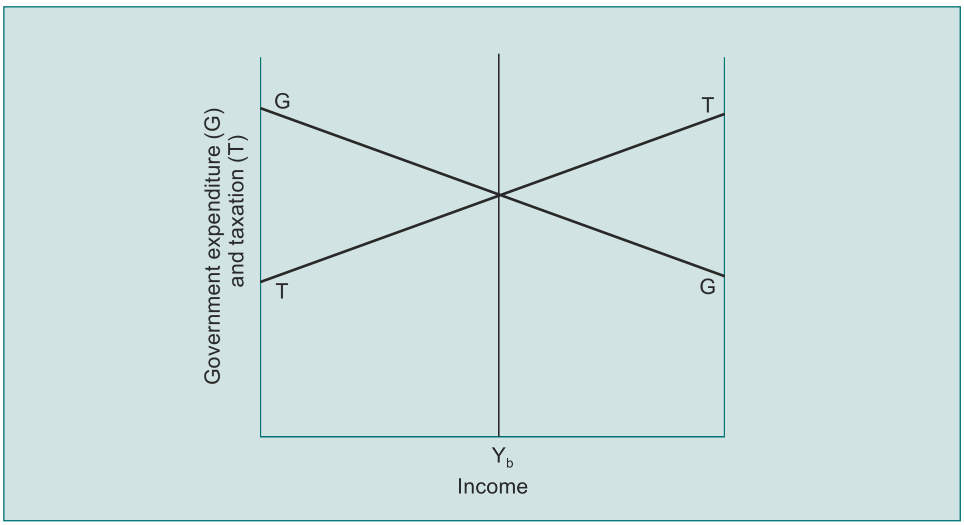


Figure 15.10 The effect of in-built stabilisers

The effect of in-built stabilisers is demonstrated in Figure 15.10. At income level Y_b the budget is in balance, with government expenditure equal to taxation. At income levels below Y_b a budget deficit emerges as transfer payments increase due to increased unemployment benefits, welfare payments, and price support; and taxation revenue falls with falling income. At levels above Y_b, a surplus emerges as taxes rise and transfer payments fall.

The properties of stabilisation built into government expenditure and taxation regimes may be desirable or undesirable. They will be desirable if the economy has a tendency to show relatively small fluctuations around full-employment income, avoiding both heavy unemployment and a high rate of inflation. In these circumstances, in-built stabilisers will assist in maintaining a high level of activity without inflation. In-built stabilisers are likely to be undesirable if the economy has departed substantially from full employment and price stability and is experiencing either high unemployment or a high rate of inflation.

Some economists believe that there are strong forces at work that tend to restore a market economy towards full employment. Because of this, they believe that a *discretionary* fiscal policy is damaging and therefore advocate a non-discretionary fiscal policy that relies on in-built stabilisers to apply fiscal stimulus and restraint where these are necessary. They propose that the aim of fiscal policy should be to establish a balanced budget at the highest level of employment that is consistent with price stability. This *full-employment budget balance* can be represented by Y_b in Figure 15.10. If this is then combined with the sensitive use of in-built stabilisers, movements away from full employment, i.e. booms and recessions, would automatically result in countercyclical surpluses and deficits, reinforcing the tendency of market forces to restore full employment.

Non-discretionary fiscal policy is advocated by economists called 'monetarists', who also advocate a non-discretionary monetary policy. The monetarists' view is that discretionary fiscal policy will aggravate rather than reduce the fluctuations in economic activity that exist in a market economy. They believe that explicit action to attempt to stabilise the economy, by manipulating aggregate demand through government expenditure and taxation, will increase the amplitude of cyclical fluctuations if a discretionary fiscal policy is applied. 'Keynesians' are of the opposite persuasion; that is, they believe that discretionary fiscal policy has a stabilising influence.

Both monetarists and Keynesians would agree that in certain circumstances in-built stabilisers may be undesirable. If an economy has high unemployment or high inflation, the effect of *fiscal drag* may produce unplanned and undesirable results. If, given heavy unemployment, the level of aggregate demand increases, then in-built stabilisers will reduce the size of the multiplier and hence the expansion in income that results from the initial stimulus to aggregate demand. Withdrawals through government taxation increase, and government expenditure will fall as income increases and may prevent the economy regaining full employment.

High inflation combined with unindexed taxes can produce a rapid, unplanned shift of income from private households to the government. An *unindexed tax system* is one in which tax thresholds are set in money terms and are not adjusted for changes in price levels. In consequence, if money incomes and prices rise in line with each other, households may find their tax liabilities increasing although their real income has not changed. *Real disposable income*, i.e. income net of tax and transfer and adjusted for price changes, will fall. This feature of the tax system has been particularly noticeable in periods of high inflation. It has led to attempts in both the United States and the UK to provide some method of tax indexation so that tax thresholds are adjusted upwards to offset the effects of inflation. Without some form of tax indexation, a progressive tax system combined with a high rate of inflation tends to reduce real personal disposable income and shift resources to the public sector of the economy.

In summary, in-built stabilisers are desirable where an economy has a pronounced tendency to oscillate around full employment without price inflation. In these circumstances in-built stabilisers reinforce the stabilising forces within the system. Where the economy departs substantially from full employment and exhibits either high unemployment or high inflation, in-built stabilisers may prevent the economy returning towards full employment with price stability, or may result in an unplanned transfer of income from private households to the government sector.

15.5 The International Sector

So far we have assumed in our simple models that a *closed economy* exists, i.e. no international economy, no imports and no exports of goods and services, no international movements of resources and no international movement of currencies. In our expanded model we shall now include the first mentioned, international trade in goods and services.

Thus households, besides spending income on domestically produced goods and services, saving income, and paying taxes can now buy imported goods and services, which we shall designate Z. Similarly domestic firms and our government can purchase goods and services produced abroad (also Z). This variable Z, of course, will be produced by foreign resources. Flowing in the other direction are a stream of goods and services produced by our resources and exported to other countries. Export sales (X) of goods and services create income and employment for factors engaged within the domestic economy.

Imports are analogous, in their effect on the circular flow of income, to household savings and taxation, and exports are analogous to investment and government expenditure. Imports constitute a demand for goods and services produced in other economies and hence constitute a withdrawal from the circular flow of income as far as the domestic economy is concerned. On the other hand, exports are sold to the residents of other economies and the income resulting accrues to factors of production within the domestic economy and constitutes an injection into the domestic circular flow of income.

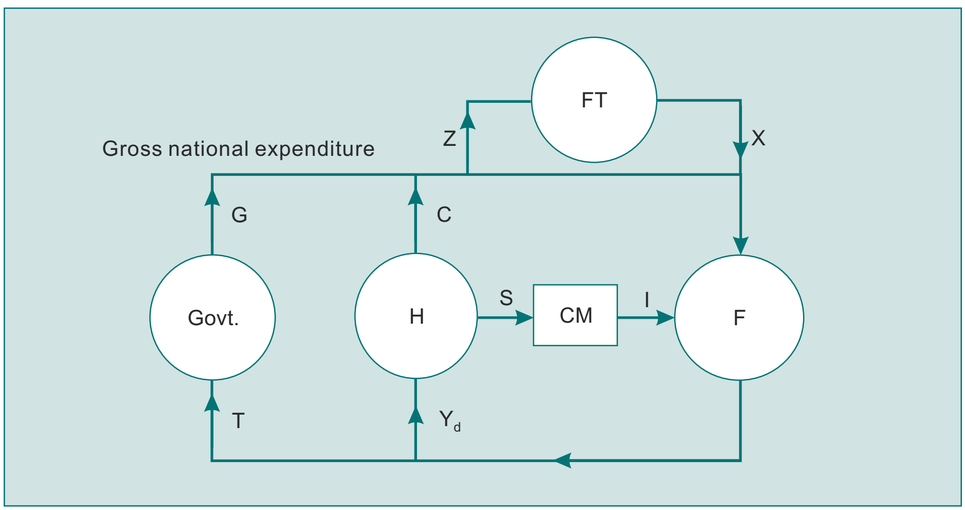


Figure 15.11 The circular flow of income and expenditure: impact of the foreign sector

Notation: Z=Imports, X=Exports, FT=Foreign trade sector, others as Figure 15.5.

The effects of imports and exports on the circular flow of income are shown in Figure 15.11 by extending the circular flow (Figure 15.5) to include an international sector. Given the government sector and international trade, withdrawals now consist of household savings, taxation and imports, and injections of business investment consist of government expenditure and exports, with equilibrium income being achieved where S + T + Z = I + G + X. For equilibrium, the sum of planned withdrawals and injections must be equal, but equality between the individual items is not necessary. In the previous section we saw that, with no international trade, equilibrium income can be realised if {S > I, T < G} and {S < I, T > G}. With the introduction of imports (Z) and exports (X), then if (S + T) > (I + G) then Z < X, and if (S + T) < (I + G) then Z > X.

In the former case, the foreign exchange reserves of the domestic economy will be increasing and in the latter case they will be falling. In either case, corrective action may be necessary. The purpose of foreign exchange reserves is to meet some temporary deficit on the balance of payments account. After some point, the continued accumulation of reserves follows no useful purpose and, indeed, serves only to hold down living standards. Similarly, a continued diminution of foreign exchange reserves will also require corrective action, if only because all holding of foreign exchange reserves are finite, and consequently a continuing deficit cannot be accommodated indefinitely. Thus, it may be necessary to pay particular attention to the relationship between Z and X, and that issue will be addressed later. For the moment, it is sufficient to emphasise the fact that imports and exports have implications for the circular flow of income and that changes in Z and X will produce changes in equilibrium income.

The relationship between changes in injections and in withdrawals, and consequent changes in national income, can be summarised as follows:

(a) A given change in household savings or taxation or imports will have exactly the same effect on national income. In each case the initial change causes national income to change in the opposite direction.
(b) A given change in business investment or government expenditure or exports will have exactly the same effect on national income. In each case the initial change causes national income to change in the same direction.
(c) The change in income, as the result of some initial change in withdrawals (S, T or Z) or some initial change in injections (I, G or X) will be some multiple of the initial change, the size of the change being given by the relevant multiplier.

We thus have to modify the multiplier concept to take account of the withdrawal for the circular flow of income due to imports and injection due to exports. The national income identity, including the international sector, becomes

$$Y \equiv C + I + G + X - Z$$

As before we shall assume previous relationships, namely

$$\begin{aligned} C &= bY_d \\ Y_d &= Y(1 - t) \\ I &= I \\ G &= G \end{aligned}$$

We shall assume X is autonomous, i.e. independent of Y, but that imports Z, are, not unexpectedly, dependent on the level of Y. The simple linear relationship we shall assume for imports is

$$Z = \eta Y$$

where η is the marginal propensity to import. Thus

$$\begin{aligned} Y &= C + I + G + X - Z \\ &= bY(1 - t) + I + G + X - \eta Y \\ \therefore Y - bY(1 - t) + \eta Y &= I + G + X \\ \therefore Y[1 - b(1 - t) + \eta] &= I + G + X \\ \therefore Y &= \frac{I}{[1 - b(1 - t) + \eta]} + \frac{G}{[1 - b(1 - t) + \eta]} + \frac{X}{[1 - b(1 - t) + \eta]} \end{aligned}$$

This formula demonstrates that the size of the multiplier is positively related to the marginal propensity to consume (b). It is negatively related to the marginal tax rate (t) and the marginal propensity to import (η), both of which are leakages from the circular flow of income. It is clear from the formula that a change in X of 100 units will have the same impact on Y as would a change in G of 100 or a change in I of 100. It is extremely complex to attempt to show the impact of the international sector diagrammatically.

If X and Z were both autonomous and equal, then the aggregate demand curve in a closed economy with a government sector would be exactly the same as one with an international sector, i.e.

$$Y = C + I + G$$

would equal

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

With X = Z, X would raise Y by the amount of exports, Z would lower Y by the same amount. Similarly, if X > Z, Y would be raised by the difference, and if Z > X, Y would be lowered by the difference. If $Z = \eta Y$, the slope of the aggregate demand curve would be less steeply inclined, indicating a smaller multiplier effect due to the import leakage.

International trade causes the national income of any trading nation to be linked through exports and imports to the national income of other countries. The major factor affecting the level and rate of growth of a nation's exports is likely to be the level and rate of growth of real incomes in the rest of the world. If, in the rest of the world, real incomes are high and increasing rapidly, then any nation's exports are likely to be high and increasing, other things being equal. Conversely, the level of imports of a nation is likely to be closely related to the level and rate of change of that nation's national income. In short, a change in the level of national income in one economy will, through international trade, have consequences for other economies, and these consequences will be greater, the greater the level of national income in the economy considered and the higher the proportion of national income that enters into international trade.

In the past four decades, most economies have become more 'open', i.e. exports and imports have accounted for an increasing share of national income. Part of the reason has been a general reduction in the barriers to international trade, especially tariffs and quotas. Figure 15.12 shows exports as a percentage of GNP for a selection of economies.

Figure 15.12 and Figure 15.13 show that exports as a proportion of GNP for the USA and Japan are lower than for all other major capitalist economies. The sheer size of the US and Japanese economies make their international sectors play a very important part in the economics of all nations with which they trade. In contrast to the USA, many small economies such as Belgium, Ireland, the Netherlands and the Asian Tigers are highly dependent on international trade. In the late 1990s China entered the international arena as one of the world's major trading nations. In 2003, China replaced Japan as the leading exporter to the US. Economic policy makers must take the trade effect into account when conducting economic policy. To this topic we shall return in a later module.

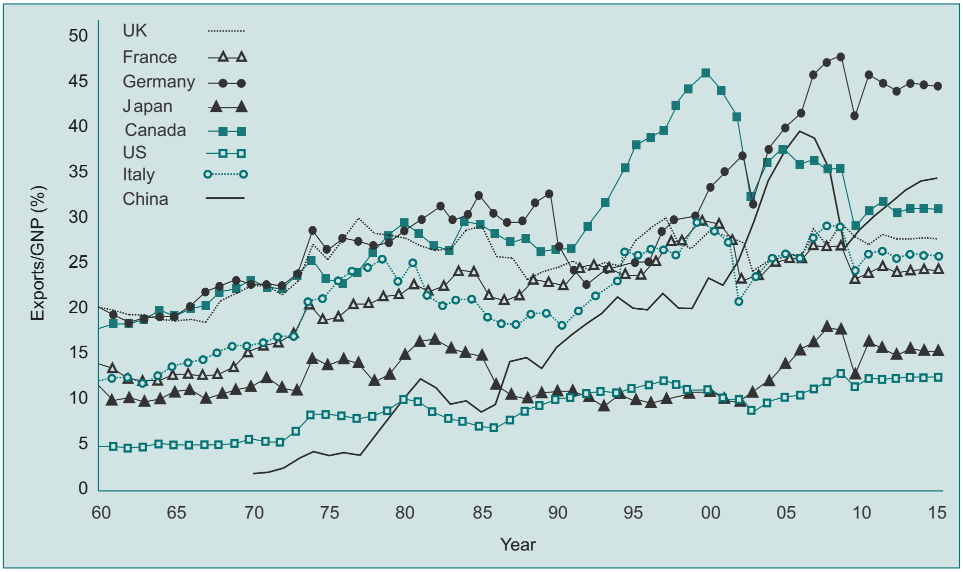


Figure 15.12 Exports as a percentage of GNP for the G7 economies and China, 1960–2015

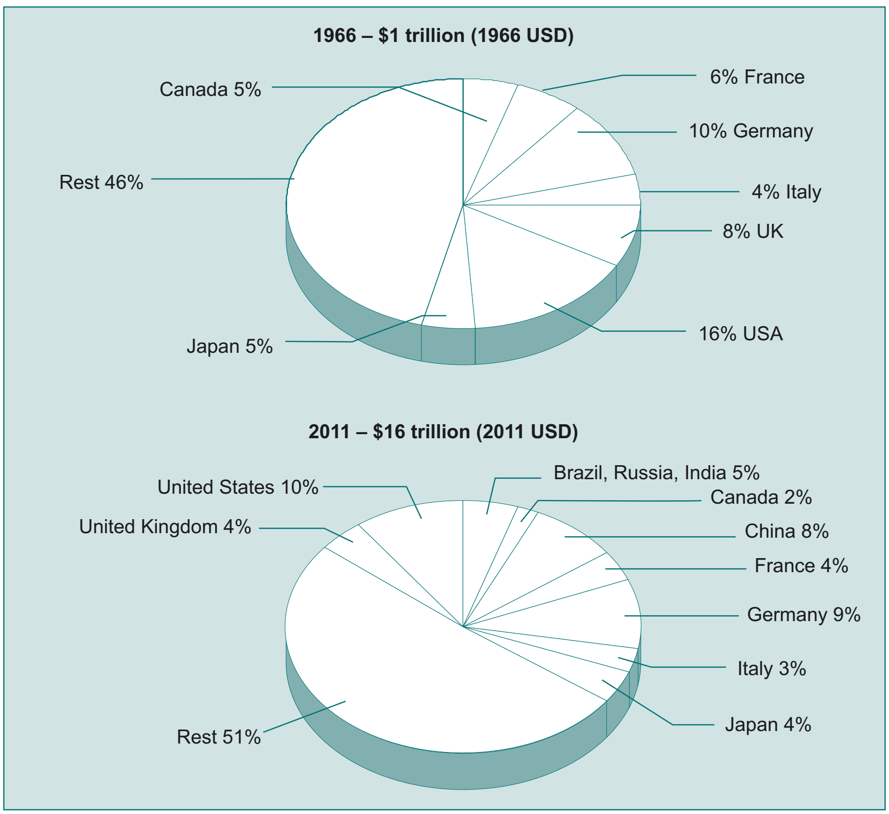


Figure 15.13 Market shares in world trade, 1966 and 2011

15.6 Business Savings

We started off our analysis by assuming that all factors of production are owned by households and all factor incomes are received by households as personal income. Clearly, this is not an accurate description of developed economies. Whereas all factors of production are owned by households, this ownership is often indirect. In consequence, some of the income that is generated in the process of production is kept by firms in the form of *business savings* or *undistributed profits* and not distributed to households, although each firm is owned collectively, often by thousands of households. Through the business decision-making process, the households that own businesses may choose to keep part of their income in their businesses to spend collectively.

The gross profit of businesses is divided into three parts:

(a) capital consumption allowances (depreciation);
(b) undistributed profits; and
(c) dividends.

Gross profits less capital consumption allowances equal net profits. Capital consumption allowances plus undistributed profits make up earnings retained by business. Dividends are an element of disposable personal income.

Undistributed profits are equivalent in their effect on the circular flow of income to household savings, representing a withdrawal from the circular flow of income that, if not offset by some planned injection, will result in a decline in national income.

Having introduced the government sector, international trade and undistributed profits, we are now able to build a complete model of the circular flow of income that approximates to the situation we observe in a developed economy. Part of gross national income is paid to the government in the form of taxes. Total tax revenue is composed of households' taxes (primarily personal income taxes), direct business taxes (primarily corporate profits taxes), and indirect business taxes (primarily sales taxes). Some of the income received by the government is passed on to households in the form of transfer payments (including welfare payments). The difference between tax revenue and transfer payments is *net taxes*. Gross national income consists of disposable personal income, business retained earnings, and net taxes. Thus

(a) disposable personal income = wages, salaries, rents, interest + dividends − taxes + transfer payments;
(b) business retained earnings = net profit + capital consumption allowances − direct business taxes − dividends; and
(c) net taxes = indirect business taxes + direct business taxes + personal taxes − transfer payments.

Put another way, the disposable income of households equals gross national income, less business retained earnings and net taxes. This distribution of spendable income among households, businesses and government can be seen by studying the income half of the circular flow diagram, as shown in Figure 15.14. Figure 15.15 shows the breakdown of total expenditure, the output side, added to the total income flow for Figure 15.14.

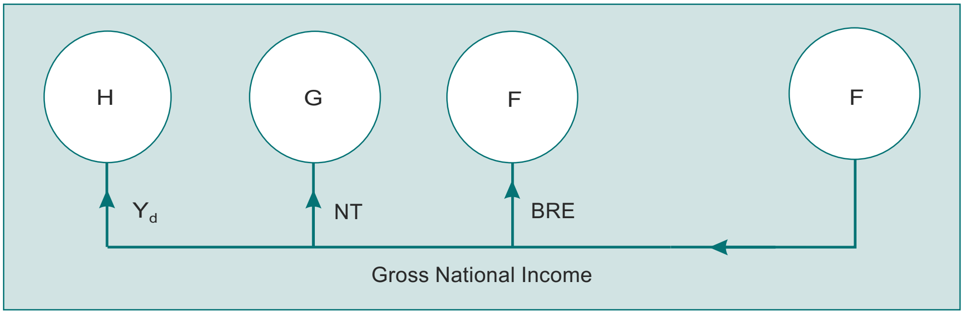


Figure 15.14 The flow of income

Notation: BRE = Business Retained Earnings; NT = Net Taxes.

Figure 15.15 summarises the circular flow of income and expenditure. Income accrues to business in retained earnings, to government in taxes, and to households in personal disposable income. In turn, firms undertake expenditure on investment (I), which is enhanced from private savings and business retained earnings or by borrowing from the capital market; government expenditure is G; and household expenditure is C. Some part of these expenditures is on goods and services from foreign producers (imports, represented by Z) and foreigners purchase goods and services from domestic producers (exports, represented by X).

The sum of expenditures (GNE) = Consumption + Investment + Government Expenditure + (Exports − Imports)

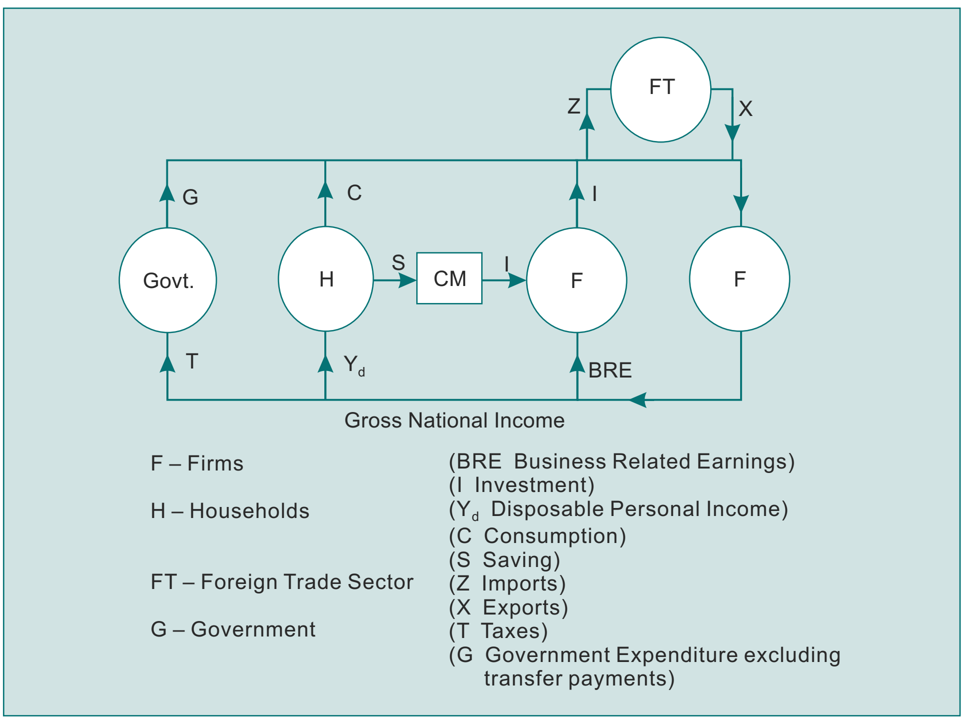


Figure 15.15 The circular flow of income and expenditure: the complete model

The following data for a hypothetical economy show the expenditure income equality.

Expenditure		Income	
Consumption	75	Disposable personal income	100
Investment	25	Net taxes	15
Government	27	Business retained earnings	10
Exports − Imports	−2		
GNE	125	GNI	125

Thus, the flow of factor incomes to the various decision-making units in the economy – businesses, government and households – and the flow of expenditures – consumption, investment, government expenditure and exports less imports – can be summarised within the single figure of the circular flow of income (Figure 15.15).

Learning Summary

You are now able to distinguish between the simple model of income determination and the expanded, more realistic, model. The concepts understood include the marginal efficiency of investment, the cost of financing investment, expected rates of return on investment over time, uncertainty and risk, and the accelerator principle.

You understand the importance of the government sector as it affects the circular flow of income and how the government, by altering various taxes, transfers and expenditures, can affect output and unemployment. You are familiar with in-built stabilisers and their impact on the economy and know how to calculate the government budget.

You can analyse the impact of imports, exports and business savings on the circular flow of income, and are thoroughly familiar with the income–expenditure equality.

Review Questions

Multiple Choice Questions

Questions 15.1–15.3 are based on the following data for a hypothetical economy:

$$\begin{aligned} C &= 200 + 0.8(Y - T) \\ I &= 125 \\ T &= 25 \\ G &= 250 \end{aligned}$$

where

C = consumption expenditure
I = investment expenditure
G = government expenditure
Y = national income
T = taxes

15.1 Which of the following is correct?

A. Y = C + I + G
B. Y = C + I + G − T
C. Y = C + I + G + T
D. Y = (C + I + G + T) / 0.8

15.2 Which of the following is the equilibrium level of national income?

A. 2400.
B. 2750.
C. 2775.
D. 2875.

15.3 Which of the following is true of the given simple economy?

A. The chief cause of unemployment is insufficient spending.
B. The equilibrium level of national income coincides with the full-employment level.
C. Inflation and employment are negatively related.
D. Tax cuts are more powerful than government spending in increasing employment.

15.4 'In public discussions about the current unemployment, policy makers have failed to appreciate the "multiplier effect", namely the theory that $1000 million spending by government ultimately creates $2500 to $3000 million increase in GNP. But if this theory is true, it must work in reverse. The net effect of a $1000 million cut in government spending today would be equivalent to a $2500 million to $3000 million increases in taxes.'

Is the economic analysis in this quotation correct or incorrect, and why?

A. Correct, because government spending has a multiplier effect on consumption.
B. Incorrect, because an increase in taxes also has a multiplier effect.
C. Correct, because the value of the marginal propensity to consume determines the size of the multiplier.
D. Incorrect, because a $1000 million cut in government expenditure would decrease national income more than a $2500 to $3000 million increase in taxes.

15.5 Which of the following is the basis of the accelerator principle?

A. Income is a function of investment.
B. Consumption is a function of income.
C. Saving is a function of income.
D. Change in net investment is a function of the rate of change of national income.

Questions 15.6–15.8 are based on Figure 15.16.

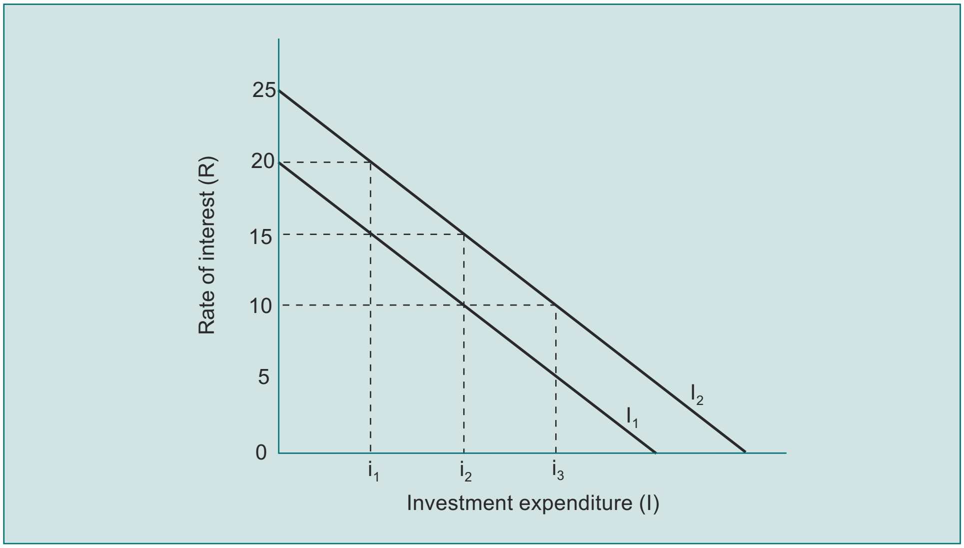


Figure 15.16

Notation: i_1 is the investment level for period 1; i_2 is the investment level for period 2.

15.6 Which of the following could have caused the shift in the investment level from I_1 to I_2?

A. A decrease in the rate of interest.
B. An increase in the business tax rate.
C. An increase in the rate of interest.
D. Labour-saving innovations.

15.7 In period 1 investment was $0i_1$ and it remained at that level in period 2, despite the shift in the investment function. Which of the following could account for this?

A. The demand for replacement investment increased.
B. The rate of interest increased from 15 per cent to 20 per cent.
C. Firms delayed investment by the amount of i_1i_2.
D. Imports of investment goods increased.

15.8 The rate of interest was 15 per cent in period 1 and 10 per cent in period 2. Which of the following represents the change in investment between the two periods?

A. From $0i_1$ to $0i_3$.
B. From $0i_2$ to $0i_3$.
C. From $0i_3$ to $0i_1$.
D. From $0i_2$ to $0i_1$.

15.9 Which of the following will tend to cause an investment expenditure curve to shift downwards?

A. A new technological advance that cuts the price of primary steel by one half.
B. A significant decline in the rate of interest.
C. A decline in wage rates.
D. An increase in corporation tax.

15.10 The assumptions made in constructing the simple accelerator model are

I. businessmen ignore expectations in making investment decisions.
II. businessmen never allow their capital stock to become excessive or deficient in meeting their current demand for output.
III. small adjustments in capital stock are made by firms in facing minor output fluctuations.

Which of the following is correct?

A. I only.
B. II and III only.
C. I and II and III.
D. Not I nor II nor III.

15.11 There is considerable unemployment in the economy. The government is proposing to finance a $200 million increase in expenditure for goods and services with a $200 million increase in income taxes. If interest rates were kept unchanged and if consumers always spend 90 per cent of their after-tax income, such a scheme would do which of the following:

A. raise national income.
B. leave national income unchanged.
C. lower national income by $280 million.
D. lower national income by $180 million.

15.12 A government wishes to balance its budget, achieve full employment and distribute Y among C, I and G in certain proportions. The Chief Economic Adviser has advised that at least one of the policy objectives should be dropped. Upon which of the following is his professional advice based?

A. The three objectives have different levels of importance.
B. Only by remote chance will values of Y, T, C, I and G occur that will achieve the three objectives.
C. Balanced budgets are incompatible with full employment.
D. It is not possible to influence the distribution of GNP between C, I and G.

15.13 If the government were to adopt a policy that led to households saving more of their income, which of the following effects would be felt on national income in the short run? National income would

A. increase because saving equals investment and an increase in investment expenditure would lead to an increase in national income.
B. decrease because an increase in the willingness of households to save is not automatically matched by an increase in the willingness of businesses to invest.
C. increase because households will save more only if their incomes increase.
D. decrease because additional saving would force up interest rates which in turn would decrease investment expenditure.

15.14 Which of the following explains why a progressive tax acts as an 'in-built stabiliser'?

A. When GNP changes, the government changes tax rates.
B. When GNP changes, tax rates change automatically.
C. Tax collections increase as government spending decreases.
D. When GNP changes, tax receipts change automatically in the same direction without any change in tax rates.

15.15 Which of the following is correct? The function of an 'in-built stabiliser' is to

A. ensure that the same amount of tax is collected each year.
B. balance the budget each year.
C. reduce fluctuations in economic activity.
D. ensure that government expenditure is the same each year.

15.16 Which of the following is correct? The importance of the international sector for the level of national income is that

A. an excess of exports over imports results in a leakage from the circular flow.
B. unless exports are greater than imports, the national income will continuously decline.
C. the excess of exports over imports determines the rate of growth of the economy.
D. imports are leakages from and exports injections to the circular flow.

15.17 In order to calculate the value of the government expenditure multiplier, it is essential to know

I. the marginal propensity to import.
II. government expenditure.
III. the marginal tax rate.

Which of the following is correct?

A. I and II only.
B. I and III only.
C. II and III only.
D. II only.

15.18 Consider the following statements.

I. An increase in import expenditure reduces GNP.
II. A reduction in export expenditure increases GNP.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

15.19 In an economy it was found that the marginal propensity to consume is 0.8. The government economic adviser therefore calculated the multiplier to be 5 (i.e. $1/(1 - 0.8)$). An increase in exports of $10 billion took place and it was found that national income increased by $20 billion. National income did not increase by the full value of the multiplier times the increase in exports, i.e. by $50 billion, because

I. the marginal propensity to import may have been greater than zero.

II. the marginal propensity to save was not included in the information, and may have been greater than zero.

III. imports and exports became equal after national income was increased by $20 billion.

Which of the following is correct?

A. I only.

B. I and II only.

C. I and III only.

D. II and III only.

15.20 An economy was experiencing a decreasing GNP, and many people blamed the increased competition of imports and the increased willingness of consumers to buy imported goods. However, when the national accounts for the year were published, it was found that both GNP and imports had fallen. Which of the following could account for this?

A. Exports had increased.

B. Investment demand had increased.

C. Fewer imports were bought in total because of a reduction in aggregate demand.

D. The marginal propensity to consume had increased.

Module 16

Fiscal Policy

Contents

16.1 Introduction

So far we have demonstrated that the limit to national income in the short run is determined by the production possibilities frontier, which in turn is determined by the available supply of factors of production and technical knowledge. Changes in the supply of factor inputs and technological know-how take place slowly through time. Given this fixed frontier in the short run, we developed a theory of national income/output determination and showed that national income is determined by the level of aggregate demand. We further showed that the level of national income could differ from capacity or potential income.

In the real world, in all major capitalist economies, potential output (Q) tends to grow steadily over time, reflecting an increase in the quality and quantity of the capital stock due to positive net investment and technological change. Technological change makes its presence known by being embodied in new capital goods. The growth in Q is also caused by increases in the quality and quantity of the labour force, the quality being enhanced by education and training. Today for the major capitalist economies the growth rate in Q varies between 2 and 4 per cent. Actual output (Y) fluctuates wildly compared with Q.

Figure 16.1 shows the annual percentage change in GNP for the 'Big Seven' in the past 55 years. It can be seen that, as the years passed, the dispersion in GNP growth rates among the 'Big Seven' decreased; in the 1970s these growth rates started to rise and fall together much more than they had in the 1960s. Many of the countries experienced negative GNP growth in the mid-1970s, early 1980s, early 1990s and late 2000s.

Given a steady increase in Q and significant fluctuations in Y, it follows that the gap between Q and Y for these countries has not been constant over time. There have been periods – characterised by persistent unemployment of the factors of production – when $Y < Q$, other periods of relatively full employment of factors when $Y = Q$, and yet other periods of overfull employment with economies operating below the full-employment unemployment rate and prices rising substantially.

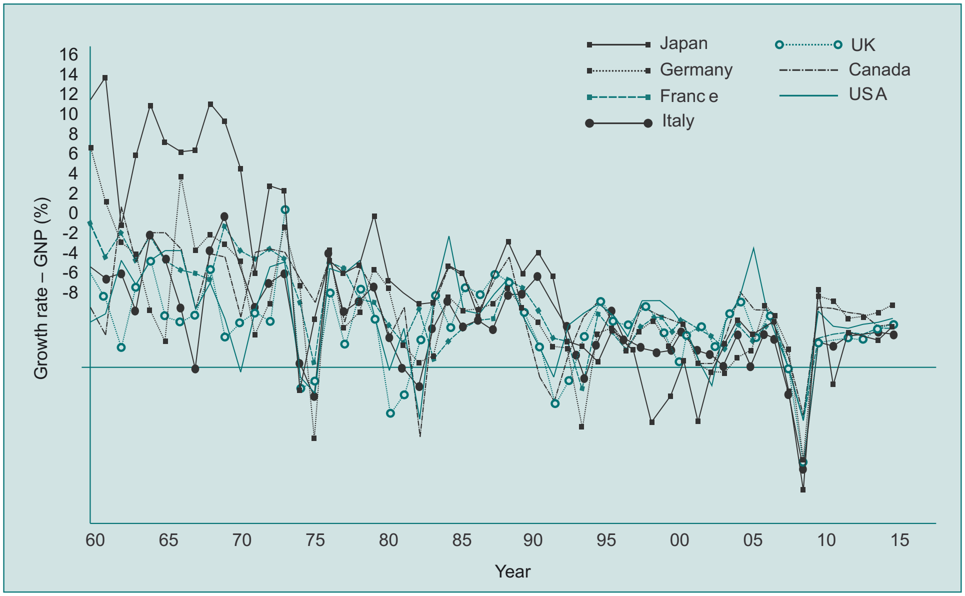


Figure 16.1 Annual growth rates as a percentage of real GNP for the G7 economies, 1960–2015

Figure 16.2 Unemployment rates for the G7 economies, 1960–2015

Figure 16.2 shows how unemployment rates have varied over the same period. Not surprisingly, increases in unemployment rates can be observed occurring when GNP growth rates approach zero or become negative.

This module is mainly a revision or 'pulling together' unit so as to provide a general explanation of the real-world data described above and to ensure you understand the interrelationship between aggregate demand and potential output and the manner in which they determine whether an economy operates below, at, or above, full employment.

16.2 Deflationary and Inflationary Gaps

Given our assumption that capacity output is fixed in the short run by the supply of factors of production and by technical knowledge, 'real' or physical factors set the limit to output that is considered to be constant in the short run. In other words, aggregate supply cannot be increased beyond full employment or capacity national income, which is denoted by Y_f in Figure 16.3.

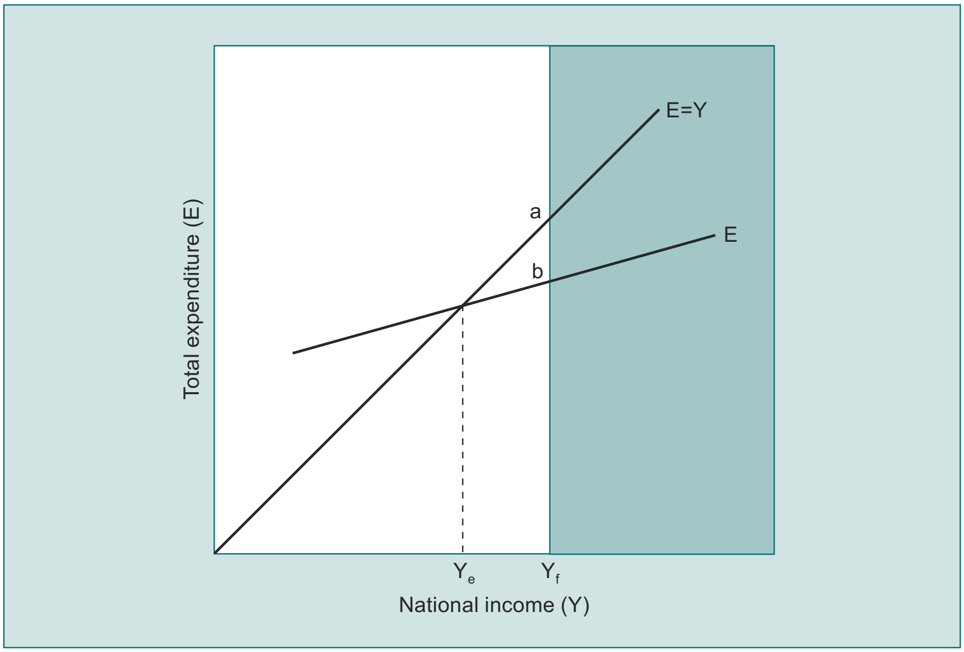


Figure 16.3 The deflationary gap

Given that Y_f sets the limit to actual output, the level of national income realised depends upon the level of aggregated demand, which consists of the following expenditures:

1. household consumption (C);
2. business investment (I);
3. government expenditure (G);
4. exports (X) less imports (Z), i.e. $X - Z$

Thus aggregate demand equals $C + I + G + (X - Z)$.

Remember, imports are deducted from aggregate demand because part of consumption, investment and government expenditures takes the form of purchases of imports, rather than home-produced goods, and these purchases of imports create factor incomes in other economies and not in the domestic economy. To allow for this leakage of expenditure from the circular flow of income within the domestic economy, the value of imports must be deducted to obtain those expenditures that accrue as factor incomes to households *within* the domestic economy.

Aggregate demand consists of all those expenditures that make a claim on the output of the domestic economy and that therefore create employment and income for domestic factors of production. The higher the level of aggregate demand, the higher will be the level of income and employment, and vice versa. Increases in aggregate demand can give rise to an increase in the output of goods and services only as long as there are unemployed factors of production. When capacity output is reached, then the limit to any further increase in output is set by the factors that determine the rate of growth of capacity output over time.

If aggregate demand is just sufficient to maintain capacity output, then full employment and capacity income would be realised. The model we have been developing suggests that this is one possible solution for short-run equilibrium income, but it is not the only solution that is possible. Indeed, the level of aggregate demand realised might not be sufficient to reach full-employment income, so that unemployment will result; or, in the opposite case, the level of aggregate demand may be greater than the level necessary to obtain full-employment income and excess demand will result in price inflation.

These outcomes are represented diagrammatically in Figure 16.3, Figure 16.4 and Figure 16.5. The figures are based on similar principles to those underlying the figures we have previously constructed, although in the present case expenditure is measured along the vertical axis. It follows that the 45° construction line is the locus of points for which expenditure (E) equals income (Y).

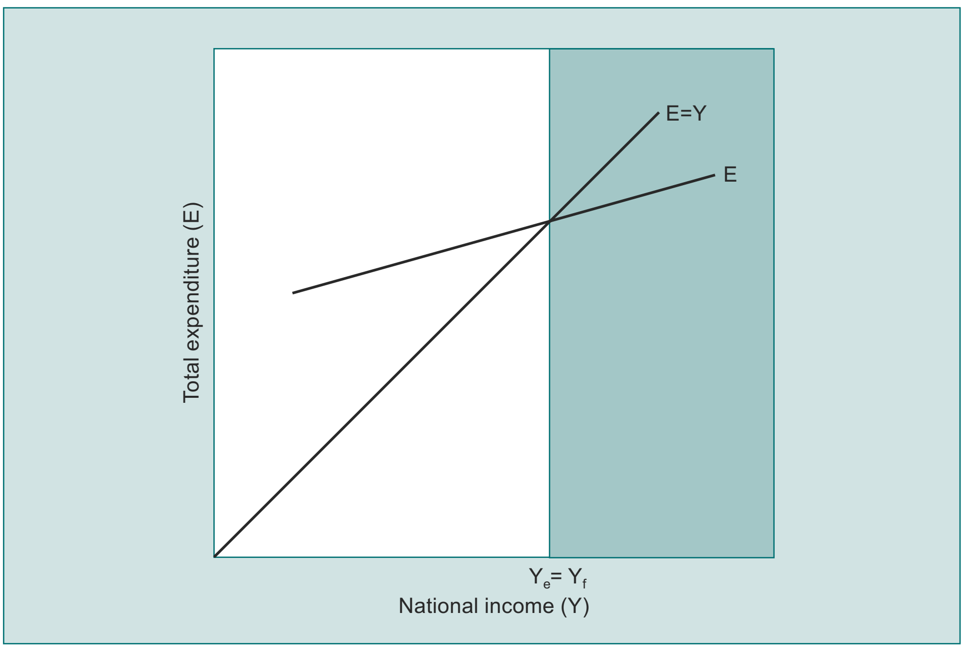


Figure 16.4 Full-employment equilibrium

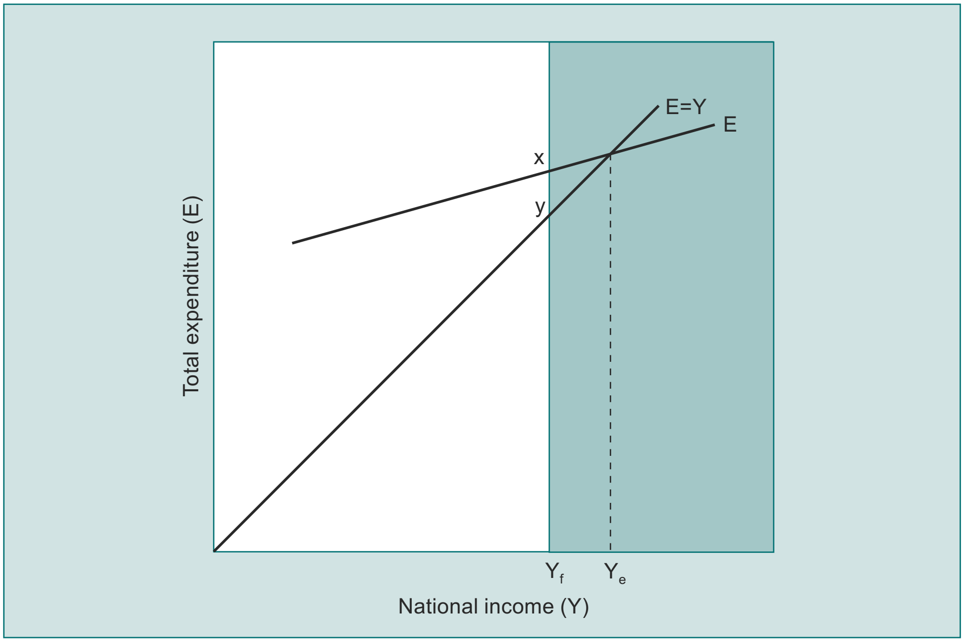


Figure 16.5 The inflationary gap

In Figure 16.5 aggregate expenditure is given by $E = C + I + G + (X - Z)$ and is measured along the vertical axis, with national income (Y) along the horizontal axis. In each figure full-employment income is given by Y_f. The shaded area to the right of Y_f represents levels of income achieved after full employment has been realised.

In Figure 16.3 the aggregate expenditure function (E) cuts the 45° construction line to the left of Y_f, so that equilibrium income (Y_e) is less than Y_f. Thus, equilibrium income is a level at which there is unemployment of factors of production. Y_e is equilibrium income because only at this level of income are planned withdrawals equal to planned investment so that $S + T + Z = I + G + X$. Looking at this rather differently, Y_f cannot be equilibrium income because at Y_f the level of aggregate demand would not be sufficient to maintain full-employment income (i.e. planned withdrawals would be greater than planned injections). There would be insufficient aggregate demand to maintain full-employment income, so that the level of income would fall from Y_f until Y_e was reached. This insufficiency of aggregate demand compared with the level of aggregate demand necessary to obtain and sustain Y_f is known as the *deflationary gap* and is measured in Figure 16.3 by ab, i.e. the vertical distance between actual level of aggregate demand that would result at full-employment income and the level of aggregate demand necessary to sustain full-employment income.

In Figure 16.5 the aggregate expenditure function cuts the 45° construction line to the right of Y_f, so that $Y_e > Y_f$, and there is no unemployment due to deficient aggregate demand. Planned withdrawals and planned injections are brought into equilibrium at a level of national income greater than Y_f. This immediately raises a problem. Y_f represents capacity output (income), namely the physical limit to the volume of goods and services that can be produced by the economy. The points to the right of Y_f cannot represent higher levels of income because, by definition, Y_f sets the limit to output. Hence, Y_e cannot represent a higher volume of income than Y_f. How then can equilibrium national income be greater than Y_f?

National income consists of a flow of final goods and services, the various outputs being represented by $q_1, q_2, q_3 \ldots q_n$. To add these dissimilar units of output together, it is necessary to apply a common denominator, or measuring rod. In a monetary economy the measuring rod is money, so that there is a series of price weights $p_1, p_2, p_3 \ldots p_n$ corresponding to the quantity weights. Hence national income in current prices (that is, the prices ruling in the current time period) is given by $Y = \sum (p_1q_1 + p_2q_2 + p_3q_3 + \ldots + p_nq_n)$.

Over time, national income measured in current prices can change for two reasons: a change in the output weights representing a change in the volume of goods and services produced; and/or a change in the price weights, which does not represent a change in the volume of goods and services but a change in the value of the measuring rod, money. Clearly, if national income changes because of changes in the output weights, then the volume of goods and services available for consumption and investment purposes has changed, i.e. 'real' national income has changed. Alternatively, if national income has changed because the price weights have changed, with prices having risen or fallen, then the volume of goods and services has not changed: real national income has remained constant and the change in money national income is due to the change in the price weights.

In Figure 16.5, Y_f sets the limit to real national income, so that all points to the right of Y_f represent increases in national income due to prices and *not* to quantities. Thus, where Y_e is greater than Y_f, the result must be price inflation, as real national income cannot be increased beyond Y_f.

In Figure 16.5, Y_f cannot be equilibrium income because at that income the level of aggregate demand would be greater than the level necessary to maintain full employment (i.e. planned injections would be greater than planned withdrawals). There would be excess demand, so that the level of national income would rise from Y_f (as prices increased) until Y_e was reached. The excess aggregate demand over and above the level of aggregate demand necessary to obtain and sustain Y_f is known as the *inflationary gap* and is measured in Figure 16.5 by xy, i.e. the vertical distance between the actual level of aggregate demand that would result at full-employment income and the level of aggregate demand necessary to sustain full-employment income.

Finally, Figure 16.4 represents the situation where there is neither a deflationary gap nor an inflationary gap. The level of aggregate demand is just sufficient to obtain and sustain equilibrium income equivalent to full-employment income, so that neither unemployment nor price inflation results.

16.3 Real and Money GNP: Index Numbers

In order to measure national income, there must be a common measure of value to apply to a wide range of goods and services. This common measure of value, or measuring rod, is (as described in Section 16.2) provided by money. The problem with money as a measuring rod is that its value, or purchasing power, is not constant but changes over time.

At a given point in time a measure of national income can be obtained by valuing the output of different goods and services by applying the prices ruling in that period. National income can therefore be measured at current prices and is given by

$$Y_t = \sum(a_tP_{at} + b_tP_{bt} + \cdots + n_tP_{nt})$$

$$Y_{t+1} = \sum(a_{t+1}P_{at+1} + b_{t+1}P_{bt+1} + \cdots + n_{t+1}P_{nt+1})$$

Over time, national income as measured in current prices may change because the quantity weights change and/or because the price weights change. It is desirable to distinguish between these different sources of change because material standards of living are affected by quantity changes but not by price changes. For example, if national income, measured in current prices, doubles because all quantity weights double, then there is an increased flow of goods and services available to satisfy material wants. If, on the other hand, national income doubles because all price weights double, then there is no change in the flow of goods and services available to satisfy material wants. In the former case, real and money incomes have doubled; in the latter case, money income has doubled while real income has not changed.

Money income is simply income measured by the price weights currently ruling. *Real income* is measured by adjusting money income to allow for changes over time in the value of the 'measuring rod' of money. Thus, changes in money income include changes in price and quantity weights; changes in real income reflect changes only in quantity weights.

To calculate real national income, it is necessary to apply a fixed set of price weights, so that any changes in national income, as measured, reflect only changes in the quantity weights. For example, to measure changes in the national income in the 1990s, the price weights ruling in 1990 could be applied to value the quantity of goods and services produced in each successive year. National income would then be measured in terms of constant 1990 prices. Changes in income so measured would reflect changes in quantity weights, i.e. changes in real income. Hence, national income in real, or constant, 1990 prices is given by

$$Y_{90} = \sum(a_{90}P_a + b_{90}P_b + \cdots + n_{90}P_n)$$
$$Y_{91} = \sum(a_{91}P_a + b_{91}P_b + \cdots + n_{91}P_n)$$
$$Y_{98} = \sum(a_{98}P_a + b_{98}P_b + \cdots + n_{98}P_n)$$

In practice, real national income is usually calculated by applying an index number of prices. An index number is an attempt to reflect a complex set of price changes in one representative measurement. It necessarily involves some degree of error.

Price changes are seldom uniform. As the general price level rises or falls, the price changes for particular goods and services will show substantial variation. Even during a period in which most prices are rising, it is usually possible to find some prices that are falling and vice versa. A price index is a representation of price changes, an attempt to measure the average or representative price change. This is obtained by selecting a *typical basket of goods and services*, observing the price changes of these goods and services and weighting these price changes to allow for the differing economic importance of the different goods and services. It is possible to obtain an index number that covers a wide or a narrow range of goods and services. The price index used to obtain real GNP is known as the *GNP deflator*, and there are other price indexes to measure changes in retail prices, wholesale prices, capital goods prices, etc. The function of all these price indexes is to measure changes in the measuring rod, money, in order to establish the real changes that have occurred in the economy, in other words the changes that have involved a change in the quantity of goods and services produced.

16.4 Fiscal Policy

The theory of income determination developed so far shows that equilibrium national income can diverge from full-employment income, at least in the short run. There is nothing necessarily desirable or undesirable about the level of national income that results, as it can produce unemployment or price inflation, as well as full employment without price inflation. The theory is consistent with the historical evidence, in that it predicts that the level of national income realised may differ from full-employment income, with resulting unemployment or inflation. An important question then arises: is the system self-regulating in that the departures from full-employment income will be temporary, with a tendency for the system to return to full employment, or will the system tend to settle down with equilibrium income substantially different from full-employment income, with continuing unemployment or inflation?

Our model developed so far has no mechanism for ensuring that in the long run there is a tendency for Y_e to equal Y_f. On the contrary, it suggests that planned injections are only brought into equality by changes in the level of national income, so that any level of national income is possible. Keynes considered that a private-sector economy might have a tendency, failing 'outside' intervention, to settle down at a level of income substantially below capacity output, with consequent unemployment. This view reflected the experience of the late 1920s and 1930s, when sustained unemployment was experienced in the UK. Hence, the *General Theory*, published in 1936, was primarily concerned with the possibility of a deflationary gap emerging, such as depicted in Figure 16.3, and with prescribing policy measures that would counteract such a possibility. Yet the basic analysis of the *General Theory* can be applied to analyse inflationary situations such as depicted in Figure 16.5, as well as deflationary ones, in that both are seen to be determined by the relationship between aggregate demand and aggregate supply.

The essence of Keynesian analysis is for the government to intervene to ensure that a position such as in Figure 16.4 is realised, rather than a situation such as in Figure 16.3 or Figure 16.5, i.e. full employment is realised rather than unemployment or inflation. This can be achieved by letting the government 'interfere' with the circular flow of income. The government can influence the circular flow of income through taxes (T), a withdrawal from the circular flow of income, or through government expenditure (G), an injection into the circular flow. A budget deficit ($T < G$) would have an expansionary effect on the economy; a budget surplus ($T > G$) would have a deflationary effect on the economy.

A budget deficit ($T < G$) would be appropriate where the economy, left to itself, would tend to settle down in a position such as that shown in Figure 16.3, whereas a budget surplus ($T > G$) would be appropriate for an economy represented by Figure 16.5. In a situation such as that represented by Figure 16.4, which is presumably desirable, the government should neither stimulate nor deflate the economy. This proposed operation of fiscal policy is known as *functional finance*, meaning that there is no single automatic rule that should be followed regarding the relationship between government expenditure and government taxation. Instead, fiscal policy should be discretionary. The relationship between G and T should deliberately be varied to reflect the underlying conditions in the economy, the goal of fiscal policy being to reach the position shown in Figure 16.4 rather than the positions shown in Figure 16.3 and Figure 16.5.

You will not be surprised to learn that not all economists agree with the policy of functional finance. Indeed, Oscar Wilde cynically observed that 'if you laid all the economists in the world end to end, they would not reach a conclusion'. It is a fact that actual and potential output can depart from each other in the short run, but it does not necessarily follow that functional finance is required to close the gap. Would not the economy tend to move towards full employment without inflation in the absence of government intervention? This was the view of the neo-classical economists of the 1870s, who find a modern echo amongst the monetarists. The neo-classical view was that, given flexible prices, flexible interest rates and flexible money wages, there would be a tendency for planned savings and planned investment to be brought into equilibrium at, or close to, full-employment income. Departures from full-employment income did occur, often because of political instability or monetary disorders, but it was believed that these would be temporary and that, so long as prices were flexible, there would always be a tendency to revert towards a situation of full employment without inflation. Given this view, the prescription for fiscal policy was that the government should pursue a balanced budget, thus minimising its effects on the circular flow of income.

We can distinguish these different schools of thought by presenting the implications of the underlying analysis in the extreme, or limiting, case. This extreme, or naive, representation is an oversimplified description of rather complex arguments, but it will help to demonstrate the essential and important differences.

The naive Keynesian view is that the components of aggregate demand are independent of each other so that, for example, an increase in government expenditure (G) does not have any adverse effect on any other component of aggregate demand, namely C, I or X. In this case, the creation of a budget deficit by raising G compared with T will increase aggregate demand by the full extent of the deficit thus created. Conversely, the creation of a budget surplus by lowering G relative to T will lower aggregate demand by the full amount of the surplus thus created. Clearly, if this were so, changes in the budget deficit (surplus) will represent a net injection into (withdrawal from) the circular flow of income and would have a substantial impact on the level of aggregate demand, and therefore on national income and employment.

The naive monetarist view is that the components of aggregate demand are interdependent, so that changes in the budget deficit (surplus) are offset by equivalent changes, of opposite sign, in other components of aggregate demand. If this were so, the creation of a budget deficit by raising G relative to T would not raise aggregate demand in that the higher government expenditure would be at the expense of lower private expenditure. This is known as *the crowding-out effect*.

In short, the naive Keynesian view is that the crowding-out effect is zero, while the naive monetarist view is that it is unity. Between lie intermediate cases, where the crowding-out effect lies somewhere between these two extreme values. At full employment, of course, the crowding-out effect must become unity. Given that output cannot be increased any further, a rise in government expenditure must be at the expense of a fall in private expenditure. It follows from this that the closer the economy is to full employment, the greater the crowding-out effect is likely to be, so that the naive Keynesian position can only hold where there is substantial unemployment. It should be recalled, however, that Keynes was writing during the Great Depression, when unemployment rates in capitalist countries reached more than 20 per cent.

This theme will be taken up when we develop the monetarist analysis at greater length. For the present we shall develop the implications of the naive Keynesian view in the interests of simplification, but we do so on the understanding that it assumes a zero crowding-out effect, which is clearly unrealistic. Suppose that in the absence of discretionary fiscal policy a situation such as that depicted in Figure 16.3 will result, namely a deflationary gap. The deflationary gap measures the shortfall of aggregate demand over the level necessary to obtain and sustain full-employment income, and in order to move towards the position represented by Figure 16.4 it is necessary to raise the level of aggregate demand. This is done by raising government expenditure and/or lowering taxes, thus creating a budget deficit. It should be noted, however, that the size of the deficit necessary to create full employment will depend upon the size of the multiplier and the balance between expenditure increases and tax cuts.

Generally speaking, the greater the proportion of the budget deficit created by expenditure increases, the lower the deficit necessary to obtain full-employment income, and vice versa. Although this may sound complicated, the process is simple enough to understand. An increase in government expenditure of $1 million results in an increase of $1 million in aggregate demand, less the government's propensity to import goods and services. The effect of a tax reduction of $1 million on aggregate demand will depend on the public's marginal propensity to save plus its marginal propensity to import. If the sum of these two marginal propensities is greater than the government's marginal propensity to import, which appears likely, a $1 million increase in government expenditure will raise aggregate demand more than a $1 million decrease in taxation revenue.

It follows from this that the balanced budget does not have a neutral effect on the circular flow of income, as $1 of government expenditure is more 'high powered' than $1 of taxation. Thus, listing fiscal policies in descending orders of expansionary effects, we obtain:

(a) a budget deficit created by raising G relative to T;
(b) a budget deficit created by reducing T relative to G;
(c) increasing both G and T by the same amount.

To obtain a given increase in aggregate demand will require a smaller budget deficit if the deficit is created by raising G, and a larger budget deficit if the deficit is created by lowering T. Alternatively, the same increase in aggregate demand could be obtained by raising both G and T, but where a substantial stimulus to the economy is required, it would necessitate a very large increase in the size of the government sector to produce the same effect as a budget deficit created by (a) or (b) above.

Given the situation depicted in Figure 16.5, i.e. excess demand, the appropriate fiscal policy would be the reverse of that described above. An inflationary gap measures the excess of aggregate demand over the level necessary to maintain full-employment income. To move towards the desired position shown in Figure 16.4, it is necessary to reduce the level of aggregate demand. This can be done by lowering government expenditure and/or raising taxes and thus creating a budget surplus. Again, the size of the surplus necessary to reach full employment without inflation will depend on the multiplier and on the manner in which the surplus is created. Generally speaking, the greater the proportion of the budget surplus created by expenditure reductions, the lower the budget surplus necessary to obtain full-employment income, and vice versa.

Up to this point we have captured the basic elements of Keynesian economics, still much-loved by many politicians. We have paid only fleeting attention to prices and ignored money completely. That situation must now be rectified through the ensuing modules.

Learning Summary

This module summarises the contents of Modules 11–15. It is important at this stage that you can distinguish between potential and actual output, and understand that when they are not equal to each other, an inflationary or deflationary gap emerges. You can explain the implications of both types of gap and understand what the government can do to bring about equality.

You are able to distinguish between real and money GNP, and know the difference between current and constant prices.

You understand the policy implications of the 'naive' monetarist view versus the 'naive' Keynesian view.

Review Questions

Multiple Choice Questions

16.1 It is argued that government can always make national output higher by fiscal and/or monetary policy. This is not the case in practice because of which of the following?

A. The marginal propensity to consume is always less than one.
B. The labour force is not a fixed percentage of the population.
C. There is an upper limit to national output determined by the economy's resources.
D. The ratio of capital to output increases as national output increases.

16.2 Which of the following accounts for the fact that there is a limit to the level of national output in the short run? In the short run

A. the price of factors of production increases when the demand for them increases.
B. aggregate demand is relatively fixed.
C. it is not possible to substitute between the factors of production.
D. factors of production are limited in supply.

16.3 Aggregate demand in a closed economy is the sum of consumption, investment and government expenditures. The upper limit to national output is determined by

I. the stock of capital.
II. the quality and quantity of labour available.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

16.4 A treasury publication describing the performance of the UK economy over the past decade includes a graph in which, for two months, actual GNP exceeded full-employment (potential) GNP. Which of the following must be true? The graph was

A. wrong because it is not possible for actual GNP to exceed potential GNP.
B. wrong because during periods of full employment actual GNP causes increases in potential GNP.

C. correct because potential GNP is not an upper limit to output in each and every period.
D. correct because during periods of full employment actual GNP is always greater than potential GNP.

16.5 Over the past five years the rate of growth of actual GNP in an economy has been 5 per cent per annum, whereas the rate of growth of potential GNP has been only 4 per cent per annum. It can be concluded that

A. an inflationary gap exists.
B. a deflationary gap that existed five years ago is decreasing.
C. high unemployment currently exists.
D. aggregate demand currently exceeds potential GNP.

16.6 Assuming no offsetting changes in other components of aggregate demand, a deflationary gap between actual and potential output would be narrowed by

I. increased investment expenditure.
II. increased imports.
III. balanced increase in government expenditure and tax revenue.

Which of the following is correct?

A. I only.
B. II only.
C. I and II only.
D. I and III only.

16.7 Recently a government announced decreases in income tax and government expenditure. From these policy decisions it can be inferred that the government concern was

I. the deflationary gap.
II. the inflationary gap.
III. the current high unemployment rate.

Which of the following is correct?

A. I only.
B. II only.
C. I and III only.
D. Not I nor II nor III.

16.8 Which of the following policies would be appropriate for dealing with a deflationary gap?

A. Increase income taxes.
B. Increase subsidies to exporting industries.
C. Decrease unemployment compensation.
D. Decrease government expenditure.

16.9 Which of the following policies would be appropriate for dealing with an inflationary gap?

A. Decrease income taxes.
B. Increase tariffs and quotas on imports.
C. Decrease government expenditure.
D. Increase unemployment compensation.

Questions 16.10 and 16.11 are based on Figure 16.6.

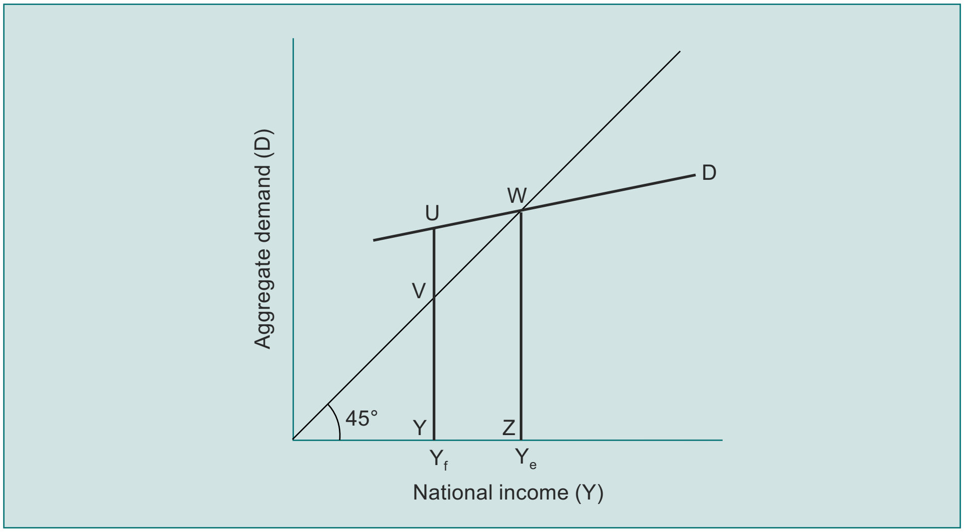


Figure 16.6 A sample aggregate demand graph

Notation: D=Aggregate demand, Y_f =Full-employment income, Y_e =Equilibrium income.

16.10 Which of the following does Figure 16.6 show?

A. a deflationary gap given by UV.
B. a deflationary gap given by Y_fY_e.
C. an inflationary gap given by UV.
D. an inflationary gap given by Y_fY_e.

16.11 At points to the right of Y_f in Figure 16.6, which of the following is correct?

A. Real national income is increasing because more factors of production are being utilised.
B. Real national income is falling because the economy is departing from full employment.
C. Money national income is rising because output is rising.
D. Money national income is rising because of rising prices.

16.12 In certain periods during the past five years, money incomes have increased in the UK but the standard of living has decreased. Which of the following can account for this? Increases in

A. prices have been less than increases in money incomes.
B. prices have been greater than increases in money incomes.
C. money incomes do not lead to increases in spending.
D. money incomes cause real GNP to decline.

16.13 The following table shows money GNP and real GNP in an economy over a three-year period.

Year	Money GNP ($m)	Real GNP ($m)
t	113	113
t + 1	115	112
t + 2	120	114

Which of the following statements must be true?

A. The columns for money and real GNP have been switched, because real GNP always increases.

B. The data are wrong because money and real GNP can never be equal.

C. Although prices increased over the three years, actual output decreased during one year.

D. Money GNP was the better measure of the performance of the economy over the period.

16.14 A given increase in government expenditure does not always lead to a greater increase in real national output because

I. there may not be a sufficient quantity of unemployed resources to produce such an increase in real national output.

II. increased incomes may be spent on imports.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

16.15 Faced with a deflationary gap, the government is considering either increasing government expenditure by $5 million or decreasing taxes by an equal amount. Assuming sufficient unemployed resources, which of the following is correct?

A. The increase in government expenditure will decrease the deflationary gap more than the decrease in taxes.

B. The decrease in tax will decrease the deflationary gap more than the increase in government expenditure.

C. Both policies will produce the same result.

D. The decrease in tax will reduce unemployment more than the increase in government expenditure, but the effect on the deflationary gap is unknown.

Module 17

Money, the Central Bank and Monetary Policy

Contents

17.1 Introduction

Up to now, we have made no reference to the role of money in the economic system. This is unsatisfactory because we live in a monetary economy in which money is the medium of exchange. The value of goods and services is reflected by their *prices* but the general price level is seldom stable. Historically, there have been periods where rising prices have been common, and periods where prices have tended to decline. Historically, periods of rising prices, called *inflationary periods*, have been longer-lasting and more numerous than periods of falling prices, i.e. *deflationary periods*.

A complete macroeconomic theory must be capable of explaining this varied historical experience. It must include an explanation of the behaviour of the general price level and no such explanation is possible without explicit consideration of the role of money in the economic system because prices are quoted in money terms. Because a complete theory must attempt to explain the behaviour of the general price level, it is unsatisfactory to develop a model that excludes any consideration of money. Moreover, money may have an importance beyond the determination of the price level alone, in that monetary factors may also influence the value of 'real' variables such as income, output and employment.

In some circumstances, monetary factors clearly influence real variables. This was most evident in the nineteenth century, when many major recessions were associated with major financial crises involving a loss of confidence in banking and other financial institutions. Keynes believed that monetary factors were important in explaining the persistent depression of the inter-war period, thus synthesising 'real' and monetary variables that previously had often been treated separately. Keynes argued, in effect, that monetary policy could not be relied upon to raise the level of effective demand in a recession, so that the burden of corrective action had to be placed on fiscal policy. This intellectual heritage was subsequently reflected in the economic policies of the capitalist countries, which emphasised the importance of 'functional finance' and largely ignored monetary policy, at least until the 1970s. Then the experience of continued and accelerated price inflation, combined with the analysis of the 'monetarist school', together with an apparent failure of the UK economy to respond to Keynesian measures, forced many policy makers to reassess their strategies.

Before tackling monetary policy and its impact on the economy, it is necessary to consider the functions and types of money and to describe some of the major features of our banking system.

17.2 The Function of Money

Money is anything that is generally acceptable for the settlement of debts. Money need not be created by fiat, i.e. government decree; for instance, cigarettes acted as money in prisoner-of-war camps during World War II because they were generally accepted as a means of settling debts, not because cigarettes were accorded that role by a central monetary authority. In many economies the single most important source of money is commercial bank deposits, which are simply the debts of private, profit-making financial institutions and are not legal tender. Bank deposits are money because they are generally acceptable in the settlement of debts and not because they have any special legal authority.

There are three types of money: coins, banknotes and bank deposits. Coins and notes (cash) are legal tender; bank deposits are not. This means that if individual A offers to settle a debt owed to B in cash, then B cannot demand payment in some other form, whereas if A offers to settle a debt with B by a cheque drawn on a bank deposit, then B can demand payment in cash. Nonetheless, there are many forms of transactions for which a cheque drawn on a bank deposit is generally acceptable as a means of settling debts – for example, in many commercial and business transactions, and in the payment of salary and wage cheques. For this reason, bank deposits are part of the *money supply*, indeed the single most important form of money in a modern economy.

Money has three functions. It acts as:

(a) a medium of exchange;
(b) a unit of account; and
(c) a store of wealth.

17.2.1 Money as a Medium of Exchange

Money acts as a medium of exchange because of its convenience compared with the alternative of exchanging goods through barter. The process of barter is wasteful as it depends on a coincidence of wants. For example, if A wants to trade a suit of clothes, then A must find someone who wants and fits these clothes and who also has something that A wishes in return. Barter is only possible when exchange is relatively uncommon, and this occurs only in an underdeveloped economy where each family unit is largely self-sufficient and there is little specialisation of labour. Specialisation of labour, which has been a key factor in economic growth, requires that exchange is efficient and cheap and this, in turn, requires money. Money is one of mankind's great inventions, making a contribution to human welfare comparable to the discovery of fire, the wheel, and malt whisky.

In a modern economy the flow of goods and services, and the rewards to the factors of production, are determined by the flow of money. Factor incomes – wages, salaries, rents, interest and profits – accrue to households in money and are then used to buy goods and services. The major function of money is its function as a medium of exchange and, to satisfy this function, money must possess the following characteristics:

- it must be widely acceptable;
- it must have a high value:weight ratio;
- it must be divisible to settle debts of differing denominations; and
- it must not be easily produced, counterfeited or debased in value.

17.2.2 Money as a Unit of Account

Money provides a standard measure of value. It is, in effect, a measuring rod providing a measure of the value of one good or service relative to another. If the going price of a new Mercedes is $60 000 and you can buy a hamburger for $1, then a new Mercedes is worth, and might be exchanged for, 60 000 hamburgers.

17.2.3 Money as a Store of Wealth

Money acts as a store of wealth in that a household (or firm) can sell its factor services or goods and store the money received until it has decided what to do with it. In a modern society, almost every household and firm holds a sum of money – it is, indeed, the most common form in which to store wealth. Most individuals have some money, even if they do not own any other form of wealth.

To act as a satisfactory store of wealth, the value of money must be reasonably stable over time. For example, if prices are rising rapidly, then the value of money held as a store of wealth is depreciating quickly. In these circumstances there will be a tendency to hold as little money as possible and to use other stores of wealth such as precious metals, stamps, paintings and other works of art. It is no coincidence, therefore, that the prices of these latter commodities have increased rapidly on more than one occasion in the past. These increases reflect the inflationary pressures causing a falling value of money, which has induced a movement 'out of money and into real goods'. As we shall see subsequently, this very process is likely to increase the rate of inflation and further contribute to a decline in the value of money.

17.3 A Brief History of Banking

As we have seen, money has three principal functions – medium of exchange, unit of account, and store of wealth – and there are three types of money – coins, banknotes and bank deposits. We now consider how these forms of money developed as a prelude to examining the process of credit creation.

In most societies coins made from precious metal developed as money. Compared with the possible alternatives, precious metals were available, divisible, and cheap to store and transport. In addition, precious metals were used in the manufacture of coins because it was believed that since money was a form of wealth it should consist of a substance which was itself valuable, for instance gold or silver. This was logical enough in less developed civilisations, but in modern societies the commodity value of the coinage is much greater than the coin's intrinsic value. This illustrates that money does not require to be made of a substance which is intrinsically valuable, so long as it is generally acceptable in the settlement of debts. This is most evident in the case of banknotes, which are simply pieces of paper with rather nondescript pictures, having no intrinsic worth, and not convertible. Historically, precious metals had intrinsic worth. This historical development of banknotes and bank deposits to form the two most important types of money underlines the importance of the acceptability of the means of exchange, and also the extent to which the commercial banks act as 'manufacturers of money' in a modern banking system.

The first bankers were goldsmiths, skilled artisans who worked with gold and who provided a safe-keeping for merchants to store gold. In return for depositing gold, the merchant would retain a receipt as written evidence of having made a deposit, the receipt being convertible into gold on demand. The receipt was simply evidence of the merchant's ability to pay and in the first stage of banking development the receipt (or banknote) was a temporary substitute for money. Thus when the merchant had to settle a debt, the receipt was converted into gold and the gold itself was used in payment. Such a system of banking is known as *cloakroom banking*, indicating that the bank simply acted as a place of safe-keeping, but did not add, through its own activities, to the money supply. The bank simply accepted one form of money, gold, and issued in return for that deposit a receipt for the same value. All bank receipts (banknotes) were therefore backed by the equivalent value of gold.

The second major stage of banking was reached when the public developed greater confidence in the ability of the banks to turn receipts into gold on demand. Once such confidence had developed, the banknote, the written evidence of a bank's liability, became 'as good as gold'. Banknotes were passed from hand to hand in the settlement of debts, being more conveniently stored and carried than gold. In any period only a small proportion of banknotes were presented to the banks to be converted into gold. In consequence, the banks found that they could issue banknotes in excess of their holdings of gold and yet maintain *convertibility* – that is, the banks began to issue banknotes that were in excess of the gold deposited with them. When this happened, the banks had become 'manufacturers of money' and had moved from cloakroom banking to *fractional reserve banking*. Fractional reserve banking, as its name implies, is a system where the deposit liabilities of the banks exceed their holdings of gold (nowadays cash). The pyramid of credit created was therefore much greater than the gold reserves of the bank.

Any fractional reserve banking system depends on its ability to maintain confidence and convertibility. The power of the commercial banks to issue banknotes in excess of their gold reserves was sometimes abused and, when confidence weakened, the banks' actions on occasions resulted in a failure to maintain convertibility. This led to the collapse of some financial institutions, and eventually state regulation of the issue of banknotes was introduced.

Today the Bank of England £5 note carries an inscription 'I promise to pay the bearer on demand the sum of five pounds'. The Federal Reserve Note (USA) states 'This note is legal tender for all debts public and private'. These inscriptions are reminders of the previous convertibility of banknotes into gold. Thus, the holder of a banknote could exchange the note for gold at a fixed rate and the total money supply was tied to the value of the gold reserves. If the gold reserves rose, the money supply automatically increased, and vice versa. This was known as *the Gold Standard*, a system abandoned by most countries in the 1920s and 1930s. Currency is no longer convertible into gold or, indeed, into anything. Banknotes are used not because they are intrinsically valuable, and not because they can be converted into something else that is intrinsically valuable, like gold; banknotes are used simply because they are an accepted means of settling debts. Money does not need to consist of something that is in itself intrinsically valuable; all that is necessary is that it is accepted and used in the settlement of debts.

In the nineteenth century the commercial banks in England were deprived of the right to issue notes, the right passing to the Bank of England, and the commercial banks therefore turned to deposit banking. Most advanced countries today have similar systems, with one main bank: the Federal Reserve Bank in the USA, the European Central Bank and the Bank of Japan. But although the mechanics of credit creation have changed, the basic principle remains the same. The banking system is a fractional reserve system, with the banks adding to the money supply not by printing paper money but by creating deposits in excess of their cash reserves. Since deposits are widely accepted as a means of settling debts, the commercial banks remain 'manufacturers of money'. The banks have to be able to convert deposits into cash on demand, but through long experience they have found that if cash reserves are some defined percentage of deposit liabilities, then this is usually sufficient to maintain convertibility.

However, such a system is not foolproof: some banks have gone bankrupt when they have made doubtful loans to foreign countries and/or concentrated loans in industries that have slumped. When Northern Rock, a small British bank with large interests in the UK mortgage market, approached the Bank of England for emergency funding in late 2007, many depositors panicked and attempted to retrieve their savings from high-street branches the next morning. Given that Northern Rock was holding only a small percentage of its total assets in cash, it was unable to allow the majority of depositors to withdraw their funds. News images of queues of pensioners waiting for hours in vain did nothing to restore confidence, and eventually the bank was nationalised by the British government to prevent bankruptcy and the spread of panic that would have resulted.

The US investment bank Lehman Brothers was heavily invested in the US subprime market, which consisted of mortgages that were considered to be very high risk. As more and more individuals defaulted on their mortgages, the assets of the bank began to reduce in value to the point in late 2008 where they were unable to meet their obligations and had to seek help from the US Treasury department. Help in this instance was not forthcoming, and the bank was allowed to collapse. This resulted in a global crisis of confidence as many banks discovered that many of their mortgage-related assets were in fact worthless. Credit markets around the world seized up as banks attempted to rebuild their now poor cash positions: a global contraction of the money supply, the effects of which will be discussed later in this chapter and again in Module 21.

17.4 The Fractional Reserve System, Credit Creation and the Credit Multiplier

Under a fractional reserve system the commercial banks are subject to two competing pulls:

- *The pull of profitability*: cash earns no interest, so the lower the proportion of the banks' assets that are held in cash and the higher the proportion held in interest-bearing assets, the greater the profitability of the banks.
- *The pull of liquidity*: the banks have to keep sufficient assets in liquid form to meet the demands of their customers for cash.

Therefore, part of the commercial banks' assets are kept in cash and part in liquid assets, which can quickly be turned into cash. Cash yields no interest and, generally speaking, the more liquid an asset, the less interest it yields. The pull of profitability leads to a desire to minimise cash holdings. The pull of liquidity necessitates that the banks always stand ready to meet their customers' demands for cash. In the subsequent analysis, we shall assume that if banks hold 10 per cent of their assets in cash, this is sufficient to meet demands for cash. Ten per cent will be the *cash ratio*.

To illustrate the process of credit creation, we shall make the following simplifying assumptions:

(a) there is a single monopoly commercial bank;
(b) the bank has a cash ratio (r) of 10 per cent ≡ 0.1, and
(c) members of the public do not wish to hold any more cash than they currently have but wish to hold all additional money they receive in the form of bank deposits. Hence, if members of the public obtain any cash they will immediately go to a bank and deposit the cash – i.e. the public's marginal propensity to hold cash is zero and there is no cash leakage from the bank to the public.

If we suppose that on day one the monopoly bank opens its doors for trading and that a member of the public deposits $100 in cash with the bank, then its balance sheet will read as follows:

Monopoly bank: Day one: Balance sheet

Assets		Liabilities	
Cash	$100	Deposit	$100

The above cannot represent an equilibrium position as all assets are held in cash, whereas the monopoly bank wishes to hold only 10 per cent of assets in cash. To reach the desired position, therefore, the monopoly bank will create deposits actively, for example by buying bonds and making loans to customers.

What are bonds? There are two principal ways that major corporations raise large sums of money that are not available to households or small firms. They issue stock or sell bonds. People who buy *stock* become stockholders and own part of the firm. If the firm prospers, the value of stock increases and, in addition, stockholders may receive income each year in the form of *dividends*, i.e. a share of the profit. When a corporation sells a *bond* for cash, it issues a piece of paper (a bond) that specifies a *redemption date* – a future date when the bond will be returned for cash – and also a *redemption value*, i.e. the amount of cash that will be returned to the bondholder. This is a legally enforceable obligation to pay the sum specified on the bond. In addition, the bond also specifies another sum of money, which will be paid each year until the redemption date. That sum of money paid each year is known as the *coupon payment*. The purchase price of a bond is determined in the bond market by the forces of demand and supply. As well as corporate bonds, there are government bonds or government securities – no government issues stock! It is those bonds or securities that constitute the major items in the balance sheets of the central banks.

Now back to our monopoly bank, a buyer and seller of government bonds. Notice that because we have assumed a monopoly bank and that the public's propensity to hold cash is zero, this process of active deposit creation cannot reduce the cash reserves of the commercial banks. There is no leakage of cash to other banks because, by assumption, there are no other banks for members of the public, because any cash received by the public is immediately redeposited with the monopoly bank. In addition, if the monopoly bank, in acquiring bonds or other assets, pays for these assets by a cheque drawn on itself, this cheque will immediately be returned to the bank to credit the account of the institution from which the bonds had been purchased, i.e. the bank's assets, as represented by the value of the bonds, will increase and will be matched by an equivalent increase in deposit liabilities. In consequence, if we assume that the monopoly commercial bank has purchased \$600 of bonds and made loans to customers of \$300 on day two, in the new equilibrium position the cash reserves of the monopoly bank will still be \$100 but deposit liabilities will have increased to \$1000. The balance sheet will now read:

Monopoly bank: Day two: Balance sheet

Assets		Liabilities	
Cash	\$100	Deposit	\$1 000
Bonds	\$600		
Loans	\$300		

This is an equilibrium position because the cash ratio is at the level (10 per cent) to maintain desired liquidity. Any further expansion of deposit liabilities will result in a cash ratio too low to maintain sufficient liquidity; any reduction in deposit liabilities will result in an unnecessary reduction in profitability. The change in deposits (ΔD) is given by the formula:

$$\Delta D = d \times \Delta C$$

where d is the credit or deposit multiplier and ΔC is the initial change in the cash reserves of the commercial bank. The size of the deposit multiplier is given by the reciprocal of the cash ratio, i.e. $d = 1/r$ where r is the cash ratio. In our example

$$D = \frac{1}{0.1} \times \$100 = \$1000$$

In terms of algebra, the credit multiplier is exactly analogous to the investment multiplier. The investment multiplier k gives the change in income (ΔY) that will be induced by a change in investment (ΔI), i.e.

$$k = \frac{\Delta Y}{\Delta I}$$

The credit multiplier gives the change in deposits (ΔD) that will be induced by a change in the cash reserves of the commercial banks (ΔC), i.e.

$$d = \frac{\Delta D}{\Delta C}$$

Clearly, the smaller is the cash ratio (r) the greater is the size of the credit multiplier and the greater is the volume of bank deposits created on a given cash base, and vice versa.

The effect of relaxing the initial assumptions can now be considered. If the assumption of our monopoly bank is abandoned, then the process of deposit creation becomes more complicated but the end result, in terms of the volume of deposits created, is the same. In a multibank case, as one bank creates deposits actively on a given cash base, that active deposit creation must reduce its cash reserves. For example, as a bank creates loan facilities for its customers, which they then draw upon by cheque, some of these cheques will be paid into accounts of customers with other commercial banks. This will result in a cash drain upon the bank that is actively creating deposits. Hence, it will not be possible for the bank receiving initial deposits of cash (\$100 in our example) to pyramid credit on this basis to reach the situation described in the Day Two Balance Sheet above, for the process of active deposit creation must reduce its cash reserves.

Yet, while in the multibank system active deposit creation by one bank will reduce its cash reserves, it does not reduce the cash reserves of the banking system as a whole: the drain is to other banks in the system and not to the general public. So, as long as our other assumptions hold, the initial increase in deposits will still be 10 times the initial increase in cash reserves. In the multibank case, the new deposit liabilities will be distributed across a number of banks in the system, but the increase in total bank deposits will still be the same as in the case of the monopoly bank.

Whereas the relaxation of the monopoly held by our single bank does not change the end result, this is not true for our other assumptions. If these assumptions do not hold, i.e. the cash rate is no longer 0.1 or leakages occur to the public, the size of the deposit multiplier will be affected and therefore the amount of deposit creation that will occur on a given cash base will change.

The credit multiplier can only be 10 if all the banks in the system keep to a 10 per cent cash ratio. If the initial bank, say Bank A, decides to add the initial deposit of cash to its cash reserves and not to create any new deposits actively, then there will be no change beyond the initial \$100 deposited. Cash reserves will rise by \$100 and deposit liabilities by \$100, but that will be the end of the process. The credit multiplier will be equal to 1. There will be no addition to the supply of money. All that will have happened is that the public will hold less cash (\$100) and an equivalent (\$100) larger amount of bank deposits. In most countries the commercial banks are required by law to hold a minimum cash ratio, although no maximum is established. For example, the minimum cash ratio in the US is fixed by the Federal Reserve Bank and has legal force. In Britain there is no legally determined minimum, but by long convention the commercial banks work to a minimum cash ratio of 8 per cent.

The final assumption was that the public's propensity to hold additional cash was zero. The relaxation of this assumption provides the most important qualification to the credit multiplier concept that we have developed. This assumption is likely to be more unrealistic. The public uses both cash and bank deposits to settle debts, and thus it seems more reasonable to suppose that the public's demand for cash is likely to rise as the volume of bank deposits increases, and vice versa. If the public wishes to hold part of any increase in the money supply in cash, then as banks create deposits actively, a cash leakage from the banking system will occur as cash finds its way into the hands of the public and is held in additional cash balances. If this is so, the ability of the banks to create deposits is limited, and is more limited the higher is the marginal propensity to hold cash on the part of the public.

In the example developed, assuming that the public's propensity to hold additional cash was zero, we saw that an increase in the cash reserves of the commercial bank of \$100 would, on a 10 per cent cash ratio, give a \$1000 increase in the volume of bank deposits. Such an increase was only possible because we assumed that all cash would return to the banks. The only limitation on the ability of the banks to create deposits was the need to maintain liquidity – in this case to meet the 10 per cent cash ratio. If the public's propensity to hold additional cash is greater than zero, then some part of the initial increase in cash reserves will finish up in the hands of the public, not in the vaults of the commercial banks. As the banks create deposits actively, some part of their initial cash reserves will find their way into cash balances retained by members of the public. The more this happens, the smaller will be the credit multiplier.

The creation of bank deposits is therefore limited by two factors:

(a) the banks' propensity to keep cash for liquidity purposes, as represented by the cash ratio; and
(b) the public's propensity to hold additional cash.

17.5 The Central Bank and Monetary Policy

All monetary economies have a central bank – the Federal Reserve Bank in the USA, the European Central Bank, and the Bank of England in the UK, for instance. The Bank of England was founded in 1694 and was not nationalised until 1947. Nationalisation merely recognised the pre-existing situation in that the Bank, for a long period prior to nationalisation, had acted as the country's central bank.

The function of a *central bank* is to control the commercial banks around it in such a way as to support the monetary policy of the economy. The central bank conducts its business by acting as the bankers' bank and lender of last resort, and also as the government's bank and manager of the public debt. In this way the central bank is responsible for the conduct and operation of monetary policy that attempts to influence the level of economic activity and that involves regulating the supply of money and the cost and availability of credit.

In most countries, the central bank has at its disposal a number of powers and instruments of control, such as the monopoly of the note issue, the ability to dictate the cash ratio and the reserve requirements that must be observed by the commercial banks, credit controls (e.g. the level of initial payments and repayment periods on hire purchase contracts), and, ultimately, the power to issue direct instructions to the commercial banks and other financial institutions. The most important instrument of control is open market operations, which involve the central bank in buying or selling government bonds in the open market. If the central bank buys bonds, the effect is to increase the supply of money and reduce the cost of borrowing, while the sale of bonds reduces the supply of money and increases the cost of borrowing.

If the central bank buys bonds, it pays by a cheque drawn on itself and payable to the seller, say a private citizen. The seller will then deposit the cheque with his/her bank. The commercial bank will present the cheque for payment to the central bank, which will credit the deposit of the commercial bank. Bankers' balances have therefore increased by the value of the cheque. As we have seen, bankers' balances can be turned into cash on demand and are therefore considered to be part of the cash reserves of the commercial banks. As cash reserves have increased, the commercial banks will create deposits actively, so that there is an expansion of the money supply that is some multiple of the initial increase in cash reserves, the size of the multiple expansion depending on the size of the credit (deposit) multiplier.

Open market operations influence the cost of borrowing as well as the supply of money. If the central bank purchases bonds, the cash reserves of the commercial banks rise and the commercial banks will wish to lend more both to private customers and to business concerns. Using normal supply and demand analysis, it can be seen that if the supply of money increases, the cost of borrowing (interest rates) must fall so long as there is no offsetting shift in the demand for money. Conversely, if the central bank sells bonds, this reduces the supply of money and (other things being equal) will increase the cost of borrowing (interest rates).

It is now possible to illustrate how monetary policy may be used to influence the level of aggregate demand and therefore the level of income, output and employment. An expansionary monetary policy, buying bonds, will cause the money supply to rise and interest rates to fall and, from our earlier analysis, this should raise the level of aggregate demand. So long as private investment is interest-elastic, a reduction in interest rates will make some investments profitable that were previously unprofitable at the higher cost of borrowing. A reduction in the interest rates will result in a movement along a given marginal efficiency of investment curve so that investment rises. Furthermore, a reduction in interest rates may cause an increase in other components of aggregate demand, e.g., household purchases of consumer durables such as cars or televisions. Central and local government will find borrowing cheaper, prospective house purchasers will be able to obtain cheaper house mortgages, hire purchase facilities will become more readily available and less expensive, even exporters – finding borrowing cheaper – may encourage foreign sales by offering easier and longer credit terms to prospective foreign buyers.

Hence, open market operations, through increasing the supply of money and lowering the structure of interest rates, may increase aggregate demand by raising private and public investment, consumption (particularly of consumer durables), and even exports, and this, in turn, will raise the level of income, output and employment. Conversely, open market operations, when they reduce the supply of money and raise the structure of interest rates, may reduce aggregate demand by lowering private and public investment, consumption and even exports; and this, in turn, will lower the level of income, output and employment. In general, investment, particularly private investment, is likely to be more sensitive to changes in interest rates, but other components of aggregate demand may also be affected through monetary policy.

The manner in which monetary policy can influence the level of aggregate demand, and therefore the level of economic activity, can be summarised as having five sequential stages:

1. The central bank acts on the cash reserves of the commercial banks.
2. Commercial banks act to change bank deposits and therefore the money supply.
3. A change in the money supply influences the cost of borrowing.
4. A change in the availability and cost of credit influences aggregate demand.
5. A change in aggregate demand leads, through the multiplier process, to some multiple change in income, output and employment.

Let us now consider these steps in greater detail.

1. The central bank influences the activities of the commercial banks through open market operations, or through changing the cash ratio and reserve requirements of the commercial banks. In open market operations, the central bank, if it wishes to engage in an expansionary monetary policy, will buy bonds, the effect of which is to increase the cash reserves of the commercial banks. If the central bank wishes to follow a restrictive monetary policy, it will *sell* bonds, the effect of this being to reduce the cash reserves of the commercial banks.
2. The change in the cash reserves of the commercial banks will have, through the credit multiplier, a multiple impact on the volume of bank deposits. A reduction in cash reserves will lead to a multiple contraction of bank deposits and therefore to a multiple contraction of the money supply. An increase in cash reserves will lead to a multiple expansion of bank deposits and therefore to a multiple expansion of the money supply.

3. Changes in the money supply will influence the ease and cost of borrowing. An increase in the money supply will make it easier to obtain credit and also make it less expensive, and so interest rates will fall. Conversely, a reduction in the money supply will make borrowing more difficult and hence more expensive, and interest rates will rise.
4. Changes in the availability and the cost of credit will influence the level of aggregate demand. If the money supply is increased so that credit is more easy to obtain and interest rates are lower, this will have a favourable effect on the volume of private and public investment, consumption and even exports. Conversely, if the money supply falls, so that it is more difficult and costly to obtain credit, this will have the effect of lowering the level of private and public investment, consumption and even, possibly, exports. As private investment is likely to be more sensitive to changes in interest rates, the effect on aggregate demand in terms of this component of demand can be analysed in terms of the previous discussion of investment decisions, i.e. a change in interest rates causes a movement along a given marginal efficiency of investment curve, so that a fall in interest rates results in an increase in private investment and vice versa.
5. The change in the volume of aggregate demand will cause national income, output and employment to change in the same direction. An initial increase in aggregated demand will lead to a multiple process and vice versa. An expansionary monetary policy, by lowering interest rates and raising aggregate demand, will lead to a multiple expansion in national income, and a contractionary monetary policy will have the reverse effect. The difference between expansionary and contractionary monetary policies is shown in Figure 17.1.

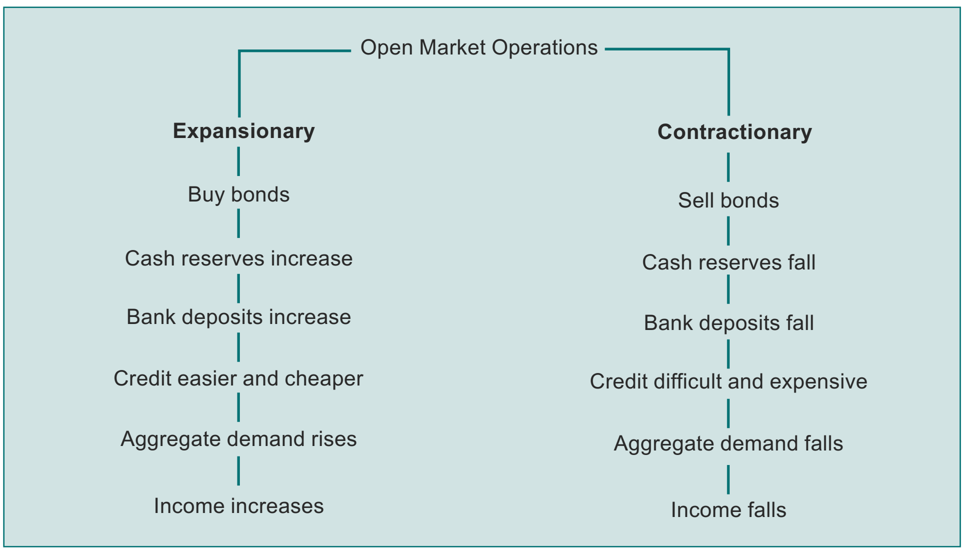


Figure 17.1 Open market operations

We are now in a position to analyse competing theories concerning the role and influence of money in the economy.

Learning Summary

You now know what money is, what characteristics it must possess and what its three main functions are in the economy. You understand how the banking system operates and how it developed from the 'gold' days to the modern fractional reserve system. You understand the system of credit and a bank's balance sheet.

You are now aware of the function of the central bank in controlling commercial banks to support the monetary policy of an economy. You also understand how monetary policy can affect the level of aggregate demand, and consequently, national income.

Review Questions

Multiple Choice Questions

17.1 Money is a medium of exchange; as a result the money supply

I. and disposable annual income are the same thing.
II. and total annual spending are the same thing.
III. will only be held for transactions purposes.

Which of the following is correct?

A. I only.
B. II only.
C. III only.
D. Not I or II or III.

17.2 Which of the following is a correct definition of the value of money?

A. The cost of producing money.
B. The buying power of money.
C. The value of gold and silver held by the Treasury.
D. The interest rate.

Questions 17.3 and 17.4 are based on the following information. A commercial banking system has $100 million of outstanding demand deposits and cash reserves of $35 million. The desired cash ratio is 1:5.

17.3 In the above example, by which of the following amounts could the banking system expand the money supply?

A. $15 million.
B. $75 million.
C. $170 million.
D. $300 million.

17.4 The reasons why the situation in the above example persisted and the money supply did not expand are

I. the reserve ratio was too high.

II. there was no additional demand for credit.

III. banks were unwilling to lend more to the type of customer who wished to borrow.

Which of the following is correct?

A. I only.

B. II only.

C. I and III only.

D. II and III only.

17.5 Faced with a deflationary gap, a government has resorted to monetary policy. The appropriate policy action includes

I. increasing commercial bank reserve requirements.

II. purchasing government securities.

III. raising interest rates to attract foreign capital.

Which of the following is correct?

A. II only.

B. III only.

C. II and III only.

D. I, II and III.

17.6 Which of the following tends to diminish the ability of UK commercial banks to lend money?

A. The raising of interest rates by the Bank of England.

B. The releasing of Special Deposits by the Bank of England.

C. The reduction of the reserve requirement.

D. The occurrence of large-scale net withdrawals of deposits from the commercial banks.

17.7 Mr Smith receives a tax rebate of \$100, which he deposits in his account at his bank, which is part of a banking system operating on a cash ratio of 10 per cent. The maximum possible increase in money supply that can result from this transaction, including his original deposit, is

A. \$190.

B. \$910.

C. \$1000.

D. \$1100.

Module 18

The Quantity Theory and the Keynesian Theory of Money

Contents

18.1 Introduction

In Module 17 we analysed how the central bank controls the activities of the commercial banks and thereby determines the money supply, and the availability and cost of credit. A complete analysis requires some explanation of the factors that determine the demand for money. Households and businesses hold money balances, and we have to consider the factors that underlie this demand for money.

Equilibrium in the money market requires that the supply of money is equal to the amount of money that households and businesses wish to hold in money balances. If this is not so, then the money market will not be in equilibrium. Clearly, if the demand for money balances exceeds the supply of money available, then households and businesses will attempt to obtain more money to add to their money balances. Conversely, if the demand for money is less than the supply of money, then households and businesses will try to run down their money balances.

To examine the manner in which monetary policy may influence income and employment, it is necessary to look at the demand for money in greater detail. For example, if money is held for transactions purposes only, in order to finance the ordinary day-to-day transactions of households or businesses, then the link between changes in the supply of money and changes in aggregate demand will be direct and immediate. An increase in the supply of money will result in a situation where households and businesses will have more money than they wish to hold in transactions balances, and they will spend the excess, thus resulting in an increase in aggregate demand. Alternatively, a reduction in the supply of money will result in a situation where households and businesses will have less money than they wish to hold in transactions balances, and in order to build up their money balances they will reduce their expenditure, so that aggregate demand will fall. This line of reasoning leads to the *Quantity Theory of Money*, which suggests a close and immediate link between the money supply and the level of aggregate demand.

At the opposite extreme lies the Keynesian theory, which suggests that changes in the demand for and supply of money are reflected immediately only in the market for securities, which for simplicity will be assumed to be represented by the market for government bonds. Keynesian theory suggests that if households and businesses have excess money balances, they will use the excess to purchase bonds and, alternatively, if they wish to add to their money balances, they will sell bonds. If this is so, changes in the balances between the demand for money and the supply of money will not be reflected directly in the level of aggregate demand but will be reflected primarily in the market for securities and will affect the level of aggregate demand only indirectly. As we shall see subsequently, changes in the demand for and supply of bonds will affect the price of bonds and interest rates, and changes in interest rates may result in changes in aggregate demand.

In certain circumstances, monetary policy may be unable to raise aggregate demand, even indirectly. If investment opportunities are considered to be poor, because the level of economic activity is low (there is excess capacity and business expectations are unfavourable), then an expansionary monetary policy may be ineffective in encouraging private investment because the commercial banks may be unable to induce businesses to take up new loans. In such circumstances the desire to hold money may be strong, households and businesses preferring to hold money to holding interest-bearing assets. If in these circumstances the supply of money is increased, it may be held in the idle-money balances instead of being used to purchase bonds. The increase in the money supply will not, then, lower interest rates, so that even this indirect method of raising aggregate demand will prove to be ineffective. In order to examine these ideas in greater detail, we will look first at the contribution of the Quantity Theory of Money.

18.2 The Quantity Theory of Money

The Quantity Theory of Money in its 'traditional' or 'naive' form suggests a direct relationship between changes in the money supply (M) and the level of prices (P). The Quantity Theory is developed from the following identity:

$$MV \equiv PT$$

where M is the money supply (determined exogenously by the monetary authority), V is the velocity of circulation, which is simply the number of times each unit of money is spent in a given period (say, one year), P is the general level of prices, and T is the number of transactions in a given period (say, one year). MV is often referred to as the *monetary side* of the equation and PT is often called the *commodity side*. Notice the identity sign, which indicates that MV must necessarily equal PT, given the manner in which the various terms are defined.

The monetary side MV (which is the supply of money times the average number of times each unit of money is spent) gives the sum of expenditures in a given period. On the commodity side, T is the number of transactions and P represents the price at which goods are sold. Hence, PT gives total receipts. The identity $MV \equiv PT$ therefore simply states that, in any period, the sum of expenditures must equal the sum of receipts. In order to turn identity into a theory, certain assumptions have to be made about the behaviour of one or more of the variables. It is these assumptions that turn the identity into a theory or hypothesis capable of being refuted by an appeal to data.

Early proponents of the Quantity Theory believed that the velocity of circulation of money would be stable and would change very little over time. In short, the velocity of circulation depended on habit and institutional arrangements. It depended on the system of banking, the extent to which people gave credit, the form in which wages and salaries were paid, and on similar factors that were not subject to rapid change. For this reason it was believed that the velocity of circulation could be regarded as constant, at least in the short run.

On the commodity side, the number of transactions, T, could be regarded as proportional to the level of real income, Y, so that there would be a close and direct relationship between the number of transactions in any given period and the level of real income. By this reasoning, Y can then be substituted for T in the above equation. Moreover, if the equilibrium level of national income is always at full employment – an assumption made by economists who believed free-market forces would lead to full employment – then Y is also fixed at full-employment income.

Hence, V and Y, one term on the monetary side and one term on the commodity side of the identity, are assumed to be constant through time. This being so, the identity can be rewritten as:

$$MV = PY$$

With V and Y constant, it follows that any change in M will be reflected, through a linear relationship, in a change in the same direction in P; that is, the price level, the value of money, is seen as dependent on the supply of money. An increase in the supply of money, M, will cause an increase in the price level, P, and therefore a fall in the value of money, and vice versa.

The 'naive' Quantity Theory suggests that changes in the supply of money, M, affect only the price level, P, and can have no effect on the level of real income, Y. The naive Quantity Theory can be modified, however, to yield what is called the 'modern' Quantity Theory. The modern Quantity Theory suggests that changes in the supply of money can indeed influence the level of real income, Y, as well as the price level, P, the effect on these two variables depending on how close, or far away, the economy is from full employment.

Suppose we begin from a position of substantial unemployment with unused factors of production, so that Y is below the level of real income that would be produced at full employment. In these circumstances, what now happens if the money supply, M, is increased? The modern version of the Quantity Theory would argue that households and businesses will find that they have more money balances than they need for transaction purposes and they will spend that excess. This additional demand leads to a rise in income, output and employment, i.e. to a rise in real income, Y. Hence, if there is sufficient unemployment, the main effect of an increase in M is to increase real income. As the economy approaches full employment, however, further increases in the supply of money will begin to affect the price level more than the level of real income. Finally, after full employment is reached, any further increase in the supply of money, will influence the price level only.

The modern version of the Quantity Theory attempts to show that changes in the money supply will have an effect on total spending, or aggregate demand; that a change in the money supply will directly affect the level of aggregate demand and therefore the level of money income, PY, the exact balance between price changes and changes in real income depending on the relationship between the current level of income and full-employment income.

18.3 The Keynesian Theory

The Quantity Theory stresses the role of money as a medium of exchange. However, money has another function, namely as a store of wealth. The Keynesian theory stresses money in this role.

Let us begin discussion of the Keynesian theory with an examination of the demand for money. There are three types of demand for money balances: the transactions demand for money, the precautionary demand for money, and the speculative demand for money.

18.3.1 The Transactions Demand for Money

The transactions demands for money arises from the fact that people need money to finance current transactions. Households and firms hold money balances to bridge the gap between the receipt and expenditure of income. Wage and salary earners are paid weekly or monthly, but they spend money throughout the week or the month. In similar fashion, firms receive income from the sale of goods and services, but this is not synchronised with their payments to factors of production, purchases of raw materials, and capital expenditure.

The amount of money held for transaction purposes will be closely related to the level of national income. Clearly, a national income of $100 billion annually will require many more transactions than a level of national income of $50 billion. The price of each transaction will affect also the amount of money required for transaction purposes. If in some given time period all prices double, the demand for money for transaction purposes will double also; thus the inflation rate will be correlated positively with the demand for money. National income, expressed in current prices, of course, captures both the quantity of transactions and price changes effects.

The demand for money for transaction purposes is also likely to be influenced, however, by the rate of interest. If the rate of interest – the return on interest-bearing assets – is high, there will be strong motivation to economise on money held for transactions purposes so as to yield funds for investment in interest-bearing assets. The extent to which individuals and business firms can switch between money and interest-bearing assets will depend, of course, on transaction costs, but given the level of transaction costs the demand for money for transactions purposes will be high when interest rates are low, and vice versa. Clearly, at high interest rates, high-income families and big-business firms in particular will have a strong incentive to minimise transactions balances in order to purchase interest-bearing assets.

18.3.2 The Precautionary Demand for Money

As the name suggests, precautionary balances consist of money held to meet the sudden arrival of unforeseen circumstances. A business firm selling goods on credit cannot be certain when the buyer will make payment; payment may be slower than anticipated; they have to make allowance for bad debt; they cannot foresee with certainty the timing of expenditure. Similarly, a household will wish to hold some money to cover the possibility that it will incur unexpected debts, or that it will have to settle existing debts more quickly than expected.

Again, the main factor affecting the amount of money held for precautionary purposes is likely to be the level of income but, as with the transactions demand for money, the precautionary demand for money will vary inversely with the rate of interest.

18.3.3 The Speculative Demand for Money

The transactions and precautionary demand for money emphasise the role of money as a medium of exchange, whereas the speculative demand for money emphasises the role of money as a store of wealth. To understand the speculative demand for money, it is useful to begin with the following rhetorical question: 'Why do individuals and institutions not hold all their wealth in the form of money?' After all, money is easily carried around or stored; it can be used to buy all kinds of goods and services. The answer to the question is, of course, that wealth kept in the form of money does not yield any return or any income. Holding wealth in the form of money has an opportunity cost – the opportunity cost is the income forgone, what would have been earned if the money had been used to buy an asset that yielded income, such as a government bond.

Hence households and business firms will want to keep some part of their wealth in the form of money because of its convenience as a means of exchange, but they will not want to keep *all* their wealth in this form because money yields no return. Economists say that '*liquidity preference is not absolute*', because households and businesses do not want to hold all their wealth in money. What, then, determines the extent of a person's or firm's liquidity preference? Part of the answer lies in the transactions demand for money, which has already been analysed. In addition, some money is held for speculative purposes, the speculative motive being the desire to hold money in order to take advantage of developments in the capital market. Speculative balances are simply money held to speculate on the course of future events.

Speculative activity occurs whenever there is uncertainty about future events. Suppose an individual is considering purchasing a consumer durable that has a fairly long life expectancy, for example, a car. If the individual considers that the price of the consumer durable is likely to rise in the immediate future, that individual is likely to make the purchase quickly before the anticipated price rise occurs. Conversely, if the individual considers the price of the car is likely to fall in the immediate future, then purchase is likely to be deferred in anticipation of the fall in the price. Speculative activity can apply to all goods, services and financial assets that are bought and sold, but to simplify the subsequent analysis we will confine our attention to the purchase and sale of government bonds, although a similar analysis can be applied to all financial assets.

If households and firms expect the price of bonds to fall in the future, then they will be likely to sell their existing holding of bonds and to defer purchasing new ones until the anticipated price fall has taken place. Naturally, both these factors will tend to make the price of bonds fall. Conversely, if households and firms expect the price of bonds to rise in the future, then they will defer selling bonds now and are likely to want to purchase additional bonds. Again, both factors are likely to have the effect of raising the price of bonds. In the former circumstance, the speculative demand for money will be strong, as households and firms will wish to hold cash in anticipation of a more favourable opportunity to purchase bonds in the future. In the latter situation, the speculative demand for money will be low as both households and firms will wish to move out of money into bonds.

The price of existing bonds and the market rate of interest are inversely related. If the price of bonds is expected to fall, this is the same thing as saying that the rate of interest will rise. Conversely, if the price of existing bonds is expected to rise, this is the same thing as saying that the market rate of interest will fall.

Example

Suppose that individual A, in period 1, buys an undated government bond for $100, that bond yielding an income of $5 per annum. Suppose that, in period 2, the market rate of interest rises to 10 per cent and that A now wishes to sell the bond to B. The market price of the bond will no longer be $100, for if B buys the bond at that price, and the return obtained is only $5 per annum, which is equivalent to a rate of interest of 5 per cent, this is less than the current market rate of interest at 10 per cent. In order to sell the bond, A has to accept a lower price, which will yield a rate of return on the asset equivalent to the market rate of interest of 10 per cent. The price of the bond in period 2 will be $50, because a fixed sum of $5 per annum on a capital sum of $50 is equivalent to the market rate of interest of 10 per cent.

In this example, what has happened is that a rise in the current rate of interest from 5 per cent to 10 per cent has resulted in a fall in the price of existing bonds. Conversely, if the current rate of interest had fallen after the purchase of the bond, then the market price of the bond would have risen. In short, the current rate of interest and the market price of existing bonds vary inversely with each other. When the current rate of interest rises, the market price of bonds falls, and vice versa.

Returning to the speculative demand for money, it is evident that speculative balances arise from the existence of uncertainty. If there were no uncertainty, then no wealth would be held in the form of money for speculative purposes because all individuals and firms would prefer to hold their wealth in the form of bonds, which yield a rate of return. Speculative balances are simply wealth held in the form of money in preference to interest-bearing assets, because of the possibility that the price of interest-bearing assets may change. Where households or firms hold money, they are expressing a desire for liquidity. *Liquidity preference* relates to the demand to hold assets as money rather than as interest-bearing bonds. The strength of the demand for liquidity preference varies inversely with the rate of interest: when the rate of interest is low, liquidity preference is high; and when the rate of interest is high, liquidity preference is low.

The reasoning underlying this assumption is straightforward. If the rate of interest is low, then – other things being equal – it will be expected that in the future an increase in the rate of interest is more likely than a further reduction in the rate of interest below the already existing low level. An increase in the rate of interest is equivalent to a fall in the price of existing bonds. For this reason, households and firms will be likely to sell existing bonds and defer the purchase of new bonds, i.e. liquidity preference will be high. Conversely, if the rate of interest is high, then – other things being equal – future changes in the rate of interest are likely to be in a downward rather than in an upward direction. A fall in the rate of interest is equivalent to a rise in the price of existing bonds. For this reason, households and business firms will defer selling their existing bonds and will wish to purchase new bonds in order to benefit from the high level of interest rates and from the expected appreciation in the price of existing bonds when interest rates fall.

18.3.4 Summary of the Demand for Money

Summarising this discussion of the transactions, precautionary and speculative demands for money, it follows that the demand for money balances will vary directly with the level of income (Y) and inversely with the rate of interest (R). The higher the level of income, the greater the demand for money for transactions and precautionary purposes; and the lower the level of income, the smaller the demand for money for transactions and precautionary purposes. The higher the rate of interest, the greater the incentive for households and businesses to economise on transactions and precautionary balances, because of the high opportunity cost of holding money, and the greater the demand to hold bonds instead of money in speculative balances. Conversely, the lower the rate of interest, the lower the opportunity cost of holding money for transactions and precautionary purposes and the higher the demand to hold money instead of bonds. Thus, when income is high, the demand for money for transactions and precautionary purposes will be high, and vice versa. When the rate of interest is high the demand for money for all purposes – transactions, precautionary and speculative – will be low, and vice versa.

18.3.5 Keynes versus the Quantity Theory

The Keynesian theory of money, in contrast to the Quantity Theory, suggests that changes in the money supply do not lead directly to changes in aggregate demand. Instead, monetary policy affects aggregate demand indirectly through the cost of borrowing (interest rates), thus influencing those components of aggregate demand that are sensitive to changes in the cost of borrowing. A reduction in the supply of money leads to an increased rate of interest and this, in turn, leads to reduced aggregate demand. Conversely, an increased supply of money lowers interest rates and tends to increase aggregate demand.

Further factors have to be introduced to complete the analysis, and these will be examined in Module 19. In the meantime, it should be noted that Keynesian theory suggests that monetary policy may be ineffective in dealing with a deep recession. When the rate of interest is low, the demand for liquidity may become insignificant. Further increases in the money supply will not be able to reduce the rate of interest and so this indirect method of influencing aggregate demand will be ineffective. Keynes suggested that in a deep recession, with substantial spare capacity and unfavourable business expectations, very low rates of interest might be necessary in order to encourage a higher level of investment – but it might be impossible to force the rate of interest below a certain floor: the *liquidity trap*. If a lower rate of interest were required in order to obtain a sufficient volume of investment, then monetary policy would be ineffective. Thus, monetary policy may be able to restrain the level of aggregate demand in an inflationary situation, but it may be ineffective in raising the level of demand in a deep recession.

We are now ready to investigate the real goods and monetary sectors of the economy in the next module.

Learning Summary

You can now understand the central role that money plays in the Quantity Theory and the Keynesian theory, and the resulting possible impact on aggregate demand. You are also able to explain the role of interest rates in affecting investment expenditure and aggregate demand, and why monetary policy may be ineffectual under specific assumptions.

Review Questions

Multiple Choice Questions

18.1 According to the Quantity Theory, if real GNP is increasing at 2 per cent per annum, if potential GNP is increasing at 4 per cent per annum, and if the money supply is increasing at 4 per cent per annum, the rate of inflation per annum will be which of the following?

A. negative.
B. zero (approximately).
C. 2 per cent (approximately).
D. 4 per cent (approximately).

18.2 If V and Y are held constant and M is increased, it follows that there will be which of the following?

A. an increase in real national output.
B. a reduction in real national output.
C. an increase in prices.
D. a reduction in prices.

18.3 The Quantity Theory does not predict short-run changes in the economy because of which of the following?

A. Prices are fixed.
B. The government controls M.
C. All four variables in the equation can change.
D. The equality does not always hold.

18.4 Which of the following is an accurate description of 'liquidity preference'?

A. The desire to hold assets that can readily be converted into money at a fixed or near-fixed price.
B. The amount that businesses wish to borrow at a given interest rate.
C. The desire to save out of income as a protection against future uncertainties.
D. The desire to hold money as an asset in preference to any form of interest-bearing assets.

18.5 Individuals will tend to move out of money and into long-term interest-bearing assets if they think which of the following is true?

A. The rate of interest is going to fall.
B. The rate of interest is going to rise.
C. The rate of interest is not going to change.
D. The price of goods and services is going to fall.

18.6 Individuals hold cash, which yields no interest, in preference to assets, which do yield interest, because of

I. the needs of day-to-day transactions.

II. expectations concerning the future rate of interest.

III. a desire to maintain the level of savings out of income.

Which of the following is correct?

A. I only.

B. I and II only.

C. III only.

D. I and II and III.

18.7 In times of recession, interest rates tend to fall even when the Central Bank takes no specific action to affect interest rates. Which of the following accounts for this?

A. The lower level of income that accompanies recession decreases the transactions demand for money.

B. The money supply automatically rises as national income decreases.

C. The velocity of circulation systematically follows the business cycle.

D. Individuals attempt to increase their savings during recessions.

18.8 In response to increased levels of unemployment, a politician is reputed to have said: 'The problem with the economy today is deficient aggregate demand. We need to decrease personal taxes in order to stimulate spending, and increase the rate of interest in order to make investing more attractive for businessmen.'

Would these measures reduce unemployment?

A. Yes, since both measures will result in increased spending.

B. No, since the speaker confuses increases in employment with increases in the deflationary gap.

C. No, since the proposals will tend to cancel each other out.

D. No, since both proposals aim to reduce spending instead of increasing it.

18.9 Faced with a substantial deflationary gap, a government used expansionary fiscal policy, i.e. it reduced taxes and increased government expenditure, kept the money supply constant, and discovered that the net outcome of such a policy was a decrease in private investment but no significant changes in either GNP or employment. This outcome arose because

I. the economy was at full employment.

II. interest rates rose.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

18.10 The goals of a government are to reduce interest rates, stimulate private investment, and decrease unemployment. The policy strategy that would guarantee the achievement of all of those goals is

I. increasing the money supply.
II. increasing government expenditure.
III. reducing taxes.

Which of the following is correct?

A. I only.
B. II or III only.
C. I, II and III.
D. Not I nor II nor III.

Module 19

Integration of the Real and Monetary Sectors of the Economy

Contents

19.1 Introduction

We have now reached the stage when we can attempt to pull together the different sectors of the economy that we have been analysing separately. In the model-building process that we have followed in macroeconomics, we started off with a simple closed economy with no government and no money. As we relaxed the simplifying assumptions, the analysis became more complex because we were dealing with an increasing number of independent variables. Up to now, however, the sectors have remained 'semi-isolated'.

In analysing how the equilibrium level of output/income was attained, we 'pulled out of a hat' the rate of interest (R), which affected investment and consumption expenditure and consequently national income (Y), the employment level, and concomitantly the unemployment rate (U). In analysing how the interest rate (R) is determined in the monetary sector, we 'pulled out of a different hat' the level of income (Y), which affected the demand for money. The demand for money and the supply of money (M), controlled by the monetary authority, determine the interest rate (R). Thus R and Y appear in both sectors, the real goods sector and the monetary sector. It follows that the equilibrium rate of interest (R) and the equilibrium level of income/output (Y) are determined simultaneously by the intersection of real and monetary factors.

The story, however, does not end there. In the microeconomics modules, we saw how market forces determined equilibrium prices and quantities in both goods and factor markets; and remember the definition of equilibrium in a market – the price at which the market is cleared of excess demand and excess supply, with no unsatisfied buyer, no unsatisfied supplier, and no tendency for the price to change. Thus the question arises: if we achieve equilibrium in the real goods and monetary sectors of the economy, shall we simultaneously achieve equilibrium in the markets for the factors of production that produce national output? In other words, is it possible to reach equilibrium in the real goods and monetary sectors of the economy and at the same time attain equilibrium in all factor markets, having full employment of all factor inputs with stable prices and consequently neither inflation nor deflation? Historically, there have been periods (no matter that they were short-lived) when such a goal was attained or close to being attained. The question must then arise: why is such attainment the exception rather than the rule?

Be advised in advance that the analysis following is complex, and conventional teaching methods such as texts like this, lectures and seminars are not ideal mechanisms for dealing with such complicated issues. We have developed a teaching tool to assist in overcoming this problem. The software entitled 'Running the British Economy', which is available via the Economics course website (http://coursewebsites.ebsglobal.net), is a simulation of the UK economy that enables you to put monetary and fiscal policy into (theoretical) practice over a 10-year time period. The policy decisions you make in this simulation have an immediate impact, just as in real life, on a host of interrelated macroeconomic variables in the economy, but these variables are also affected, just as in real life, by past policy decisions and the current state of the economy. This simulation has been used in company seminars, management and executive programmes, colleges, universities and business schools throughout the world, and experience suggests that trying to teach complex macroeconomics without such a simulation is like trying to teach chess to a child without a chess board and pieces.

In this module we shall integrate the real and monetary sectors and in the next module worry about inflation and unemployment; but, first, a quick summary of the real and monetary sectors.

In Module 18 the rate of interest was analysed as if it were a purely monetary phenomenon, determined by the interaction of the demand curve for money and the supply of money fixed by the central monetary authority. This is the view taken by Keynes in his *General Theory* and has the important implication that monetary factors may have an influence on 'real' variables such as the level of income and employment. The existence of a liquidity trap may prevent the rate of interest falling below a certain floor, so that it may be impossible to use monetary policy effectively to counteract a recession; in such a situation, further increases in the money supply would be unable to reduce the rate of interest so that there would be no effect on the level of aggregate demand. Using the language of the Quantity Theory, in the liquidity trap region, any increases in the money supply would find their way into speculative balances. The velocity of circulation of money (V) would fall in consequence, so offsetting any increase to the money supply (M).

This insight is valuable in that it concentrates attention on the factors underlying the demand for money, but the analysis so far developed is incomplete. Just as monetary factors may affect real variables, so the reverse also applies: real factors such as the shape of the consumption function and the marginal efficiency of investment curve may influence the rate of interest.

Assuming initially a private-sector closed economy, equilibrium income occurs when planned savings equal planned investment. The relationship between planned savings and income is given by the consumption function. The marginal efficiency of investment curve shows that the volume of investment is a diminishing function of the rate of interest, rising when the rate of interest falls, and vice versa. Putting the two sides of this sector together, it emerges that fluctuations in incomes and interest rates will cause fluctuations in planned savings and planned investment. Equilibrium is achieved when the level of income and the interest rate stand in a relationship to each other that produces equilibrium between the plans of savers and the plans of investors.

The consumption function and the marginal efficiency of investment comprise the two parts of the 'real sector'. However, it has been seen that the rate of interest, and hence the level of income, may be influenced by monetary factors, in other words by the demand for money and the supply of money. The demand for money is a function of the level of money income and the rate of interest. The transactions demand for money will be high when income is high, and vice versa. The speculation demand for money will be high when the rate of interest is low, and vice versa. The supply of money is determined by the monetary authorities. If the supply of money is reduced, then, as microeconomic analysis suggests, the cost of holding money, represented by the rate of interest (R), will rise. Conversely, if the supply of money is increased, R will fall.

The two sectors, the 'real sector' comprising the consumption function and the marginal efficiency of investment, and the 'monetary sector' comprising the supply of and demand for money, interact together to influence the level of income and the rate of interest. The equilibrium national income and the equilibrium interest rate depend on simultaneous equilibrium in the 'real' and 'monetary' sectors. This means that simultaneous equilibrium requires that planned savings equal planned investment and the demand for money equals the supply of money. If these plans are not consistent, equilibrium cannot exist.

Any change in the factors determining the consumption function, the marginal efficiency of investment, the demand for money and the supply of money, may have implications for the equilibrium level of income and the equilibrium interest rate. Thus, a change in real factors will have implications for the equilibrium interest rate as well as for the equilibrium level of income. A change in monetary factors may influence the level of income as well as the rate of interest. The extent to which such changes influence real or monetary factors depends on the characteristics of the functions of consumption, the marginal efficiency of investment, the demand for money, and the supply of money. Let us first consider equilibrium in the monetary sector.

19.2 Equilibrium Interest Rates and National Income Levels – The Monetary Sector

In Module 18 it was seen that, given the level of income, the intersection of the demand curve for money and the supply curve of money would determine the equilibrium rate of interest, but this does not explain what determines the level of income. In order to complete the analysis, it is necessary to analyse the relationship between the demand for money and the level of income. When this is done, money can be fully incorporated in the model of national income that was developed earlier.

Module 18 also considered the demand curve for money given a certain level of income. Now, if the level of income is allowed to vary, then the demand for money for transactions plus precautionary purposes will be directly related to the level of income, being high when income is high and low when income is low. For simplicity's sake, we shall assume a linear relationship between the demand for money transactions plus precautionary purpose (M_{T+P}) and the level of national income (Y). This is shown in Figure 19.1.

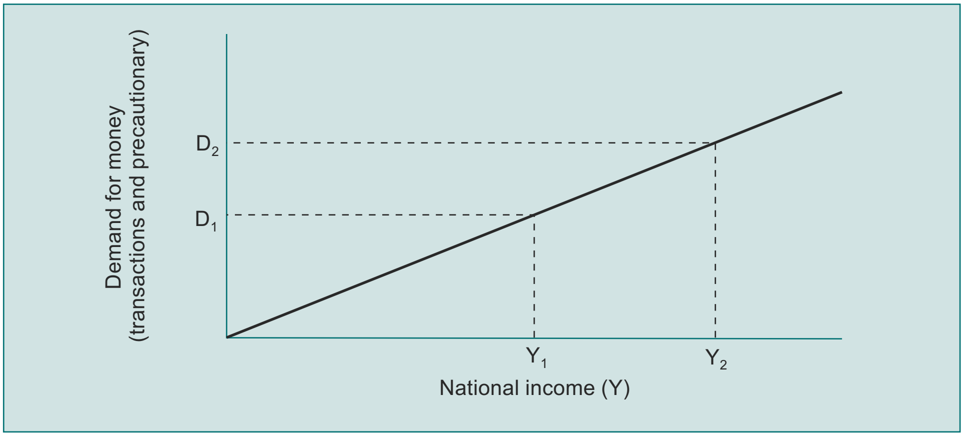


Figure 19.1 Demand for money (transactions and precautionary) and national income

When Y $=$ $Y_1, M_{T+P} = D_1$
When Y $=$ $Y_2, M_{T+P} = D_2$

Let us now consider the demand for money for speculative purpose M_S. The higher the rate of interest (R), the smaller will be the amount held for speculative purposes. Why? Because money held for speculative purposes yields no immediate return; people only hold it waiting for R to rise. But holding money has an opportunity cost, namely the interest forgone; therefore, the higher is R, the less money will be held. When R is low, however, people will be unwilling to commit their speculative holdings since money committed will not be available if R rises. Thus at some low rate we shall assume all speculative money is held in idle balances. Such a situation is shown in Figure 19.2. For interest rates greater than R_3, no speculative balances exist and all money is committed to bonds. For an interest rate slightly lower than R_3, for instance R_2, a few individuals hold money for speculative purposes (M_{S2}) hoping R will rise. For interest rates lower than R_1 in our diagram, all speculative money is held in idle cash balances.

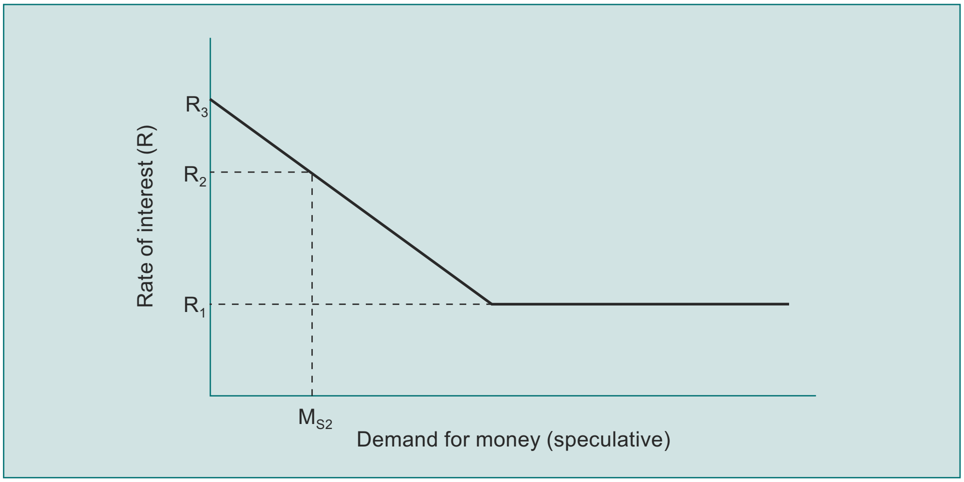


Figure 19.2 Demand for money (speculative) and rate of interest

Now consider Figure 19.1 and Figure 19.2 together. If we know the level of national income/output (Y), we can calculate from Figure 19.1 the demand for money for transactions plus precautionary purposes (M_{T+P}). If, further, we know the supply of money, M, as determined by the monetary authority, we shall know the amount of money available for speculative purposes: it will be M – (M_{T+P}). But from Figure 19.2 we know that each interest rate (R) – with the exception of the horizontal 'trap' – is associated with a given demand for M_S, e.g. R_2 is associated with M_{S2}. Thus for any given Y there is one and only one R for which the total demand for money, M_{T+P} + M_S, equals the money supply M. This relationship is shown in Figure 19.3.

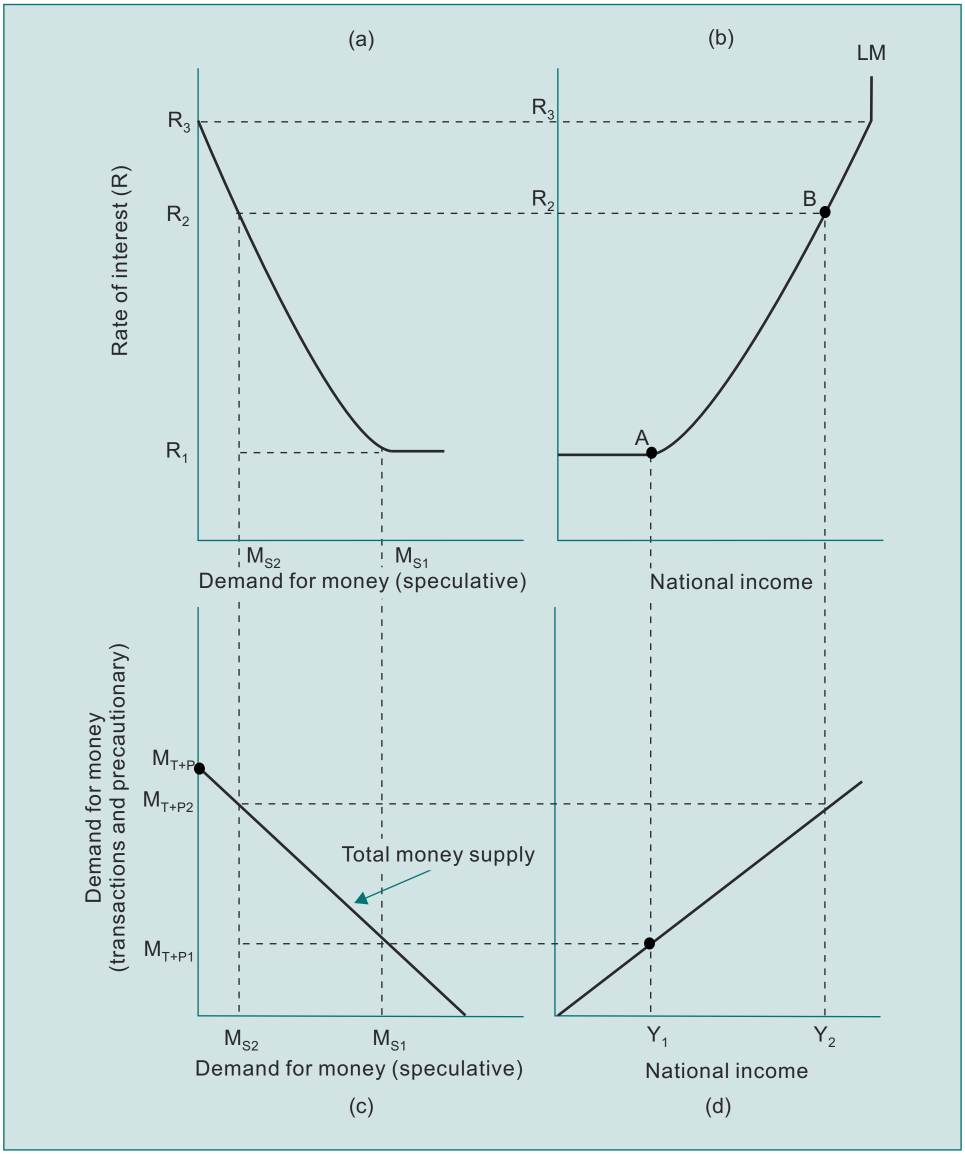


Figure 19.3 The liquidity–money (LM) curve

At income level Y_1, the amount of money required for transactions and precautionary purposes is M_{T+P1}. The figure in the lower-left quadrant (Figure 19.3c) represents the money supply. Since M_{T+P1} is the amount required for transactions plus precautionary purposes, the amount left for speculative purposes is M_{S1}. The only interest rate (R) that equates the demand for speculative money with the available supply M is R_1 (Figure 19.3a). Therefore, point A in Figure 19.3b represents the rate of interest R_1 that produces equilibrium in the monetary sector, where the level of income is Y_1. Similarly point B represents equilibrium for Y_2 and R_2. This figure in the upper right-hand quadrant (Figure 19.3b) is known as the *liquidity–money curve* or the *LM curve*.

For interest rates > R_3 the LM curve is a vertical line, indicating that all money is used for transactions purposes and no money is held for speculative purposes. For interest rates < R_1 all money not used for transactions and precautionary purposes is held as idle balances.

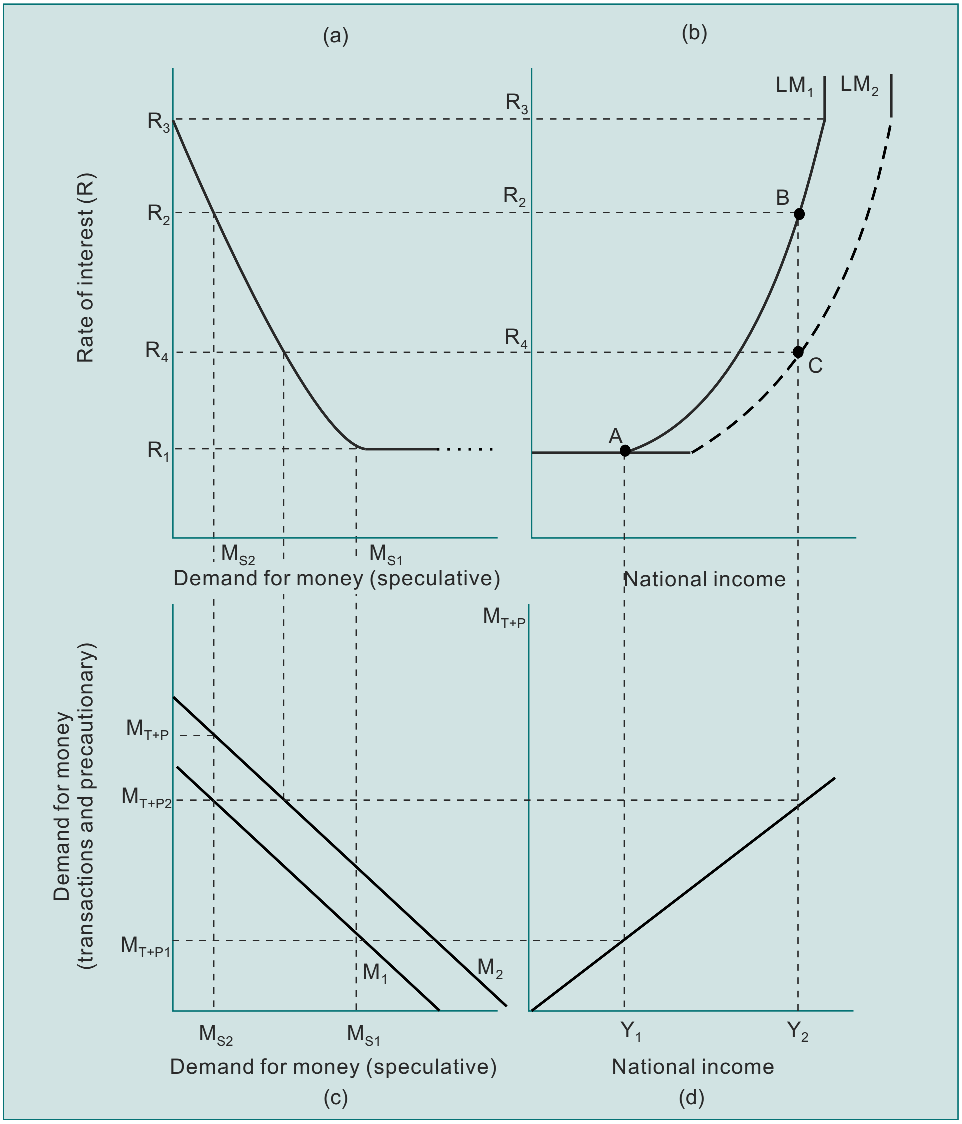


Figure 19.4 Increase in the money supply: a shift in the LM curve

What happens when the monetary authority increases the money supply? This is shown in Figure 19.4. The shift in the money supply is represented by a right shift in the 45° line in the lower left-hand quadrant (Figure 19.4c). Point B in Figure 19.4b is taken from Figure 19.3 with the money supply set at M_1. The increase in the money supply from M_1 to M_2 produces a new set of equilibrium points such that it shifts the LM curve. LM_2 traces out the new equilibrium points between Y and R for a money supply M_2; the new equilibrium value of R for income level Y_2 is R_4, indicated by point C on Figure 19.4b. R_4 is less than R_2. Note that the increase in the money supply does not lower the interest rate for low levels of income. This is the Keynesian liquidity trap, where there is a limit below which the monetary authority cannot lower the interest rate.

Let us concentrate on the LM curve from Figure 19.4b, which is reproduced in Figure 19.5. The initial curve is LM_1, and R_1 and Y_1 represent the equilibrium values of the rate of interest and income. An increase in the money supply shifts the liquidity–money curve to LM_2. If the level of income is held fixed at Y_1, then the rate of interest must fall to R_2 to maintain equilibrium in the money market. Alternatively, if the rate of interest is held fixed at R_1, then the level of income must rise to Y_2 to ensure equilibrium in the money market.

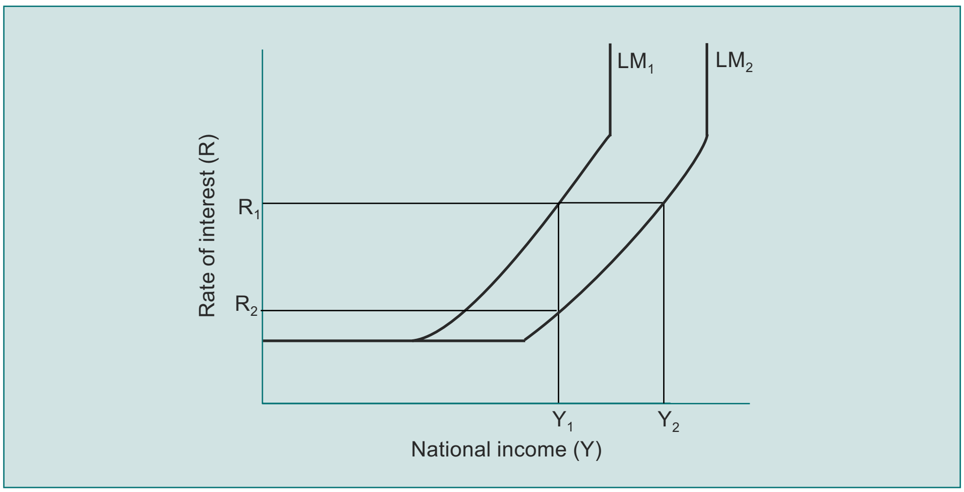


Figure 19.5 The LM curve shifted

Given a particular LM curve, the level of income is determined once the rate of interest is known, or the rate of interest is determined once the level of income is known. But to determine the rate of interest and the level of income simultaneously a further curve is necessary; the LM curve provides one half of the picture. The investment saving (IS) curve, which we are about to develop, provides the missing portion. The intersection of the two curves determines the equilibrium rate of interest and the equilibrium level of national income that are consistent with equilibrium in the money market and with equilibrium in the circular flow of national income.

19.3 Equilibrium Interest Rates and National Income Levels – The Real Goods Sector

Our earlier discussion provided some clues as to the nature of the analysis now developing. This suggested that low interest rates would tend to be associated with high levels of aggregate demand, particularly through encouraging private investment, and would result, through the multiplier process, in an expansion of national income (Y) which was some multiple of the initial increase in demand. On the other hand, high interest rates would restrain aggregate demand and be associated with low levels of income. This type of relationship underlines the IS curve.

This IS curve is developed initially by returning to the model of the private-sector closed economy. The analysis will be extended later to incorporate the government sector and international trade, but since these introduce complexities we shall ignore the government and international sectors in this initial exploratory stage. The introduction of the government sector and international trade does not change the fundamentals of the analysis now presented.

In a private-sector closed economy, it was shown that the equilibrium level of the national income is achieved when planned investment (I), equals planned savings (S). The level of investment is inversely related to the rate of interest, as shown by the marginal efficiency of investment curve that slopes downwards from left to right, while the level of savings is an increasing function of the level of income, as shown by the consumption function. Up to now we have concentrated on the consumption function in our national income/output Figures. A similar result, however, can be achieved by using the savings function.

In our simple model there were only two things that households could do with income: spend it or not spend it. The act of not spending is saving, and then $Y \equiv C + S$. Recall also $C = bY$, where b is the marginal propensity to consume (MPC). The MPC is the proportion spent of an additional unit of income. Thus, if MPC = 0.75, you would spend \$0.75 of an additional \$1 of income. It follows that you would save \$0.25 of that additional dollar and so 0.25 would be the value of your marginal propensity to save (MPS): MPC + MPS must equal 1. And, since $S \equiv Y - C$ and $C = bY$, it follows that $S = (1 - b)Y$, i.e. $b(\text{MPC}) + (1 - b)(\text{MPS}) = 1$.

Figure 19.6 shows how to derive the various equilibrium points that exist between national income and the rate of interest in the real sector. Figure 19.6a shows the relationship between the level of income and the level of savings. This is simply derived from the consumption function and shows, for each level of income, the amount not spent. Thus at each level of income, the level of savings is given by $S = (1 - b)Y$ where b is the marginal propensity to consume. In Figure 19.6b the level of savings is equated to investment by the line $S = I$. Following through to Figure 19.6c, the rate of interest consistent with each level of investment is given by the MEI curve. Finally, in Figure 19.6d the level of income is plotted against the rate of interest. The line IS therefore shows the various combinations of R and Y that will make the planned investment of businesses consistent with the planned savings of households. Hence the label 'IS': investment–saving.

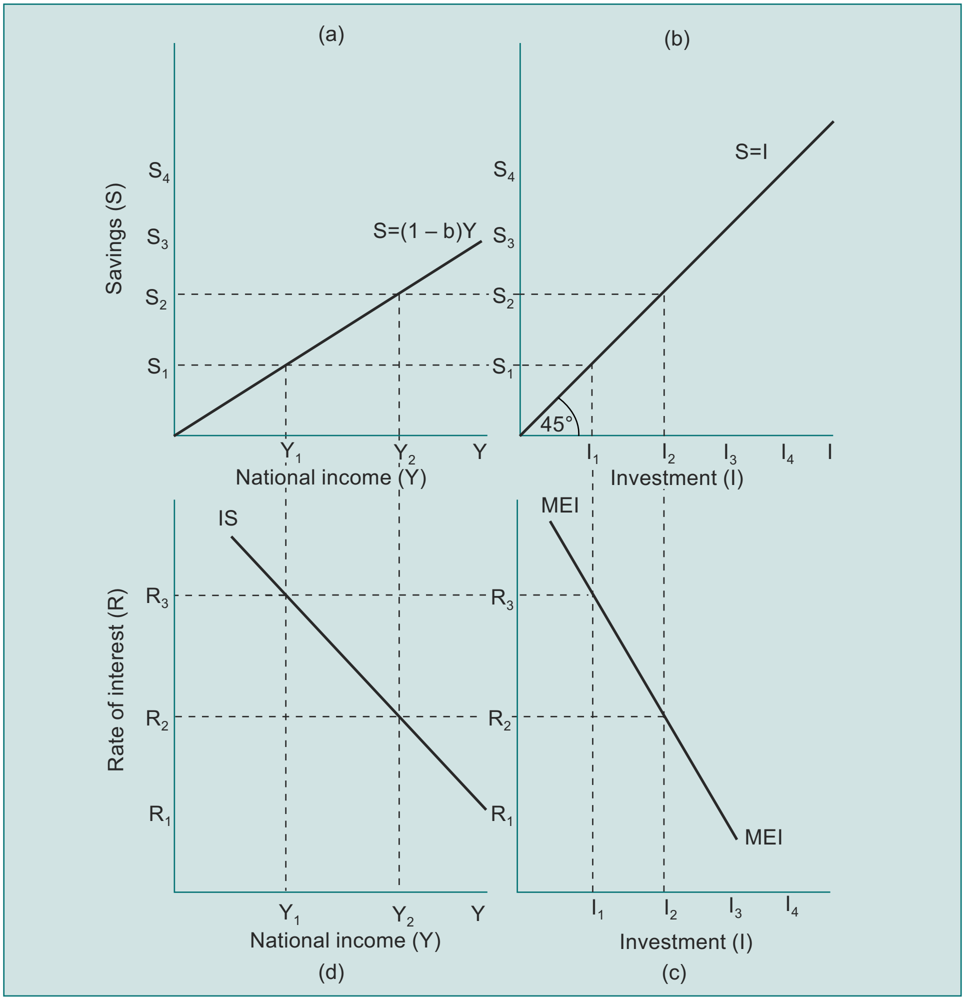


Figure 19.6 The IS curve

The IS curve is interpreted as follows. At low levels of income (Y) the volume of planned savings is low and a high rate of interest is necessary to ensure equality between planned investment and planned savings. Conversely, at high levels of income (Y) the volume of planned savings is high and a low rate of interest is necessary to ensure a high volume of planned investment so as to maintain equilibrium. Looking at this from a rather different angle, the lower the rate of interest the higher the equilibrium level of national income, and the lower the level of income the higher the equilibrium rate of interest.

Underlying the IS curve are 'real' factors, namely the marginal efficiency of investment and the consumption function. For example, if the marginal efficiency of investment were interest-inelastic so that changes in interest rates brought relatively small changes in the volume of investment, and if the marginal propensity to consume were low, then the IS curve would be relatively steep. On the other hand, if the marginal efficiency of investment were interest-elastic, so that changes in interest rates brought forth relatively large changes in the volume of investment, and the marginal propensity to consume were high, that combination would yield an IS curve that was relatively flat.

The two alternatives are depicted by IS_1 and IS_2 in Figure 19.7. For any given change in the rate of interest, the steeper curve, IS_1, shows a much smaller change in income (Y) than the flatter curve IS_2. The reasoning underlying this is that if the marginal efficiency of investment curve is interest-inelastic and the marginal propensity to consume is low (that is the multiplier is small) then a given change in the rate of interest will produce only a small change in the volume of investment and, with a small multiplier, only a small change in income. The more elastic the marginal efficiency of investment and the higher the multiplier, the flatter the IS curve, i.e. the greater the change in investment that will accompany any change in interest rates and the greater the subsequent change in income because of a high multiplier.

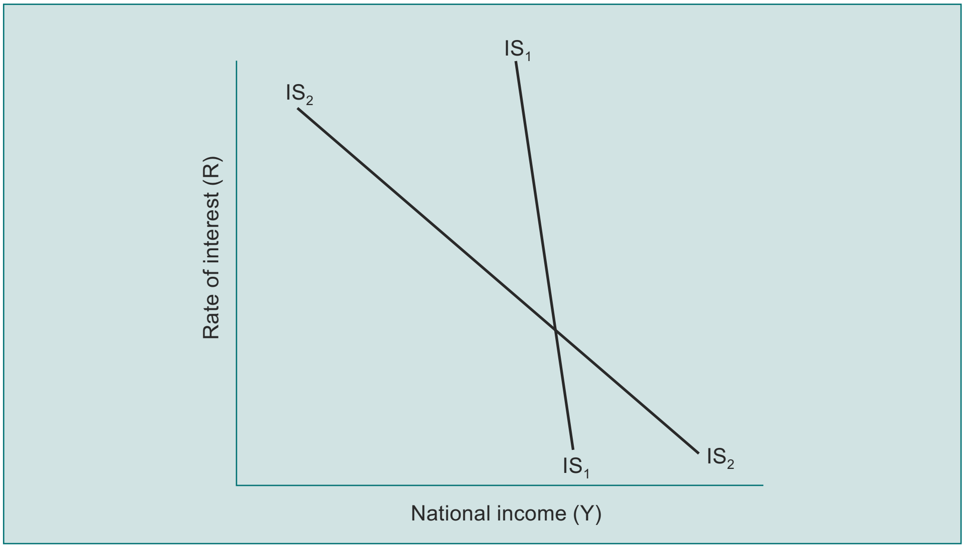


Figure 19.7 The IS curve : alternative forms

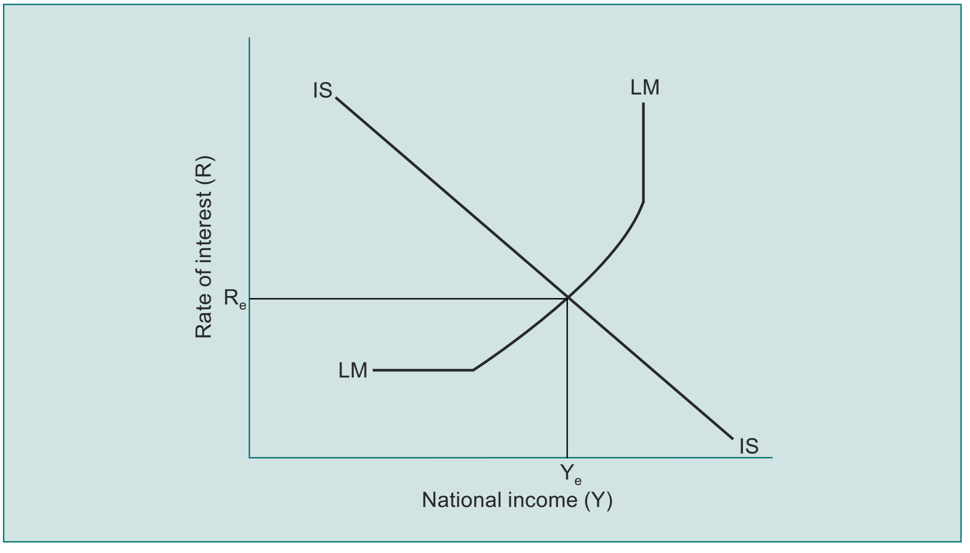


Figure 19.8 IS/LM: equilibrium in the real and monetary sectors

The LM curve shows the equilibrium level of R given Y, but it cannot show what determines Y. The IS curve shows the equilibrium level of Y given R, but it cannot show what determines R. Now, it can be shown that equilibrium R and equilibrium Y are determined at the intersection of the LM and IS curves. This is demonstrated in Figure 19.8. The LM curve slopes upward to the right, showing that at high levels of income higher rates of interest are necessary to ensure that the demand for and supply of money are in equilibrium. The IS curve slopes downward to the right, indicating that at higher levels of income low rates of interest are necessary to ensure that planned savings and planned investment are in equilibrium. At the point of intersection, we shall have a level of national income Y and a rate of interest R where the demand for money (liquidity) and the supply of money are in equilibrium (equilibrium in the monetary sector), and planned savings and planned investment are in equilibrium (equilibrium in the real goods sector). Equilibrium in both these sectors means that no forces exist to change the rate of interest (R_e) or the level of income (Y_e).

It is now a relatively straightforward matter to expand our IS curve to include a government sector and an international sector.

19.4 The Expanded Model: Shifting the Curves

Figure 19.9 shows the derivation of the IS curve to include government expenditures. The only differences between Figure 19.6 and Figure 19.9 occurs in parts (b) and (d) respectively.

Savings (S) must now accommodate private investment expenditure (I) and government expenditure (G). Income level Y_2 generates savings of S_2. Since S now equals G + I, the amount of savings available for private investment is S – G. Given S = S_2 and G = G, the amount of investment expenditure associated with an income level Y_2 is I_1. From the MEI schedule we see from Figure 19.6c that I_1 is associated with an interest rate R_3. Thus, (Y_2, R_3) is one point on the new IS curve IS_2. IS_2 traces out the set of equilibrium points relating Y to R when G = G. In other words, the increase in government expenditure from 0 to G shifts the investment–savings curve from IS to IS_2.

If we were to include taxes, the dotted line S = G + I would be lowered towards the 45° (S = I) line of Figure 19.9b, and the value of G would reduce. Similarly if we were to add a trade surplus, where X > Z, the distance G would increase, whereas if X < Z the distance G would decrease. To these factors we shall return. What is important for now is that you understand how the IS curve is derived and what it represents.

Let us now investigate how we can use our new-found wisdom and tools to 'manipulate' the economy. You are about to become a policy maker!

The IS/LM approach integrates 'monetary' factors (the demand for and supply of money) and 'real' factors (consumption, investment, government expenditure and net exports) within one analysis. This allows an investigation of the interaction between monetary and real variables, which together determine the equilibrium rate of interest and the equilibrium level of national income. This basic analysis may be used to investigate the effects of policy instruments.

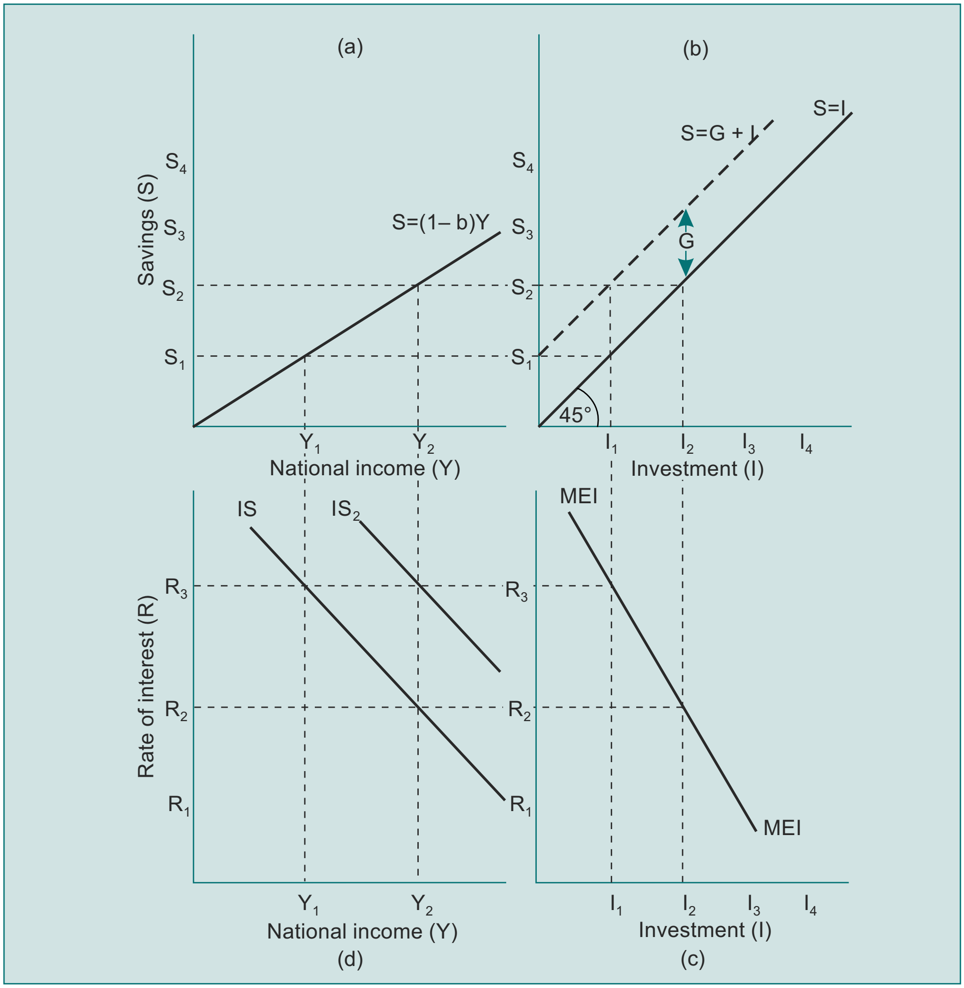


Figure 19.9 Shift in the IS curve

Suppose in Figure 19.10 the IS/LM curves are initially IS_1 and LM_1 and intersect to give an equilibrium level of interest of R_e and an equilibrium level of income of Y_e. Suppose also that this equilibrium level of income is less than full-employment income, which is given by Y_f. That is, there is a deflationary gap and the equilibrium level of income is less than full-employment income, with resulting unemployment and a level of income less than that which could be achieved at capacity output.

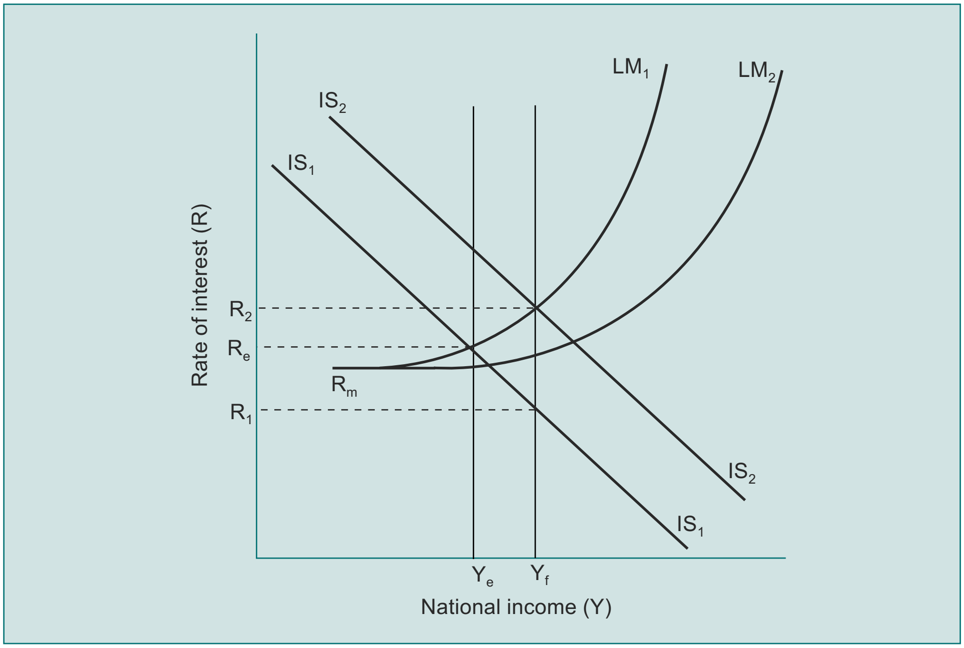


Figure 19.10 IS/LM: the liquidity trap

In Figure 19.10, it would not be possible to reach full-employment income through the use of monetary policy alone because the equilibrium rate of interest consistent with full-employment income is R_1, which is below the minimum rate of interest R_m set by the horizontal portion of the LM curve. This is the situation envisaged by Keynes where the liquidity trap prevents monetary policy from pushing the rate of interest low enough to encourage sufficient investment to obtain full-employment income.

However, given LM_1, it would be possible to reach Y_f by a shift in the IS curve from IS_1 to IS_2. The new equilibrium income, now Y_f, would necessitate, given a fixed supply of money and a fixed LM curve, a higher equilibrium rate of interest of R_2. Such a shift in the IS curve could be achieved in a number of ways: first, if there were a rightward shift in the marginal efficiency of investment curve, causing a rise in the volume of investment at any rate of interest. Second, it could be achieved by an upward shift in the consumption function, which would increase the level of consumption at any level of income. Finally, a rightward shift in the IS curve could be achieved by creating a budget deficit through increasing government expenditure or cutting taxes.

Notice that if the rightward shift in the IS curve were accompanied by a rightward shift in the LM curve from LM_1 to LM_2 due to an increased supply of money, then Y_f could be achieved with a less marked rightward shift in the IS curve because the increase in the money supply would tend to lower the rate of interest associated with any level of income and hence encourage a higher volume of investment and a rightward movement along a given IS curve.

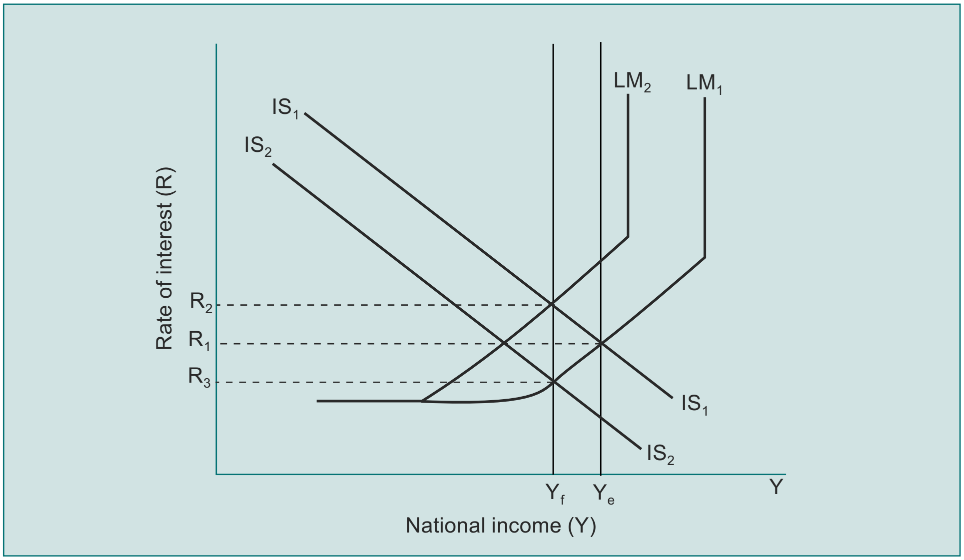


Figure 19.11 IS/LM: dealing with an inflationary gap

An inflationary gap is depicted in Figure 19.11. In this case, the initial IS and LM curves, IS_1 and LM_1, intersect to give an equilibrium rate of interest of R_1 and Y_e, the equilibrium level of income being to the right of full-employment income Y_f. Thus there is an inflationary gap, with the level of aggregate demand greater than that necessary to sustain full employment, with consequential effects in raising the level of prices. If it is desired to reduce the aggregate demand to obtain a level of income of Y_f, then this can be done by a leftward shift in the LM curve and/or a downward shift in the IS curve. For example, if the chosen instrument is monetary policy, then a reduction in the money supply will shift the LM curve to the left. If the curve shifts from LM_1 to LM_2 then, given the IS curve of IS_1, the equilibrium level of income will fall from Y_e to Y_f, this being accomplished by an increase in the equilibrium rate of interest. As the rate of interest rises, this reduces the volume of investment, lowers the level of aggregate demand, and causes a leftward shift along a given IS curve. If the decision is instead made to use fiscal policy, then a decrease in government expenditure and/or increase in tax, creating a budget surplus, will shift the IS curve to the left. The curve would also shift leftwards, of course, if there were a leftward shift in the marginal efficiency of investment curve. If the IS curve shifts downwards to the left to IS_2, then the equilibrium income falls from Y_e to Y_f and, with a given LM curve such as LM_1, this is achieved at an equilibrium interest rate of R_3.

Clearly, shifts in the IS and/or LM curves will have implications for the equilibrium levels of income and interest, and it is worth summarising the factors that might cause these shifts and their consequential effects. An upward shift in the marginal efficiency of investment curve (giving a higher volume of investment at each rate of interest) and/or an upward shift in the consumption function will shift the IS curve to the right and give higher equilibrium values of both Y and R in the intermediate range of the LM curve. Conversely, within the intermediate range of the LM curve, a downward shift in the consumption function will shift the IS curve to the left and give lower equilibrium values of both Y and R. The same results will flow from the application of functional finance, a budget deficit being equivalent to a rightward shift in the IS curve and a budget surplus to a leftward shift in the IS curve.

How effective fiscal and monetary policy will be will depend critically on the shapes of the IS and LM curves and the starting position, i.e. where the level of Y is and what R is when the policy decision is made. Figure 19.12 shows the potential impact of fiscal policy in three scenarios. In each diagram the IS curve shifts to the right from IS_1 to IS_2 because of, say, an increase in government expenditure.

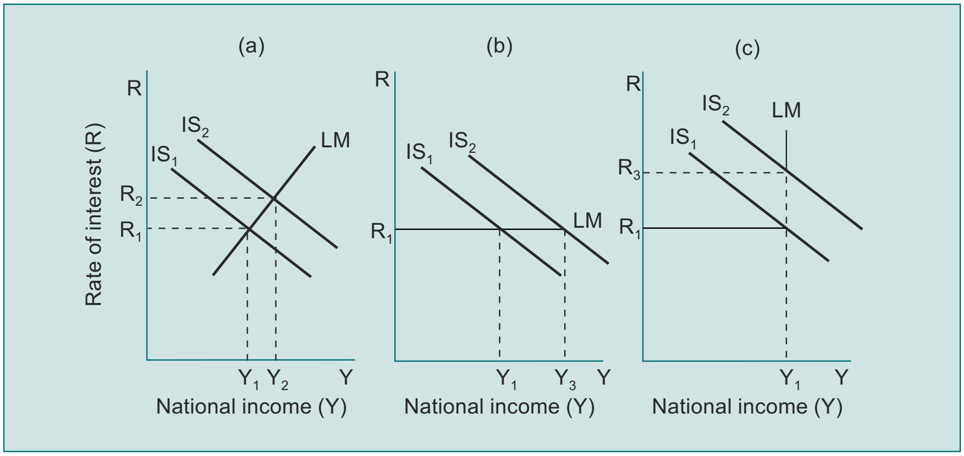


Figure 19.12 Fiscal policy: impact on national income/output

The effects of national income/output (Y) and the rate of interest (R) depend on the slope of the LM curve. In Figure 19.12b – the liquidity trap of Keynesian fame – national output expands by the full multiplier effect, but R remains constant at R_1. In Figure 19.12c, R increases from R_1 to R_3 but Y remains at Y_1. How can this happen, given the increase in G? The increase in G has been offset by an equal and opposite decrease in investment (plus consumption) expenditure, where there has been crowding out due to the increase in R. Figure 19.12a, the normal case, can be analysed in slightly greater detail, and is repeated as Figure 19.13.

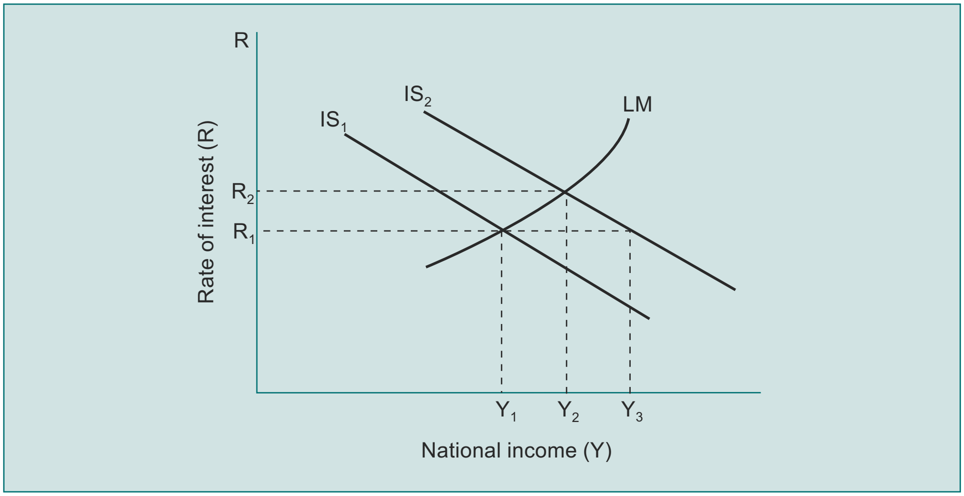


Figure 19.13 Figure 19.12a (expanded): the standard case

The impact of the shift in the IS curve from IS_1 to IS_2 is to raise both R and Y. R increases from R_1 to R_2; Y increases from Y_1 to Y_2. The increase in Y from Y_1 to Y_2 can be considered in two parts: Y_1 to Y_3 and Y_3 to Y_2. The increase in Y from Y_1 to Y_3 is the increase that would have occurred had R remained at R_1 (the case of Figure 19.12b with a horizontal LM curve). Thus the decrease in Y from Y_3 to Y_2 must be the decrease due to R increasing from R_1 to R_2, i.e. a partial crowding out.

Let us now consider the potential impact of monetary policy in three scenarios as shown in Figure 19.14. In each of Figure 19.14a, b and c, the LM curve shifts to the right from LM_1 to LM_2 because of an increase in the money supply. The effects on national income/output (Y) and the rate of interest (R) depend on the slope of the IS curve.

In Figure 19.14b, the horizontal IS curve would represent the very odd case of complete interest elasticity with respect to investment expenditure, such that firms undertake all their investment expenditures at interest rate R_1 (but no other). Thus the increase in the money supply leaves R unaffected but increases Y from Y_1 to Y_2. While this example may stretch your credibility, you could imagine a very small decrease in R bringing forth an enormous increase in I and consequently a large increase in Y (i.e. Y_1 to $Y_{2)}$.

Figure 19.14c is easier to visualise as a real-world example. Investment expenditure is completely inelastic with respect to R. The shift in the LM curve causes R to decrease from R_1 to R_2, but since no additional investment (or consumer) expenditure is forthcoming there is no change in Y.

Figure 19.14a represents the normal situation, where the decrease in R from R_1 to R_2 stimulates investment expenditure (and consumer expenditure) and, through the multiplier process, causes an increase in Y from Y_1 to Y_2.

Figure 19.14c belongs to the extreme Keynesians; Figure 19.14b belongs to the extreme monetarists.

This just leaves us with the price level and the employment level or, if you prefer, inflation and unemployment, to consider. These are the topics of Module 20.

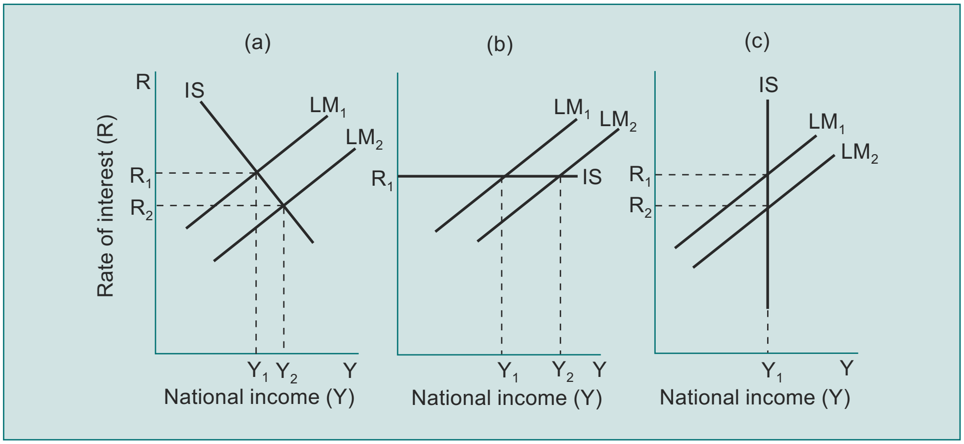


Figure 19.14 Monetary policy: impact on national income/output

Learning Summary

This module pulls together much of the analysis of earlier modules. You are now in a position to understand how equilibrium is determined in the monetary sector and how each level of national income is associated with a unique interest rate. You can trace this connection through the separable demand functions for money and money supply. In a similar vein, you should understand the connection between different levels of income and interest rates in the real goods market and are able to trace the interconnection through the savings function and marginal efficiency of investment function.

By combining the monetary and real goods sectors, you can see how an interaction between the two produces a level of income and a rate of interest that together achieve equilibrium in both sectors simultaneously.

The macroeconomic model is then expanded and you can analyse how the policy tools at the disposal of governments cause the IS and LM curves to shift, producing new equilibrium outputs. You have also become aware that the relative efficiency of fiscal and monetary policies is critically affected by the existing state of the economy, that is, by the existing positions of the IS and LM curves.

Review Questions

Multiple Choice Questions

19.1 Which of the following statements describes the liquidity trap?

A. An increase in the money supply lowers the interest rate and stimulates investment expenditure.

B. An increase in the money supply has no effect on the interest rate and aggregate demand.

C. An increase in government expenditure increases aggregate demand and household savings and lowers the interest rate.

D. A decrease in government expenditure decreases aggregate demand and household savings and increases the interest rate.

19.2 It can be stated that

I. the LM curve represents equilibrium values of the rate of interest and national income.

II. at any point on the LM curve, the government budget is balanced.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

19.3 If the IS curve is almost vertical, it means that

I. the marginal propensity to consume is large.

II. the marginal efficiency of investment curve is elastic.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

19.4 Assuming the IS and LM curves do not intersect in the horizontal or vertical portions of the LM curve, which of the following would happen to the interest rate (R) and the level of national income (Y) if an increase in the money supply were accompanied by a decrease in taxes?

A. R would decrease; Y would increase.

B. R would decrease; Y could increase, remain constant or decrease.

C. R could increase, remain constant or decrease; Y would increase.

D. R could increase, remain constant or decrease; Y would decrease.

Questions 19.5–19.7 are based on Figure 19.15.

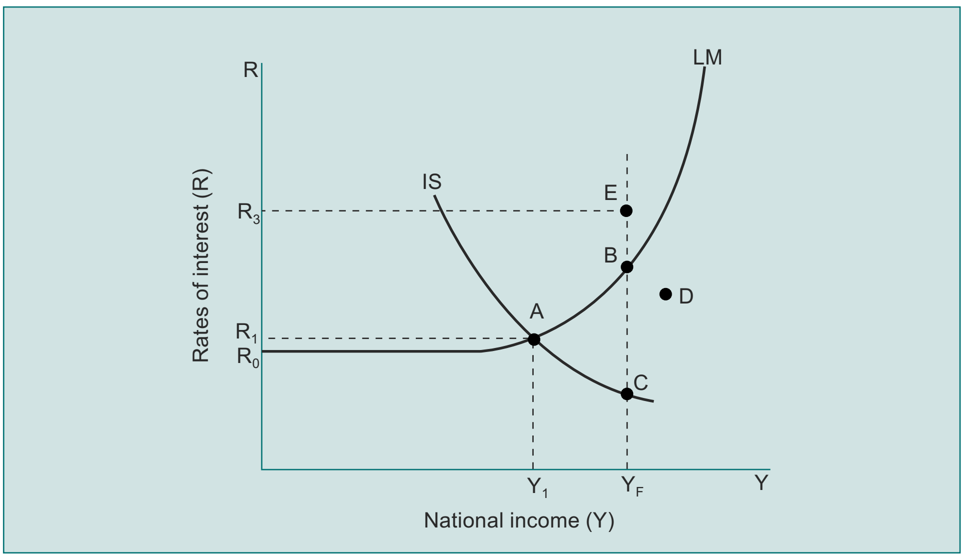


Figure 19.15 Sample IS and LM curves

Notation: The curves intersect at A, with a national income level Y_1 and interest rate R_1. R_0 is the liquidity trap of interest and Y_F is the full-employment level of national income.

19.5 Which of the following is correct in relation to Figure 19.15?

A. To achieve full employment using only monetary policy, the money supply would have to be increased by the amount BC.

B. To achieve full employment using only monetary policy, the money supply would have to be increased by the amount AC.

C. To achieve full employment using only monetary policy, the money supply would have to be decreased by an amount BC − AB.

D. It is not possible to achieve full employment just by increasing the money supply.

19.6 Which of the following statements is correct with respect to point D in Figure 19.15?

A. Point D cannot be reached because it implies a level of national income greater than full-employment national income.

B. Point D can be achieved, but it would produce an inflationary gap.

C. Point D can be achieved by a combination of tax cuts and increases in government expenditure with the money supply constant.

D. Point D can produce equilibrium in both the money and real goods markets by an increase in the money supply with no changes in fiscal policy.

19.7 On the basis of Figure 19.15, the government desires to achieve full employment Y_F, to raise the rate of interest to R_3, and to have equilibrium in both the real goods and money markets. Which of the following policies could achieve all of these goals?

A. An increase in government expenditure and increase in the money supply.

B. A decrease in taxes and a decrease in the money supply.

C. An increase in government expenditure and decrease in taxes.

D. A decrease in the money supply.

19.8 Government economists calculate that the value of the government expenditure multiplier is 4. The government increases expenditure by $1 million but the resultant increase in national income is only $3 million. The shortfall can be explained by

I. an increase in the demand for money.

II. an increase in potential output.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

19.9 The following can be stated:

I. An increase in the money supply must always cause national income to rise.

II. A decrease in taxes must always cause the rate of interest to rise.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

19.10 An economy is in equilibrium in both the real goods and money markets. The government decides to increase both taxes and government expenditure. At the same time trade, which was balanced, now shows a trade surplus. What is the impact of these events combined on the IS and LM curves?

A. IS shifts to the right, LM shifts to the left.

B. IS shifts indeterminately, LM curve unaffected.

C. IS shifts to the left, LM curve unaffected.

D. IS shifts indeterminately, LM shifts indeterminately.

Module 20

Inflation and Unemployment

Contents

20.1 Introduction

The IS/LM approach has provided us with a tool for integrating the monetary and real goods sectors of the economy. The demand for and supply of money, underlying the LM curve, interact with real factors – the consumption function, the marginal efficiency of investment curve, net government expenditure, and net exports – to determine equilibrium rates of interest and equilibrium levels of income.

But something else almost 'tumbles out' of the analysis, and that is the level of employment or, conversely, the rate of unemployment. If we know the equilibrium level of national income (Y) and also the potential output (Q), i.e. the capacity national income, then we also know the unemployment rate. If Q = Y, the unemployment rate will be the full-employment unemployment rate, where unemployment will be confined to frictional, structural and seasonal unemployment. If Y < Q, there will be demand-deficient unemployment, which can be caused by fiscal and/or monetary policy. Similarly, if Y > Q, deflationary policies of reduced government expenditure (G), increased taxes, and/or a reduction in the money supply (M) can take us back to full employment. Of course, since Q generally grows over time in the real world of dynamic economies, we may only be talking about a decrease in the rate of growth of G or M in the Y > Q example. What does *not* 'tumble out' of the analysis, however, is the price level – which is, of course, the rate of inflation.

Economists who adhere to the Keynesian school believe that monetary factors are critical in determining equilibrium income/output and equilibrium interest rates, but they do not ascribe a central role to monetary factors in determining the price level. For many economists, the Keynesian model had a great revival and became 'flavour of the month' throughout the early 1960s with the advent of the Phillips Curve. The Phillips Curve was the missing link in the Keynesian structure since it purported to show a fixed relationship between the unemployment rate and the inflation rate, albeit with a lag. Thus, given the equilibrium level of national income, the unemployment rate was determined; given the unemployment rate, the inflation rate was determined. The Phillips Curve, therefore, presented policy makers with a painful choice: inflation or unemployment – for there was a trade-off. This is shown in Figure 20.1 for a hypothetical economy.

Figure 20.1 A Phillips Curve

According to the Phillips Curve in Figure 20.1, you as policy maker can choose where you would like the economy to be. If you want zero inflation, you have to suffer an 8 per cent unemployment rate. If you want full employment, say a natural unemployment rate of 4.5 per cent, then you will have to live with a 2 per cent rate of inflation. If you drive the economy's unemployment rate down to 3 per cent, you will have to live with an inflation rate of 10 per cent. The Phillips Curve has had great appeal for many politicians and government decision makers because economists, at last, had produced a diagram that most politicians could understand!

As you might imagine, not all economists agreed that the nation's macroeconomic problems had been solved. Those who argued that there are strong equilibrating forces in market economies, which, if left unhampered by governments and other disruptive forces, would lead to full-employment national income, maintain that money is of prime importance, *not* in determining real variables such as the level of employment and national income (the province of market forces) but in determining the price level. Not surprisingly, these economists are known as monetarists, disciples of Milton Friedman, a long-term opponent of the Keynesian school.

We shall now examine the quite separate and distinct explanations of price determination that have emerged in the past five decades, because they have important implications for the efficacy of fiscal and monetary policy.

20.2 Causes and Effects of Inflation

In the past five decades, all of the industrialised economies of the Western world have experienced inflation – although the rate of inflation has varied substantially over time and, in any given period, between different economies. The persistence of inflation and a tendency for the rate of inflation to rise for substantial periods has resulted in a situation where great weight has been placed on the prevention or reduction of inflation. In most capitalist economies, the reduction of inflation had become of greater priority than the maintenance of a low rate of unemployment. This immediately raises the question: why should the avoidance of inflation be a goal of economic policy and what effects will flow from the continuation of 'high' inflation?

Price inflation may have undesirable social and economic effects because it impairs the efficiency of markets and redistributes income and resources in an unexpected and largely capricious fashion so long as the inflation is unanticipated. If inflation continues, then presumably behaviour patterns will anticipate that inflation and so mitigate the redistribution of income and resources. Yet some redistribution will still take place because of imperfect knowledge and because some social and economic groups do not have the same possibilities for adjustment to inflationary pressures. Moreover, the more successfully inflation is anticipated, the more rapid the rate of inflation is likely to become.

First, inflation impairs the efficiency of the price mechanism and raises the costs of buying and selling because money becomes less reliable as a standard of value. The greater the general rate of price inflation and the greater the shift in relative prices, the less is the reliance that can be placed on money as a standard of value. For example, suppose that in a period of rapid general inflation and substantial shifts in relative prices, a shopper observes that commodity X has risen by 20 per cent in price. In order to determine whether to buy X, the shopper would have to collect information on all relevant commodity prices then ruling so as to arrive at a view as to whether the observed 20 per cent increase in the price of X represents general inflation or the fact that X has become more or less expensive relative to substitute goods. These additional search costs are likely to become particularly important where long time periods elapse between decisions, such as normally occurs with consumer durables, such as motor cars, TV sets, and household appliances. Similarly, the seller is likely to have increased difficulty in ascertaining the price of possible substitute goods and the relationship between selling price and the price of required inputs. Without the certainty that money has a stable value, the search costs of buying and selling rise sharply.

Second, inflation penalises people on 'fixed' incomes and favours those whose money incomes adjust quickly to price changes. The former group includes pensioners, university students, and many salary earners, while most wage and profit earners fall into the latter category. Where household incomes are composed of transfer earnings determined by the state, it is always possible to link incomes to price changes so as to protect real income; but the more successfully this is accomplished, the greater the inflationary bias in the economy.

Third, inflation favours borrowers and penalises lenders so long as it is unanticipated. Thus, if interest rates are fixed in money terms in the anticipation that the level of prices will remain constant, an increase in prices will reduce the real cost of borrowing. In extreme cases, the real cost of borrowing may even become negative where the increase in prices is greater than the nominal interest rate. While this is generally favourable to particular types of investment, the anticipation of continued inflation will lead to an upward adjustment of the money rate of interest. Where businesses are considering financing investment with funds borrowed from external sources, this creates uncertainty because, for loans negotiated at high nominal rates of interest, the real cost of borrowing rises if interest rates decline.

Fourth, given a system of unindexed taxes, namely one where tax thresholds are specified in money terms rather than real terms, inflation will redistribute resources from the private to the public sector. For example, it was estimated that for the UK, a 10 per cent increase in all money incomes and prices – which left real incomes unchanged – raised taxes by 15 per cent, so that real disposable personal income fell and tax revenue increased in real terms. Such an effect can be prevented by indexing taxes, which presents no major technical difficulties, but indexation is relatively rare in capitalist countries. If full indexation were attempted, it would increase the inflationary bias of the economy.

Fifth, a continuing higher rate of domestic inflation than that experienced in other economies can lead to increased imports and reduced exports and can create potential problems for stable exchange rates.

Price inflation therefore has important consequences for the distribution of incomes, for the operation of markets, for savers and investors, for the distribution of resources between private and public sectors, and for foreign trade and the balance of payments. If the inflation is unanticipated, then these effects are often capricious and unintended. Continued inflation will lead to an adjustment in behaviour patterns – which can mitigate some of the distributional consequences, but inflation can never be fully anticipated. Full anticipation requires not merely information concerning the aggregate rate of inflation, but requires that all economic agents should have perfect knowledge of all the price movements affecting their decisions. Thus, economic agents have to anticipate the *relative price changes* that accompany any general price inflation, and a failure to anticipate these relative price changes will have distributional effects and impair the efficiency with which markets operate.

Up to the outbreak of World War II, most industrialised economies experienced periods of rising prices, associated with the upswing of the business cycle, and periods of stable or falling prices, associated with the recession phase of the business cycle. This experience was punctuated by examples of high, sustained inflation as a consequence either of major gold discoveries (e.g. Spain in the fifteenth century) or by increased government expenditure financed by printing large sums of paper money (e.g. the German hyperinflation of 1923). Against this background, the inflation experience in the second half of the twentieth century is unique in that there is no historical parallel to such sustained inflation across the whole of the developed industrialised world. Whereas the historical experience has been of periods of rising prices alternating with periods of stable or falling prices, from 1945 there has been sustained inflation. The rate of inflation has varied from year to year, but the rate has (with minor exceptions) always been positive. It is only in the relatively recent past that inflationary fears have subsided as a result of sustained low rates. Figure 20.2 shows inflation rates for some of the major industrialised economies in the past five decades.

The emergence of persistent, widespread price inflation led to a major re-examination of the theory of price inflation and to a number of conflicting explanations of the observed phenomena. At the most fundamental level, the basic distinction is between *demand–pull* and *cost–push* models of price inflation, which lead to quite differing policy conclusions as to how inflation might best be combated.

The demand–pull explanation sees price rises as a consequence of excess demand for goods and services, namely demand that exceeds the capacity output of the economy at current prices. As real output cannot increase significantly beyond potential output, excess demand 'pulls up' the prices of final goods and services. Moreover, as employers compete with each other to obtain scarce factors of production, this is reflected in the bidding up of the prices for factor services, so that money incomes rise. In this process, the source of inflation is excess demand in the final market for goods and services, and this results in rises in wages, salaries and other factor incomes.

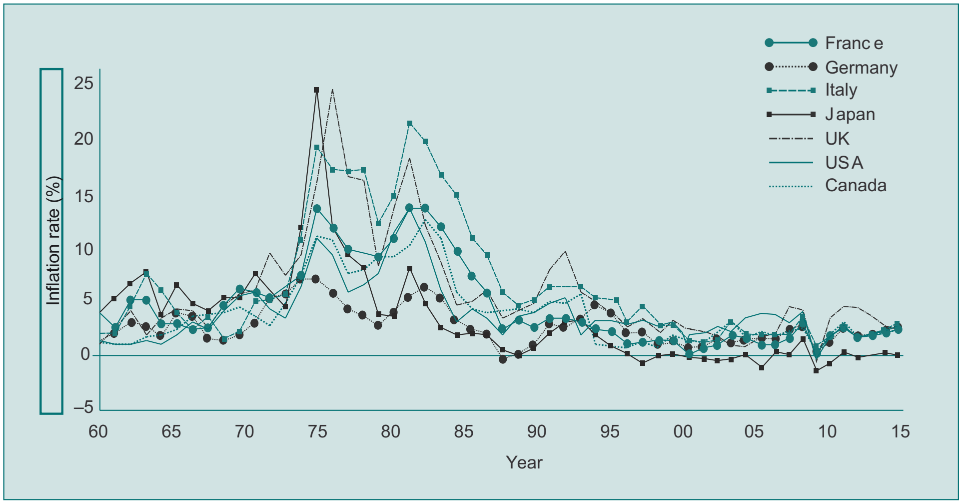


Figure 20.2 Inflation rates for the G7 economies, 1960–2015

The income–expenditure model is represented diagrammatically by the inflationary gap diagram shown in Figure 20.3. The aggregate expenditure function, E, consisting of private consumption, private investment, government expenditure and exports less imports (C + I + G + (X − Z)), gives an equilibrium level of income, Y_e, that is greater than Y_f, full-employment income. It is not possible to reach Y_e by increasing real output beyond Y_f, and the consequence is that the excess of desired real expenditure over the real capacity output (income) of the economy results in price inflation.

There are a number of difficulties with this 'Keynesian' approach to inflation. The analysis does not incorporate monetary factors that may be responsible for causing inflation (as in the major inflations associated with increased gold supplies or the failure to control the supply of paper currency), nor does it deal explicitly with the possible role of monetary policy in curbing inflation, and these omissions appear serious in the light of the development of a 'modern' or 'sophisticated' Quantity Theory of Money. Initially, however, another omission appeared even more serious, namely that the income–expenditure model regarded money wages and salaries as essentially passive, reacting to, rather than causing, price inflation. Wage and salary earners were seen as reacting to price changes by bargaining higher money wages to maintain the value of the real wage, the adjustment in money wages lagging behind the changes in prices, rather than being the cause of the initial change in prices.

However, in response to the development of more centralised wage and salary bargaining machinery in many capitalist economies, a new school of thinking developed that elevated labour markets from a subordinate defensive role to a primary causal role in price determination. Instead of reacting defensively to the consequences of price inflation, so that money wages and salaries lagged behind price changes, this new school saw the labour market as the primary source of the inflationary process. The cost–push explanation of inflation sees price rises as a consequence of bargains struck in the factor market, which raise the production costs of employers, who then pass on higher costs in the form of higher prices. Most cost–push models incorporate the elements described next.

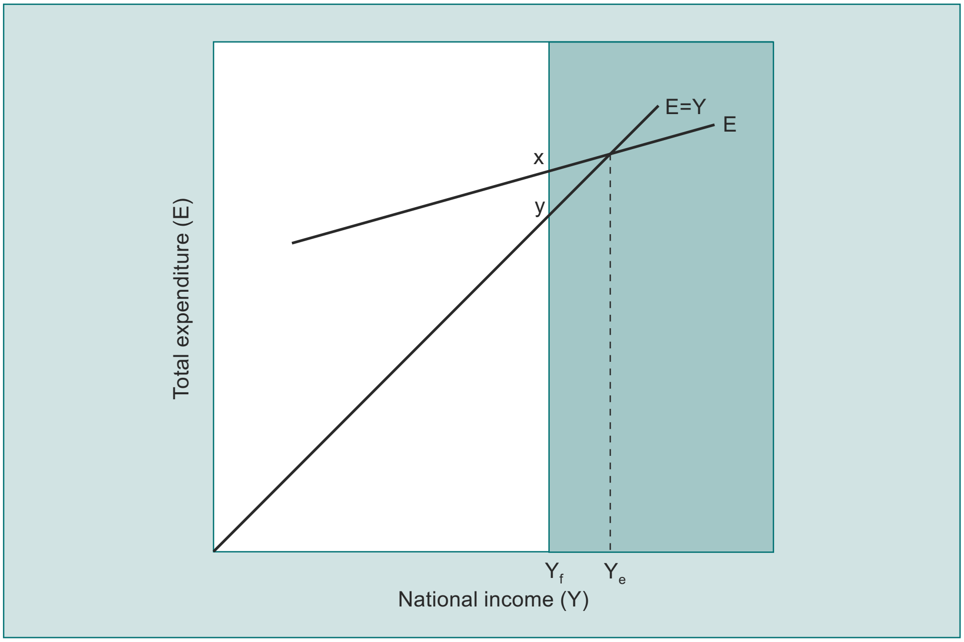


Figure 20.3 Inflationary gap

Prices and costs are 'administered' rather than responsive to the market forces of demand and supply. Outside some competitive markets, such as those for agricultural products, most markets for final products are considered to have strong anti-competitive elements in that one producer, or a small group of producers, has a significant influence in determining price. Similarly, wages and salaries are seen as 'administered' prices, determined by bargains struck between employers and trade unions, rather than as prices responsive to demand and supply factors. And final product prices are 'administered' on the basis that business firms set prices on a cost-plus basis, prices reflecting the full cost of production plus some mark-up for profits. Hence, if costs rise, businesses will attempt to pass on the higher costs in the form of higher prices, so that the economy is placed on a cost-plus basis. Any factor that pushes up costs is likely to result in a subsequent increase in final prices.

In such a system of administered prices, bargaining over money wages and salaries is given primary emphasis as the 'motor' of inflation. The reasons for this are:

(a) the existence of trade unions, whose ostensible objective is to bargain for better pay and money conditions for their members;
(b) the highly centralised nature of much collective bargaining between trade unions and employers' associations; and
(c) the importance of wage and salary costs in the total costs of production.

Consequently, if a union secures an increase in money wages it will, because of the existence of centralised wage bargaining, affect most employers in the relevant industry at the same time and in much the same degree. Business firms finding their profit mark-up eroded will act in unison to pass on the higher costs in the form of higher prices. The increase in prices will erode the real value of the initial change in money wages and may then lead to further demands for increases in money wages. Cost–push inflation originates with higher wage costs, which then push up prices. Price inflation is then generated in the factor market, particularly the labour market, and is then transmitted to the product market.

Cost–push inflation is more likely to occur in economies where prices and wages are not flexible downwards, a feature of many modern economies. It has long been recognised that trade unions and wage groups will resist particularly any cut in money wages. This being so, business firms, faced with a fall in the demand for their product or service, may be reluctant to adjust prices downwards, because the stickiness of money wages would have the result that price reductions would be mainly at the expense of profit margins. In these circumstances, reductions in demand for a product or service will not lower prices but will cause reduced output and employment.

Whereas deficient demand may not cause prices to fall, excess demand will be reflected in higher prices (and wages). In short, it is quite possible that the reaction of wages and prices is asymmetrical. Excess demand will raise prices and wages, but deficient demand will not lower prices and wages. If this is so, change in the distribution of demand, without an accompanying change in the level of aggregate demand, could cause the general price level and the general wage level to rise. So price inflation would occur as a consequence not of excess aggregate demand but of excess demand in particular markets, and the failure of prices and wages to adjust downwards where demand was deficient in particular markets.

The above arguments suggest that excess demand in some markets, and deficient demand in others, will give rise to an upward drift in the general price level as prices are bid up in the markets of excess demand and remain stable in the markets of deficient demand. The argument can be extended to suggest that the inflationary pressures in a modern economy will be even stronger. Thus, if the excess demand in particular markets generates price increases in those markets, which then spill over into other markets with deficient demand, price inflation will intensify. Such spillover or linkage effects are a prominent feature in most explanations of cost–push inflation. Thus, if wage settlements are interlinked, so that employees and their representatives the trade unions attempt to obtain similar settlements in all or most wage bargains, wage rises in a sector with excess demand will encourage employees in other sectors to attempt to match that rise, even if there is no excess demand in their sector of the economy.

Underpinning such a mechanism is the suggestion that equity and justice are extremely important in wage and salary settlements, and this leads to an attempt to maintain the existing structure of wage and salary differentials. This will be reinforced if *coercive comparisons* are important, that is, if employees judge their success and the success of their representatives, by their ability to obtain wage increases matching those obtained in other sectors of the economy. Wherever such emulation occurs, relative wages may not be responsive to differences in demand and supply conditions in different sectors of the labour market. The origin of wage increases would still arise in those sectors with excess demand, but the end result would be a similar increase across all sectors of the labour market, irrespective of demand conditions prevailing in the labour market. Of course, such a process would have an adverse effect on employment in sectors with deficient demand, but in terms of the effects on cost levels, and subsequently on prices, the effect would be to increase the inflationary pressure in the economy.

There are a number of variants on this general theme of wage leadership and emulation. For example, wage leadership could arise in sectors of the economy with excess demand; or, alternatively, it has been suggested that wage leadership might originate in those sectors of the economy where productivity is rising most rapidly, or in the public sector of the economy. The rationale for this is that where productivity is rising rapidly, firms can better afford wage increases, and further that public sector enterprises are protected by their monopoly position or by the possibility that the state will underwrite losses. Once again, wage increases generated in a leading sector of the economy, whatever the origin of this leading sector, may be passed on to other sectors of the economy.

Cost–push factors can only create a continuing stimulus to inflation if they are accompanied by a permissive monetary policy that allows an expansion of the money supply – that is, higher wages that result in higher prices must raise the money value of output (given by the price level times the level of output) unless offset by a reduced level of output and resultant unemployment. If the money supply is fixed, it is necessary for the velocity of circulation of money to rise to generate the higher level of monetary demand consistent with the higher money value of output. To sustain a continuing inflation with successively higher price levels, the velocity of circulation would have to rise continuously. As the velocity of circulation is heavily influenced by existing habits and institutions, it is unlikely that it could adjust sufficiently and, indeed, the empirical evidence suggests that it demonstrates considerable stability over long periods.

In short, faced with an increase in money wages, the monetary authorities can either hold the money supply constant or prevent it rising at the same rate as wages, with consequent falls in output and more unemployment, or they can increase the money supply to allow a sufficient level of monetary demand to sustain the same output at higher prices. A strong version of the cost–push hypothesis is that the latter policy will be chosen because monetary authorities prefer higher prices to higher unemployment – that is, the money supply is considered to be endogenous, reacting to the wage levels set in the labour market and allowing these to work through to higher prices without reducing output and employment. In these circumstances, monetary policy is said to be 'permissive'. Instead of controlling the money supply to prevent inflation, the monetary authorities condone cost–push inflation by allowing the money supply to increase and thus 'ratify' higher price levels without imposing costs on employers and employees through lost output and employment.

Two additional factors can be added to the cost–push hypothesis; imported inflation and expectations inflation.

First, if the demand for imported goods is relatively price- or income-inelastic, then a rise in the prices of imported goods will be reflected in higher inflation rates. For example, the OPEC cartel caused the world price of oil to rise substantially in the 1970s. For most users of oil and oil-related products, demand is price-inelastic, especially in the short run. In the long run, households and firms can substitute fuel-efficient cars for 'gas guzzlers' and switch from oil heating systems to alternative sources of energy; but they can't in the short run. Thus the large oil price increases of the early 1970s and 1980s affected the inflation rates of oil-importing countries.

Second, economists have long been aware of expectations but it is only comparatively recently that they have incorporated expectations into economic models. The three most common assumptions used in building models including expectations are:

(a) expectations are static;
(b) expectations are adaptive; and
(c) expectations are rational.

The assumption behind static expectations is that tomorrow will be just like today. If today's inflation rate is 10 per cent, we shall assume next year's inflation rate will also be 10 per cent. Many short-term forecasts use static expectations. Adaptive expectations provide a more flexible version of static expectations, in that expectations are deemed to change gradually when previous expectations have proved to be wrong. For example, if you expect next year's inflation rate to be 5 per cent and it turns out to be 10 per cent, then you will revise your expectation of future inflation to be (say) 6 per cent, rising to 7 per cent, and only eventually to 10 per cent if the inflation rate remains at 10 per cent. Rational expectations, the present 'in' phrase of economists, assumes that everyone in the economy has, and uses, the same information as the economic policy makers. Suppose, for example, the government cuts taxes and significantly increases the money supply to reach full employment despite the fact that this is expected to cause the inflation rate to increase from 2 per cent to 10 per cent. Under adaptive expectations theory, people would not vary their 2 per cent prediction of the inflation rate until the rate actually rose. Under rational expectations they would immediately assume that the inflation rate would be 10 per cent and make current decisions with this assumption in mind, e.g. they would assume a 10 per cent inflation rate when bargaining for wage increases.

The favourite example of the advocates of rational expectations is provided by the stock market, which they assume absorbs any information instantaneously, with this information reflected immediately in stock prices. Thus, if you read in today's newspaper that Hewlett-Packard, the electronics company, is expected to have a 'good year' there is no point in rushing out to buy Hewlett-Packard stock. There is no point in taking advice from a stockbroker or studying the financial returns of companies before you invest in the stock market – the going prices of stocks reflect all the information known and predicted for all companies. The only way you can make 'real' money in the stock market is by having *insider information*, i.e. information no one else has (or very few have) – and that, as many people know to their cost, is a dangerous game, for in the UK at least it is a criminal offence and giving away insider information, or using it for personal gain, can result in prison sentences and fines. Thus boards of directors of companies, knowledgeable about a take-over, cannot legally take advantage of this information for personal gain. The theory incorporating rational expectations is known as the *Efficient Market Hypothesis*, where markets work efficiently in the sense that all information is used to clear any market of excess demand or excess supply, so that the market price reflects equilibrium.

How, then, do expectations affect the inflation rate? If you are making a deal with a firm to supply goods in a future time period, you must agree today on a price for future delivery. Suppose the good in question is a new cargo ship, which you would like to take delivery of today but which, because of a backlog of orders, cannot be built for three years. Today's price of a new vessel is, say, $60 million – but remember that such a vessel was ordered around 2½ years ago; last year's price was $58 million. Somehow you have to negotiate a price today with the Japanese shipyard. You are told next year's price will be $65 million. Suppose you ultimately agree on a price of $80 million for a vessel delivered in three years' time but this vessel is going to be an input in your business. It is going to be one of your factors of production at a cost of $80 million. This factor will have a bearing on the prices you negotiate with your customers in making, say, a three-year deal to carry cars. If enough future contracts are made in such a fashion, a portion of the future price level will be determined today. These future contract prices will be based on expectations.

Thus when a government is faced with an inflation rate of, say, 10 per cent, it would be useful to know how much of the 10 per cent is due to demand–pull factors and how much is due to cost–push expectations and imported factors. Why? The answer is because the policies appropriate for dealing with one case of inflation may be very different from those appropriate for other cases. To this topic we now turn.

20.3 Anti-Inflationary Policies

The importance of the distinction between demand–pull and cost–push inflation can be illustrated by considering the appropriate policy response to each type of inflation.

If inflation is considered as cost–push in origin, arising from institutional factors enshrined in the process of collective bargaining, then the only method of combating inflation is to change the institutional framework within which wages and prices are determined. If expectations are the cause of inflation, then expectations have to be changed in the long run. Both the income–expenditure and the monetarist approaches to inflation suggest that inflation can only be curbed by reducing demand. While the Keynesian and monetarists' prescriptions agree on this point, they disagree as to how demand should be constrained. In general, Keynesians would emphasise the importance of fiscal policy in restraining demand, whereas monetarists would emphasise the need to control the money supply.

Although it is not difficult to point to the importance of distinguishing the cause of inflation, it is much more difficult to determine which explanation the evidence better supports. Thus, observing wage and price changes is a rather 'chicken and egg' process in which it is difficult to say which comes first, the price change or the wage change. Even if one could identify the timing of price and wage changes, this would not necessarily be conclusive evidence as to what caused inflation. For example, an inflation due to excess demand could initially affect the level of wages in the labour market and only subsequently be transmitted to the product market. Thus, the existence of wage increases in advance of price increases would not necessarily support the argument that the cause of inflation was cost–push. Again, whatever the cause of inflation, that inflation would in practice be accompanied by increases in money wages, so that it is difficult to conceive of any practical situation in which money wages would fail to rise.

The difficulty of determining the cause of inflation can be illustrated by reference to the Phillips Curve shown in Figure 20.4. Given an unemployment rate of 2–3 per cent and a rate of wage change of 2 per cent, price stability will result. The rate of change of prices is shown along the right-hand axis. If productivity is rising at 2 per cent, then it follows that a 2 per cent increase in money wages will be consistent with price stability. Given a Phillips Curve as shown, such a rate of increase of money wages would occur at a level of unemployment of 2–3 per cent.

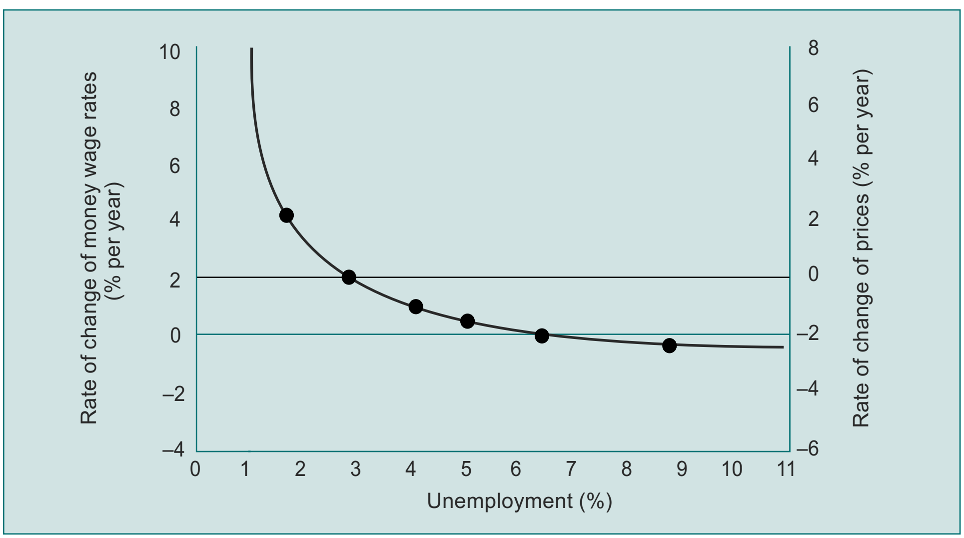


Figure 20.4 The Phillips Curve: unemployment with price stability

As previously mentioned, one weakness of the Phillips Curve is that it is quite possible to interpret the empirical results as supporting either a demand–pull or cost–push explanation of inflation, and results have been so interpreted by the differing schools of thought. As a demand–pull explanation, three steps were necessary to produce the observed relationship between unemployment and changes in money wages. These were:

1. a stable, inverse relationship between the unemployment rate (U) and the excess demand for labour, where U was therefore a proxy for the excess demand for labour, rising when the excess demand for labour fell and falling when the excess demand for labour rose;
2. a stable, positive relationship between the excess demand for labour and the rate of change of money wages, W, rising when the excess demand for labour rose, and vice versa;
3. if steps 1 and 2 had been taken, then there would be a stable, inverse relationship between unemployment and the rate of change of money wages, money wages rising more rapidly when unemployment fell and rising more slowly when unemployment rose.

As a cost–push explanation of wage inflation, it was simply necessary to suggest that, at low levels of unemployment, unions would be more militant in demanding increases in money wages, and employers more willing to pass on increases in money wages in higher prices in the expectation that price increases would have little adverse effect on sales. Again, the observed tendency of the Phillips Curve to flatten out at high rates of unemployment, without producing actual falls in money wages, was consistent with the view that unions and employees would strongly resist cuts in money wages so that wages would be inflexible downwards even in the presence of deficient demand.

Although the empirical evidence demonstrates a short-run trade-off between the rate of change of money wages and the level of excess demand for labour as represented by unemployment, it is not clear whether this relationship persists in the longer run. Indeed, it has been suggested that the trade-off between unemployment and the rate of change of money wages is a transitory phenomenon caused by a failure of expectations to adjust immediately to price changes. However, while expectations do not adjust instantaneously, they do after some time lag, and when expectations have so adjusted, the temporary trade-off between unemployment and changes in money wages disappears entirely.

This analysis has been given added weight after the 1960s – the day of the Phillips Curve – by the evident breakdown in the historical relationship between the level of unemployment and the rate of change in money wages. Indeed, the 1970s and 1980s witnessed rising money wages, rising inflation rates and the presence of a rising unemployment rate. Advocates of the Phillips Curve argue that the relationship postulated by Phillips still held: the Phillips Curve had shifted upwards due to cost–push elements and expectations.

Figure 20.5 exhibits the annual inflation rate and unemployment rate data for the USA from 1960 to 2010, with five Phillips Curves 'interpreting' the data. Upward displacement of the Phillips Curve yields an explanation for the events observed in the 1970s and 1980s, with the curve heading back towards the position of the 1960s Phillips Curve in the 1990s as cost–push elements and inflationary expectations receded. However, monetarists such as Friedman argue that there is a more fundamental objection to the Phillips Curve. This analysis, which denies long-run trade-off between unemployment and money wages, argues that labour is concerned with the behaviour of the real wage and not with the behaviour of the money wage. If this is so, a long-run relationship between the rate of unemployment and the rate of change of money wages will only hold when the rate of price inflation is expected to be zero or negligible. As soon as a significant rate of price inflation is expected, the short-run relationship between unemployment and money wages will shift to the right – and will shift further to the right the higher is the expected rate of price inflation. In the long run, there is no trade-off between unemployment and the rate of change of money wages, so that the policy maker is quite unable to choose different combinations of money wage changes (and therefore inflation rates) and unemployment. Macroeconomic policy cannot affect the long-run rate of unemployment, only the long-run rate of price inflation. This being so, the policy maker should choose a zero rate of inflation.

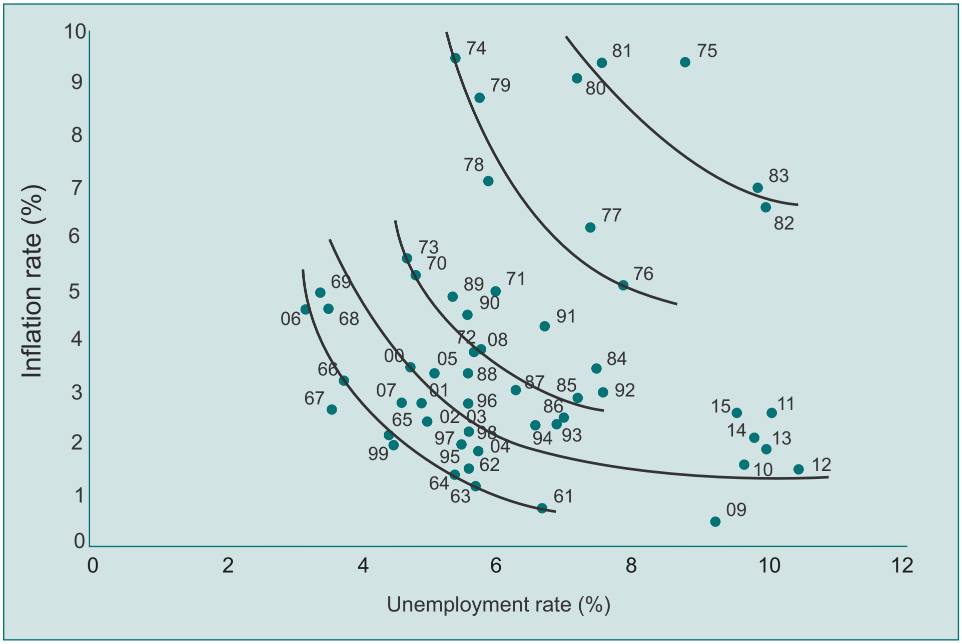


Figure 20.5 The inflation rate/unemployment rates, USA, 1960–2015

Friedman suggests that the rate of unemployment can only be reduced below the natural rate of unemployment by creating inflation, and that it can only be held below that level of unemployment by accelerating inflation. As we saw earlier, the natural rate of unemployment is dependent on the properties of the labour market, for example the flow of information in the market, the rate of structural change in the economy, and the costs involved in obtaining information and undertaking labour mobility. In simplified terms, we may regard the natural rate of unemployment as unemployment that is not due to deficient demand. The rate of unemployment can be reduced below the natural rate permanently only by measures that reduce the amount of structural and frictional unemployment. For example, anything that improves the flow of information in the labour market is likely to lower the natural rate of unemployment. But the rate of unemployment cannot be pushed below the natural rate of unemployment permanently by demand management policies.

Suppose initially that the rate of unemployment is at the natural rate, so that the economy is at full employment. If the level of aggregate demand is now raised, prices will rise; and if money wages lag behind prices, the real wages will fall. As real wages fall, employers will hire more labour and then unemployment will fall. However, this situation can only be sustained as long as 'money illusion' exists. In time, as expectations adjust to the fact of inflation (and believers in rational expectations would have them adjust immediately), labour will demand increases in money wages to offset the increase in prices. Real wages will return to the initial level and employers will discharge labour, so that the rate of unemployment will return to the natural rate. Notice, however, that this process could produce a short-run trade-off between the rate of change of money wages and the level of unemployment. Thus, as aggregate demand rises and prices rise, the competition among employers for labour would raise money wages. So long as the increase in money wages was less than the increase in prices, the real wage would fall and the increase in money wages would be accompanied by a fall in the unemployment rate as employers hired more labour. Of course, this trade-off between the rate of money wages and unemployment would be temporary and would disappear as expectations adjusted to the fact of inflation.

In summary, the Friedman view is that employees are concerned with the real wage and that behaviour will adjust, possibly after some time lag, to changes in prices. In consequence, there is no long-run trade-off between the rate of change of money wages and the rate of unemployment. This being so, no menu of policy choices between unemployment and the rate of change of money wages (and therefore price inflation) faces the policy maker. Instead, the rate of unemployment cannot be forced much below the natural rate of unemployment, and a freely working market economy will tend to create a situation where unemployment will tend towards the natural rate. In these circumstances, the only proper objective of macroeconomic policy is to produce price stability. The rate of unemployment can only be reduced below the actual rate by microeconomic policies aimed at improving the efficiency of the labour market. We can now incorporate this information into the views of the monetarists compared with the Keynesians.

20.4 The Modern Quantity Theory versus Keynes: Policy Implications

The Quantity Theory of Money captured in the famous equation MV = PY, was initially based on the assumption that V was constant and Y was determined by capacity output, with the result that changes in M were reflected directly in changes in the price level P. This 'naive' version of the theory encapsulates, in its purest form, the view that changes in the money supply cause changes in the price level.

Modern Quantity Theory restates this conviction in a more elaborate and sophisticated form, but the conclusion remains much the same – namely that control of the money supply is the key to the control of prices, so that if price inflation occurs, it originates in the failure of the central monetary authority (rather than in the activities of organisations such as trade unions, as advocated by supporters of the cost–push hypothesis).

The essential difference between Keynesian economists and the monetarists may be summarised as follows:

(a) Monetarists consider that the demand for money is a stable function of a number of variables, such as the level of income, the expected rate of return on money and other assets, and the rate of change in prices. The importance of this assumption of stability is that, once accepted, it implies that the velocity of circulation of money, V, will be constant or will be subject only to modest and predictable fluctuation.

    The Keynesian view is that changes in the velocity of circulation may offset and frustrate monetary policy, in that when M rises V will fall, and vice versa. Hence, in a depression, increases in the money supply find their way into speculative balances and the velocity of circulation will fall, so that the change in the money supply has no effect on the level of aggregate demand.

(b) Monetarists consider that the demand for and supply of money are in large degree independent of each other; that is, the supply of money is held to be determined exogenously to the economic system by the central monetary authority. It follows that, given a stable demand function for money, exogenous increases in the supply of money will increase aggregate demand for goods and services as individuals spend their excess cash balances. The resulting price inflation is thus due to the increased supply of money, and is not due to shifts in the function describing the demand for money.

The neo-Keynesian view is that the money supply can be determined endogenously; that is, it may be responsive to economic variables. For example, if the level of money wages rises due to cost–push factors and these higher costs are passed on in higher prices, the central monetary authority may expand the money supply in order to 'underwrite' the new higher level of prices. Alternatively, the commercial banks may expand the money supply to satisfy an increased demand for credit. In this case, the money supply would respond to changes in economic variables and would not be determined exogenously to the economic system.

It has also been suggested that the central monetary authority may be unable to control the money supply effectively. Thus, private financial institutions may be able to frustrate the wishes of the central authority by finding efficient substitutes for money, or by finding more effective means of using the existing money supply. For example, recent decades have seen the development of credit as a major and important substitute for money, as well as developments such as charge accounts and PAYE taxation, which synchronise the receipt and expenditure of money and therefore economise on the need for money balances.

From these two sources, the basic differences between Keynesianism and monetarism arise. The monetarists maintain that changes in the money supply have been the chief cause of substantial fluctuations in national income, causing both major inflations and major recessions. The monetarists' view is, therefore, that monetary policy *can* have an impact on the level of real income and employment, but it should not be concluded from this that a discretionary monetary policy is recommended. It is no part of the monetarists' policy prescription that the money supply should be lowered or raised to regulate demand and therefore the level of real income and employment. Instead of advocating 'functional' monetary policy along the lines of 'functional' finance, the monetarist view is that monetary policy should follow an automatic rule, allowing an annual change in the money supply to match the long-run growth rate of the economy. This policy prescription is based on the proposition that strong equilibrating forces exist in a market economy, which tend to maintain or restore full-employment output. This being so, changes in the money supply have only a short-term effect on real income and employment, but no lasting long-run effect. Instead, in the long run, changes in the money supply influence only the level of prices, so that the proper and only role of monetary policy is price stability, which is achieved by the application of the one simple, automatic, rule.

A review of the empirical evidence suggests three conclusions that would be acceptable to almost all economists:

1. There is no recorded example of any major inflation occurring without an accompanying substantial increase in the money supply.
2. There is no recorded example of any substantial increase in the money supply which has not been accompanied by a major inflation.
3. Given 1 and 2, a high rate of inflation cannot be sustained unless the money supply is expanded significantly.

These observations mean that, as compared with the neo-Keynesian alternative, the monetarist explanation better fits the experience of hyperinflation (Germany 1923) or sustained high rates of inflation as have been experienced in many Latin American countries more recently. Yet, this does not necessarily imply that milder inflations are a purely monetary phenomenon, as it can be argued that in these circumstances changes in the money supply are permissive and not causal. The money supply may be endogenous in that it reacts to prior changes in other economic variables and simply permits mild price inflation to occur, the genesis of that inflation being due to factors that lie outside the control of monetary policy.

The manner in which such an inflation may be generated can be understood by referring back to our earlier discussion of the circular flow of income. Suppose that trade unions are successful in pushing up money wages and salaries in the labour market and that businesses, in order to protect their profit mark-up, pass on the higher money wages and salaries in the form of higher prices. In terms of the circular flow of income, it should be noted that all of the increase in the money value of output must accrue as income to the factors of production, so that the incomes accruing to the factors of production must equal the money value of output. In short, the higher the level of prices the higher the level of money output, accompanied by a parallel and equivalent increase in money income. However, if the money supply is exogenous and fixed by the monetary authority, then it does not rise in line with the increase in money incomes. The higher level of money incomes will result in an increased demand for money for transactions and precautionary purposes and, given a fixed supply of money, the money available for speculative purposes will fall. Consequently, to ensure that households and businesses are content to hold this reduced supply of money for speculative purposes, the rate of interest must rise; and this higher rate of interest will discourage investment, resulting in a fall in aggregate demand.

In short, unless there is an increase in the money supply or greater economy in the use of money balances, then the process of passing on higher costs in the form of higher prices cannot be sustained indefinitely without a major adverse effect on investment and aggregate demand. Unemployment will rise as a consequence of increases in money wages, unless there is an accommodating increase in the money supply. If this accommodating increase in money supply is not forthcoming, there can be no continuing inflation. Put the other way round, continuing inflation is impossible to sustain in the face of a sustained and determined attempt to restrain the money supply.

An increase in the money supply is therefore a necessary condition for continuing inflation. However, neo-Keynesians argue that an increase in the money supply may simply be an enabling condition for continuing inflation. The cause is cost–push in the factor market, as a consequence of the exercise of trade union bargaining power. Higher wage costs lead to higher prices and these can only be sustained if the money supply is increased. This position is accepted by both Keynesians and monetarists: if the money supply does not increase, the result will be higher unemployment. Where Keynesians and monetarists disagree is in the assessment of the amount and duration of unemployment that would be necessary to contain inflation, and of the long-term benefits relative to the short-term costs of a restrictive monetary policy.

Keynesians would argue that modern governments are regarded as having a major responsibility for ensuring full employment. A departure from full employment would have heavy economic, social and political costs in the short term, which must be weighted more heavily than possible long-term benefits. In these circumstances, trade unions and similar organisations may bargain more aggressively for higher wages and employers may be more willing to accede to higher wage demands, as both parties would expect the monetary authorities to allow the money supply to expand in order to ratify higher wage and price levels. When such inflationary expectations are current, any attempt to control inflation through controlling the money supply would require a long and costly process of adjustment, within which employees and employers would only slowly modify their expectations and behaviour in the light of actual experience.

Keynesians might be said to operate with shorter economic and political time horizons than monetarists. Any sharp increase in unemployment would be unacceptable, socially undesirable, and a waste of resources. The long-run possibility of greater price stability is not weighted very heavily. Of course, because it is not weighted very heavily, inflation may be more likely, in that people may take action based on their belief that the government will be more anxious to avoid any increase in unemployment rather than maintain price stability. In summary this means that, under a proposed 'Keynesian' or 'full-employment' government, it may be much more difficult to revise inflationary expectations downwards.

Monetarists might be said to operate with longer economic and political time horizons. They believe that in an inflationary situation a restrictionist monetary policy will have short-run costs and that these costs will be higher if inflationary expectations are widely held. If employers and employees adjust their behaviour only slowly in the light of experience, then losses of output and employment will be more substantial and prolonged. On the other hand, monetarists believe that such output and employment losses are temporary phenomena and that the long-run benefit will be price stability.

Moreover, monetarists believe that any attempt to push the rate of unemployment below the natural rate by demand management accompanied by an expansionary monetary policy will be self-defeating; unemployment, they believe, can be held below the natural rate only by a continuous increase in the rate of inflation. If the natural rate of unemployment is set by factors that govern the efficiency of the labour market, and if the market system has strong equilibrating tendencies, then macroeconomic management should not attempt to influence the rate of unemployment; its only target should be the maintenance of price stability. Indeed, monetarists would argue that, in the long run, price stability will reduce the natural rate of unemployment itself, as price stability improves the efficiency with which markets operate.

The differences between Keynesians and monetarists are partly based on different technical assessments and partly on value judgements. They are technical insofar as they rest on different views of how markets operate and how expectations and behaviour adjust to different situations. They are partly value judgements insofar as they reflect different social and political priorities. In principle, technical differences can be resolved by an appeal to the evidence, although the evidence is often less decisive than may be commonly supposed. Differences in value judgements cannot be resolved in this fashion as they reflect basic attitudes and perceptions about which there can be no ultimate 'scientific' judgement. The debate will continue!

We cannot leave macroeconomics as though the only important policy goals were low unemployment and low inflation. There are others that we must address.

20.5 National Goals: Policy Implications

In the world of macroeconomics a list of 'goods' and 'bads' would contain many of the following items.

The 'goods' would certainly include consumption expenditure: the higher our consumption of goods and services, the better off we are in material terms.

Investment expenditure must also be a 'good': with zero investment, our capital stock would decline through depreciation and also the future flow of goods and services. That flow can be enhanced by investment expenditure. By embodying the latest technological innovations into new capital goods, investment today means a larger flow of goods and services in the future.

What about government expenditure? 'Good' or 'bad' column? No matter what your political leanings, a certain level of government expenditure is necessary if only to establish the 'rules of the game' for the market economy and see that those rules are obeyed. To deal with market failures, such as those associated with public goods and externalities, collective action or government, it is also necessary. Thus government expenditure has got to be a 'good'. But should the government's budget be balanced every year, i.e. should total government outlays equal total government revenues? Would you prefer a budget surplus, the surplus being used to reduce the *national debt*, which is the summation of past deficits? There are no right or wrong answers to this question.

International trade has to be a good because, as we have seen, it makes possible higher living standards among trading nations. However, is it better for a country to have exports (X) equal to imports (Z), or would you prefer X > Z, a sign of strength as in Japan and Germany, or would you prefer X < Z, using foreign resources? Perhaps you are more concerned with stability of your currency on foreign exchange markets rather than your trade balance. If so, then you might assign 'bad' to balance of payments, deficits or surpluses.

What about the 'bads'? Unemployment is an obvious 'bad', for unemployment of labour is normally associated with unused capacity, resulting in a lower GNP than could be achieved.

What about inflation? If inflation makes current GNP or future GNP lower than it otherwise would be, then inflation can be added to the 'bad' column. If inflation produces a less desirable income distribution than the income distribution without inflation, then inflation enters the 'bad' column twice. If inflation makes the political system in market economies unstable, then it may be entered in the 'bad' column three times. (Karl Marx, incidentally, predicted that high unemployment would prove to be the downfall of the capitalist system: Lenin argued that the Achilles heel of capitalism was inflation.)

If we return to our 'good' column, any 'good' not achieved would enter the 'bad' column. For example, if balancing the government budget exactly was a 'good', then a budget deficit or surplus would be a 'bad'.

Having identified the 'goods' and the 'bads' we now have to weight them, by assigning degrees of importance to each. Since many of the 'goods' and 'bads' are directly related, assigning such weights will produce trade-offs. A political platform on which people vote for a specific party is an attempt (albeit not well defined) to identify and weight 'goods' and 'bads'.

Consider the following 'welfare function' where W = national welfare.

$$\mathrm{W} = \mathrm{C}^{0.6}\mathrm{I}^{0.2}\mathrm{G}^{0.2} - \mathrm{U}^2 - \mathrm{INF}^3 - 10|\text{Budget Surplus}|$$

This function explicitly states that consumption (C) investment (I) and government expenditure (G) are all 'goods' – they yield plus points. But the form of the function $C^{0.6}I^{0.2}G^{0.2}$ states something else: it states that for any given GNP equal to C + I + G, the optimal proportions of the three measures

$\frac{C}{GNP}, \frac{I}{GNP}$, and $\frac{G}{GNP}$

are 60 per cent, 20 per cent and 20 per cent respectively. For example, suppose GNP = 100 = C + I + G. Table 20.1 shows what happens to 'welfare' points when the ratios stray from the optimal.

Table 20.1 Welfare points with different ratios of C, I and G

C + I + G	C	$C^{0.6}$	I	$I^{0.2}$	G	$G^{0.2}$	$C^{0.6}I^{0.2}G^{0.2}$
100	80	13.86	0	0	20	1.82	0
100	70	12.80	10	1.58	20	1.82	36.92
100	60	11.67	20	1.82	20	1.82	38.66 (max)
100	50	10.46	25	1.90	25	1.90	37.89
100	20	6.03	75	2.37	5	1.38	19.74

Thus to maximise W, the naive answer appears obvious: make GNP as large as possible, allocating GNP among C, I and G so that the ratios 60 per cent, 20 per cent and 20 per cent respectively prevail, have a zero unemployment rate, a zero inflation rate and a balanced budget. The answer is naive because a zero unemployment rate is not possible; in addition, if you believe a Phillips Curve exists, you have to make some difficult calculations given that negative points are deducted for both inflation and unemployment. The unemployment rate is squared; thus if U = 8 per cent you lose 64 points. The penalties for inflation are higher since INF is cubed: if the inflation rate is 8 per cent you lose 512 points. However, your trade-off is not only between inflation and unemployment. The higher is the unemployment rate, the lower is GNP and consequently the smaller is the positive points score from $C^{0.6}I^{0.2}G^{0.2}$. Finally, whatever policy measures you take you may end up with an unbalanced budget, which could prove costly since the welfare equation multiplies by 10 any imbalance in the budget.

To the uninitiated, balancing the budget may appear a simple process: governments can reduce outlays, increase taxation, or employ a judicious blend of the two. It is, unfortunately, never quite so simple. Governments are constrained in a variety of ways, from restrictions imposed by international agreements to the need to be re-elected. Let us assume that the only avenue remaining to our government is to raise revenue through income taxes. What should the government do? The obvious response would be to raise the tax rates. As in so many areas, the obvious may not be the right decision.

Consider Figure 20.6 below, known as the Laffer Curve after its inventor Arthur Laffer. Laffer argued that if the tax rate were 100 per cent, no one would be willing to work since all wages and salaries would go in taxes; government tax revenue would then be zero. If the tax rate were zero, government tax revenue would also be zero. According to the Laffer Curve, maximum tax revenues would be collected at the tax rate of x per cent.

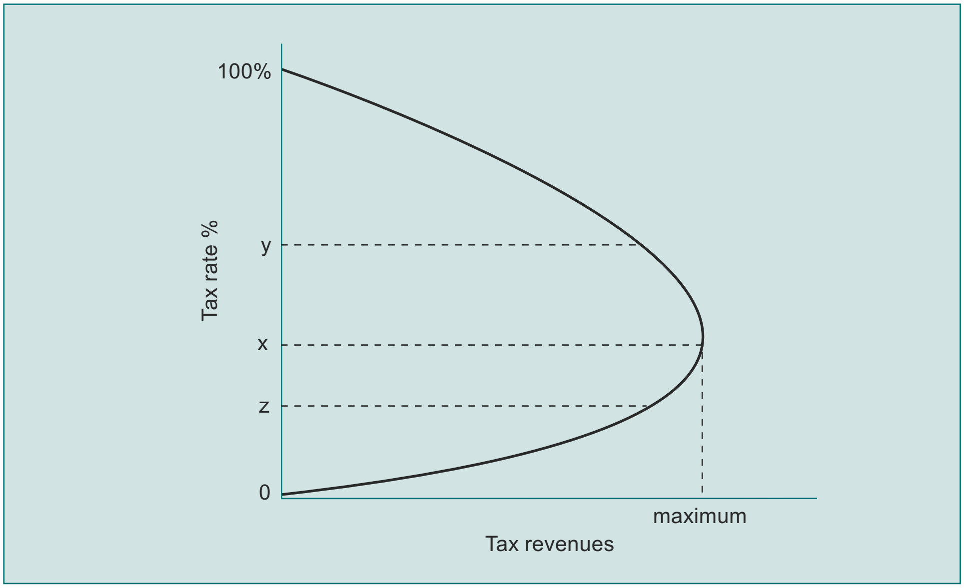


Figure 20.6 The Laffer Curve

If the tax rate were y per cent, a decrease in the tax rate would increase tax revenues. On the other hand, if the tax rate were z per cent, an increase in the tax rate would increase tax revenues. Thus if a policy goal of the government were to balance the budget and suppose this necessitated increasing tax revenues, the government would have to decide where on the Laffer Curve the economy actually was before it decided whether the correct policy was to increase *or* decrease the tax rate.

One further point deserves mention. Suppose you are concerned with national welfare over a time span of several years (*n*) and you are going to measure success by summing the annual values of W over that period. That is

$$S = \sum_{t=1}^{n} W_t$$

Then the problem becomes even more complicated because maximising S does not necessarily imply maximising W each year. This is because there are lags in the economy – for instance, if this year's inflation rate is one major determinant of next year's inflation rate and a second determinant is this year's unemployment rate, you may score highly this year by driving down the unemployment rate but cause problems for your economy in subsequent years as inflationary forces gather.

Different political parties implicitly have different *national welfare functions*; the items in these functions may be the same but the weights may differ greatly. Evidence suggests that the state of an economy is a major factor determining which party gets elected, and as a result a government in power may enact policy measures to have the economy in apparently 'good shape' during an election year and 'solve' economic problems caused by such policies in the intervening years between elections.

Learning Summary

You now understand that the IS/LM analysis produces equilibrium only in the real goods and money sectors; that is, it tells us what the level of national output and what the interest rate will be. Knowing the levels of actual and potential output tells us what the unemployment rate will be.

What remains to be determined is the inflation rate. The two schools of thought, the Keynesians and the monetarists, have different explanations of what determines the inflation rate; you have been made aware of these differences. For the monetarists, inflation is a monetary phenomenon alone; for the Keynesians, many factors can affect the price level. The importance of the different schools of thought arises when policy issues are discussed, since the appropriate policy to deal with inflation will depend on the type of inflation.

You are also aware that low inflation and low unemployment rates are not the only two desirable macroeconomics goals, but increasing the number of goals increases the complexities of policy making. Again, the differences between the monetarists and the Keynesians emerge. You understand why the monetarists believe that there are equilibrating forces at work in the economy and why the Keynesians believe there is no inherent tendency for an economy to return to full employment with price stability.

Review Questions

Multiple Choice Questions

20.1 The following data, in index form, refer to an economy.

Year	Real GNP	Money GNP	Potential GNP	Labour force
t	100	100	100	100
t + 1	104	106	100	103

It can be inferred that

I. average output per worker decreased between Year *t* and Year *t* +1.

II. prices rose on average by about 2 per cent between Year *t* and Year *t* +1.

III. the capital stock decreased between Year *t* and Year *t* +1.

Which of the following is correct?

A. II only.

B. II and III only.

C. I and III only.

D. I, II and III.

20.2 Which of the following explains why considerable unemployment can exist in a market economy?

A. At full employment, national income is not always sufficient to purchase all output produced.

B. Many product and factor prices respond very slowly when supply exceeds demand.

C. The rate of productivity growth is not always great enough to permit sufficient growth of actual output.

D. The growth of productive capacity always outstrips the growth of consumers' wants.

20.3 The Chancellor of the Exchequer wishes to use fiscal policy to control aggregate demand and also wishes to reduce the public sector as a proportion of GNP. In the event of a recession, which of the following would be an appropriate policy to adopt?

A. Reduce the level of taxation.

B. Increase transfer payments.

C. Increase government spending on goods and services.

D. Leave both the tax structure and government spending unchanged.

20.4 The following policies would be appropriate to help reduce inflationary pressures in an economy:

I. lengthening the repayment period on consumer loans.

II. increasing the down-payment requirement on the purchase of consumer durables.

III. selling securities by the central bank.

Which of the following is correct?

A. I and II only.

B. II only.

C. II and III only.

D. III only.

20.5 Faced with an inflationary gap, the government decides to increase the money supply. They argue that this would lower interest rates, which in turn would stimulate investment, increase potential GNP, and consequently decrease the inflationary gap. The likely outcome of such a policy would be

I. as predicted by the government in the long run.

II. to worsen the inflationary gap in the short run.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

20.6 An inflationary gap exists; which of the following would tend to decrease that gap?

A. An increase in investment expenditures caused by optimistic forecasts regarding the world economy.

B. An increase in savings.

C. Increased orders for our trucks and buses abroad.

D. An increase in our defence commitments to NATO.

20.7 Optimistic forecasts concerning world trade have boosted business confidence, with the result that private investment is expected to increase next year. Without any policy changes, the outcome of the increase in investment must be

I. a decrease in consumption expenditures.

II. an increase in imports.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

20.8 If an economy were at full employment prior to, during, and after a period of unanticipated inflation, the impact of that inflation would have been to do which of the following:

A. increase unemployment.

B. redistribute income.

C. decrease the average price level.

D. increase potential GNP.

20.9 An economy that was recently operating at full employment with a high rate of inflation has been severely deflated by substantial increases in tax rates and cuts in government expenditure. However, although unemployment is now much higher than before, inflation has not decreased. The factors that could account for this are

I. some goods and factor markets do not respond immediately to changes in demand and supply.

II. the policies have generated demand–pull inflation.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

20.10 If the goal of a government is to reduce the inflationary gap, the appropriate policy is

A. increase the money supply to stimulate investment.

B. increase the money supply to stimulate private saving.

C. decrease the money supply to increase interest rates.

D. decrease the money supply to decrease interest rates.

20.11 Faced with a deflationary gap, the British government adopts a policy of lower tax rates and an increase in the money supply. The following unanticipated changes then occur:

I. an increase in tariffs in the US and Japan.

II. an increase in the marginal propensity to consume.

Those unanticipated changes helped the government solve the deflationary gap problem.
Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

20.12 For an economy operating at full employment, the imposition of domestic tariffs and quotas is expected to reduce imports by 50 per cent next year. Which of the following is the appropriate policy for the government to adopt in order to maintain equality between potential and actual output in light of the decrease in imports?

A. Increase in government expenditure.

B. Decrease in taxes.

C. Decrease in the money supply.

D. Subsidy to exporters.

Module 21

Economics and the World Today

Contents

No multiple choice questions are provided for this module as the nature of the discussion does not lend itself to this form of testing.

21.1 Introduction

With your new-found mastery of economics, you are ready to tackle some of the problems facing the world today. The answers to many of these problems are not simple – if they were, these problems would have been resolved by now. What economics can give you is a framework within which you can analyse these problems and reach a reasoned, rather than emotional, conclusion. Nowhere is the need for reason rather than emotion greater than in discussions of the economies of developing countries.

21.2 Developing Countries

The economic problems of low-income countries must now be given detailed examination. Let us begin by considering the following series of figures.

Figure 21.1 shows a savings function for a low-income country. Savings are a function of income, $S = f(Y)$, but because of poverty most resources are allocated to the production of consumer goods and services.

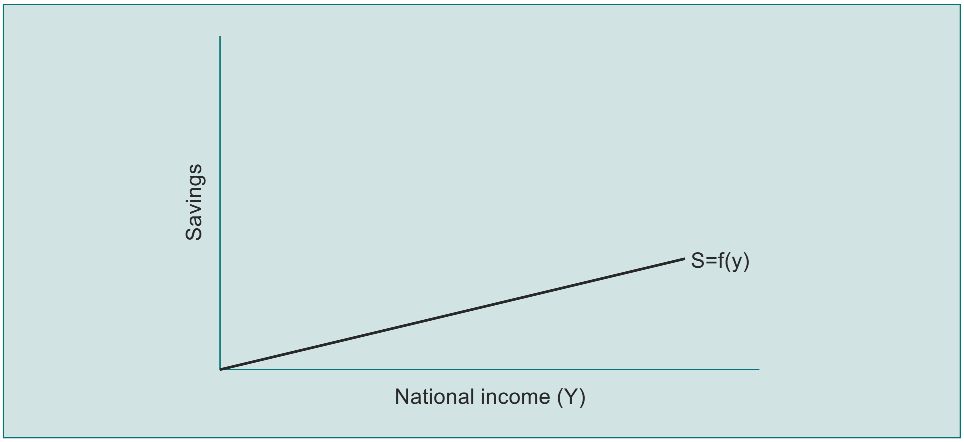


Figure 21.1 Savings function

Figure 21.2 shows the relationship between the capital stock of a nation and national output (GNP). Since output is a function of both the capital stock and the labour force, Figure 21.2 assumes a fixed relationship between labour and capital in producing output.

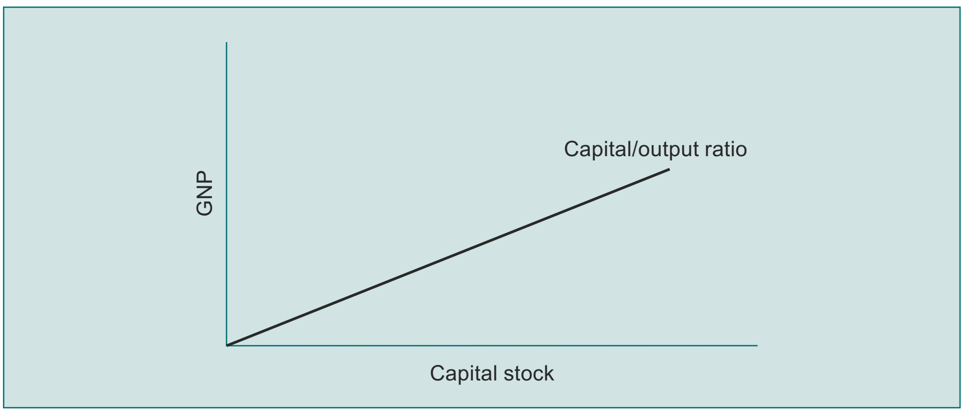


Figure 21.2 Basic capital/GNP relationship

It may appear paradoxical that less-developed countries with 'large' populations have labour shortages, but population and the labour force can be meaningfully separated. Most developing countries face acute labour shortages in the semi-skilled and skilled categories. Such shortages can be shown diagrammatically, as in Figure 21.3. The curve demonstrates that additions to the capital stock beyond C_1 produce no additional output because of a labour constraint.

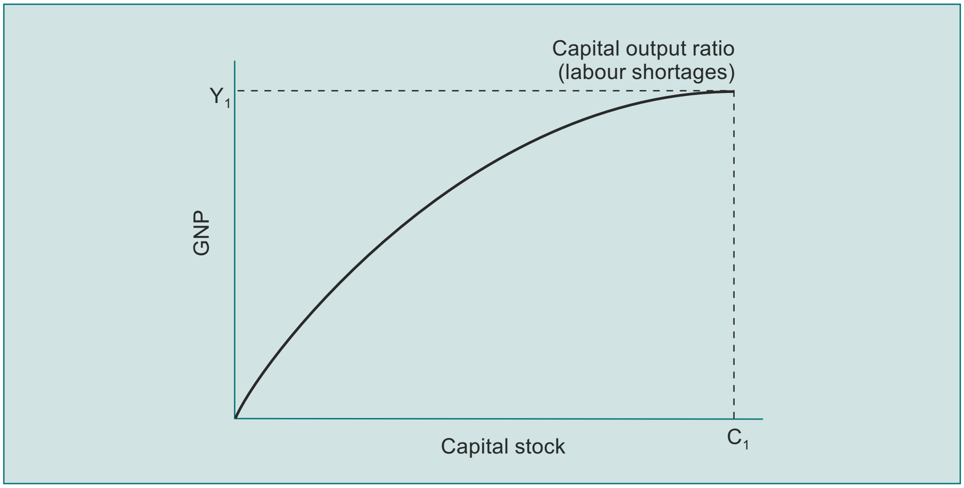


Figure 21.3 Capital/GNP relationship with labour shortages

Figure 21.4 shows how GNP could be increased for any given amount of capital if the labour force were expanded and/or if the existing labour force were better trained and consequently more highly skilled. For a given amount of capital, C_1, the larger/better-trained labour force can now produce a GNP level of Y_2 rather than Y_1.

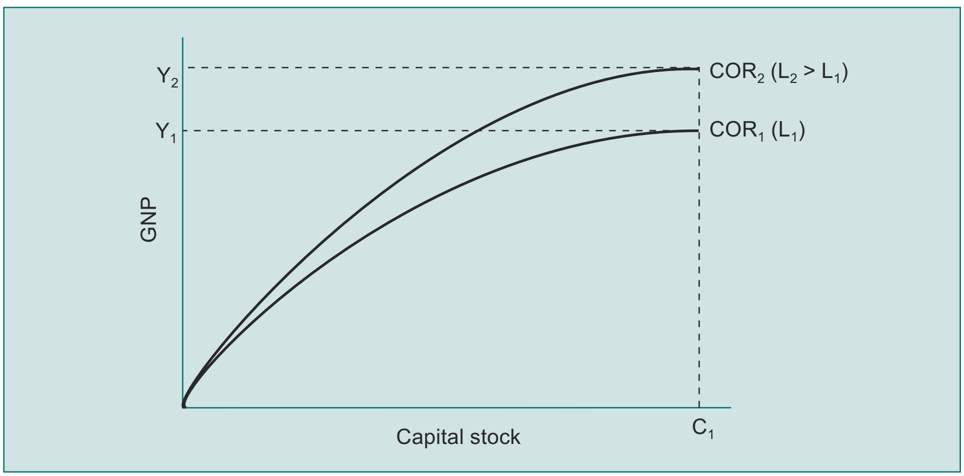


Figure 21.4 Capital/GNP relationship with large/better-trained labour force

We can now put Figure 21.1 and Figure 21.3 together, with two connecting figures, to exhibit economic growth, as shown in Figure 21.5.

Figure 21.5a comes from Figure 21.1, and Figure 21.5c comes from Figure 21.3. In Figure 21.5b, the 45° line indicates that national income (Y) equals national output (GNP).

Let us start off in quadrant (d) in Figure 21.5 with an initial capital stock of C_1. The capital/output ratio shows that a capital stock of C_1, given the labour force, will yield an output of GNP_1 (Figure 21.5c). National output must equal national income. Thus $GNP_1 = Y_1$ (Figure 21.5b). However, Y_1 generates savings of S_1 (Figure 21.5a). Assuming no depreciation in this simple model, and further assuming all savings are devoted to the production of capital goods, the capital stock C_1 will be increased by an amount S_1. Thus C_2 (Figure 21.5c) $= C_1 + \Delta C$, where $\Delta C = S_1$.

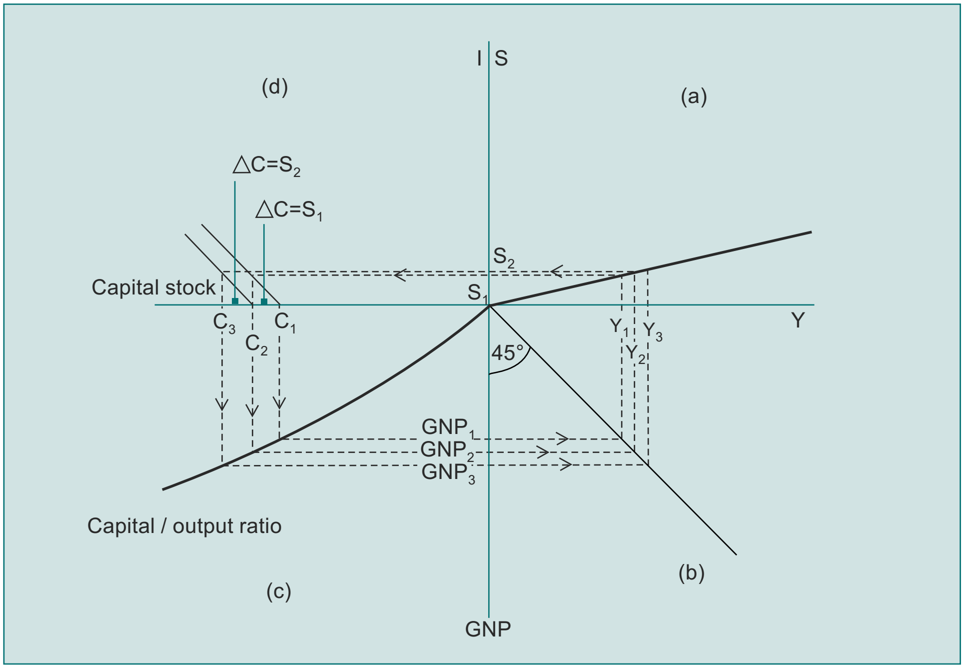


Figure 21.5 Economic growth model

C_2 produces GNP_2, Y_2 and S_2, yielding further growth in the capital stock, the GNP, national income, and savings; therefore we have economic growth. The problem with most less-developed countries is that the growth in Y is matched by growth in population; as a result, GNP per capita remains approximately constant.

How can this problem be resolved? How can living standards (GNP per capita) in these countries be increased? Consider again Figure 21.5a. Why not raise the savings function? That means lowering the consumption function for the current population – and, for many of these countries, current living standards are already extremely low. In addition, the practical problems of raising savings and lowering consumption in primarily rural societies are not easily resolved – for example, the government removing output to sell abroad for capital goods requires an infrastructure for moving such produce. In the Soviet Union, Joseph Stalin adopted such a strategy in the 1930s by reallocating resources from agriculture to capital goods production; this policy raised the Soviet Union's growth rate, which was Stalin's goal, but led to millions of deaths through lack of food, clothing and shelter.

China was able to achieve low family size because of the command nature of its economy and the social pressures brought to bear on each husband and wife to be responsible citizens and have only one child. In most other low-income countries such pressures are non-existent. When India introduced vasectomies in return for three months' wages, the only 'volunteers' were old men. Efforts to limit family size fail largely because of the principal–agent problem – remember the conflict of interest? In this case the principal is the government, seeking a slower population growth as a way of increasing GNP per capita, savings, investment and more rapid economic growth. The agent, in this example, is households – to have large families is a form of social security, with lots of children who survive helping to take care of the parents in their old age; additionally there is the possibility, albeit remote, that one child may develop into a world-famous athlete or make it in the world of business and become the breadwinner for the whole family.

The view of the World Bank, provider of loans to developing countries, can be captured by reference to quadrants (d) and (c) in Figure 21.5. The philosophy behind, and the form of, many of the loans is increasing the quality and quantity of the capital stock and the labour force. Such loans would move the economy farther out from the origin along the capital stock axis and also shift the capital/output ratio, as shown in Figure 21.4. This would cause both an increase in the level of Y and also an increase in the growth rate of Y, the eventual goal being the self-sustaining growth experienced in the leading market economies.

21.3 The Changing Global Economy

A useful rule of thumb in macroeconomics is that if you take the number 70 and divide it by the annual growth rate of GNP you will get the number of years it takes GNP to double. In the first decade of the twenty-first century the GNP growth rates of the developed capitalist economies have been low, averaging less than 2 per cent per annum, while the growth rates of China and India have been averaging 9 per cent per annum. This means that the GNP of these countries will more than double for every decade that these growth rates continue.

Given the disparity in growth rates, the ranking of countries in terms of GNP has changed significantly since the start of this century and continues to change at a rapid rate. The USA continues to be by far the largest economy in the world, with a GNP more than twice that of China, at number two, which is closely followed by Japan, at number three.

You will recall from Module 9 that dividing GNP by population yields GNP per capita. It has already been discussed that this merely indicates average income levels and fails to consider how a country's income is distributed. Indeed, when a highly skewed income distribution exists, for example the top 1 per cent of income recipients taking over 30 per cent of national income, the GNP per capita figure may yield little information on how poor or otherwise individuals in a country actually are. In per capita GNP comparisons the USA leads Canada, Germany, France, the UK, Japan, South Korea and Italy by a large margin. Norway, with its sales of oil and gas, and Singapore, with large foreign investments and rapidly growing exports (and complete lack of crime), feature high up on the per capita stakes.

While China does not rank highly in the GNP per capita tables, its economic potential is enormous. On the non-economic front China, with its authoritarian government, has political and social stability, features lacking in many developing countries. The government accepts the basic role markets play in allocating resources and producing goods and services and has assigned important roles to the non-public sector, self-employment and private ownership. These are necessary conditions, many would argue, for sustained economic growth.

On the economic front China has a large, high-quality, inexpensive labour force and has become an attractive home for foreign capital. Figure 21.4 captures the former and Figure 21.5 the latter if the capital stock is adjusted for imported capital, i.e. is not dependent upon domestic saving. Such factor endowments will give China a competitive advantage in many economic activities for the foreseeable future.

A few decades ago China had a comparatively even income distribution and the pursuit of individual wealth was not a socially acceptable goal. Rapid economic growth, however, both requires and produces entrepreneurs, and entrepreneurs command economic rewards for their talents and activities. China is facing, and will continue to face, rising income gaps between the rich and the poor. Deciding how national income should be allocated is a value judgement the government must face. The imposition of the relevant tax and transfers is an action the government must take.

China's population of over 1.3 billion currently exceeds that of India, and China's GNP per capita of around $7000 is about twice that of India. But India also is attracting foreign investment again, in part due to its large, inexpensive labour force. India is a democracy like Brazil, which has a substantially higher level of GNP per capita but a much smaller population. Russia, with its wealth of natural resources and large population, is the fourth member of the 'Big Four' developing countries that have the potential to become the largest economies in the world by the mid twenty-first century. These four are commonly referred to as the BRIC group (Brazil, Russia, India, China), a term that has grown in use in the last decade as their economic importance has increased. Other economies that are sometimes included alongside the 'Big Four' are South Africa, Nigeria, Turkey and Vietnam. Despite the various methods of grouping developing countries together, none has the same economic potential as China and India.

21.3.1 The Age of Prosperity, the Financial Crisis and the Great Recession

In the majority of the high per capita income countries, i.e. the mature capitalist countries, growth rates are now low by historical standards, causing severe problems for macroeconomic policy makers. The question is, why are those economies in trouble and what can they do to return to happier times? With your understanding of macroeconomics, you can understand the problems. The solutions are less obvious.

In the late twentieth century three of the world's most powerful leaders, Ronald Reagan (USA), Margaret Thatcher (UK) and Mikhail Gorbachev (Russia), expressed their belief in the free market as the most efficient mechanism for producing and delivering goods and services, and in democracy as the most stable form of government. Undemocratic governments in Latin America, Portugal and Spain fell, and government controls in the private sector were reduced. The removal of impediments to the free enterprise spirit was strongly resisted in some sectors; the coalminers' union in the UK was rendered powerless after bitter strikes, the Soviet Union collapsed after Gorbachev's reforms and only in China were democratic impulses suppressed. Globalisation and increasing levels of European integration ended a century of hostilities in Europe, and an age of rapid economic growth and prosperity began in the last decade of the twentieth century and lasted almost 20 years, until the financial crisis of 2008.

In the latter part of the last decade a number of factors combined to precipitate the worst recession since before World War II, which some economists are now referring to as the 'Great Recession'. Historically low interest rates encouraged more people to buy homes, and investors searching for a decent level of return to look towards the property market. Banks with surplus cash established 'subprime' mortgages, whereby buyers with poor credit history and no proof of income were able to take out a mortgage on property that they could ill afford. From 2000 to 2005 in some countries house prices doubled, and the argument used by lenders was: 'The future is going to be just like the present – get in on the boom now.'

Naive borrowers were told they could buy a property today that would appreciate significantly over the following years and could finally be sold at a significant profit.

Banks in some countries borrowed money at high rates to secure the cash to sell mortgages to people who could not afford them. Other banks packaged up their mortgages and sold them on to other banks and institutions. Credit Ratings Agencies approved these 'collateralised debt options' as AAA-rated – a sure investment. The stability of the whole system rested on the assumption that house prices would keep on increasing: eventually the bubble burst and house prices started to tumble. In extreme cases, such as the Republic of Ireland, house prices fell by 50 per cent from their peak. Many banks across the Western world, especially those with interests in the mortgage market, found that much of their asset base became worthless overnight. Banks stopped lending to one another, and to firms and individuals, in an attempt to rebuild their shattered capital bases. This was effectively a large and sharp reduction in the global money supply, resulting in slowing growth rates and decreased liquidity across the world. Over the same period the war in Iraq, rising demand from the developing world and later uprisings in the Middle East affected the oil market; significant increases in the price of crude oil and resultant increases in prices at petrol/gas stations increased inflation rates and resulted in a reduction of real income in economies already struggling with rising levels of consumer and business pessimism.

Some banks were so heavily involved in mortgage-backed securities that they went bankrupt. Bear Stearns, a large US investment bank, collapsed in 2008 and was acquired by JPMorgan Chase. Later that year, Lehman Brothers, another US investment bank, was allowed to go bankrupt. This resulted in turmoil across global stock markets as banks' shares around the world plummeted. Eventually governments had to step in to either rescue the banks with massive loans or let them go bankrupt. This meltdown further affected stock markets, whose fall added to the deteriorating economic situation.

Unemployment started to rise and continued to do so for some years. Many individuals who bought houses in the boom period on promises that prices would always continue to increase found themselves with mortgages significantly higher than current values of their properties. Some of these individuals also found themselves unemployed. This resulted in a situation in which no buyers existed and unemployed householders could not afford to seek employment elsewhere because all their capital was tied up in negative equity! This constraint on labour mobility affected – significantly – an early return to economic growth.

The 2008 financial crisis was blamed on failures in regulation, corporate mismanagement and reckless risk taking by financial institutions. As you learned in microeconomics, in a perfectly competitive market system, market failures in the form of public goods, externalities and economies of scale prevent the marginal equivalency conditions being attained. In 1776 Adam Smith pointed out that for markets to operate efficiently and for contracts to be valid and upheld, the rule of law must prevail. (Indeed, one of the largest impediments to greater economic growth in emerging countries is 'corruption' – lack of enforceable contracts, lack of competition, and bribery.) There also has to exist a reliable currency to both facilitate trade in the present and encourage borrowing and lending today for future economic activities.

However, financial markets need to be regulated. Banks have to have sufficient reserves and margin requirements to maintain stability. Banks developed complicated instruments such as derivatives and credit default swaps designed primarily to manage risk. But these instruments could also be deployed to take risks. For example, Bank A could bet with Bank

B that a certain currency would rise steeply in foreign exchange markets. Successful gambling led to large gains, many of which resulted in large bonuses for the winning managers. The outcome of such recklessness was that when people tried to get this money out of banks many were found to be bankrupt. Derivatives have been described as 'financial weapons of mass destruction'.

By 2011 the European Banking Authority had 'stress-tested' 90 banks and concluded that some failed to meet the necessary reserves to avoid risk and to restore confidence; many of the large leading banks passed the risk test but passed it very narrowly. They were encouraged to raise more capital, get rid of risky assets, add to their asset bases and begin lending to other banks rather than relying on the European Central Bank for funds. One of the side effects of taking such advice is unwillingness of banks to lend money to non-established borrowers. Even though interest rates were at an all-time low, there was little lending being made to first-time house buyers, and individuals with no borrowing records wishing to start new small businesses were being turned down by banks. The reduced supply of money in developed economies that commenced in 2007 when house prices started to fall continued. Thus a traditional source of economic growth and employment opportunities – entrepreneurship – was destroyed.

21.3.2 Government Reaction to the Financial Crisis

In the past, macroeconomic policy makers were concerned primarily with unemployment and inflation rates and trade-offs between the two – the Phillips Curve generation. But macroeconomic goals other than full employment and a low inflation rate have grown in importance over time. The correct balance between consumption, investment and government expenditure has become a concern with many politicians. Some economists now argue that during the age of prosperity G was too large, C was far too large and I was too small across the developed world.

One of the features of the late twentieth and early twenty-first centuries was concern for those individuals who were not sharing in the age of prosperity. Welfare programmes take many forms, some cash (children's allowances, unemployment compensation) but many 'in kind': free/subsidised housing, food stamps, free medical care and prescription drugs, free elementary education, free/subsidised university education. All of this, coupled with higher pensions, meant a significant increase in government outlays in the form of expenditure and transfers. Indeed, it was alleged in the UK that the most lucrative occupation of 'unemployable' youths was having as many children as possible with multiple partners to collect all the children's benefits. In addition, state and local, as well as national, governments were starting many projects and hiring consultants, often with little regard to marginal analysis.

The financial crisis, with major banks facing ruin and requiring large bailouts from central governments, brought the age of prosperity to an abrupt halt. Economic growth rates fell below zero in many countries, bankruptcies soared and business pessimism was rife. The initial response by many governments was to run an expansionary fiscal policy – i.e. a large budget deficit – coupled with an expansionary monetary policy, resulting in historically low interest rates in the USA, the UK and Europe. With your knowledge of the IS/LM model, you will be aware that monetary policy itself is ineffective in a severe recession – the Keynesian liquidity trap. Various economists argue that many of the major economies of the world entered the liquidity trap in the late 2000s. As you now know, the existence of the liquidity trap renders monetary policy ineffective in a severe recession, so many governments in the West turned towards fiscal policy as the tool of choice.

You will recall from Module 15 that the central government budget, like household budgets, is composed of two elements: outlays (expenditure) and income (receipts). But there, despite the claims of some politicians, the similarities end. Households are constrained on what they can spend in any year by their income and/or borrowing powers, which, for most people, is dependent upon expected income and/or wealth holdings. When the government runs a deficit, as most governments do, it has to borrow, and this it does by issuing government bonds – promises to return the sum borrowed at some time in the future plus an annual interest rate. The accumulation of all past deficits is the national debt, the interest payments on which are a significant part of government outlays.

One danger countries with large budget deficits face is creditworthiness. To finance deficits, governments have to borrow, and if lenders believe countries are not addressing deficits in a responsible fashion then the cost of borrowing will rise or borrowing opportunities may disappear. After the financial crisis, various nations had their credit ratings downgraded – with severe consequences. Greece, Ireland and Portugal all had to be bailed out by other European countries when they found themselves unable to meet the obligations they had built up in the past. A condition of these bailouts was reduced expenditure and/or raised taxes – referred to in the media as 'austerity measures'. These countries had little choice but to agree to such measures, and, with interest rates at an all-time low and companies and consumers pessimistic about the future in the face of high and rising unemployment, the medium-term economic outlook was depressing.

With the potential consequences of high levels of national debt in mind, many other nations also shifted their macroeconomic goals towards budget deficit reduction. Unless some factor is operating in an offsetting direction, budget deficit reduction requires reducing government expenditure or raising taxes, or both. What could an offsetting factor be? Increasing levels of GNP would suffice, but with pessimism in the business sector the level of private investment would be decreasing. With potential output growing faster than actual output, unemployment would increase, which in turn could cause consumption to decrease; the result, a stagnant economy often in the liquidity trap rendering any monetary policy ineffectual. With all capitalist economies in a similar position, no relief from increasing exports would be likely. One possible benefit of decreasing government expenditure is forcing the government sector to become more efficient and shed workers whose productivities are low or zero but who have been protected by contracts or convention. In universities, for example, a reduction in faculty numbers may mean the disappearance of classes with fewer than 10 students in each.

The countries that were most affected by these events were the mature capitalist countries with, unfortunately, the slow economic growth rates. The focus of those nations on budget balancing is a long-term strategy, and this must mean sacrifice of other desirable economic goals, such as low unemployment, in the short run. In the meantime the rapidly growing economies such as China will play an increasingly important role in world economics and politics.

21.4 The Pessimists versus the Optimists

What is the world going to be like two or three years from now? You might predict quite confidently that, in the absence of a world war, the world will not be very different from the world of today. If current trends continue, world potential output will grow between 2 per cent and 3 per cent per annum.

What about the year 2030 – any significant changes predicted? The answer may depend upon your age and what you mean by 'significant'. Life expectancy has increased dramatically in the past 50 years. While a healthier diet and attention to exercise are cited by some as the causes, those factors are in fact minor. Mortality rates have declined, meaning that more people survive to old age: heart drugs, heart operations, treatments of certain cancers, and technological changes in hospital emergency services have all contributed to significantly longer life spans. (Personal genetic make-up, however, remains the major determinant of how long one can expect to live – it explains around two-thirds of the variance in life expectancy. If you are male, choose a long-lived father, and if female a long-lived mother!) Unfortunately, the increased longevity has not been accompanied by increases in health or in perceptions of physical well-being.

Many of the large developed economies in the world today feature an aging population. These individuals are large consumers of medical services no longer producing goods and services but receiving 'large' retirement pensions ranging from 50 per cent to 80 per cent of highest salary/wage when employed. Pension funds established when life expectancies were around 70 (three score years and ten) have been forced to raise premiums for existing workers, delay payment of pensions and/or persuade workers to postpone retirement and reduce future pension payouts. This aging population is accompanied by smaller families and, without immigration, a small labour force to support the old, the sick and the lame. In contrast countries such as China and India have a much higher proportion of the population in working age brackets.

Why are aging populations limited to the developed countries? Go back only three generations and families of six to twelve were not unusual. Now many families in the developed world have two children or fewer. Again economics can help explain why. Suppose all goods are classified as consumption or investment goods; where would you place children? If as parents you hoped to depend eventually upon your children for support in your old age, or suspect one may become a soccer superstar or movie idol, then those children are investment goods, and the more you have the higher the probability of high investment returns. If, on the other hand, you consider your children consumption goods, then the 'prices' of those goods will be a major determination of numbers. As rational consumers you will economise on expensive consumption goods, and as your income and wealth increase over time so will your opportunity costs. Children can consume lots of their parents' time, and time is scarce and therefore expensive. Save scarce time: have fewer children.

Technological change in the medical field will continue to lengthen life expectancy, which in turn will increase the demands on resources for the care of the elderly. By the year 2030, 20 per cent of the population of the major capitalist countries will be over the age of 65 and a significant number will be 80 years of age or older. One in five children born in the UK in 2011 will live past the age of 100. A substantial portion of the GNP of 2030 will be required to be spent on the personal care of these people. Looking after the old is not cheap; in the UK today, the cost of providing care for an elderly person in a private room in a nursing home without any medical treatment exceeds the annual average wage. You might be well advised to do some financial planning now for your own old age.

Robert Malthus, an English economist writing 200 years ago, is credited with giving economics the name of the 'dismal science' because he predicted a world of continuous growth in population outstripping the growth of resources – and consequently the growth in goods and services. Today the Malthusian spirit survives, and pessimists abound. The world population of six billion in 2000 is predicted to double before the year 2100, and that estimate for the pessimists spells mass-starvation and environmental disaster; the 10 per cent of the world's population today that are living in dire poverty will expand, and continue to expand. This message was given to the world in forcible terms by the Club of Rome in the late 1960s and was based on the simple economic model of demand and supply. The Club of Rome predicted a world of ever increasing scarcity and rising prices and urged nations to concentrate on technology, to make infertile land fertile, to develop new resources and to limit population growth; these suggestions would increase supply and slow down the growth in demand.

The pessimist's model is very persuasive. We know the world's resources are finite; there is a finite amount of land, a finite amount of coal, copper, iron ore and oil. We also know the world's population continues to grow: three billion in 1960 reads around seven billion today. Yet, strangely, world prices of land, coal, copper and food are not soaring; indeed, compared with manufactured goods, many of these prices have been falling. Farmers in the mid-West of the USA, who believed the Club of Rome story and who cultivated additional acres, have gone bankrupt; agriculture in most of the world is not very active in advocating the purchase of stocks in basic raw materials. What is wrong with the basic economic model?

Consider an alternative hypothesis. World demand for goods and raw materials is growing, but the world supply of resources is infinite. You may accept the first half of that hypothesis but certainly not the second, the *infinite supply assumption*, and yet it appears to be consistent with the facts. Enter the optimists. How big, relative to the world, is the world's population? Malthus's prediction of world population growth led to the theoretical possibility of each person having sufficient room only to rotate with arms outstretched before he or she touched other people's outstretched arms, i.e. about 2 square metres of surface area on which to exist. If we perform the outstretched arms test today, what is the smallest country with standing room only for today's population of roughly seven billion? We shall need 8500 square miles, i.e. a country only of the size of El Salvador.

The next time you are on a plane, take a window seat and look down. What do you see? You see lots of water – 70 per cent of the earth's surface consists of oceans, the economic potential of which, outside of being a rubbish dump and a fish farm, many contend that the human race has not really started to exploit. You see the sun – a future source of substantial energy? You also see lots of land, mostly uncultivated. What you do not see are people; there essentially appears to be no one down there. There are certainly not enough people there to consume all the goods and services the world is capable of producing now or in the near future. But the United Nations Population Fund informs us that 10 per cent of the world's population is starving, and the power of television has brought such misery into our homes. Why? The answer is not that we lack the technology and resources to feed the world's population many times over; the answer is that, for a great variety of reasons, we cannot arrange the world so that suppliers have sufficient incentives to get food to those who are starving.

Consider food. Genetic engineering, development and efficient application of pesticides, adding trace elements to infertile ground, the control of rodents, improved handling and storage methods, refrigeration, specialisation and production, matching crop strains to prevailing climatic conditions – all these are processes that have gathered pace in the past few decades and have led to enormous gains in agricultural productivity. Many argue that recent developments are the 'tip of the iceberg' and will pale in comparison with developments to come. The serious problem facing many nations involved in food production – and the source of much friction in the European Union – is what to do with agricultural surpluses and redundant farmers with their agricultural land. Paying farmers not to produce, burning crops, or ploughing unwanted produce into the ground have all been features of nations involved in food production in the past few decades. Why not give these surpluses to the poor nations? Some transfer has taken place, but the impact on the plight of the starving nations has been negligible. There are two principal reasons for this sad state of affairs; first, the existing infrastructure of most impoverished nations, and their political instability, tend to inhibit the distribution of the world's 'surplus' foods to the needy; second, the proportion of GNP that the high-income nations are prepared to vote away to the poorer nations is minuscule. Do you know of any political party in today's developed countries with a political platform that includes an increase in income taxes of, say, 10 per cent, the proceeds of which are to be allocated to reducing the income differentials between the poor and the rich nations? Would such a party stand a good chance of being elected? Humanitarianism does not appear to be a profitable business to be in, for politicians or resource owners.

You might accept these arguments for potential food supplies, especially if you enquire about the latest milk yields from cows; the latest size of beef cattle and pigs; the dramatic increase in farmed fish; and the millions of eggs that can emerge annually from a small farm with two henhouses, one farmer with an assistant, and technology controlling temperature, light, food, cleaning, egg collection, inspection and packaging. But what about oil? Surely there we have a potential world shortage. Are oil stocks therefore a good buy? The answer appears to be 'no'. We keep on finding more proven oil reserves almost daily, and we keep on improving our ability to extract additional oil from existing fields. The very low costs of extraction of oil from the Middle East states for many decades, and their seemingly endless supplies, have been disrupted by only three events to date: the OPEC oil cartel, conflicts and wars among the Middle East states and the popular uprisings of 2011, causing oil prices to soar in the short run. Exploration to find alternative sources has proved to be extremely fruitful, with vast fields discovered in the Caucasus, Alaska, Australia, Russia and the Gulf of Mexico. The major difference between those new finds and Middle Eastern oil is that extraction costs are much higher.

The fears that the ancient Greeks had about running out of forest as a source of wood fuel were echoed in the nineteenth century concerning coal supplies, and are echoed in the twenty-first century about oil supplies. Come the twenty-second century, coal and oil may be museum relics of a bygone era, with other energy sources by then in regular use.

21.5 Globalisation

The goal of commercial companies is profit making; those profits are a return to the stockholders who own the companies and may constitute also incentives to the executives/managers who run the companies.

Part of the duties of managers is finding the best source of factor inputs including labour. This search has taken many managers of companies in industrialised nations to foreign countries to fulfil labour requirements in situ, i.e. establishing factories abroad, where labour is cheap. In some instances factories in the home country are closed down and complete operations established abroad; in other instances companies will carry on business both at home and abroad. This movement of business abroad is known as *globalisation*.

In macroeconomic terms it means potential output domestically is lower than it otherwise would have been and potential output abroad is higher than it otherwise would have been. In microeconomic terms there are winners and losers. Those who gain are the company shareholders and purchasers of the goods produced, since lower costs of production mean competitively produced goods are cheaper than they otherwise would have been. Foreign workers now have jobs and gain also. Those who lose, however, are the workers whose jobs have gone abroad and who cannot find alternative employment. Indeed, if such workers cannot find any other employment and stop seeking jobs because none are available, they are no longer classified as unemployed in some countries because they are no longer considered to be in the labour force. What is particularly painful about this process is that, if the company is large and dominant in a town/city or even state, not only do thousands of people become unemployed but also the multiplier process – less income, less consumption – can cause a high level of secondary unemployment, leading to depressed areas.

Is this any different from a foreign country having a competitive advantage in the production of a good leading to all the benefits of free trade? In one sense the answer is ‘no’. The goods produced abroad, even if the only foreign factor input is cheap labour (capital, management, technical know-how, all supplied by the home country), are exports from the cheap labour country and imports to the home nation. One major difference, however, can be the shock effect of the suddenness by which an announcement is made that the local factory employing 5000 people is moving offshore. The larger such employers are in a given area or state, the greater the depressive impact. The impact of increasing imports tends to be a more gradual process and in most cases does not completely replace domestic producers.

21.6 The Environmentalists

The food/raw materials/population pessimists have been joined recently by the environmental pessimists concerned with damage that economic development is causing the natural world. The main focus of concern is currently global warming. While accurate records of world temperature changes are non-existent prior to this century, there is clear evidence that the world, in the past, has experienced and withstood ice ages and warm periods, with calculations that a mere 5 °C average annual temperature separates such conditions. Some environmentalists argue that the observed temperature increase witnessed since 1900 is no longer a cyclical trend but has been caused by increasing pollution levels in the atmosphere attributed to the burning of fossil fuels. Climate models are now in vogue, some of which predict an increase in world average temperature of between 3 °C and 6 °C by the year 2100. If such predictions prove to be true, water levels, as ice melts, would rise significantly throughout the world, wiping out vast areas of low-lying land and destroying part of the capital stock of many nations. The developed world would suffer least (Holland has the resources to build higher dykes); the less developed world, lacking such resources, would become ever poorer.

The climate models, however, are subject to even greater uncertainty than economic models! Short-term weather forecasters resemble mothers-in-law in the joke category. Long-term climate models must embrace the disciplines of physics, chemistry and biology, with no reliable empirical evidence to guide the model builders (for instance, is vegetation capable of removing more carbon dioxide from the atmosphere as pollution levels rise?). If we believe the environmentalists who say that carbon emissions must be reduced by 60 per cent to stabilise pollution levels, why aren't all governments enacting legislation to reduce emissions by the amount advocated? The answer, of course, is largely that such reductions are expensive in the short term. The opportunity cost of pollution reduction is all the goods and services that would have to be surrendered in the fight against pollution. If the environmentalists are wrong, then such actions would be a waste of resources. In addition, in the shorter run, governments undoubtedly would encounter severe political pressure by announcing to their electorates that current plans for better housing, more goods and services, more hospitals and schools, a better health service or better provisions for the old should be sacrificed for a better environment for generations as yet unborn and currently with no voting rights.

In addition, the politicians are encouraged by the optimists who argue that the world can support a much larger population given today's technology. Indeed, they go further and warn of a potential labour shortage – a current concern of Germany and Hungary. Living standards, the optimists argue, have improved because of increases in the population and labour forces. The doubling of the world's population in the second half of the twentieth century has been accompanied by increased GNP per capita throughout the world, decreased infant mortality rates, and increased life expectancy. The optimists point out that, as income per capita rises in countries, birth rates fall – this is empirically verifiable. Why? The economists' response is that in their world there are only two types of goods – consumption and investment. What are children? For much of the less developed world children are investment goods – one may turn out to be a genius and earn sufficient income to support the whole family – or, collectively in the family system, they are a form of insurance for taking care of aged relatives. Among high-income families looking for no long-term financial return, children are consumption goods – there to be enjoyed (at least that's the theory!).

Children, however, are expensive: they take up valuable time in playing with them, talking to them and taking them to tennis tournaments. As wages rise, time becomes more valuable, and consequently the cost of children rises. You will recall from the microeconomics modules that you buy less of a good the higher the price; therefore rationality dictates society has fewer children as incomes rise. This model predicts that the bulk of the increase in the world's population in the next century will occur in the developing countries, and that the 19 per cent of the world's population in North America, Europe and Japan producing 83 per cent of the world's GNP will experience a falling indigenous population. In those developing countries in which GNP increases, population growth will slow. Developed countries will continue to allow immigration to meet labour requirements.

The United Nations predicts that the world population will stabilise at around 12 billion people. However, demographers, like any other scientists involved in the prediction game, are not always correct. Several predicted in the late 1950s that Japan, a nation producing tawdry toys for export, overcrowded and with few natural resources, faced a bleak economic future!

Political scientists joining the debate argue that the growing disparity between the rich and the poor nations, the great imbalance between distribution of the world's goods and services, and the differences in population pose the greatest threat to long-term world stability and security. If the developed world refuses to share its largesse with the poor, it is offering itself as the ideal blackmail victim. The past three decades have shown the havoc that can be caused by small dedicated groups of terrorists armed with conventional weapons located in strategic positions such as aircraft. The medical profession further warns of a non-human threat. Despite the tremendous medical advances in combating bacteria, no effective antidote to the viral threat is on the horizon. More than several maintain that the real battle for the world will be between mankind and viruses, and to date it is not clear who will win.

Those are major long-term issues facing the world, many of which, it is argued, require immediate action. Such action, it is further argued, requires a degree of cooperation between the governments of the leading capitalist countries never seen yet in a non-war situation.

Appendix 1

Practice Final Examinations

This section contains two practice final examinations, which are indicative of the type and level of material appearing in the Heriot-Watt University final MBA degree examination.

The duration of the examination is three hours. The points value of each section is shown. Within the total time of three hours, students may allocate their time among sections as they see fit. The pass mark is 50 per cent.

There is no choice in the selection of questions to be answered.

In the multiple choice questions, there are no points deducted for wrong answers.

For each question, a solution is provided that will allow students to assess their performance. The examination serves two purposes: to test understanding of the course and to provide information on standards required to pass the university final degree examination.

The rationale for providing two examinations is that students who have worked through the course, have taken the first practice examination and, on the basis of their performance in that examination, are not satisfied that they have attained mastery of the material, will be able to study the course again and have a second opportunity to test themselves. Where the first examination is satisfactory, the second may, of course, be used for additional practice.

Practice Final Examination 1

The examination is in three sections

Section A: Multiple Choice Questions

Section B: Case Study

Section C: Essay Questions

Section A: Multiple Choice Questions	
30 questions each worth 2 points	
Total points available in Section A	30 × 2 = 60
Section B: Case Study	
Total points available in Section B	= 20
Section C: Essay Questions	
2 questions each worth 40 points	
Total points available in Section C	2 × 40 = 80
Total points available for examination	= 160
Pass mark = 50 per cent of 160	= 80

Multiple Choice Questions

Choose only one option amongst A–D in each case, and make sure that option is clearly marked.

1 The labour force in a car plant has approached management with the following proposition: 'We shall raise average labour productivity (cars per worker) by 5 per cent if you agree to increase our wages by 5 per cent.' Assuming all additional cars produced can be sold at current prices, what would happen to the payments to all other factors of production if management agreed? They would

A. increase by 5 per cent.

B. be unchanged.

C. decrease by 5 per cent.

D. decrease by more than 5 per cent.

2 Suppose that the highly competitive market for electronic calculators is in equilibrium today. If the government decided that every school pupil should be given a calculator, bought at government expense, and at the same time the average total cost of producing calculators were to decrease because of another technological breakthrough in the electronics industry, what would be the joint effects of the government decision and the technological breakthrough?

A. An increase in the equilibrium price of calculators and an increase in the equilibrium quantity.

B. An increase in the equilibrium quantity of calculators; the equilibrium price might rise, fall or be unaltered.

C. A decrease in the equilibrium price of calculators; the equilibrium quantity might rise, fall or be unaltered.

D. A decrease in the number of firms in the industry, and an increase in the equilibrium price of calculators.

3 It has been said that the fundamental fact of scarcity is no longer applicable to the UK economy. Which of the following is correct? This statement is

A. true, because the welfare system in the UK ensures that the basic needs of the poor are met.
B. true, because goods such as tea and coffee, which are not produced in the UK, can be imported from abroad.
C. false, because not all wants of all UK citizens are fully satisfied.
D. false, because a significant number of families in the UK have incomes below what is generally accepted as the poverty level.

4 The table below shows how the factors of production may be combined to produce 100 units of a certain good. Using only the information in the table, which one of the methods can be dismissed by the producer on the grounds of technical inefficiency?

A.

Land	Units of labour	Capital
10	20	10

B.

Land	Units of labour	Capital
10	30	6

C.

Land	Units of labour	Capital
10	40	6

D.

Land	Units of labour	Capital
10	50	4

5 Which one of the following quotations is *not* an example of the 'law' of diminishing marginal utility?

A. 'The second glass of beer tastes just as good as the first.'
B. 'There's a limit to how much ice cream you can eat in a day.'
C. 'After a while, even TV gets boring.'
D. 'All work and no play makes Jack a dull boy.'

6 A consumer's typical demand curve for a good is downward-sloping to the quantity axis because at higher prices the consumer

I. encounters a leftward-shifting supply curve for that good, resulting in a lesser equilibrium quantity.
II. buys less of the good in place of other goods, which are now relatively less expensive.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

7 A local government has been exploring schemes to make its region more attractive to industry. It has produced the following estimates.

Project	Description	Increase in employment
Housing estate	Construction of 300 houses	500 jobs
Industrial estate	Construction of 10 advanced factories	2 000 jobs
Motorway	Construction of internal motorway system	500 jobs

To achieve economic efficiency, which of the following policies should be adopted?

A. Undertake only project 'Industrial estate'.

B. Undertake project 'Motorway' or 'Housing estate' but not 'Industrial estate'.

C. Insufficient information exists to determine which, if any, of the projects should be undertaken.

D. Undertake all three projects.

8 Perfect weather conditions in Brazil have produced a bumper coffee crop this year. Although the whole crop could be sold at a price that would more than cover the cost of storing and delivering it, a Brazilian economist has suggested that sufficient coffee be burned to prevent the price of Brazilian coffee beans from falling. Which of the following would result from acting upon this suggestion?

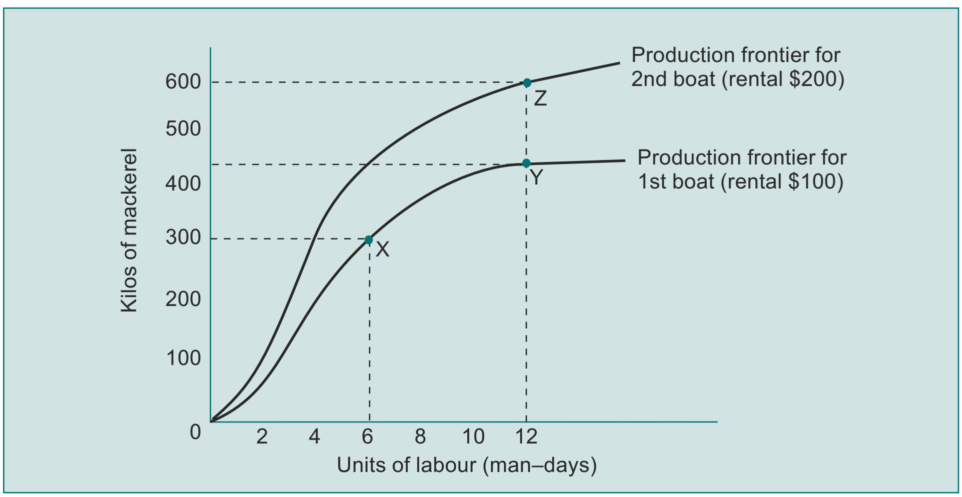
A. A reduction in Brazilian coffee growers' real income.

B. A reduction in world real income.

C. An increase in Brazilian coffee growers' real income.

D. An increase in Brazilian real income.

9 The diagram shows two production frontiers for catching mackerel, in other words the maximum amount of fish that can be caught with two different amounts of a fixed factor of production (boat) and with different amounts of a variable factor of production (labour). The rental costs of the boats are $100 and $200 respectively. All units of labour are identical.



I. A movement from X to Y exhibits increasing marginal returns to labour.

II. A movement from X to Z exhibits increasing returns to scale.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

10 People buy more fuel when prices are reduced and less when they are raised. Suppose that advertising does not increase the amount of fuel bought but does influence the choice of brand. Suppose also that fuel of any one grade is for practical purposes a homogeneous product across brands and that any change in industry cost is passed on in part to the public in higher or lower fuel prices. Which of the following would necessarily be true?

A. If all fuel refiners increase advertising expenditure, refining industry total profits will increase but the general public will be worse off.

B. If all fuel refiners increase advertising expenditure, refining industry total profit will decrease and the general public will be worse off.

C. If one fuel refiner increases his advertising expenditure, that refiner's profit will decrease but other refiners' profits will increase.

D. If one fuel refiner increases his advertising expenditure, that refiner will have higher profit, higher sales and lower costs per gallon sold than other refiners.

11 When travelling to an airport from a city centre, most people use the airport bus or travel by taxi. In response to rising fuel costs, taxi fares have risen although the bus fare has remained unaltered. If some travellers change their mode of transport in response to the change in fares, what impact will this have on taxi and bus revenues?

A. Taxi revenues will rise and bus revenues will remain unaltered.

B. Taxi revenues will fall and bus revenues will remain unaltered.

C. Bus revenues will rise and taxi revenues may rise, fall or remain unaltered.

D. Bus revenues will rise and taxi revenues will rise.

12 For several decades the Eastfield iron foundry has been processing iron ore and has been an important employer in the local area. Until recently, no one complained about the ash particles, the bulk of which fell on vacant land. However, a farmer who bought the vacant land and is now cultivating it has complained to the government that the ash particles are polluting his fields and causing wheat yields to be less than they would otherwise be. If the government wishes to promote efficient resource use in the area. which of the following policies should it adopt?

A. Force the iron foundry to stop polluting the area or put it out of business.

B. Allow the iron foundry to continue as it is but charge the owners for any resultant decrease in the value of the farmer's output.

C. Refuse to intervene and let market forces resolve the conflict between the farmer and the foundry owners.

D. Provide the farmer with an alternative field and compensate him for any loss of income incurred.

13 Approximately one-half of a city university's students live in rooms/flats that they rent in the private market. The others live in university rooms or flats. Average housing expenditure does not vary significantly between the two groups. In the light of government cutbacks in university finance, the university has decided to convert its student accommodation to offices and seminar rooms. The 3000 students who will be affected have been told they will have to find their own accommodation in the private housing market in the city. Other things being unchanged, such as population, what effect will the university's decision have on the market for private rental housing when the university housing service ceases?

A. An increase in rent, followed later by an increase in the quantity of private rental housing supplied.
B. A decrease in the number of non-student tenants and simultaneously a decrease in rent.
C. A decrease in rent as home owners and landlords anticipate the increased student demand.
D. No effect, because the students who occupy university housing will not be able to afford private housing.

14 It is rumoured that the government intends to increase the sales tax on alcohol and tobacco from 5 per cent of the selling price to 20 per cent and to use this tax revenue to subsidise gas and electricity production, i.e. to lower the prices of these commodities. If the government carries out such a plan

I. the average consumer will have a lower real income because the cost of living will be higher.
II. income inequality will be increased if people with lower incomes spend a greater proportion of their incomes on alcohol and tobacco but no higher a proportion on gas and electricity than higher-income groups.

Which of the following is correct?

A. I only.
B. II only.
C. Both I and II.
D. Neither I nor II.

15 A government announces that it does not intend to launch any more satellites; the ten it has launched so far are more than adequate for its needs. On average each satellite cost $1 million; they are all monitored from a single station, whose running costs are $50 000 per annum. A group of businesses with plants abroad and a private university with isolated students have requested permission to use one of the satellites to 'bounce' messages to their plants and students. Such a use of the satellite will not affect any government use and will not incur any additional government cost. Given that the government goal is efficient allocation of resources, how much should the government charge the businesses and the university for using the satellite each year?

A. Zero.
B. Zero to the university and a share of the $50 000 to each of the businesses.
C. A share of the $50 000 to both the university and to each of the businesses.
D. A share of the $50 000 and a share of the $1 million to the university and to each of the businesses.

16 One year, the serfs in a feudal manor produced only $500 worth of ploughs, together with $2000 worth of wheat, all of which they used to produce $3000 worth of flour, all of which in turn they used to produce $4000 worth of bread. The feudal lord took half the bread as his dues and exchanged it for other goods from outside the manor such as cloth, meat and drink.

Which of the following represents the gross output (gross product) of the manor for the year?

A. $4500.

B. $6500.

C. $9500.

D. $11 500.

17 The following table shows Money GNP and Real GNP in a hypothetical closed economy for three years:

Year	Money GNP ($m)	Real GNP ($m)
t	113	113
$t+1$	115	112
$t+2$	120	114

Assuming a constant number of workers, it follows that

I. the supply of money increased by $7 million between Year t and Year $t + 2$.

II. output per worker was higher in Year $t + 2$ than in Year t.

Which of the following is correct?

A. I only.

B. II only.

C. Both I and II.

D. Neither I nor II.

18 In any given year, some businesses sell more goods and services than they produce, some sell exactly what they produce while others sell less than they produce. However, in the national income accounts in a closed economy, total expenditure for final goods and services is always equal to total output.

Which of the following is correct?

This apparent disparity is explained by the fact that

A. what is produced must be sold eventually.

B. in the national income accounts, total expenditure is defined to exclude depreciation.

C. in the national income accounts, total expenditure is defined to include any change of inventories.

D. the businesses that sell more than they produce always cancel out the businesses that sell less than they produce.

19 A government hopes this year to balance its budget (i.e. make government expenditure (G) and tax receipts (T) equal), achieve full employment, and distribute gross national product (GNP) among consumption expenditure (C), investment expenditure (I) and government expenditure (G) in the proportions 60 per cent/20 per cent/20 per cent. The Chief Economic Adviser has warned the government not to make a public statement to this effect. On which of the following is his professional advice based?

A. The setting of 20 per cent of GNP as being devoted to government expenditure is incompatible with a balanced budget.
B. Only by remote chance will values of GNP, C, I, G and T occur that will achieve the three objectives.
C. Balanced budgets are incompatible with full employment.
D. It is not possible to influence the distribution of GNP among C, I and G.

20 In a hypothetical economy, a group of economists estimated that the value of the marginal propensity to consume was two-thirds. An increase in government expenditure of $10 000 million resulted in an increase in real national income of $20 000 million. The reasons why the increase was not $30 000 million are

I. the marginal propensity to import decreased.
II. the economy was close to full employment when government expenditure was increased.
III. the increase in demand for money caused interest rates to rise and private investment to fall.

Which of the following is correct?

A. II only.
B. II and III only.
C. I, II and III.
D. Not I nor II nor III.

21 Suppose a significant decrease in the GNP of other European Union countries led to a decline in United Kingdom exports; and suppose that the government, because of this, increased the interest rate from 8 per cent to 14 per cent but made no other change in policy.

Which of the following is correct?

From the policy adopted it can be inferred that the government was more concerned about

A. the rising unemployment rate than the rising inflation rate.
B. the falling value of sterling than the rising unemployment rate.
C. the falling investment rate in export industries than the rising inflation rate.
D. the rising unemployment rate than the falling value of sterling.

22 ‘The more money there is in the economy, the more people spend. The more people spend, the higher is GNP. Therefore, the greater the supply of money, the better off people must be.’

Which of the following is correct?

A. The statement is correct because increases in the money supply lead to decreases in interest rates, which in turn lead to increases in investment.
B. The statement is incorrect because real GNP is limited by the economy’s potential to produce goods and services.
C. The statement is correct because increases in the money supply lead to increases in interest rates, which in turn lead to increases in consumer saving.
D. The statement is incorrect because the supply of money has no effect on consumer, government or investment expenditure.

23 'Since there is a limit to the amount of goods and services a nation can produce in a given year, any expansion of a nation's currency will mean a diminution in the real value of the currency.' Which of the following is correct? The quotation is

A. incorrect, since even in the short run the exchange rate determines the value of a nation's currency.
B. incorrect, since an increase in the money supply may lead to increased demand for goods and services and, given available resources to an increase in real output.
C. correct, since even in the short run all money supply increases in the UK have been associated with corresponding price increases.
D. correct, since a nation's wealth or resources are in fixed supply at any point in time.

24 Unemployment reached an all-time high in the United Kingdom economy in the early 1930s. Many people have argued, that had World War II never happened, the Great Depression might have continued. In the absence of the war, the following policies would have been appropriate for decreasing unemployment:

I. a decrease in the money supply to increase interest rates and stimulate savings.
II. an increase in income taxes to decrease consumption and increase investment.
III. a decrease in government expenditure to free more resources for use in the private sector.

Which of the following is correct?

A. I only.
B. II only.
C. I and III only.
D. Not I nor II nor III.

25 If at the beginning of a year an economy were operating at full employment and if during the year aggregate demand were to remain constant but potential output (full-employment output) were to increase, which of the following must occur?

A. An increase in the inflation rate.
B. An increase in the unemployment rate.
C. An increase in interest rates.
D. An increase in the balance of payments deficit.

26 Between March and September the exchange rate of the pound against the dollar fell from \$1.50 to \$1.40. At the same time, the interest rate in the USA rose from 8 per cent to 12 per cent. If there were no inflation in the USA, which of the following would be the effect on a UK holidaymaker who went for a week's holiday to the USA with £500 in March compared with one who went with £500 for a week's holiday in September?

A. The March holidaymaker would be able to buy more for her £500.
B. The September holidaymaker would be able to buy more for her £500.
C. The effect would depend on whether the £500 was exchanged for dollars in the UK or in the US.
D. It is impossible to say who was better off because of the rise in the interest rate.

27 The German economy was operating at full employment. Its high wage rates attracted employed skilled workers from Turkey to work in Germany. These workers sent home the bulk of their wages, which were spent by their families in Turkey.

The effect of the migration of labour to Germany was to

I. increase potential (full-employment) output in Germany.
II. reduce potential (full-employment) output in Turkey.
III. increase real output in Germany.

Which of the following is correct?

A. I only.
B. II only.
C. I and III only.
D. I, II and III.

28 A government decided to abandon import controls and as a result there was a sudden increase in the value of imports. Which of the following represents the impact that the increased imports would have on the country's National Income? It would be

A. increased by an amount equal to the value of the additional imports.
B. reduced because imports are leakages from the circular flow of income.
C. unaffected because the amount paid for the additional imports would be exactly equal to their value.
D. increased by an amount equal to the value of the additional imports times the multiplier.

29 A government increases expenditure on investment in the coal and steel industries by $40 million and discovers that this forces up interest rates, which in turn leads to a decrease in private investment expenditure of $40 million. Assuming no other offsetting changes in the economy, what is the impact of the increase in government expenditure?

An increase in

A. real GNP.
B. the budget deficit (reduction in the budget surplus).
C. potential (full-employment) GNP.
D. the money supply.

30 A simple 'textbook' economy is described by the following:

C = 20 + 0.5Y, where
C = consumption expenditure
Y = national income

I = 10 where	I = gross investment expenditure
G = 10 where	G = government expenditure
X = 10 where	X = exports
Z = 10 where	Z = imports
M = 300 where	M = money supply

What is the value of the equilibrium level of national income?

A. 80.
B. 100.
C. 120.
D. 300.

Case Study

1 The only time that fishing boats can enter or leave a particular harbour is at morning high tide, which occurs between 6.00 and 7.00 am. In practice, boats have 24 hours both to travel to and from the fishing grounds and to catch fish. The fishing grounds are 30 miles from the harbour. The faster the boat travels to and from the fishing grounds, the more time is available for fishing and the greater is the amount of fish caught. However, the faster the boat travels, the greater the amount of fuel consumed. This relationship between speed and fuel consumed is non-linear, as shown in the table below.

Speed (miles/hr)	Gallons for 60 mile voyage
6	10
7	11
8	13
9	16
10	20
11	25
12	31
13 (max speed)	38

When the boat reaches the fishing ground it drifts, thus using no fuel. If each hour of fishing yields fish worth $10, if the fuel cost is $1 per gallon and the goal of the boat owner is profit maximisation, how fast should the boat sail to and from the fishing grounds? Explain your answer in detail.

Essay Questions

Essay 1

If too much or too little economic activity is undertaken by, or regulated by, government in a market economy, neither economic efficiency nor an equitable distribution of income will be achieved. Discuss.

Essay 2

In many of the capitalist nations of the world today, a significant output or employment gap exists, i.e. actual GNP is significantly less than potential GNP. Economics textbooks suggest that monetary and/or fiscal policy can eliminate output gaps. Resolve this apparent paradox.

Practice Final Examination 2

The examination is in three sections
Section A: Multiple Choice Questions
Section B: Case Study
Section C: Essay Questions

Section A: Multiple Choice Questions	
30 questions each worth 2 points	
Total points available in Section A	30 × 2 = 60
Section B: Case Study	
Total points available in Section B	= 20
Section C: Essay Questions	
2 questions each worth 40 points	
Total points available in Section C	2 × 40 = 80
Total points available for examination	= 160
Pass mark = 50 per cent of 160	= 80

Multiple Choice Questions

Choose only one option amongst A–D in each case, and make sure that option is clearly marked.

1 A recent government document states that UK society would be better off if UK universities produced more engineers and fewer social scientists (e.g. economists). Which of the following is correct? Assuming the government is right, it follows that

A. the marginal product of engineers is positive and the marginal product of social scientists is negative.
B. the benefit to be gained from using one more unit of resources in order to produce engineers is greater than the benefit of using it to produce social scientists.
C. the opportunity cost of engineers is greater than the opportunity cost of social scientists.
D. at going wage rates, there is excess demand for engineers and excess supply of social scientists.

2 Using the notation shown, an economics professor wrote the following equations on the blackboard. Which equation is incorrect?

AVC	= average variable cost
f	= is a function of/depends upon
TVC	= total variable cost
Q	= output
ATC	= average total cost
FC	= fixed cost
MC	= marginal cost

ΔTC = change in total cost
ΔQ = change in output
ΔTVC = change in total variable cost

A. $\mathrm{AVC} = f(\text{Average Product of the Variable Factor Input})$
B. $\mathrm{ATC} = \mathrm{FC} + \frac{\Delta \mathrm{TVC}}{\Delta \mathrm{Q}}$
C. $\mathrm{MC} = \frac{\Delta \mathrm{TC}}{\Delta \mathrm{Q}}$
D. $\mathrm{MC} = \frac{\Delta \mathrm{TVC}}{\Delta \mathrm{Q}}$

3 In an electronics company, all assets other than labour, i.e. factories, machinery etc., are constant and valued at £1 million. The annual wage bill for the labour force is £0.25 million. If both output per worker and wages increase at 4 per cent per annum, what happens to the return to all other assets?

A. It remains constant.
B. It increases by 1 per cent per annum.
C. It increases by 3 per cent per annum.
D. It increases by 4 per cent per annum.

4 In perfectly competitive markets, whenever a firm's short-run profit is being maximised or its short-run loss is being minimised

I. the difference between average cost and average revenue is greatest.
II. marginal cost equals marginal revenue.
III. average cost is at its minimum.

Which of the following is correct?

A. II only.
B. I and II only.
C. II and III only.
D. I, II and III.

5 Consider the following average cost schedule for frozen turkeys. These frozen turkeys can be sold in unlimited amounts for £10 each.

Output	Average cost (£)
50	8.00
60	7.00
70	6.50
80	6.75
90	7.00
100	7.30
110	8.00

To maximise profit, which of the following should be the output level, and why?

A. 60, because the decrease in average cost is maximum.
B. 70, because average cost is minimum.
C. 100, because marginal revenue equals marginal cost.
D. 110, because it is the largest output given for which average cost is less than its price of £10.

6 You have been hired by the local shovel works to recommend what actions to take to maximise profit. Shovels sell for £5.00 each. You can have the 50 employees do overtime; each hour of overtime costs £300 and its marginal product is 50 shovels. Or you can employ more workers; each additional worker hired costs £30 per day and marginal product is 5 shovels per day.

Which of the following should your recommendation be so as to maximise profit?

A. Work overtime and hire more workers.

B. Work overtime but do not hire more workers.

C. Hire more workers but do not work overtime.

D. Make no change in working hours or size of labour force.

7 In the woollen industry, labour is a variable factor of production and land and capital are fixed factors of production. Improved relations between our country and Russia have led to a promise of large annual woollen garment sales to Russia for the foreseeable future, commencing immediately.

In the short run

I. more factor inputs will be employed in the woollen industry.

II. profits in the woollen industry will increase.

III. land now devoted to other agricultural uses will be switched to the woollen industry.

Which of the following is correct?

A. I only.

B. II only.

C. I and II only.

D. I, II and III.

8 'In the USA the labour market works like any highly competitive market; the demand for and supply of labour are like textbook demand and supply curves; when unemployment appears, wages fall and the economy tends towards full employment continuously.'

Assume the above situation exists. Which of the following must happen in the US labour market if a minimum wage (i.e. a wage below which no firm could legally hire workers) were introduced that was significantly higher than the lowest wage rate for unskilled workers?

A. Total wages payments to all workers would rise.

B. Total wages payments to unskilled workers would rise.

C. Some people would lose their jobs.

D. The supply curve of labour would shift to the right.

9 In order to make the most efficient use of a country's telephone system, how should charges for telephone calls vary?

A. They should be relatively low during business hours to reduce costs for the large number of people making calls.

B. They should be relatively high during business hours to ensure that people who value the service most highly can make calls.

C. They should be the same at all times to avoid making calls at inconvenient times to save money.

D. They should be relatively low during business hours because that is when the poor make calls to businesses, employment agencies, doctors etc.

10 Suppose all Brazilian coffee growers agreed to restrict the supply of coffee and thereby force up the price. The coffee growers as a result will experience an increase in income only if

I. the demand for Brazilian coffee is price-inelastic.

II. world income is rising.

III. there are no economies of scale in growing coffee beans.

Which of the following is correct?

A. I only.

B. III only.

C. II and III only.

D. I, II and III.

11 If the government were to increase sales tax on new cars from 15 per cent to 25 per cent, the effect on the second-hand car market would be

I. the average price of second-hand cars would increase.

II. the quantity of second-hand cars sold would decrease.

Which of the following is correct?

A. I only.

B. II only.

C. I and II.

D. Neither I nor II.

12 An imperfectly competitive firm discovers that at its present level of output its average total cost curve is at its minimum point and equal to £6.00; average revenue equals £5.00, marginal revenue equals zero. What should the firm do to maximise profit?

A. Decrease price and increase output.

B. Decrease price and decrease output.

C. Increase price and decrease output.

D. Nothing; profit is already being maximised.

13 In capitalist economies, market forces play a significant role in determining wages. In the absence of market imperfections, it is argued that wages would equal the value of the marginal product of labour. In communist economies, planners sometimes pay attention to market forces in setting wages, but wages are often determined by central committees.

Which of the following is inconsistent with marginal productivity theory?

A. The imposition of a minimum wage rate without guaranteed employment for all job seekers.

B. From each according to his ability, to each according to his needs.

C. If productivity gains are 5 per cent and workers are paid an additional 5 per cent, nothing is left for the state.

D. Ice hockey players, politicians and ballet dancers are often paid high salaries in command economies.

14 Which of the following would make the UK balance of payments deficit larger than it otherwise would be?

A. Purchase of a UK government bond by a foreigner.

B. An additional UK government grant to a foreign country.

C. Dividend payments to UK stockholders by foreign companies.

D. A decrease in imported goods to the UK from EU countries.

15 Suppose that in India and the UK the costs of producing wool and cars measured in man/days per unit are as given in the table.

Country	Wool	Cars
India	3	10
UK	2	5

Which of the following is correct?
A. The UK has comparative advantage in both wool and cars.
B. The UK has both absolute and comparative advantage in wool.
C. India has comparative advantage in cars.
D. India has comparative advantage in wool.

16 Which of the following transactions would make the demand for pounds (£) higher than it otherwise would have been in the foreign exchange market?
A. A Japanese distillery buys £100 of malt whisky from a Scottish distillery.
B. A UK company pays £100 in dividends to a shareholder living in France.
C. A UK meat processor buys £100 of tin from a company in Bolivia.
D. A Norwegian boy receives £100 for Christmas from his English aunt.

17 A country has a surplus in its balance of payments and a deficit in its balance of trade. Which of the following is correct?
A. The currency of the country must be depreciating.
B. Net capital inflows must exceed the export–import gap.
C. The country must be lending more abroad than the excess of imports over exports.
D. Gross national product (GNP) of the country must be higher than it would have been if exports and imports had been balanced.

18 The 'consumption function' of an imaginary economy with no government sector is represented by the following equation:

$$\begin{aligned} C = 50 + 0.8Y \text{ where } C &= \text{consumption expenditure} \\ \text{and } Y &= \text{national income} \end{aligned}$$

Which of the following is correct?
A. The amount saved is 170 when Y is 600.
B. C is zero if Y is zero.
C. C and Y can never be equal.
D. The average propensity to save increases as Y increases.

19 The following factors all affect how large the multiplier effect will be:
I. Marginal propensity to consume (MPC).
II. Rate of unemployment (U).
III. Rate of change in the money supply (M).
In the absence of offsetting factors, the largest multiplier effects will come from which of the following circumstances?
A. Low MPC, high U and a decrease in M.
B. High MPC, zero U and an increase in M.
C. Low MPC, zero U and a decrease in M.
D. High MPC, high U and an increase in M.

20 The following would enter the national income accounts under the heading 'investment':

I. Building a new house by a family for their own occupation.
II. Purchase of a second-hand car by a family for their own use.
III. Buying shares in BP.

Which of the following is correct?

A. I only.
B. III only.
C. I, II and III.
D. Not I nor II nor III.

21 On a small island with no international trade, there are only two multi-product companies, Y plc and Z plc. Their yearly sales, raw materials costs and profits are shown below. All the island's production takes place in these two companies. Y plc buys all its raw materials from Z plc, and Z plc buys all its raw materials from Y plc.

	Y plc	Z plc
Sales	200	400
Raw materials	100	60
Profits	4	140

Which of the following is the island's gross national product?

A. 440.
B. 544.
C. 744.
D. 904.

22 A 1585 painting by Annibale Carracci depicting the Holy Family was valued at £400 by one auction house last year but sold by another auctioneer on Tuesday night for £847 000. The painting was bought in 1930 for £5 by the father of the woman who put it up for sale.

By how much will the sale of the painting increase this year's GNP?

A. The auctioneer's commission for selling the painting.
B. £847 000 − £400 − £5.
C. £847 000 plus the auctioneer's commission.
D. £847 000.

23 Gross National Product is a measure of an economy's total output of final goods and services, which

I. excludes depreciation on capital as part of total output.
II. uses prices as weights in adding together quantities of different goods and services.
III. includes unemployment compensation for the unemployed and supplementary benefits payments to the poor.

Which of the following is correct?

A. II only.
B. I and II only.
C. I, II and III.
D. Not I, nor II, nor III.

24 The data below refer to an economy in Year t and Year $t + 10$:

Year	Money GNP	Price index
t	100	80
$t + 10$	250	200

Comparing Year $t + 10$ with Year t, which of the following is true? It follows that real gross national income

A. fell by 25 per cent.
B. was constant.
C. increased by 25 per cent.
D. increased by 50 per cent.

25 If an economy is at full employment, real gross national product (GNP) must always increase when

I. potential (full-employment) GNP increases.
II. nominal (money) GNP increases.
III. the unemployment rate increases.

Which of the following is correct?

A. I only.
B. I and II only.
C. II and III only.
D. Not I, nor II, nor III.

26 The foreign trade multiplier shows the effect on gross national product (GNP) of an increase/decrease in imports. The foreign trade multiplier will be

I. larger, the lower the marginal propensity to import.
II. larger, the higher the marginal propensity to save.
III. larger, the lower the marginal propensity to consume.

Which of the following is correct?

A. I only.
B. I and II only.
C. I and III only.
D. I, II and III.

27 In its budget message, the government announced a 2 per cent tax cut and 10 per cent money supply increase, which it declared would lead to full employment without inflation. A group of research economists, reviewing the budget proposals in detail, declared the government had overestimated how much the average consumer would spend out of additional income (marginal propensity to consume, MPC) and had also underestimated how much the average consumer would spend on imports out of additional income (marginal propensity to import, MPI). If the economists are right, which of the following is correct?

To achieve the stated goals, the 2 per cent tax cut and 10 per cent money supply increase is

A. too expansionary: it will cause excess demand and inflation.
B. too expansionary if, and only if, MPC is greater than MPI.
C. not expansionary enough if, and only if, MPC is less than MPI.
D. not expansionary enough: it will lead to an employment/output gap.

28 Despite the fact that an economy was at full employment, its government decided to electrify its total railway system. It financed this project by selling government bonds to the general public. Aggregate demand increased in the economy and also the rate of inflation. After the project was complete, aggregate demand returned to its original level, but the rate of inflation remained higher than its original level, the reasons being

I. potential GNP increased.

II. the initial inflation changed expectations about future inflation.

III. the increase in the money supply caused by the government selling bonds to finance the electrification of the railways.

Which of the following is correct?

A. II only.

B. III only.

C. I and II only.

D. I and III only.

29 If the government were to adopt policies that encouraged households to save a higher proportion of their incomes, which of the following would be the effect on national income in the short run, assuming no other offsetting factors affecting the economy?

A. National income would be lower than it otherwise would have been because aggregate demand would be lower than it otherwise would have been.

B. National income would be higher than it otherwise would have been because higher savings means higher investment expenditure.

C. National income would be higher than it otherwise would have been because savings functions show that higher savings are associated with higher levels of national income.

D. National income would not change because the increase in savings (decrease in consumption expenditure) would be matched by an equal increase in investment expenditure.

30 The following data refer to a hypothetical economy:

Year	Potential GNP (full employment)	Employment index	Price index
t	100	100	100
$t+1$	103	98	102

Comparing Year $t+1$ with Year t, which of the following is correct?

A. GNP per capita increased.

B. There was inflation.

C. Real GNP decreased.

D. Money GNP increased or was constant.

Case Study

During the past few years, the price of beef in supermarkets and butchers' shops in the United Kingdom has fluctuated dramatically and is now substantially higher than it was three years ago. At the same time, for reputed health reasons, the per capita consumption of beef has declined – but not enough to prevent total beef consumption from increasing. Also, despite the much higher price of beef, cattlemen claim to have experienced a substantial decline in profit.

1 Assume the above statements to be correct. Compare the market today with the market situation three years ago and identify those factors you believe to be responsible for the rise in the price of beef. You may use diagrams in your answer.

Essay Questions

Essay 1

The price system works wells in normal times but breaks down under the stress of war or other major upheavals. Discuss.

Essay 2

Despite the sophisticated models produced by economists, governments seem incapable of running economies in a satisfactory fashion. In the past decade most European countries have experienced at one time high inflation, high unemployment, wildly fluctuating exchange rates, and imbalanced government budgets. Why are governments unable to solve such problems?

Examination Answers

Practice Final Examination 1

Multiple Choice Questions

1 The correct answer is A. The effect of selling 5 per cent more cars at the current price is to increase total revenue by 5 per cent. It is therefore possible to increase the payments to all factors of production, including labour, by 5 per cent.

2 The correct answer is B. The government decision to buy every school pupil a calculator would shift the demand curve to the right. The reduction in the costs of producing calculators, i.e. a lowering of the average total cost curve, would shift each firm's short-run marginal cost curve to the right and consequently shift the industry short-run supply curve to the right. Whether the new equilibrium price would be higher, the same or lower than the original price will depend upon the relative shifts of the demand and supply curves. Thus the joint effect on price cannot be determined from the information given. However, both curves will intersect to the right of the original equilibrium position.

3 The correct answer is C. No country today has sufficient resources to satisfy all the wants of every member of its population; this is the 'fundamental fact of scarcity'. The fact that society's basic needs, however defined, might be being met or that commodities not produced in the UK can be imported from abroad does not imply that scarcity no longer exists. Even if every family in the UK were above poverty level, however defined, scarcity would still remain.

4 The correct answer is C. Since no prices are provided, it is not possible to compare options A and B since A uses less labour and more capital than B. However, it is possible to compare B and C since they use the same quantity of capital, but C uses more labour; B is therefore more efficient than C, which can be dismissed on the grounds of technical efficiency. B and C is the only pair that it is possible to compare in this way.

5 The correct answer is A. Any statement that suggests diminishing additional utility derived from additional consumption is consistent with the 'law' of diminishing marginal utility. Diminishing marginal utility is consistent with a limit to the amount of ice cream that can be eaten in a day, with boredom watching TV after a time, and with dullness from spending too much time working. However, to say that the second glass of beer tastes just as good as the first suggests that the utility is the same for the second as for the first glass of beer, and is not consistent with diminishing marginal utility.

6 The correct answer is B. A demand curve shows the quantity of a good that a consumer would buy at different prices, *cet. par.* This is independent of the conditions of supply; therefore statement I is incorrect. As the price increases, the quantity typically decreases for a number of reasons, one of which is that other goods become relatively more attractive; this is the substitution effect. Thus statement II is correct.

7 The correct answer is C. In deciding which policy contributes most to economic efficiency, it is necessary to compare marginal costs and marginal benefits of each. Since no cost information is provided, it is not possible to draw a conclusion.

8 The correct answer is B. The Brazilian economist's advice was probably based on his belief that the demand for coffee is price-inelastic, i.e. that the bumper crop could only be sold as the result of a price reduction, which would lead to a decrease in total revenue. However, no such definite information on the market is provided. The only definite outcome of burning the coffee would be to reduce the total of consumption goods available to the world, i.e. to decrease world real income.

9 The correct answer is D. The output of the first 6 man-days, as shown by point X, was 300 kilos; the addition of a further 6 man-days, as shown by point Y, added just over 100 kilos, thus the marginal product of labour decreased between points X and Y; statement I is thus incorrect. The movement from X to Z involves doubling both the input of labour and of capital; if there were increasing returns to scale, output would more than double. However, the output of mackerel also doubled from 300 to 600 kilos, exhibiting *constant* returns to scale. Statement II is incorrect.

10 The correct answer is B. The effect of increased advertising is to increase the cost of production and alter the distribution of expenditure between brands, while leaving total expenditure unchanged. Thus, if all refiners increased their advertising, there would be a redistribution of consumer expenditure between brands, depending on how much additional advertising each refiner carried out, and industry total profit would decrease. When the increased advertising costs were passed on in price increases, the general public would become worse off. Option A is therefore incorrect and B is correct. Since the amount of brand switching in response to changed advertising is unknown, it is not possible to predict the effect on an individual refiner of a change in his advertising.

11 The correct answer is C. Since more people now use the bus (because taxi fares have risen) and the fares are unchanged, bus revenues will rise. The effect of increased taxi fares on total revenues depends on the price elasticity of demand. Since this is unknown, taxi revenues may rise, fall or be unaltered.

12 The correct answer is B. The social cost of producing iron is greater than the private cost of production by an amount equal to the decrease in the farmer's output; consequently, because of this externality, setting output at the level where marginal revenue equals private marginal cost is not economically efficient. The solution is to include the decrease in the farmer's output as a cost to the firm, i.e. to charge it according to how much output the farmer loses as a result of iron production. The firm's private marginal cost will now be equal to marginal social cost and the iron foundry will set its output at a lower profit-maximising level. Consequently B is correct. None of the other responses contributes to equalising private and social cost.

13 The correct answer is A. The impact of the university's decision will be a shift to the right of the demand curve for rented accommodation. The stock of rented accommodation is fixed in the short run; therefore this will result in an increase in the market rent. In the long run, potential landlords will be attracted into the private renting sector because of the higher rents obtainable, and thus there will later be an increase in the quantity of private rental housing supplied.

14 The correct answer is B. For any given consumer, the greater the proportion of income spent on alcohol and tobacco, the greater will be the amount of tax paid and, since the smaller will be the amount of income available for all other purchases, including gas and electricity, the less will be the amount of the subsidy received. Thus, if people with lower incomes spend a greater proportion of their incomes on alcohol and tobacco than do people with higher incomes, but not a higher proportion on gas and electricity, then the scheme will increase income inequality. Thus statement II is correct. Real income is determined by the nation's resources and you might argue that distortions caused by the tax subsidy programme would cause real income to fall. However, this fall would not be because the cost of living would be higher; this cost of living, e.g. the consumer price index or GNP deflator, could be rising, constant or falling. Thus statement I is incorrect.

15 The correct answer is A. The use to which the businesses and the university wish to put the satellite does not exclude any other use; therefore the opportunity cost is zero. Furthermore, there are no negative externalities; thus the marginal social cost is zero. A necessary condition for economic efficiency is that price equals marginal social cost, which in this case is zero.

16 The correct answer is A. The value of final or gross output in economy was \$500 worth of ploughs and \$4000 worth of bread. The wheat and flour are intermediate goods and to include them in the value of gross output would be double counting. What the feudal lord did with his share of gross output does not affect the value of gross output.

17 The correct answer is B. The fact that money GNP has risen faster than real GNP indicates rising prices. This could be associated with an increasing money supply, increasing velocity of circulation of money, or both. There is no proof that the money supply increased by £7 million. Statement I is therefore incorrect. Output per worker is real GNP divided by the number of workers; since real GNP increased between Year t and Year $t + 2$ and the number of workers did not, it follows that output per worker was higher in Year $t + 2$ than in Year t. Statement II is therefore correct.

18 The correct answer is C. Total expenditure on final goods and services includes expenditure on goods sold from inventories; however, investment expenditure is calculated net of additions to or deductions from inventories. Consequently the net calculation of total expenditure is equal to the value of final output.

19 The correct answer is B. In order to balance the budget, the government has to predict national income and, on the basis of the prediction, calculate total tax revenue from the selected tax rate. This total tax revenue must be equal to the selected government expenditure plus transfers; transfers depend partly on the unemployment rate, and hence on the

actual value of national income. However, disposable income, and hence national income itself, depends on transfers; therefore it is extremely complicated to calculate in advance the appropriate values for tax rates and government expenditure. The calculation is made even more complex if government expenditure is to be 20 per cent of national income; furthermore, consumption expenditure must be 60 per cent of national income, again adding to the complexity of the calculation. To ensure that investment expenditure is 20 per cent of national income, it is again necessary to make a prediction of national income and the rate of interest that will have the desired effect. Given that the government can only make one set of policy decisions for a given year, it is extremely unlikely that the three objectives will be attained. Thus option B is correct. Government expenditure of 20 per cent of GNP is not incompatible with a balanced budget, which in turn is not incompatible with full employment. Thus options A and C are incorrect. Since monetary and fiscal policy can influence the distribution of GNP among C, I and G, option D is also incorrect.

20 The correct answer is B. The calculated MPC of two-thirds suggests an expenditure multiplier of 3, but the actual multiplier was only 2. One possible explanation for the multiplier turning out to be less than expected is an increase in leakages – for example an increase in the marginal propensity to import. Statement I suggests a reduction in the marginal propensity to import and is therefore incorrect. Full employment or potential output sets a limit to actual output; thus if the gap between actual and full employment output were less than $30 000 million, the multiplier would not have been able to work through fully; therefore statement II is a possible explanation. The increased government expenditure caused an increase in aggregate demand and an increase in output, thus causing an increase in the demand for money. Since the supply of money was not increased, it is possible that the price of money, i.e. the interest rate, increased, causing private investment and consumption expenditures to fall, thus reducing the net effect of the multiplier; statement III is therefore also a possible explanation.

21 The correct answer is B. The decline in exports would lead to a reduction in aggregate demand, causing the unemployment rate to be higher than would otherwise be the case. Increasing the interest rate would cause a reduction in investment and consumption expenditures, which would reduce aggregate demand further and add even more to the unemployment rate. A reduction in UK exports would cause a decrease in the demand for sterling, the value of which would fall; an increase in the UK interest rate would make the UK relatively more attractive for savers, and then the demand for sterling would increase, causing the value of sterling to rise. Thus from the policy adopted, it can be inferred that the government was more concerned about the falling value of sterling than the rising unemployment rate.

22 The correct answer is B. While it is possible to increase aggregate demand by increasing the supply of money, potential output sets an upper limit to real GNP. Therefore it does not always follow that the greater the supply of money, the better off people will be.

23 The correct answer is B. An increase in the money supply will make interest rates lower than they otherwise would have been and investment and consumption expenditures higher than they otherwise would have been. If the economy is already at full employment, i.e. if actual output equals potential output, the excess demand will cause inflation and consequently a

reduction in the value of the currency. However, if there are unemployed resources, real output will increase without necessarily causing an increase in the price level.

24 The correct answer is D. An increase in interest rates would reduce aggregate demand, thus increasing unemployment; therefore statement I is wrong. The reduction in consumption arising from an increase in income tax would also reduce aggregate demand and increase unemployment; given there are substantial unemployed resources already, it is not necessary to reduce consumption in order to make resources available for investment; therefore statement II is wrong. A decrease in government expenditure would also reduce aggregate demand and increase unemployment; again, since there are unemployed resources, there is nothing to be gained from freeing additional resources for use in the private sector; therefore statement III is incorrect.

25 The correct answer is B. The economy develops a gap between actual and potential output during the year, and so some resources must have become unemployed. Consequently there will be an increase in the unemployment rate. There may be increases or decreases in the inflation rate, interest rate and balance of payments deficit, but a change in any is not guaranteed by a gap between actual and potential GNP in a given year.

26 The correct answer is A. In March the holidaymaker would be able to buy £500 × 1.50 = \$750. In September she would be able to buy £500 × 1.40 = \$700. If there were no inflation in the US, clearly the holidaymaker would be able to buy more in March with \$750 than in September with only \$700. The location at which the dollars were bought has no bearing on the value of the dollars. Since the holidaymaker went to the USA to spend money, not to invest it, the increase in interest rates in the USA would not affect her.

27 The correct answer is D. The upper limit to the potential GNP of any country is determined by the quantity and quality of the labour force, the capital stock and the level of technology. Thus if Germany were able to expand its labour force through immigrant labour from Turkey, Germany's potential GNP would increase. As the migrants were employed, actual GNP would also increase. Thus statements I and III are correct. Conversely, Turkey would experience a fall in its potential GNP when its labour force declined as workers moved to Germany. Thus statement II is correct.

28 The correct answer is B. $Y \equiv C + I + G + X - Z$, where Y = National Income, C = Consumption Expenditure, I = Investment Expenditure, X = Export Expenditure and Z = Import Expenditure. Imports are leakages from the circular flow of income because expenditure on imports result in income being generated in countries abroad, not in the domestic economy. As a result, the increase in imports will make national income lower than it otherwise would have been.

29 The correct answer is B. The increase in government expenditure has crowded out an equal value of investment expenditure; therefore the net effect on real output is zero. The budget surplus or deficit is found as follows: budget surplus or deficit = total tax revenue − government expenditure − transfers. Since tax rates and national income are unchanged, total tax revenue is also unchanged. The increase in government expenditure therefore leads to either a budget deficit or a reduction in the budget surplus.

30 The correct answer is A. $Y \equiv C + I + G + X - Z$

Substituting the values given, we have

$$
\begin{aligned}
Y &= 20 + 0.5Y + 10 + 10 + 10 - 10 \\
\therefore Y &= 0.5Y + 40 \\
\therefore 0.5Y &= 40 \\
\therefore Y &= 80
\end{aligned}
$$

Case Study

1 This is a case ideally suited to marginal analysis.

The boat will maximise profit by increasing speed up to the point at which the marginal cost of increasing speed is just equal to the marginal revenue from the increased fish catch. Thus it is necessary to calculate travel time for different speeds which, subtracted from 24 hours, will yield fishing time. Each hour of fishing time is constant and worth \$10. But each additional hour of fishing time has a cost, namely the increase in fuel costs.

The table below shows the relationship between speed, travel time, fishing time and fuel consumption.

Speed (miles /hr)	Travel time (hrs)	Fishing time (hrs)	Fuel consumption (gallons)
6	10.00	14.00	10
7	8.57	15.43	11
8	7.50	16.50	13
9	6.67	17.33	16
10	6.00	18.00	20
11	5.45	18.55	25
12	5.00	19.00	31
13	4.62	19.38	38

One laborious method of calculating the optimal speed is to estimate the total cost of fuel for each speed and the value of the total catch for each hour of fishing time. By subtracting the difference the profit-maximising speed becomes evident.

Marginal analysis is much more efficient. Choose any speed, say 10 miles per hour, and compare the consequences with a speed of 11 miles per hour. The additional fuel cost is 5 gallons × \$1 = \$5. The increase in fishing time is 0.55 hours (18.55 − 18.00). The additional 0.55 hours produces a catch worth \$5.50 (\$10 × 0.55). Therefore it is worthwhile increasing speed from 10 to 11 miles per hour.

Increasing the speed from 11 to 12 miles per hour adds costs of \$6 (\$31 – \$25). The increase in fishing time of 0.45 hours yields an increase in revenue of only \$4.50. Since the marginal cost (\$6) exceeds the marginal benefit (\$4.50), the increase in speed from 11 to 12 miles per hour will decrease profit.

Consequently, the profit-maximising speed is 11 miles per hour.

Essay Questions

Essay 1

Your discussion should describe in detail how the market system operates, the meaning of economic efficiency and the assumptions underlying the behaviour and motivation of households and resource owners. This requires analysis of utility-maximising consumers, profit-maximising firms, operation of markets and the meaning of equilibrium in both the short and long run. The crucial role of prices should be highlighted as a major determinant of household well-being, a signalling device to resource owners in the allocation of factor inputs, and as the link in the marginal equivalency conditions. The role of government, a body of citizens of the market economy representing the interests of the citizens, in establishing the 'rules of the game' in the market economy should be discussed.

Market failures involving public goods, externalities, economies of scale, and income distribution should be analysed. The analysis should include why the price mechanism will not result in economic efficiency in the presence of these factors and how government 'interference', while not guaranteeing economic efficiency, can bring the economy closer to an efficient allocation of resources.

Government action that produces prices and output that would not prevail through market forces should be compared with, and evaluated against, the market solution. Such an analysis could include price ceilings, price floors, subsidies, taxes and transfers. The role of government in using monetary and fiscal policy in the pursuit of desirable macroeconomic goals should be included in your discussion. You should also point out that the motivation of governments as vote seekers may prompt actions that benefit a vocal minority at the expense of the silent majority.

Essay 2

Your analysis should include the determinants of growth in potential GNP, i.e. the growth in the quality and quantity of the capital stock and growth in the quality and quantity of the labour force. It should also show how technological change is embodied in investment goods and training and education of the labour force, and how growth in potential GNP is independent of changes in actual GNP.

The determinants of actual GNP should be discussed, i.e. consumption, investment and government expenditure, exports and imports. Your analysis should include the circular flow of income and expenditure and should show how GNP changes as each of the components changes. The instability of some of the components of GNP can be shown to be a cause of fluctuations in GNP. The fluctuations in GNP compared with the relative stability of potential GNP produces output and inflationary gaps. Induced expenditure should be distinguished from autonomous (exogenous) expenditure, a source of economic instability.

The policy tools available to the government, consisting of fiscal policy (changes in government expenditure and taxation) and monetary policy, and their efficacy in different economic situations should be analysed. In this context you should also point out the disagreements that exist between the different schools of thought as to the appropriateness of the different policy instruments and their potential impact on output and employment. You should point out that desirable macroeconomic goals such as a low inflation rate, a balanced budget and stable exchange rates may conflict with the goal of a low unemployment rate and consequently an employment gap may be the penalty to be borne for achieving other goals. The Phillips Curve, for example, suggests a trade-off between inflation and unemployment; you should evaluate the arguments about the Phillips Curve and distinguish between short-term and long-term macroeconomic goals.

Practice Final Examination 2

Multiple Choice Questions

1 The correct answer is B. To become better off, a society should undertake any activity for which the marginal benefit exceeds the marginal cost. The government is not arguing that additional social scientists provide no benefits, nor that there is an excess supply of social scientists, but rather that the benefit to society from additional engineers exceeds the benefit to society from additional social scientists.

2 The correct answer is B. TC = FC + TVC

therefore $\text{ATC} = \frac{\text{FC}}{\text{Q}} + \frac{\text{TVC}}{\text{Q}}$

and not $\text{FC} + \frac{\Delta\text{TVC}}{\Delta\text{Q}}$

3 The correct answer is D. Consider the following simple example:

Year 1

Number of workers	=	2
Output per worker	=	£200
therefore total output	=	£400
Wage per worker	=	£50
therefore total wage bill	=	£100
therefore return on all other assets	=	£300

Year 2

Number of workers	=	2	
Output per worker	=	£208	(£200 × 1.04)
therefore total output	=	£416	
Wage per worker	=	£52	(£50 × 1.04)
therefore total wage bill	=	£104	
therefore return to all other assets	=	£312	

Because output per worker increases by 4 per cent, total output increases by 4 per cent and the return to all other factors taken together increases by 4 per cent. The return to each factor input can increase, therefore, by 4 per cent.

4 The correct answer is A. If the product of an additional unit of output adds more to revenue than it does to cost, it will pay a firm to produce that unit, i.e. it will produce additional units as long as additional revenue exceeds marginal costs. As a consequence, the profit-maximising (loss-minimising) level of output of any firm occurs where marginal revenue equals marginal cost. Thus statement II is correct.

The marginal cost curve intersects the average cost curve at its minimum point. Thus, producing at minimum average cost would maximise profit *if, and only if*, the marginal revenue curve intersected the marginal cost curve at this level of output. Thus statement III is incorrect.

If average revenue exceeds average total cost, the greatest difference between them will occur when average total cost is at a minimum, which will not be the profit-maximising output level because marginal cost will not equal marginal revenue. Thus statement I is incorrect.

5 The correct answer is C. Since the selling price of each turkey is £10, the marginal revenue, i.e. the revenue gained from one additional turkey, is £10. To maximise profit, marginal revenue must equal marginal cost. Thus marginal cost must be calculated at the different levels of output. The average cost of producing 90 turkeys is £7; therefore, the total cost of producing 90 turkeys is £630. The total cost of producing 100 turkeys is £730. Therefore, the cost of producing the additional 10 turkeys is £100 or £10 per turkey. At this output level, total profit equals £270 (£1000 − £730). The firm would achieve the same profit producing 90 turkeys, since for the additional 10 turkeys marginal revenue equals marginal cost, but an output of 90 is not an option. The same solution can be reached by calculating total revenue and total cost for each level of output.

6 The correct answer is D. Since each hour of overtime costs £300 and the additional output is 50 shovels, the marginal cost of an 'overtime' shovel is £6. This exceeds the selling price of £5. Since each additional worker costs £30 and the additional output is 5 shovels per day the marginal costs of an 'additional worker' shovel is also £6, again exceeding the selling price of £5. Thus either course of action will reduce profits. The firm should hire no additional workers and work no overtime.

7 The correct answer is C. In the short run, an increase in demand will lead to a higher equilibrium price. The short-run supply curve of a firm is its marginal cost curve, and firms will expand output by hiring additional variable factors of production, incurring additional marginal cost, and will maximise profit by producing that level of output where marginal revenue equals marginal cost. Because the average cost curve does not change in the short run, the firm will experience increased profit as it expands output and hires additional variable inputs. Thus statements I and II are correct. However, since the question specifies the short run, no additional fixed factors, such as land, will be involved. Thus statement III is incorrect.

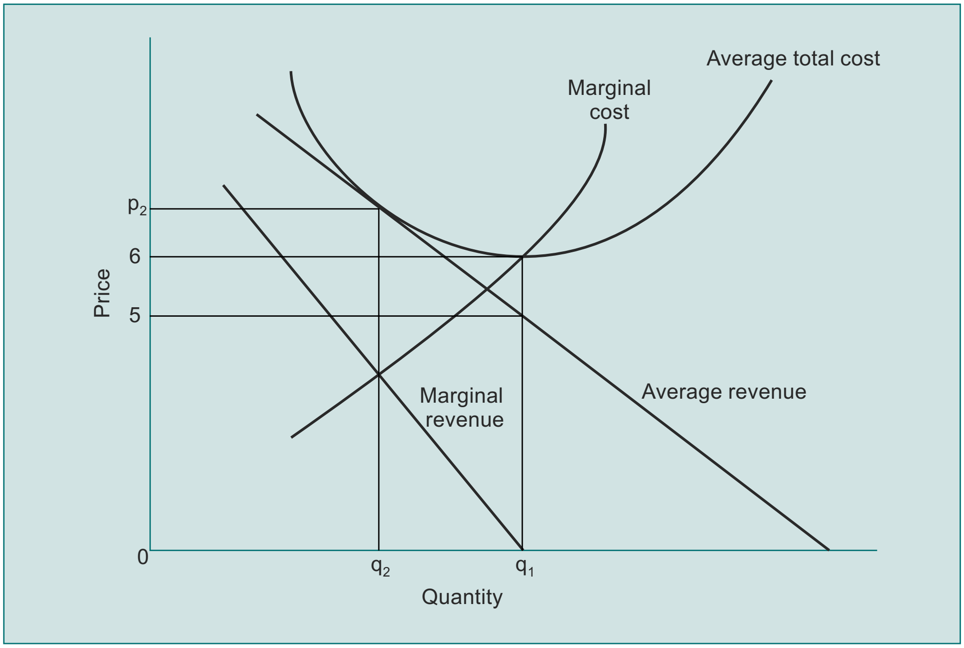
8 The correct answer is C. Given traditional demand and supply curves in an efficient market, the imposition of a price floor – in this case through a minimum-wage law – will result in excess supply of labour, in this case unemployment. There are two causes of this unemployment: a larger number of workers offering labour services at the (higher) minimum wage rate (i.e. a shift along the supply curve), and reduced demand for labour (i.e. a shift along the demand curve). Thus some people would lose their jobs.

9 The correct answer is B. In order for resources to be used efficiently, it is necessary that the prices of goods and services produced be equal to their marginal costs. The social cost of telephone calls is higher during business hours when telephone circuits are being used at or near their capacity and each user prevents other users from making calls. Thus prices should be set during peak hours to ensure no excess demand. This will ensure that the people who make calls at that time value them most.

10 The correct answer is A. An increase in the price of a good is associated with a decrease in the quantity bought. But if the demand for Brazilian coffee is price-inelastic, the decrease in quantity bought will be proportionately smaller than the increase in price, leading to an increase in total expenditure and revenue. Thus the coffee growers will be better off, and statement I is correct. Rising world income does not guarantee increased revenue to Brazilian coffee growers at higher prices; thus II is incorrect. Economies of scale affect production costs, not revenue; thus III is incorrect.

11 The correct answer is A. The increase in the sales tax on new cars would increase the demand for used cars by people who do not own any car but wish to acquire one. Simultaneously, the increased sales tax on new cars would encourage car owners to keep their current cars rather than buy new ones, thereby decreasing the supply of second-hand cars. The increase in demand coupled with the decrease in supply will cause the price of second-hand cars to rise. The quantity bought and sold, however, could increase, remain constant or decrease.

12 The correct answer is C. The diagram below exhibits the data: at output level $0q_1$ average total cost is at a minimum and equals £6, average revenue equals £5, and marginal revenue equals zero. The upward-sloping marginal cost curve must pass through the minimum point of the average total cost curve and therefore must intersect the marginal revenue curve at an output level less than $0q_1$. Since profit maximisation requires that marginal revenue equals marginal cost, the firm should reduce output to $0q_2$, and charge a higher price p_2.



13 The correct answer is B. In competitive markets it will always pay an employer to hire workers until the wage rate equals the value of the marginal product of labour. If a worker is paid less than the value of his/her marginal product, he/she will move to another firm. Thus the worker's contribution to output will determine his/her wage and not his/her wants/needs. Thus 'from each according to his ability, to each according to his needs' is inconsistent with marginal productivity. A, C and D are all consistent with marginal productivity.

14 The correct answer is B. An increase in imported goods or services and/or an increase in capital outflows will make a balance of payments deficit larger than it otherwise would have been. The only item that falls into this category is the UK grant to a foreign country.

15 The correct answer is D. Since the UK can produce both wool and cars with fewer man-days of input, the UK has an absolute advantage in the production of both. In India a car will exchange for three-and-a-third units of wool. In the UK a car will exchange for two-and-a-half units of wool. The UK will export cars to India because of its comparative advantage in cars; India will export wool to the UK because of its comparative advantage in wool.

16 The correct answer is A. When a UK company wishes to buy foreign goods or assets, it must exchange £s for the currency of the foreign country. Such an action would increase the demand for the foreign country's currency and increase the supply of £s available for exchange. Conversely, when a foreign country wishes to buy UK goods or assets, the demand for £s will be higher than it otherwise would have been.

17 The correct answer is B. The balance of payments is the sum of the trade balance plus net capital flows; thus, if a country had a trade deficit and net capital flows were zero, the country would have a balance of payments deficit. To have a balance of payments surplus in the presence of a trade deficit, net capital inflows must exceed the trade deficit.

18 The correct answer is D. $C = 50 + 0.8Y$; when $Y = 600$, $C = 50 + 0.8(600) = 530$.

Therefore $S = Y - C = 70$, and therefore option A is incorrect.

When $Y = 0$, $C = 50$, therefore option B is incorrect.

When $Y = 250$, $C = 50 + 0.8(250) = 250$, therefore option C is incorrect.

Since $C = 50 + 0.8Y$, $S = -50 + 0.2Y$ and so the average propensity to save is

$$\frac{S}{Y} = \frac{-50+0.2Y}{Y}$$

As Y increases, $\frac{S}{Y}$ increases.

19 The correct answer is D. The simple multiplier formula is:

$$\text{multiplier} = \frac{1}{1-\text{MPC}}$$

Thus the higher the value of MPC, the larger the multiplier. For the multiplier to work fully, the three following conditions must be met:

(a) sufficient unemployed resources exist to allow the multiplier process to operate;
(b) the additional demand generated by increased expenditure is spent on domestically produced goods and services; and
(c) the rate of interest does not increase; an increase in the interest rate would reduce investment and consumption expenditure, with negative multiplier effects partly or fully cancelling out the positive expenditure multiplier.

20 The correct answer is A. Additions to the capital stock are classified under the heading 'investment'. Houses are part of a nation's capital stock and thus a new house would be classified as an investment in the year in which it was built. If it were subsequently resold, it would not enter the national income accounts in the year of resale since it would already be part of the capital stock. Thus statement I is correct. By the same logic, the purchase of a second-hand car would not enter the national income accounts. Thus II is incorrect. Also by the same logic, buying a share in an existing company is merely transferring ownership of part of that company to another person or institution. Thus III is incorrect.

21 The correct answer is A. If a miller sells £100 worth of flour to a baker who uses it to make bread which sells for £200, including both the flour (£100) and the bread (£200) in the national income accounts would involve double counting of the flour. Thus the costs of raw materials in the question have to be subtracted from total sales, i.e. 200 + 400 – 100 – 60 = 440. Insufficient information is given to calculate total income to all factor inputs since only profit figures are given.

22 The correct answer is A. A good (service) enters the national income accounts in the year in which it is produced (performed). In the case of the sale of any antique or second-hand good, the only element of the sale that enters the national income accounts is the service(s) of the agent(s) facilitating the exchange between the owner and the new owner – in our question the auctioneer's services, the value of which is the commission.

23 The correct answer is A. Gross national product = $\sum (ap_a + bp_b + \dots + zp_z)$, i.e. the sum of the quantities of each good produced multiplied by its price. Thus II is correct. Gross national product − depreciation = net national product, i.e. depreciation is part of GNP. Thus I is incorrect. Unemployment compensation and benefit payments to the poor are transfer payments and not returns for factor inputs gainfully employed. Thus III is incorrect.

24 The correct answer is B. Gross national product = $\sum (ap_a + bp_b + \dots + zp_z)$.

There are three ways money (nominal) GNP can increase:

(a) The quantities of the goods and services produced can increase, prices remaining constant.
(b) The prices of the goods and services produced can increase, quantities remaining constant.
(c) Both the prices of the goods and services produced and the quantities can increase.

If (b) occurs, real GNP will remain constant. In the question, money GNP increased by the same proportion as the price index, and therefore there was no increase in real output.

25 The correct answer is D. Potential output increases when, in the absence of offsetting factors, labour inputs increase and/or the capital stock increases and/or technological change occurs. Such changes can occur independent of what is happening to real GNP; thus statement I is incorrect. Nominal GNP will increase when the average price level increases, real GNP constant; thus statement II is incorrect. The unemployment rate is determined by the gap between potential and actual GNP. Thus, if potential GNP were to increase and actual GNP were to remain constant, the unemployment rate would increase; thus statement III is incorrect.

26 The correct answer is A. The size of the multiplier is determined by leakages from the circular flow of income – the smaller the leakages, the larger the multiplier. Saving and expenditures on imported goods are both leakages from the circular flow.

27 The correct answer is D. To achieve full employment with the 2 per cent tax cut and the 10 per cent money supply increase, the government had to make assumptions about many variables in the economy, including both the marginal propensity to consume (MPC) and the marginal propensity to import (MPI). The higher the value of MPC, the larger the multiplier effect; and the higher the value of MPI, the lower the multiplier effect. Since the government overestimated the former and underestimated the latter, the fiscal and monetary policies enacted would not be sufficient stimulation to achieve the desired increase in GNP. Thus unemployment would result.

28 The correct answer is A. Two causes of inflation are excess demand and expectations about higher future inflation. The electrification of the railways while the economy was at full employment would cause inflation due to excess demand. However, when aggregate demand returned to its original level, the excess-demand cause of inflation would disappear; indeed, potential output would be expected to increase, resulting in an output-employment gap. Thus, for the inflation rate to remain higher than its original rate, expectations about future inflation must have increased. Selling bonds to finance the electrification of the railways would not increase the money supply.

29 The correct answer is A. If households were to save a higher proportion of their incomes, by definition they would consume a smaller proportion of their incomes. Thus, in the absence of offsetting factors, total consumption expenditure and aggregate demand would decrease. As a result, national income would be lower than it otherwise would have been.

30 The correct answer is B. Since no information is given about actual output (GNP), in either real or nominal terms, or the population, we do not know what happened to real GNP, nominal GNP or GNP per capita. Thus, A, C and D are not necessarily correct. The price index increased from 100 to 102, indicating an inflation rate of 2 per cent.

Case Study

1 The simplest way to tackle this case is to consider one representative firm producing cattle, make assumptions about that firm and the market in the two time periods (the base year (t) and today, three years later, ($t + 3$) in accordance with the facts given, and carry out an analysis. Assume perfect competition and assume the firm is in equilibrium in year t making normal profit.

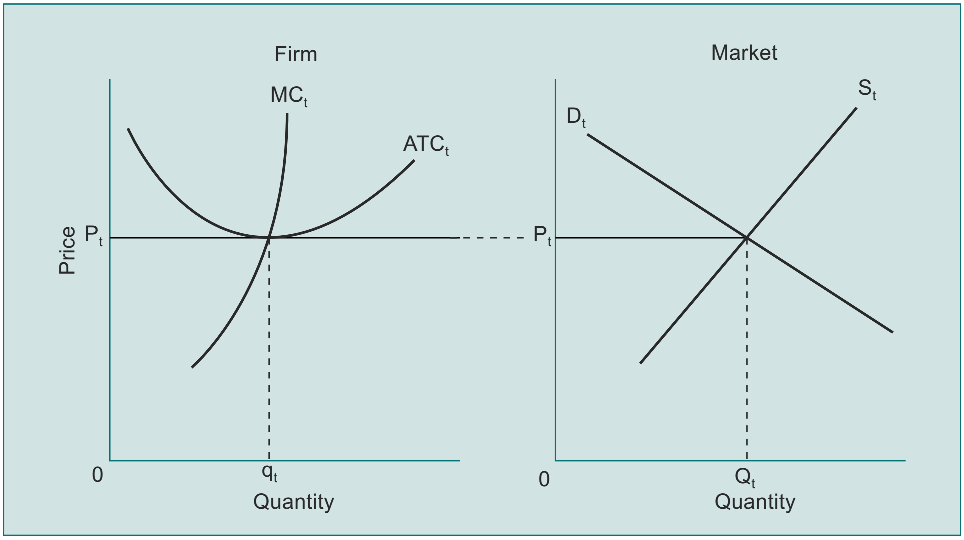


Figure A1.1 Firm and market in Year *t*

Figure A1.1 represents the situation. The market supply curve is the summation of all firms' MC curves. Demand and supply intersect at the price P_t, equilibrium output being $0Q_t$. The profit-maximising firm produces $0q_t$ and earns normal profits.

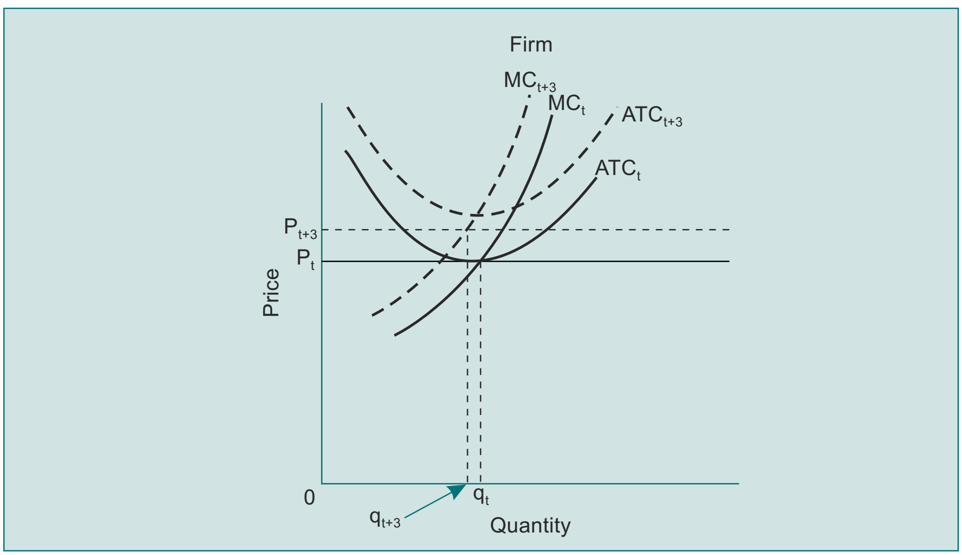


Figure A1.2 Firm in Year $t + 3$ compared with Year *t*

Figure A1.2 compares the firm in Year $t + 3$ with Year t. We are told that $P_{t+3} > P_t$ and yet there has been a substantial decline in profit. Costs, therefore, must have risen. This is shown by the upward shift of ATC and leftward shift of MC. The decline in profit means that average revenue (at price P_{t+3}) is less than minimum ATC, as shown; the firm now produces $0q_{t+3}$, which must be less than $0q_t$.

However, we are also told that despite reduced per capita consumption of beef, total consumption is higher in Year $t + 3$ than in Year t. This means there are more beef consumers in the market: the demand curve has shifted to the right. But, since each firm is producing less, as shown above, there must also have been an increase in the number of firms in the market. Figure A1.3 captures these shifts.

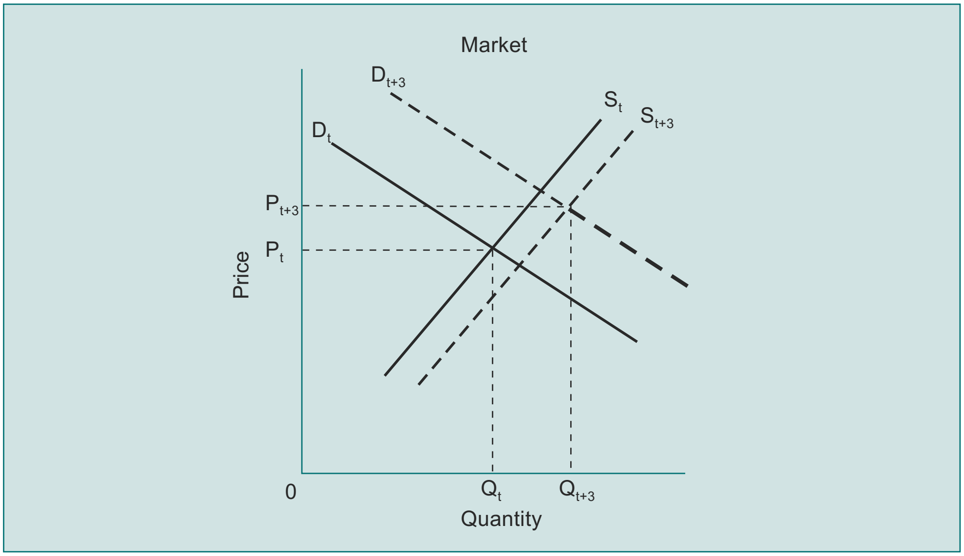


Figure A1.3 Shifts in demand and supply

Alternative Approaches

1. Equilibrium need not be assumed in Year t, but this makes the analysis more complex. It is possible to envisage 'rapid' technological change in the raising of cattle within the three-year span. This could have the effect of increasing the optimal size of the firm, shifting the supply curve to the right – but increasing factor input costs could result in losses for the firm.
2. It is also possible to analyse a market in disequilibrium via the 'cobweb' model. In Year t, demand exceeds supply, resulting in rising prices and profits. Cattlemen respond by raising more cattle but, three years later, even though both price and sales are higher than in Year t, there is excess supply, resulting in falling prices and profits.

Essay Questions

Essay 1

A really good answer must address the following issues.

(a) How does the price system work in normal times?
(b) What is meant by the expression 'work well'? Work well for whom?
(c) What conditions are necessary for the price system to operate efficiently?

(d) Given the required conditions, are there areas of economic activity in which the price system does not work well?
(e) Do wars or major upheavals impose additional constraints upon the price system so as to make it inoperative?

Part (a) should include an analysis of demand, supply and markets; an analysis of utility-maximising consumers, profit-maximising firms, and price and output determination in competitive markets, in order to show how resources are allocated in a market economy.

Part (b) should analyse the marginal equivalency conditions, showing the critical link of prices in governing consumer and firm behaviour. A concomitant analysis of factor input markets should indicate how income distribution is determined by ownership of resources and marginal productivity theory.

Part (c) should include consumer and firm rationality, information flows, established markets, technological advances, and changing tastes and preferences, and should show how such prerequisites may be threatened in war/upheaval situations.

Part (d) should analyse areas of economic activity leading to 'market failure', e.g. public goods, externalities, economies of scale, and income distribution.

Part (e) should follow up on the discussion in Part (c), and while accepting the required increase in the provision of public goods and the probability that the price mechanism in the short run may not elicit the required resources for national defence (e.g. armed forces), it should make the case that, in many instances, a war situation does not violate normal market criteria – i.e. those who can, and wish to, pay for desired goods and services will find willing suppliers. Government action, in the form of rationing, expresses a concern for income (goods and services) distribution, but economic forces make themselves manifest in terms of 'black markets' for those goods and services that the government attempts to control.

Essay 2

The following points, expanded to show a thorough understanding of the relevant macroeconomic concepts and relationships, would appear in a really good answer.

(a) Potential output (Q), a function of the quality and quantity of the labour force and capital stock, grows over time through investment in human and non-human capital and because of technological change.
(b) Actual output (Y) seldom matches potential output in the real world. When $Y < Q$, the result is unemployment. When $Y > Q$, inflationary forces are set in motion. Only when $Y = Q$ is the economy at full employment, but Q is constantly growing – a moving target.
(c) $Y \equiv C + I + G + X - Z$
With the exception of G, each of the components of Y is subject to forces outside the control of fiscal and monetary authorities. In addition, because of interdependencies, policies designed to affect one variable have secondary effects on other variables.

(d) There are differences of opinion within, as well as between, political parties on desirable macroeconomic goals and how to achieve them.
(e) Economists and politicians disagree, from a theoretical standpoint, as to which tools are most appropriate for dealing with different states of the economy.
(f) Political expediency often wins the day over long-run, macroeconomic policy making.
(g) There are time lags between governments recognising where the economy is and where they would like it to be, legislating appropriate policies and having those policies take effect.
(h) The economy is subject to exogenous shocks over which policy makers have no control but which they must nevertheless take into account when making policy decisions.
(i) It is not often possible for desirable goals to be met simultaneously. In most circumstances, trade-offs have to be made, e.g. inflation, unemployment and balancing the budget.
(j) In attempting to balance the budget, i.e. have government revenue match government outlays, estimates have to be made of national income (income tax receipts), business income (corporation tax receipts) consumer expenditure (value added tax receipts). On the outlay side, while government expenditure can be controlled directly and some transfer payments estimated fairly accurately, transfer payments to the unemployed depend not only on the rate of payment but also the number unemployed.
(k) In the international arena, factors outside the government's control include foreign national incomes, foreign inflation rates, foreign interest rates, all of which affect exports and capital flows, the balance of trade, balance of payments and ultimately the exchange rate which, with a lag, affects exports, imports, national income and employment.

Appendix 2

Answers to Review Questions

Contents

Module 1

Review Questions

Multiple Choice Questions

1.1 The correct answer is B. A resource is scarce whenever having more of it would make someone better off. The fact that one individual's wants for a resource, e.g. a television set or an automobile, can be fully satisfied does not imply that all individuals' wants are fully satisfied. A resource that cannot be increased in quantity may be scarce; so may resources that can be. Resources, by definition, are important in satisfying the wants of society, e.g. air, water and sunshine, but their ability to satisfy wants does not imply scarcity. Whenever there are *unsatisfied* wants, resources are scarce.

1.2 The correct answer is C. It is true that the US is one of the richest countries in the world and also true that the US imports many items from other countries. Neither of these facts, however, implies that the fundamental fact of scarcity no longer applies to the US. Scarcity would not exist in the US only if all wants of all US citizens were fully satisfied. Even if no family in the US were below the poverty level, scarcity would still exist.

1.3 The correct answer is D. Although it appears in each of the examples that a good, namely a tyre, stamps, an air-conditioned atmosphere, is available at zero cost to a potential consumer, each of the goods in question has utilised society's scarce resource and consequently is not a free good.

1.4 The correct answer is B. The fact that many people buy and eat hamburgers at the same time implies that these people enjoy hamburgers, not that they have identical tastes and preferences. Even if a nation's resources were fully and efficiently utilised, all wants would not be fully satisfied even if all people had similar tastes and preferences. In addition, there is no guarantee that any individual's wants would be satiated. What differing tastes and preferences imply is that free exchange of goods having a low utility for goods with a high utility can increase the level of well-being of those individuals involved in such an exchange without an actual increase in the amount of goods available.

1.5 The correct answer is B. Since Method 3 requires the same amount of resource B and less of A than Method 2, Method 3 is more technically efficient than Method 2. No conclusion can be made about the relative engineering efficiency of Method 1 compared to the other two methods because it uses less of A but more of B. Units of A and B are not additive; A could be very expensive machine tools and B could be man-days, or vice versa.

1.6 The correct answer is C. In an economically efficient society, producing more of any one good by definition means producing less of some other good. In comparing two time periods, assuming economic efficiency in both, a higher output in one time period means an increase in resources and/or superior production techniques using reduced, the same or more resources.

1.7 The correct answer is D. To maximise society's welfare, the government should allocate expenditures so that the marginal benefit from the last dollar spent on each good equals the marginal cost. If this condition did not hold, the government could increase welfare by reallocating its expenditure. Because no information is given about the costs of constructing highways or airports, it is impossible to tell whether carrying out either project or both projects would achieve an efficient use of resources.

1.8 The correct answer is B. If the government were to build the new highway system, it would have to raise the money to pay for it and resources would be required to build it. But neither of these costs is the opportunity cost of the highway system. The opportunity cost of any good is the best alternative forgone to obtain that good. Whether the opportunity cost of the new highway system is greater than the economy can afford depends on whether it will yield lower benefits than an improved public transportation system throughout the country.

1.9 The correct answer is C. An economic activity will increase society's welfare if the marginal benefits exceed the marginal costs. Each of the first additional three terminals meets this criterion. The fourth terminal should not be built because marginal costs exceed marginal benefits after three terminals. The fact that the total benefits from four terminals exceed total costs is irrelevant.

1.10 The correct answer is C. For every nation in the world today, resources are insufficient to satisfy all human wants. Thus every nation faces the problem of scarcity. If human wants will always outstrip the ability to satisfy them, scarcity will always be a problem. All labour being employed does not imply scarcity is non-existent, neither does some labour being unemployed.

1.11 The correct answer is C. Given the government's goal is efficient allocation of resources, it should increase expenditure on medical services only if the marginal benefits exceed the marginal costs. The marginal benefits are those attributable to the extra medical services; the marginal costs are the goods forgone to produce those additional medical services.

1.12 The correct answer is A. Statement I is a correct definition of the opportunity cost of a unit of land, i.e. the best alternative forgone. Statement II is false because it does not stipulate the best alternative forgone. If, to use an extreme example, the resources used to produce the good in question could have been used to produce goods no one wanted, e.g. wheelbarrows with square wheels having zero value, the opportunity cost of the good in question would be zero.

1.13 The correct answer is D. When production is being carried out at point L, i.e. $0W_1$ of wine and $0C_2$ of cheese is produced, the opportunity cost of cheese is the best alternative forgone. Since there is only one alternative to cheese in the example, i.e. wine, the opportunity cost of cheese therefore at point L is the wine forgone when producing $0C_2$ of cheese, i.e. W_1W_3 of wine. Thus statement I is incorrect. The information given in Figure 1.3 concerns production possibilities; no information is given on how society values wine and cheese. Therefore statement II is not necessarily true.

1.14 The correct answer is B. Any output combination on or within the production possibilities curve or frontier is attainable, e.g. points 0, W_1, W_2, C_1, C_2, K, L, M. With existing resources and technology, all points outside the frontier, e.g. N, are unattainable.

1.15 The correct answer is A. A society's resources are capable of producing an infinite number of alternative sets of goods and services. The best set to produce is that which maximises society's welfare, i.e. satisfies wants as fully as possible. Producing various goods at random will not guarantee the best use of resources; ensuring free goods are available to all may require many resources (e.g. taking everyone to Loch Lomond to watch a sunset), and may not be what all the people want. There is no operational distinction in economics between necessities and luxuries.

1.16 The correct answer is B. Free exchange of goods permits higher levels of satisfaction (utility) without an increase in the amount of goods. Both individuals gain by each exchanging a good with a low marginal utility for one with a high marginal utility. Marginal utilities are determined by tastes and preferences and the quantity of a good already consumed (or in this case, in one's possession to be consumed). Both individuals could have had identical amounts of apples and oranges in their baskets, but different preferences, and both could have benefited equally. Also, both could have preferred oranges to apples but one was willing to give up, say, one orange for two apples, because he had no apples in his basket and plenty of oranges.

1.17 The correct answer is B. When an economy is operating in an economically efficient manner, engineering efficiency prevails in the production of goods and services, and each is being produced using the least amount of resources. Further, the set of goods and services being produced is that which makes society as well off as possible. However, any one individual's utility could be increased by giving him/her more goods and services – although they would have to be taken from someone else, making that person worse off.

1.18 The correct answer is A. No matter how a society is organised for economic decision making, each must decide what set of goods and services to produce, how to produce them and how to allocate them among the members of that society. Not all land in command economies is owned privately, and income is not shared equally in non-market economies. Thus statements II and III are incorrect.

Case Study 1.1: Marginal Analysis and Noise Pollution

1

1. The total benefit of each option is calculated by multiplying the reduction in dB(A) by $31 000. The net benefit is calculated by subtracting total cost from total benefit:

Wall 2m high

Benefit	= 10 × $31 000 =	$310 000
Cost		$160 000
Net benefit		$150 000

Wall 3m high

Benefit	= 12 × $31 000 =	$372 000
Cost		$190 000
Net benefit		$182 000

Wall 4m high

Benefit	= 13 × $31 000 =	$403 000
Cost		$230 000
Net benefit		$173 000

Speed limit

Benefit	= 3 × $31 000 =	$93 000
Cost		$90 000
Net Benefit		$3 000

Resurface road

Benefit	= 2 × \$31 000 =	\$62 000
Cost		\$70 000
Net benefit		−\$8 000

Rebuild and sink road

Benefit	= 10 × \$31 000 =	\$310 000
Cost		\$2 500 000
Net benefit		−\$2 190 000

2. The costs of resurfacing the road and rebuilding the road are greater than the respective benefits and therefore should not be undertaken.
3. The net benefit is positive for enforcing the speed limit. The table below sets out the marginal cost and benefit of each height of wall:

Height (metres)	Total cost	Marginal cost	Total benefit	Marginal benefit
2	160 000	160 000	310 000	310 000
3	190 000	30 000	372 000	62 000
4	230 000	40 000	403 000	31 000

   The marginal benefit of the third metre of wall is \$62 000 compared with the marginal cost of \$30 000. The marginal cost of the fourth metre of wall is \$40 000 compared with the marginal benefit of only \$31 000.
4. An economic activity should be undertaken if marginal benefit exceeds marginal cost. Since the different heights of wall are mutually exclusive, building additional metres should continue as long as the condition holds. Thus a three-metre wall should be built and the speed limit enforced.
5. The mistake made by the Council was to consider total costs and benefits instead of marginal costs and benefits. While it was correct that the four-metre wall gave the highest total benefit, it did not see that the addition of the fourth metre of wall incurred more cost than benefit.
6. The total noise reduction is 3dB(A) from the speed limit and 12dB(A) from the three-metre wall, which reduces the sound level from 85dB(A) to 70dB(A). Since the residents indicated that they would be willing to pay for a reduction of noise down to 68dB(A), they could still be made better off by further noise reduction. However, since each of the remaining available options costs more than \$31 000 for each dB(A) of noise reduction, it would not be economically efficient to adopt any of them.

Module 3

Review Questions

Multiple Choice Questions

3.1 The correct answer is B. How fully a good satisfies an individual's wants will be determined by the total quantity of that good which the individual consumes. The demand curve that embodies the concept of demand yields information on how much an individual would be prepared to pay for an additional unit of the good *cet. par.*, i.e. it does not yield information on total satisfaction obtained from the good.

3.2 The correct answer is A. How fully a family's wants for food are satisfied will depend on the quantities of food available, which in turn will depend on the amount of income (budget) allocated to food purchases and on the prices of foodstuffs. Since, in the question, the budget is determined (or given), the only remaining factor that will affect the quantity of foodstuffs that can be purchased is the price of foodstuffs. Had the question been concerned with the satisfaction of all wants, II and III as well as I would have been relevant factors.

3.3 The correct answer is B. The airport bus and the taxi compete for passengers in terms of speed, privacy, convenience and price. These factors, together with the income of passengers, help determine which form of transport is chosen. The taxi fare, not the purchase price of the taxi itself, will affect the decision.

3.4 The correct answer is B. All societies have laws designed to regulate behaviour for the benefit of society as a whole or for individuals within society, e.g. compulsory education and drug control. People often disregard the law and suffer penalties if caught. The existence of laws prohibiting the purchase of goods does not imply that there is no demand for such goods; indeed, the underworld flourishes on illegal markets. Demand theory assumes that individuals know their own interests best and that if a good is bought – legally or illegally – it is because the marginal benefit to the individual exceeds the marginal cost. Economics is not concerned with making moral judgements on what is good or bad for a rational person.

3.5 The correct answer is D. Because total receipts increase when the price is lowered from \$10 to \$2, it means that ticket sales increase proportionally more than the price decrease, e.g. $100 \times \$2 > 8 \times \$10$. Thus, in price range \$2–\$10, the demand for tickets by old-age pensioners is price-elastic. However, we have no corresponding data for spectators of other ages and consequently can make no elasticity comparison. Nevertheless, the increase in gate receipts means that, as a group, old people will have less to spend on other goods and services; the increase in gate receipts also means that the loss in revenue from original older fans (now paying \$2 instead of \$10) has been more than compensated for by new fans each paying \$2.

3.6 The correct answer is A. The more an individual consumes of a good, the higher will be total utility and the lower will be marginal utility of that good.

3.7 The correct answer is C. This is a purely arithmetical question. If a utility-maximising individual in a two-good world is in equilibrium, and if the price of one good is double that of a second good, the marginal utility of the first good must also be double that of the second good. (Rework the example with the price of a glass of beer set at $1 and make sure you understand why option D would be correct.)

3.8 The correct answer is C. After the fall in the price of meat, Family F said it was better off despite having bought fewer potatoes, therefore Family F must have bought more meat. However, as more of a good is consumed in a given period, its marginal utility decreases; therefore for Family F the marginal utility of meat must have fallen. The only way in which Family G could not have benefited from the decrease in the price of meat is if Family G purchased no meat before or after the fall in price.

3.9 The correct answer is C. Income and substitution effects explain why, in economic theory, demand curves are negatively inclined to the price axis. If the price of a good bought by a consumer decreases, *cet. par.*, the consumer will experience a rise in real income. This enables the consumer to increase purchases of all goods. Further, since the decrease in the price of the one good changes relative prices, a utility-maximising consumer will alter purchases by substituting the cheaper good for the now relatively more expensive goods until equilibrium is restored.

3.10 The correct answer is C. A shift in a demand curve is caused by a change in any of the factors determining the position of the curve – income, prices of other goods, etc. A movement *along* a demand curve is caused by a change in price.

3.11 The correct answer is B. The three factors mentioned in I, II and III are likely factors determining the position of the demand curve for cigarettes. To the extent that some people pay attention to the BMA report and reduce or stop their smoking, DD will shift to the left. Similarly, as people switch from cigarettes to cigars and pipe tobacco as their prices decrease, DD will shift to the left. (For practical purposes we can ignore the extreme example of a person spending almost all his income on tobacco products, who experiences such a large increase in real income when the price of cigars and pipe tobacco fall that he is able to smoke more cigars, pipe tobacco and also cigarettes, i.e. income effect outweighs the substitution effect.) Given cigarettes are a normal good, a higher income would cause DD to shift to the right.

3.12 The correct answer is C. A rise in the price of a good will lead to a smaller quantity being purchased, an increase in the purchase of substitute goods and a decrease in the purchase of complementary goods.

3.13 The correct answer is A. As underground train fares increase, some people will switch to substitute goods – in this case buses. Thus total bus revenues will increase. Total underground fare revenue will increase if the percentage fare increase exceeds the percentage underground passenger loss, that is, if the demand for underground services in the relevant price range is inelastic. If the demand is price-elastic or of unitary elasticity, underground fare revenue will not increase as underground fares increase. (It is stated that buses and the underground are substitutes. Were they complements, for example if passengers used buses to reach underground stations, then, assuming the demand for the underground is not perfectly inelastic, bus revenue would fall as underground fares increased, because travelling on both would be reduced. Again, underground revenue would rise or fall, depending on the price elasticity of demand.)

3.14 The correct answer is A. This is a complicated application of the concept of demand. Since the membership demand is inelastic, raising the price of membership will result in a relatively small reduction in the number of members and an increase in Community Centre revenue. By raising the disco price, total revenue will fall since demand is elastic. Although we do not know the exact figures involved, this policy is on the right lines since the fall in revenue from the disco, which will no longer be overcrowded if a high enough price is charged, will be compensated for by the increased revenues from membership fees.

Case Study 3.1: The Price of Potatoes

1

1. Last year the revenue from sales was 2 million × €2 = €4 million. Both teams agree that if the price were set at €2 per bag again this year, the same number of bags would be sold. From Figure 3.12 it can be seen that Group A concluded that at a price of €1.50 a total of three million bags would be sold, an increase of one million bags over last year. Total revenue from sales would then be 3 million × €1.50 = €4.5 million, an increase of €500 000.
2. Group B concluded that 1.5 million bags could be sold at €3 per bag, as can be seen from the diagram. While this would leave 1.5 million bags unsold, the total revenue from sales would be 1.5 million × €3 = €4.5 million, again an increase of €500 000.
3. Curve A is a price-elastic demand curve in the range €1.50–€2, i.e. when price is reduced, total revenue from sales increases, and vice versa. Curve B is a price-inelastic demand curve in the range €2–€3, indicating that when price is reduced, total revenue from sales falls, and vice versa.
4. If Group A is wrong and the demand curve is Curve B, the effect of reducing the price would be to increase the number of potatoes sold to 2.2 million. Total revenue would then be 2.2 million × €1.50 = €3.3 million, a reduction of €700 000 on the revenue at €2 per bag.
5. If Group B is wrong and the demand curve is Curve A, the effect of increasing the price to €3 would be to sell no potatoes.

Advanced Note

You may wonder why the two teams picked the prices that they did once they had estimated the demand curves. On a linear demand curve there is one point at which price × quantity is a maximum; this is the point where marginal revenue (i.e. the increase in total revenue from selling one more bag) is zero. The demand curves are shown in Figure A2.1 together with the marginal revenue curves.

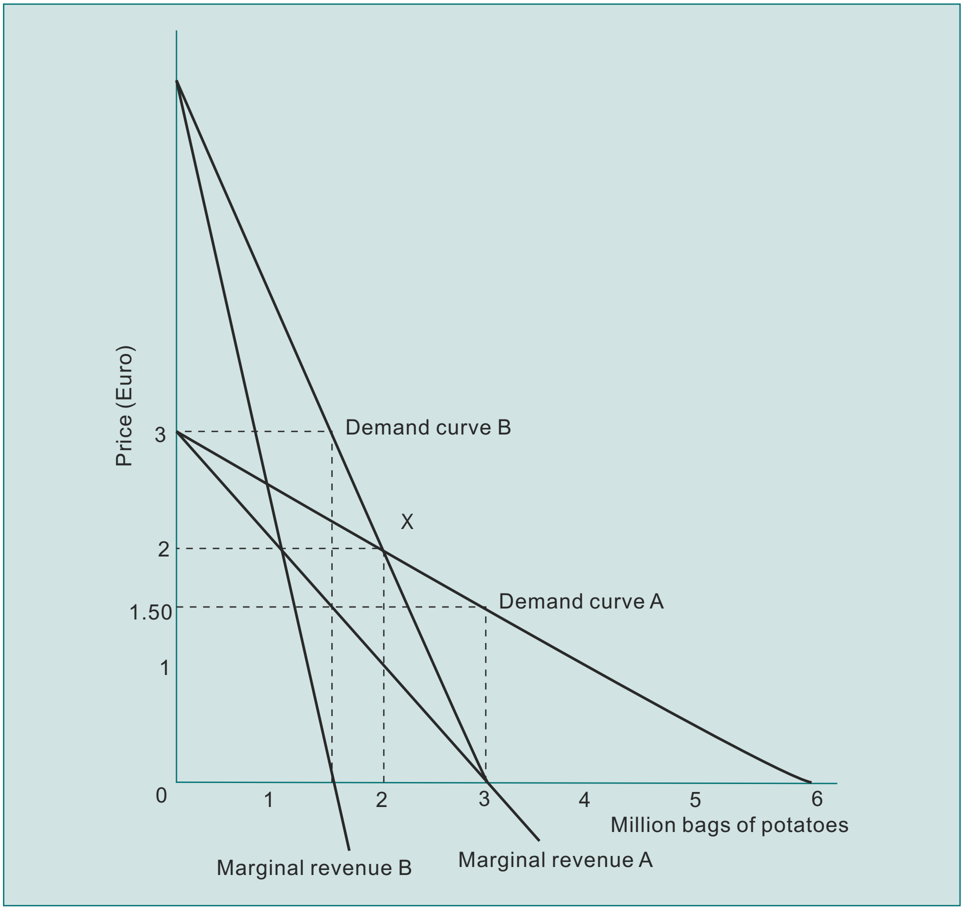


Figure A2.1 Demand and marginal revenue curves

For demand curve B, marginal revenue is zero at 1.5 million bags; if the price were greater than €3 per bag, then total revenue would be increased by reducing the price since above that price marginal revenue is positive; below €3 per bag, marginal revenue is negative and hence a reduction in price would reduce total revenue. For demand curve A, marginal revenue is zero at 3 million bags at €1.50 per bag; at prices above €1.50, total revenue could be increased by reducing the price, and at prices below €1.50, total revenue could be increased by increasing the price. Thus not only had the two groups estimated demand curves but they had also recommended the price that gave the highest revenue given the slope of that demand curve.

Case Study 3.2: Cash versus Vouchers

1

1. The underlying assumption in this analysis is that each household, knowing its own preferences, given a limited income and the prices of all goods and services, spends that income in the way that will yield maximum satisfaction; in other words, the household attempts to maximise utility given a particular budget constraint. The additional constraint faced by the recipients of vouchers is that the vouchers can be used only for certain specified goods.

There are two possibilities in analysing the likely effect of vouchers on a household's expenditure:

(a) The value of the food and clothing vouchers is less than the amount that would have been spent on food and clothing had the household been given cash. In this case the issue of vouchers would be the same as the payment of cash, since households would spend the same amount on food and clothing whether cash or vouchers was given. This is because after using the vouchers the household would be able to spend the money saved on food and clothing for anything it wanted (including more food and clothing).

(b) The value of the food and clothing vouchers is greater than the amount that would have been spent on food and clothing had the household been given cash. In this case, since the vouchers can only be used for food and clothing, the household would buy more food and clothing and fewer other goods than if it had been given cash.

In the first case there would be no difference between the cash and the voucher schemes, but in the second case people would spend more on food and clothing under the voucher scheme.

2. The evidence from the experiment suggests that the value of the food and clothing vouchers was greater than the amount that would have been spent on food and clothing had the households been given cash. Households that received vouchers spent €20 more on food, €10 more on clothing, and €30 less on all other goods, compared with households that received cash payments.
3. If the group had been split randomly into two samples, the two samples would be expected to spend their income similarly. Those who were given vouchers would have been able to increase total utility had they been given cash, because they would have spent less on food and clothing and more on other things. Hence they would have been no worse off – and might have been better off, given their preferences – with cash.
4. Since households would be no worse off under the cash scheme, and may even be better off because they would be able to spend their total income in a way that maximised utility, cash payments are likely to be a better option than vouchers; in other words, the cash scheme is more likely to attain the objective of economic efficiency.

   But many people do not believe that economic efficiency in terms of the maximisation of households' own preferences is the most important consideration. Some people believe that the government *does* know better than poor households what is best for them. Others are simply not interested in maximising the satisfaction of the poor but want to see the poor well fed and clothed, i.e. they wish to maximise their own satisfaction as 'givers'.

   These issues involve value judgements and are therefore matters of ethics rather than economics. Economics does not have much to say about ethics, but it is important to separate the ethical arguments from the economic arguments. The ethical issue is to identify whose satisfaction is to be satisfied; the economic issue is, for given preferences, how resources can be allocated efficiently. In making the actual decision about cash versus vouchers, the government must make its decision on both these matters.
5. Any attempt to help the poor involves diverting income from the better-off people in society. The loss in utility to the better-off must be balanced against the increase in utility to the poor. Economic analysis cannot conclude which distribution of income is best for society as a whole. Many people have ideas about what minimum income should be provided for the poor, but this tends to change as general living standards change. The

economic problem is to identify which form of aid is more efficient. Economic analysis cannot decide the issue of how much aid the poor should be given, but it can help to identify the consequences of helping the poor in different ways. Consequently, it can help us decide what course of action should be taken.

Module 4

Review Questions

Multiple Choice Questions

4.1 The correct answer is B. Engineering efficiency by definition prevails in the production of a good when that good is being produced using the minimum amount of resources. This does not imply that society is producing all goods in such a fashion and producing that set of goods that maximises society's utility, i.e. the conditions for *economic* efficiency. Some resources may be unemployed; and if that is true, increased output of other goods could result without decreasing the output of the good being produced in an engineering-efficient manner.

4.2 The correct answer is D. To produce the maximum amount of goods in some time period, two conditions must prevail:

(a) each good must be produced in an engineering-efficient manner, i.e. using the least amount of resources; and
(b) all resources must be employed.

Thus if the economy is not producing its maximum output, then either condition I or II or both I and II in the question prevail.

4.3 The correct answer is D. Necessary conditions for economic efficiency are:

(a) engineering efficiency – all goods and services must be produced using the least amount of resources; and
(b) all resources must be fully employed.
A necessary *and sufficient* condition for economic efficiency is
(c) the production of that set of goods and services that satisfies society's wants as fully as possible.

Since (c) subsumes conditions (a) and (b), conditions (a) and (b) above are not sufficient to guarantee condition (c).

4.4 The correct answer is D. A firm's typical supply curve is upward sloping, i.e. more is supplied the higher the price. Below some price the firm will not be prepared to supply any quantity, but this does not imply that the traditional supply curve does not exist. A firm might refrain from supplying a good in one time period if it expected the price of the good to rise, but it would not if it expected the price to fall.

4.5 The correct answer is B. The two factors that could have caused greater sales of high-priced cosmetics in Germany are higher German incomes and/or different relative prices, e.g. luxury cosmetics being twice as expensive as 'regular' cosmetics in Germany and three times as expensive in Britain. What could *not* have caused the difference in the volume of sales is the length of the time period being considered: we are not comparing, for example, annual sales in Germany with monthly sales in Britain; a common time period of a month exists.

4.6 The correct answer is D. This is a definitional question. Economists find it convenient for analytical reasons to distinguish the short run – a time period during which at least one factor of production cannot be increased or decreased – from the long run – a time period where no factor of production is fixed. For example, in deciding how large one should make a supermarket when no building has even been begun can be considered a long-run decision. Once the building is up, however, the variables that can be altered are such things as the number of checkout stands and employees – the variable factors of production.

4.7 The correct answer is D. The answer to this question uses the same rationale as in Question 4.6 above. In the long run, all the variables can be altered.

4.8 The correct answer is D. Given that at any given level of output price exceeds average variable costs, a firm will maximise profit or minimise loss by producing that level of output at which marginal revenue equals marginal cost. Maximum revenue would be obtained by a perfectly competitive firm if it were to sell an infinite quantity, and by an imperfectly competitive firm where marginal revenue was zero. Thus I in the question is incorrect. A perfectly competitive firm in equilibrium will produce at a level of output where average total cost is minimum; it is possible for this to occur under monopoly but it is highly unlikely. Thus, there is no guarantee that the output corresponding to minimum average total cost is the profit-maximising output. Thus II is incorrect.

4.9 The correct answer is A. All points on the production frontier and all points between the production frontier and the horizontal axis are attainable outputs: they can be produced with the given fixed factor input and different amounts of the variable factor input. However, only points on the frontier in Figure 4.26, e.g. points X, Y and Z are engineering-efficient outputs, i.e. the maximum outputs attainable with the corresponding amounts of the variable input. Points outside the frontier, e.g. point A, could not be achieved without more of the fixed factor, i.e. point A is not attainable in the short run. Figure 4.26 shows the outputs that *can* be produced, not what *should* be produced to maximise society's welfare. No information exists on the economically efficient level of output.

4.10 The correct answer is B. All points on the production frontier and all points between the production frontier and the horizontal axis are attainable outputs: they can be produced with the given fixed factor input and different amounts of the variable factor input. However, only points on the frontier in Figure 4.26, e.g. points X, Y and Z, are engineering-efficient outputs, i.e. the maximum outputs attainable with the corresponding amounts of the variable input. Points outside the frontier, e.g. point A, could not be achieved without more of the fixed factor, i.e. point A is not attainable in the short run – but may be attainable in the long run.

4.11 The correct answer is B. A definitional question. Total output divided by the number of units of a factor of production is the average product of that factor. The marginal product of a factor is the change in output divided by the change in the factor. The total cost of production of a good is the sum of fixed cost and total variable costs; variable cost of production is total cost minus fixed cost.

4.12 The correct answer is C. The fact that total output increases as more workers are hired results from the fact that the marginal product of labour is positive. The marginal product of labour can be greater than, equal to, or less than the average product of labour.

4.13 The correct answer is D. The long run is that time period during which all resources are variable. Since there are no fixed factors of production, there are no fixed costs; all costs are variable.

4.14 The correct answer is C. For perfectly competitive firms, price equals marginal revenue (MR). The rule for profit maximisation for all firms is to produce that level of output where MR = marginal cost (MC). For a perfectly competitive firm, this translates to price = MC; and when a perfectly competitive firm is profit-maximising, in the long run price = LMC (long-run marginal cost). But that implies it is also maximising the short-run marginal cost (SMC); if it were not, it would change output level and price would not equal long-run marginal cost.

4.15 The correct answer is D. In the long run, firms move into and out of an industry. If such movements do not affect factor input price, the long-run supply curve will be a horizontal line. The long-run supply curve is therefore unrelated to any of the cost curves of existing firms in the industry.

4.16 The correct answer is C. The position of an industry's long-run supply curve is affected by the cost of factor inputs. If the price of factor inputs increases, the curve will shift to the left, and vice versa.

4.17 The correct answer is D. Diminishing returns means that as a firm uses more of a variable factor input with a given amount of fixed factors, the marginal productivity of the variable factor decreases. Thus III is correct by definition. It follows that each additional unit of output will require ever increasing amounts of the variable factor input, and consequently each additional unit of output will be more expensive to produce.

4.18 The correct answer is A. The value of the marginal product of A is \$12, defined by the number of extra lunches times price. The value of the marginal product of factor B is \$6 (3 × \$2). The marginal cost of factor A is \$3 and of factor B is \$1. It will always pay a firm to hire more of a resource as long as MR > MC, since the hiring of more of each factor would contribute more to revenue than it would to cost. Thus as \$12 > \$3 (factor A) and \$6 > \$1 (factor B), more of both should be hired.

4.19 The correct answer is C. Suppose the factor input in question is labour, the variable factor input; then, with Q = output and L = units of labour, we have

$$AP_L = \frac{Q}{L} \qquad \text{Table A2.1}$$

and

$$MP_L = \frac{\Delta Q}{\Delta L} \qquad \text{Table A2.2}$$

Where Q = output and L = units of labour.

$$\text{Average Cost (AC)} = \frac{TC}{Q} = \frac{FC}{Q} + \frac{VC}{Q} = \frac{FC}{Q} + \frac{L \times W}{Q} \qquad \text{Table A2.3}$$

$$\text{Marginal Cost} = \frac{\Delta TC}{\Delta Q} = \frac{\Delta FC}{\Delta Q} + \frac{\Delta L}{\Delta Q} \cdot W \qquad \text{Table A2.4}$$

where W = wage rates and the other expressions refer to total, fixed and variable costs.

The expression Q/L appears in Equation 2.1 and L/Q appears in Equation 2.3. Thus AP_L and AC are inversely related. Thus II in the question is correct. $\Delta Q/\Delta L$ appears in Equation 2.2 and $\Delta L/\Delta Q$ appears in Equation 2.4. Thus MP_L and MC must be inversely related, and so I in the question is correct. $\Delta FC/\Delta Q = 0$, and thus fixed cost and MC must be unrelated. Therefore III in the question is incorrect.

4.20 The correct answer is D. Fixed costs do not vary with output levels and consequently the higher the level the lower must be average fixed costs, i.e. fixed cost per unit of output must decline as output increases.

4.21 The correct answer is B. Average labour productivity (AP_L) is calculated by dividing total output by the number of workers. The agreement with the union was that a wage rate increase would equal the increase in AP_L. Since AP_L increased by 6 per cent, the wage rate increase should also be 6 per cent, independent of changes in other factor inputs.

4.22 The correct answer is B. If the house is left unoccupied, the owner does not have to pay local utilities costing \$30 per month. Thus, the owner must reclaim at least \$30 per month in rent to make renting the house a viable economic activity.

4.23 The correct answer is B. A firm's marginal cost will shift to the right if the cost of a variable factor input decreases. Since the industry short-run supply curve is composed of the sum of firms' short-run marginal cost curves, the industry short-run supply curve will shift to the right if the price of a variable factor input decreases. Thus III is correct. It will also shift to the right as the number of firms in the industry increases, but that is a long-run situation. Thus I is incorrect.

Case Study 4.1: How to Produce Switches

1

1. Average cost is total cost divided by total output. Taking the larger machine as an example, as the number of workers hired increases, and total output increases, the fixed cost of the machine is distributed over a larger and larger number of workers, thus decreasing the average cost. However, once the machine has reached its capacity of eight workers, it is necessary to incur the cost of an additional machine before output can be increased further. Each time an additional machine is brought into operation, the average cost increases; and then it subsequently decreases until that machine reaches full capacity operation.
2. The average cost using the larger machine is €118 at 45 switches per week and €90 at 60 switches per week. The average cost using the smaller machine is much lower at €87 and €70 respectively. Since the average cost using the smaller machine is lower, three should be rented rather than one larger machine.
3. Between output levels of 135 and 150 switches per week, the situation is less clear-cut. The cost of producing exactly 135 switches per week is €72 per switch using seven small machines, compared with €81 using two large machines. However, increasing output from 140 to 145 units using the small machine requires the use of an additional machine, with a consequent increase in average cost. The average cost of producing exactly 150 switches is identical, at €73, using either eight small machines or two large machines. For any output in the range less than 150 switches, a number of the smaller machines should be rented.
4. At output levels of 165 switches per week and above, the large machine results in lower average cost. At such output levels there is no ambiguity about which machine to rent – the larger one.
5. From the data in Table 4.10, it can be calculated that the percentage of total cost accounted for by wages differs between the machines at full capacity output.

   Large machine: 800/5800, i.e. 14 per cent
   Small machine: 400/1400, i.e. 29 per cent

   Because of the larger percentage of total cost accountable to labour, any given increase in labour cost will affect total cost more for the smaller machine. For example, a 50 per cent increase in wage cost at full-capacity operation for each machine will increase total cost as follows:

   Large machine: 400/5800, i.e. +7 per cent
   Small machine: 200/1400, i.e. +14 per cent

   Since a given percentage increase in labour costs has a larger percentage effect on total cost for the smaller machine, the effect of the expected 50 per cent increase in labour cost will be to make the larger machine relatively more attractive.

Case Study 4.2: The Price of Butter

1

1. If Group A is correct, the cost of the additional five million kilos of butter would be the total cost of nine million kilos less the cost of the original four million kilos, i.e.

   $$(9\text{m} \times \$3) - (4\text{m} \times \$2) = \$27\text{m} - \$8\text{m} = \$19 \text{ million}$$

   The cost per kilo for the additional butter is then \$19m/5m = \$3.80 per kilo.
2. If Group B is correct, the cost of the additional one million kilos of butter would be the total cost of five million kilos less the cost of the original four million kilos, i.e.

   $$(5\text{m} \times \$3) - (4\text{m} \times \$2) = \$15\text{m} - \$8\text{m} = \$7 \text{ million}$$

   The cost per kilo for the additional butter is \$7m/\$1m = \$7 per kilo. This is more than three times the cost of the original butter at \$2 per kilo.
3. The difference between the two curves in Figure 4.27 is their price elasticity. Curve A is price-elastic in the range \$2–\$3, where a 50 per cent increase in price led to a 125 per cent increase in output. Curve B is price-inelastic in the range \$2–\$3, where a 50 per cent increase in price led to only a 25 per cent increase in output.
4. When the price is increased to \$3 per kilo, it is necessary to pay that price for all butter produced; this includes the butter that would have been produced at the price of \$2 per kilo: the price set on the market is the price paid for *all* butter, not just the additional butter produced.
5. When estimating the relationship between the price and the quantity supplied, it is necessary to make the *ceteris paribus* assumption, namely that all other relevant factors remain unchanged. However, between the time that the study was done and the time that the price increase had its effect, many things could have changed – for example, it might be a bad year for weather and the output of milk may decrease; the price of wheat may increase dramatically and many dairy farmers may switch over to wheat production; there may be increases in wage costs, which put some producers out of business. When analysing the output of butter the following year, it would be necessary to take into account any such changes.

Case Study 4.3: Production Designs for a Jewellery Manufacturer

1 In making short-run decisions, a firm is concerned only with variable costs; fixed costs remain unaltered in the short run and should not enter into output decisions. Also, a competitive firm has no influence over market price: it is a price taker, and should attempt to sell as many units as it wants at that price. It has no incentive to charge a lower price and will sell zero output if it attempts to charge more than the going price.

In this analysis, the information provided is presented in its most useful form for decision making in Table A2.1 below, which consists of 10 columns for each product where:

Column 1: different quantities of output.
Column 2: price per unit.
Column 3: total revenues from sales, calculated by multiplying price by quantity.
Column 4: total cost of producing each level of output.
Column 5: fixed cost of plant and equipment, which is independent of output.
Column 6: variable cost, i.e. the direct costs incurred at each level of output, calculated by subtracting fixed cost from total cost.
Column 7: average total cost, calculated by dividing total cost by the number of units produced.
Column 8: average variable cost, calculated by dividing the variable cost by the number of units produced.
Column 9: marginal cost, namely the cost of additional output where (for example) the cost of an extra 10 rings when 30 are already being produced is $200 (i.e. $900 − $700).
Column 10: marginal cost per unit produced. It is not possible to find the marginal cost of each unit except in the case of necklaces; therefore the figure derived in Column 9 is divided by the additional number of units produced – for example, the marginal cost per unit of bracelets when 300 are already being produced is $3 per unit (i.e. (850 − 550)/100).

Table A2.1 Cost and revenue information relevant to short-run output decisions

	Output per day	Price $	Total revenue $	Total cost $	Fixed cost $	Variable cost $	Average total cost $	Average variable cost $	Marginal cost $	MC per unit $
Bracelets	0	4.00	0	100	100	0				
	100	4.00	400	200	100	100	2.00	1.00	100	1.00
	200	4.00	800	350	100	250	1.75	1.25	150	1.50
	300*	4.00	1 200	550	100	450	1.83	1.50	200	2.00
	400	4.00	1 600	850	100	750	2.13	1.88	300	3.00
	500	4.00	2 000	1 300	100	1 200	2.60	2.40	450	4.50
Necklaces	0	300.00	0	800	800	0				
	1	300.00	300	900	800	100	900.00	100.00	100	100.00
	2	300.00	600	1 025	800	225	513.00	113.00	125	125.00
	3	300.00	900	1 200	800	400	400.00	133.00	175	175.00
	4	300.00	1 200	1 500	800	700	375.00	175.00	300	300.00
	5*	300.00	1 500	1 900	800	1 100	380.00	220.00	400	400.00
Rings	0	15.00	0	300	300	0				
	10*	15.00	150	400	300	100	40.00	10.00	100	10.00
	20	15.00	300	540	300	240	27.00	12.00	140	14.00
	30	15.00	450	700	300	400	23.00	13.00	160	16.00
	40	15.00	600	900	300	600	22.50	15.00	200	20.00
	50	15.00	750	1 500	300	1 200	30.00	24.00	600	60.00
Brooches	0	8.00	0	0	0	0				
	10	8.00	80	100	0	100	10.00	10.00	100	10.00
	20	8.00	160	210	0	210	10.50	10.50	110	11.00
	30*	8.00	240	330	0	330	11.00	11.00	120	12.00
	40	8.00	320	490	0	490	12.30	12.30	160	16.00
	50	8.00	400	700	0	700	14.00	14.00	210	21.00

* Current output

These data are presented in diagrammatic form in Figure A2.2 to Figure A2.5.

The short-run decision for each product can now be analysed quite simply. If the price for which the product is sold on the market, i.e. the marginal revenue, differs from the marginal cost of production, some adjustment requires to be made in the level of output. The rule for profit maximisation is, given that price exceeds average variable cost, to continue increasing output up to the point at which marginal revenue equals marginal cost. As long as marginal revenue is greater than marginal cost, each additional unit sold adds more to total revenue than it does to total cost and consequently increases profit (or decreases loss).

Bracelets

At the current level of output of 300 bracelets per day, marginal cost is less than the market price of $4. Output should therefore be increased. If it is possible to produce between 400 and 500 units of output and marginal cost continues to increase with output, the optimal level of output would be approximately 420 per day, as shown in Figure A2.2, i.e. where marginal revenue equals marginal cost.

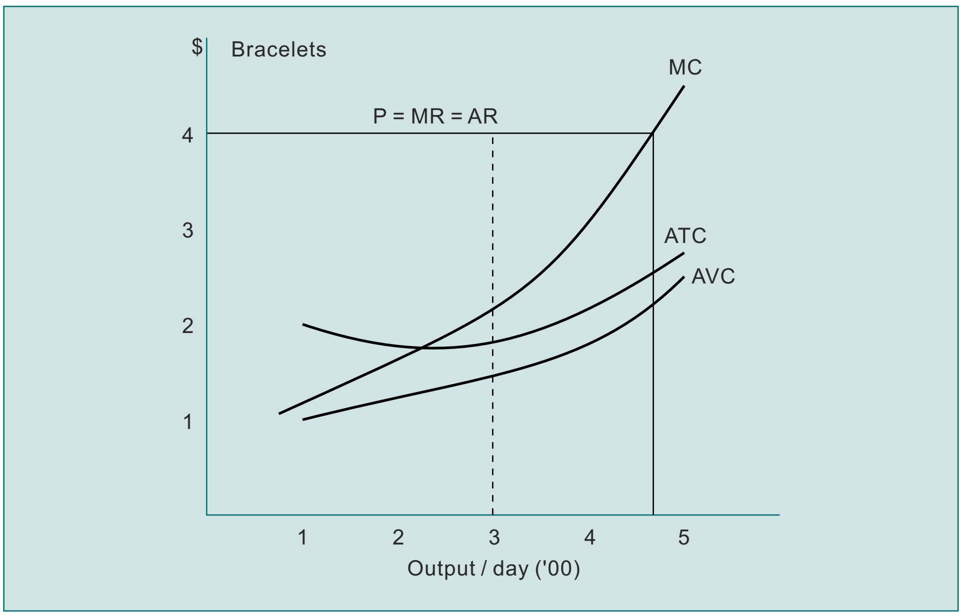


Figure A2.2 Cost and revenue curves – bracelets

Necklaces

At the current level of output of five necklaces/day, marginal cost exceeds the market price of $300. The output of necklaces should therefore be contracted as in Figure A2.3. Since the cost of producing the fourth necklace is $300 and since the firm receives $300 for each necklace, it will be indifferent between producing three and four necklaces per day. At either level of output, AR is less than ATC and greater than AVC. As a result, variable cost is covered and some revenue is available to set against fixed cost. Thus, by continuing to produce necklaces the firm loses less than it would do if it were to cease production. This is a loss-minimising rather than a profit-maximising decision.

Rings

Marginal cost is lower than price at the current level of output of 10 rings per day. The firm should increase production to approximately 20 units per day as shown in Figure A2.4. As in the case of necklaces, this is a loss-minimising decision.

Brooches

Marginal cost exceeds marginal revenue for all levels of output. At no level of output does average revenue exceed average variable cost, which equals average total cost since fixed cost is zero. This means that the firm cannot, at any level of output, cover variable costs, and consequently should cease production of brooches.

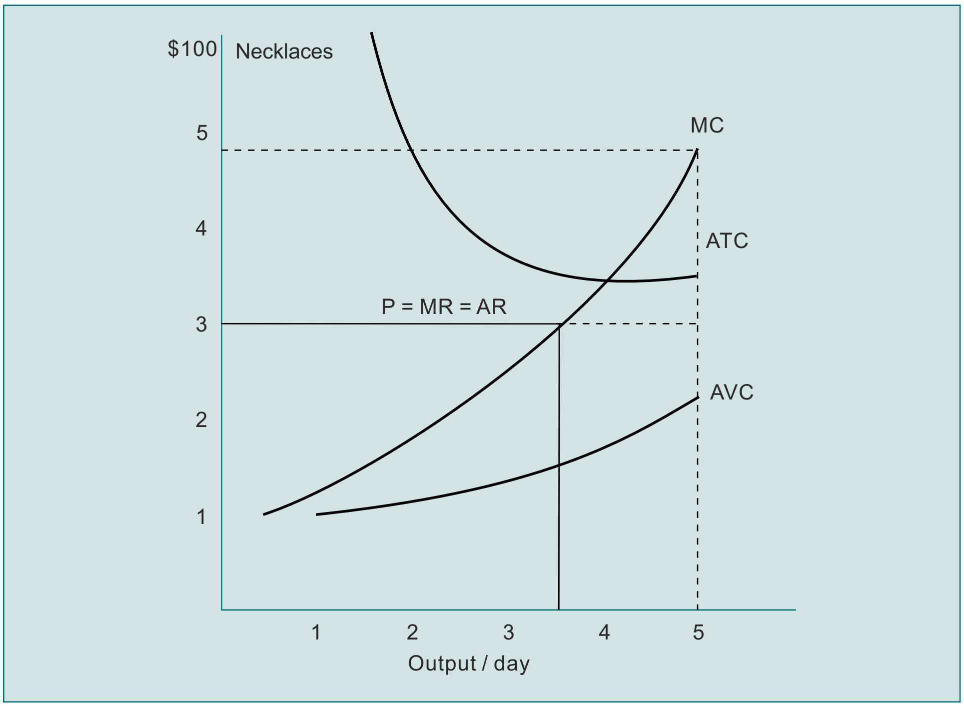


Figure A2.3 Cost and revenue curves – necklaces

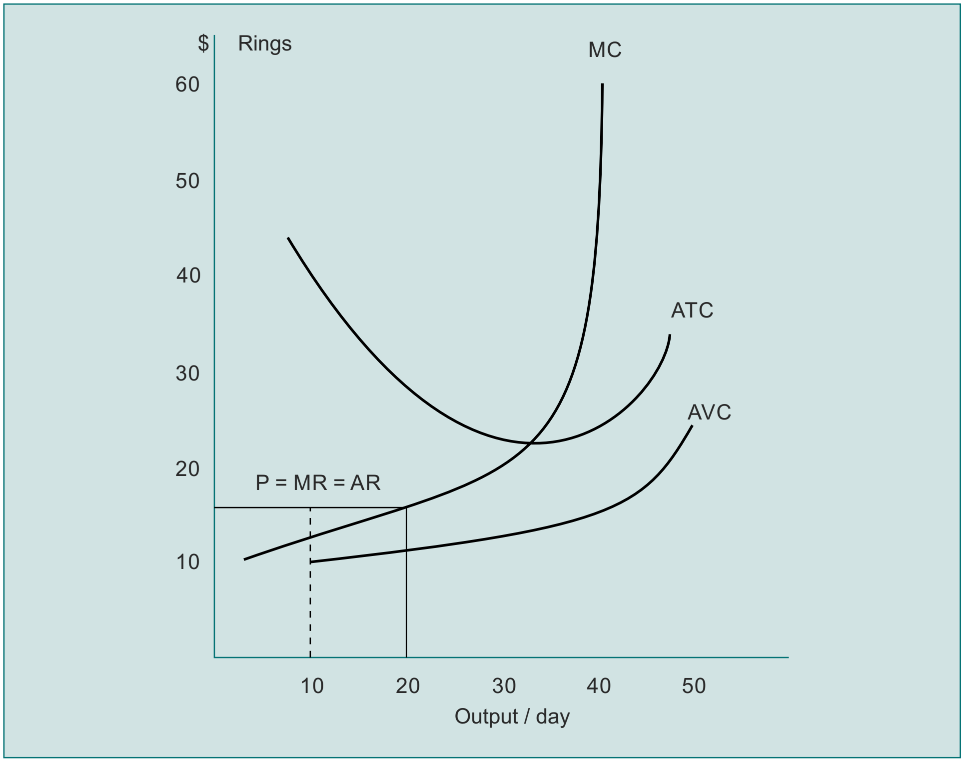


Figure A2.4 Cost and revenue curves – rings

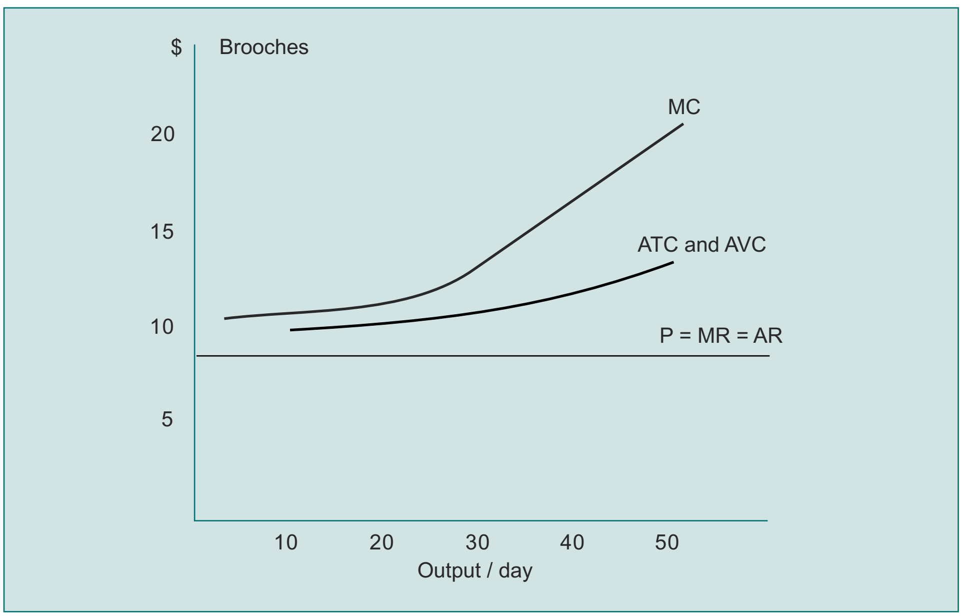


Figure A2.5 Cost and revenue curves – brooches

Case Study 4.4: The Empty House

1 To solve this case it is necessary to separate the cost of each listed item by category, as follows:

(a) Fixed costs: These are unavoidable costs, borne by the surgeon whether she lets her house, leaves it empty, or stays in it herself. They comprise mortgage, insurance and rates.
(b) Variable costs borne by tenant: These are user costs, which vary by the amount of each item consumed. They comprise electricity and gas, but since they are borne by the tenant they will not enter into the rent decision.
(c) Variable costs borne by surgeon: These are costs which vary depending upon whether the house is empty or let, and they are borne by the surgeon. They comprise depreciation, the gardener, the security firm and dog care.

Table A2.2 below summarises these costs by category in terms of 'empty' versus 'let'. Some people might take the view that the rent charged will have to be in excess of \$81 per week, but this is incorrect. It may be impossible to charge such a rent, and this will depend on the circumstances of the market, but the minimum amount that the surgeon will be willing to charge in pursuit of her own interest is significantly less. It can be seen that the weekly cost of maintaining the house is less when occupied. Thus, the surgeon would be willing to pay a tenant up to the difference between the cost empty and the cost occupied, i.e. \$81 − \$58 = \$23 per week. Thus, the rent at which she should be indifferent between leaving the house empty or letting it is −\$23 per week. This is by no means an untypical situation and explains why many people are willing to let their houses for short periods at nominal rents, zero rents, or even pay people to live in, especially if they have to take care of dogs or children.

Table A2.2 Summary of costs by category

	Empty (\$)	Let (\$)	Cost category
Mortgage	30	30	Fixed
Electricity	irrelevant	irrelevant	–
Gas	irrelevant	irrelevant	–
Gardener	10	0	Variable
Depreciation	5	10	Variable
Insurance	3	3	Fixed
Rates	15	15	Fixed
Security firm	10	0	Variable
Dog care	8	0	Variable
Totals	81	58	

Another way to tackle the problem is to ignore the fixed costs and concentrate only on those costs that vary with occupancy. The reason for ignoring the fixed costs is that the surgeon must pay them whether the house is empty or occupied, and thus they do not enter into the house-renting decision-making process given that the surgeon is going to go to the USA. It is true that if the house remains empty the surgeon will save \$5 (\$10 − \$5) per week on depreciation, but she will incur other costs totalling \$28 (gardener \$10 + security firm \$10 + dog care \$8) if the house is empty, yielding a net cost of \$23 (\$28 − \$5). Thus the surgeon would be prepared to pay someone up to \$23 per week to live in her house. If no one was prepared to accept up to \$23 to live in the surgeon's house, she should leave the house unoccupied rather than pay more than \$23.

Module 5

Review Questions

Multiple Choice Questions

5.1 The correct answer is D. A market exists when potential buyers and sellers are in communication with each other. Some transactions occur via brokers and/or on telephone lines, e.g. foreign exchange and shipping cargos. A physical building or location is not necessary. Sellers with goods unable to contact any buyers and buyers wishing goods but unable to contact sellers do not constitute markets.

5.2 The correct answer is B. The quantity of pianos being exchanged is three, and since the store manager agrees to pay for the removal of the pianos, the price is negative. The only figure showing an equilibrium quantity of three at a negative price is Figure 5.31b.

5.3 The correct answer is B. The demand curve shows no pencils are demanded at a price of \$250. The supply curve shows that no pencils are supplied at a price of \$300. Thus no exchange will occur.

5.4 The correct answer is B. When excess demand exists in a market (i.e. at a given price the quantity demanded exceeds the quantity supplied), the price of a good in the market will rise until equilibrium is reached, i.e. until excess demand or supply have been eliminated. Firms obeying marginal principles will move up their supply curves (their short-run marginal cost curves). For each additional unit of output produced for which MR > MC, profit will increase by the difference between MR and MC.

5.5 The correct answer is A. At the price set by the government, excess demand existed. Thus the price required to clear the market exceeded the set price. It should be noted that had the vaccine been sold privately at the 'equilibrium' price, the problem of vaccine distribution would have been solved differently: having set a price at which excess demand existed, the government then had to decide who was going to receive the vaccine and who, while willing to pay the set price, was going to be excluded. In a competitive market, distribution would have been determined by competition among buyers, with those able and willing to pay receiving the vaccine.

5.6 The correct answer is B. In equilibrium there are no unsatisfied buyers or sellers, i.e. no excess demand and no excess supply. This means that any buyer can buy as much as he wants at the going price; any buyer who would be prepared to pay more for the good need not. Similarly any seller can sell as much as he wants at the going price; any seller who would be prepared to offer the good at a lower price need not do so.

5.7 The correct answer is A. There is a demand for, and a supply of, private rental housing, which jointly determine the equilibrium market output and price. When the 5000 dwelling units are complete, the demand curve for private housing will shift to the left as many people move into the dwelling units. In the short run, given that the supply of private housing is fixed, the shift of the demand curve must cause the rental price to fall. This will cause landlords' profits to decrease and, in the long run, resources to move out of the private housing market.

5.8 The correct answer is A. The increase in the number of golfers will shift the demand curve for golf clubs to the right. Simultaneously, the decrease in production costs will shift the supply curve of golf clubs to the right. Thus a larger quantity of clubs will be exchanged (bought or sold) but insufficient information exists to determine what will happen to the price of clubs: it depends on the relative shifts of the demand and supply curves.

5.9 The correct answer is D. Highly price-inelastic demand means that any increase in price will reduce the quantity demanded by a relatively small amount. Thus a tax levied on a good that was highly price-inelastic would be borne primarily by the buyers and would cause the least disturbance of the existing quantity exchanged. A tax levied with highly inelastic supply would also disturb the equilibrium quantity by a small amount but the incidence would fall more on the sellers. By whom the tax is borne depends upon the relative elasticities of demand and supply.

5.10 The correct answer is C. The increase in the sales tax on new cars would have two effects on the market for used cars: it would increase the demand for used cars by people who do not own any car, and it would reduce the supply of used cars as people who would otherwise sell their present cars to buy new cars decide to keep their old cars. When the demand increases and supply decreases, the price will definitely rise, but the quantity sold might either increase or decrease.

5.11 The correct answer is C. In the market time period, it is assumed that there is a fixed quantity of salmon to sell irrespective of the price, i.e. the supply curve is a vertical line. The salmon have been delivered to the auctioneer, whose job is to sell them at the highest price he can get (that price may be zero if no one wants them). Given an increase in demand in the market period, the equilibrium quantity will not change but the price will rise until equilibrium is achieved.

5.12 The correct answer is B. The increase in wages (an increase in variable costs) will raise each tomato farmer's average variable cost curve and consequently average total cost curve and shift the marginal cost curve (the short-run supply curve) to the left. The industry short-run supply curve therefore will shift to the left, with a higher equilibrium price and lower equilibrium quantity resulting.

Case Study 5.1: Housing in Southern California

1

1. The effect of an increase in demand is to shift the demand curve to the right, while an increase in costs shifts the supply curve to the left. This is shown in Figure A2.6. The demand curve in time period 1 is D_1 and in time period 2 is D_2, and the supply curves are S_1 and S_2 in the two time periods.

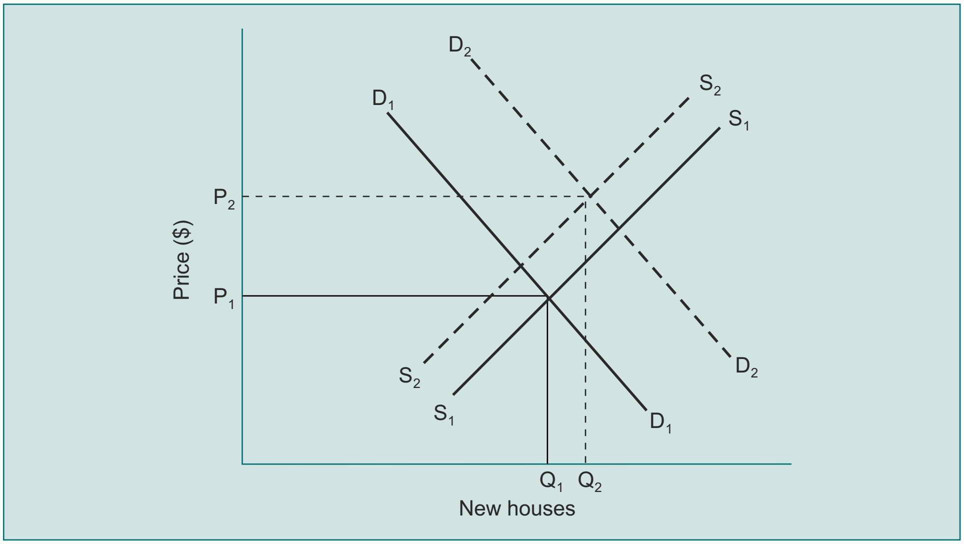


Figure A2.6 Shifts in the demand and supply curves for new houses

   The demand curve has shifted to the right because, at each price, people are now willing to buy more houses. The supply curve has shifted to the left because the cost of production has increased and therefore firms are willing to supply fewer houses at each price.

   The original equilibrium price is P_1. Either an increase in demand or a reduction in supply would have caused the price to increase; therefore a combination of an increase in both will obviously result in a higher equilibrium price, such as P_2. But the effect on quantity is not the same as that on price. In Figure A2.6 the original equilibrium quantity is Q_1 and the new equilibrium quantity is Q_2. However, although Q_2 is greater than Q_1 in this example, it could just as well have been less. For example, if the shift to the left of the supply curve had been much greater, then the new equilibrium quantity would have been less than Q_1. Thus the effect of an increase in costs and an increase in demand on the market for new houses is as follows: equilibrium price will increase but equilibrium quantity may increase, decrease or remain the same.
2. Second-hand houses are substitutes for new houses, so that when deciding on buying a house, people will compare the prices and characteristics of both new and existing houses. The effect of an increase in the price of new houses will be to increase the number of people who are willing to buy an existing house at each price. This is shown in Figure A2.7, in which the demand curve for second-hand houses has been shifted to the right as a result of the increase in price for new houses. The demand curve D_1 shifts to D_2 after the increase in the price of new houses. As a result, the equilibrium price of

second-hand houses increases from P_1 to P_2 and the equilibrium quantity increases from Q_1 to Q_2.

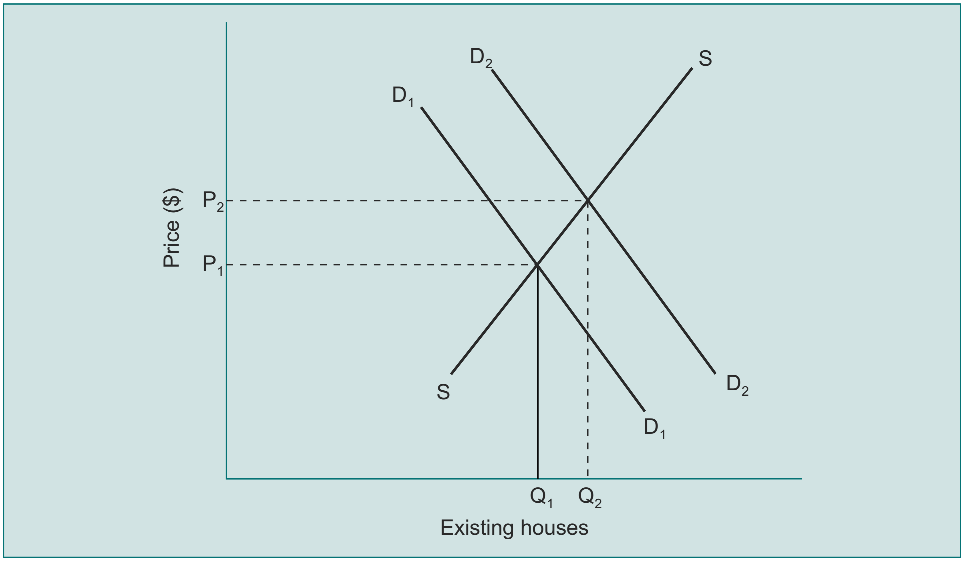


Figure A2.7 Effects of a rise in price of new houses on the market for existing houses

It is also possible that there would be a reduction in the supply of second-hand houses as people became less willing to move house due to the higher price of new houses. This would have the effect of shifting the supply curve SS to the left. While the equilibrium price would still be higher than P_1, the equilibrium quantity Q_2 could be higher, equal to or less than Q_1, depending on the extent of the shift in S.

3. The market for the 191 houses is represented in Figure A2.8. It is known that, at a price of \$115 000, 1000 houses would be sold.

   However, the supply of houses is fixed at 191, i.e. it is shown by the inelastic supply curve SS. Thus, there is an excess demand of $1000 - 191 = 809$ houses at that price. If the free market had been allowed to operate, i.e. if buyers had been allowed to bid for the houses, some higher price would have emerged such as P_e. At this price, there would have been no unsatisfied buyers and excess demand would have ceased to exist. However, it is not known what the equilibrium price P_e would be other than that it would be higher than \$115 000. The lottery has no effect on economic efficiency because the same number of houses are bought and sold whether a lottery is used or not.

   However, the distribution of income is different under the two methods of allocation. Under the lottery system the sellers are much worse off because they are selling 191 houses at \$115 000 instead of at the higher equilibrium price P_e.

Figure A2.8 **Excess demand for new houses**

Case Study 5.2: Obtaining Wimbledon Tennis Tickets

1

1. The current method of allocating Wimbledon tennis tickets and the free-market system are compared in Figure A2.9. It is assumed that Wimbledon can hold, at a maximum, $0Q_2$ people and that $0Q_2$ tickets are printed for each day of play. Figure A2.9 therefore represents the demand for and supply of tickets on any given day.
   Assume that P_1 is the ticket price set by the Wimbledon authorities. The demand curve shows that, at the price of P_1, $0Q_1$ tickets would be purchased. Since the supply of tickets is limited to $0Q_2$, excess demand for tickets of Q_2Q_1 exists at the price P_1. Were a free-market system to operate for the tickets, the equilibrium price would be (say) P_2, for at this price the market is cleared of excess demand. Everyone who wishes to buy a ticket at P_2 can do so; no ticket remains unsold.

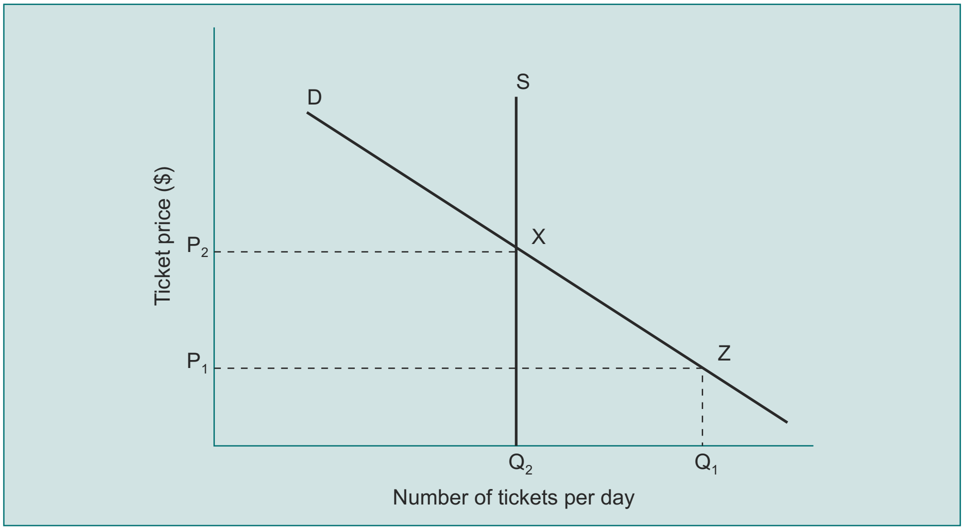


Figure A2.9 Allocation of Wimbledon tennis tournament tickets

2. In the current Wimbledon set-up, with a price of $0P_1$, the price system will not clear the market of excess demand and some non-price mechanism therefore is required to decide who gets the tickets. The two 'official' non-price mechanisms used are (a) a lottery and (b) queues. In the lottery, chance decides whose name comes out of the draw. In the queue system, first-come-first-served is the rule of the day. The lottery system requires some resources – application forms, stamps, cheques, and facilities to make the draw and to mail tickets to lucky applicants. The queue system also involves scarce resources, namely the time of people standing/sleeping in queues, blankets, hot drinks – and maybe even umbrellas!

   The people who gain in the current system are those who are allocated a ticket through the lottery at a price of $0P_1$, who would have been prepared to pay more than $0P_1$. Those who lose out in the current system are the individuals who neither receive a ticket in the lottery nor are able to queue but who would have bought a ticket at the price $0P_2$. There are other gainers/losers in the current set-up. Some people have no intention of watching tennis live at Wimbledon but apply for tickets through the lottery. If they are lucky and obtain a ticket (at price $0P_1$), they then offer it for resale at whatever price the 'unofficial' market will bear. The market may be the local pub, outside Wimbledon itself – the favourite beat of the touts – at stations on the Wimbledon line, or even at underground stations in London. People going to Wimbledon with an extra ticket or two – some companion having fallen sick and unable to go – are another source of supply. All such people who receive a price higher than $0P_1$ are gainers.

   Since Wimbledon Tennis Club does not gain from unofficial sales, the loss of revenue to the Wimbledon authorities from the current allocative policies is represented in the figure by the difference between income currently received, $0P_1 \times 0Q_2$ (price × quantity) and the income that would have been received in the open-market situation, $0P_2 \times 0Q_2$.
3. The current system, it is argued, allows tennis enthusiasts who cannot afford $0P_2$ an opportunity (if they are lucky in the lottery) to see Wimbledon live for $0P_1$. But these enthusiasts, once they have the tickets for $0P_1$, can always sell them in the unofficial market for $0P_2$. If they cannot afford $0P_2$ for a ticket, this implies that they are not willing to forgo the other goods and services that they can consume in order to attend Wimble-

don at the market price. A rational person would find it difficult to turn down an offer of $0P_2$ for a ticket.

4. Economic efficiency is promoted by the touts because they enable the exchange of tickets to the mutual advantage of both parties, i.e. both buyers and sellers are made better off in the unofficial market; if this were not the case, no exchange – tickets for cash and cash for tickets – would occur. However, since the unofficial market involves additional resources – touts, and police to regulate the touts' activities – the current system plus the unofficial market is not as economically efficient as a free market would be.

   Should society retain the law making touts' activities illegal? From the viewpoint of economic efficiency one would have to consider the marginal benefit and the marginal cost of touting. The marginal benefit to society is the benefit that accrues from any market, i.e. ensuring that those who are willing to pay for a good receive that good, and generating sufficient profit for the sellers to ensure supplies are forthcoming. The marginal cost to society is the opportunity cost of the touts; this may or may not be zero.

Module 6

Review Questions

Multiple Choice Questions

6.1 The correct answer is D. The definition of industry equilibrium is that no forces exist to cause change. This means that firms are making profits, average revenue (price) equals long-run average cost (which includes normal profit), and no incentives exist for firms to enter or leave the industry.

6.2 The correct answer is A. Given that we are dealing with a purely competitive industry, any firm earning above-normal profits (i.e. having a lower average-cost curve) would be imitated by other firms. This would result in the industry supply curve shifting to the right, a lower equilibrium price, and the re-establishment of normal profits for all firms. Resources will flow into the industry; lazy and complacent firms will be forced out.

6.3 The correct answer is C. When an industry is in long-run equilibrium and experiences a rise in consumer demand (a shift to the right of the demand curve), the short-run industry equilibrium will occur at that output and price at which the new demand curve intersects the industry's short-run supply curve (the summation of the firms' marginal cost curves). Thus there will be an increase in price and producers' profit. The above-normal profits will in the long run attract new firms, hence increasing industry supply. This in turn will force down the price until a new long-run equilibrium is established.

6.4 The correct answer is B. If a purely competitive economy were in equilibrium, by definition no force would exist to cause change. Consumers could not become better off by adjusting their spending patterns (they would all be maximising utility) and no resource could earn a higher return by moving to another industry (they would all be earning normal profits). In the real world, of course, general equilibrium is never reached because change, e.g. inventions, is continually occurring but consumers and firms aim for the state of general equilibrium – a useful analytical concept.

6.5 The correct answer is C. The short-run supply curve of a firm is its marginal cost curve above the minimal point of the average variable cost curve. A profit-maximising firm will offer an additional unit of output as long as the extra revenue (price, in a perfectly competitive industry) exceeds the extra cost of production, i.e. the marginal cost; thus the marginal cost curve determines the supply curve. For an industry as a whole, the short-run supply curve is the summation of the firm's short-run marginal cost curves.

6.6 The correct answer is C. In order for resources to be used efficiently, it is necessary that the prices of goods and services produced by those resources be equal to marginal cost. Clearly, the marginal cost of transportation services is higher during rush hour if trains and buses are filled to capacity since each traveller prevents some other potential traveller from consuming the services. The price should be set during rush hour so that there is no excess demand. This will ensure that the people who travel during rush hour are the people who value it most, i.e. the people who are willing to pay that high price. During the off-peak hours, the price should just cover the cost in terms of resources used up associated with carrying an additional passenger. Looked at from another point of view, such a pricing scheme would serve to spread the use of transportation facilities over the day, thereby reducing the amount of equipment required to handle the peak load.

6.7 The correct answer is A. The existence of conscription implies that at going wage rates there is excess demand by the government for military personnel. If there were a shift to a market system in which wages were raised to a level where personnel requirements were achieved by offering a sufficiently high wage, then the income of military personnel would rise and the taxes required to pay for the higher military budget would also have to rise. Actually, since using the price system to determine who would join the service would result in those people joining who were least unwilling to join (or who had the least desirable alternatives), it would be possible under such a system for everyone to be made better off if there were some means by which the conscripts who would not join under the price system could be made to compensate other taxpayers who face higher taxes. Since there normally would not be such a means, the conscripts have lower income and civilian taxpayers higher income under the conscription system.

6.8 The correct answer is C. A supply curve tells what quantities would be offered at different prices. If a market were in disequilibrium with excess demand existing, not all buyers would be satisfied and market forces would result in a rise in price. At the higher price a larger quantity would be offered – but there would be a movement along the supply curve and not a shift in the position of the curve.

6.9 The correct answer is A. While profits indicate where capital can most efficiently be allocated, they yield no information about the wage level that, in a private-enterprise economy, is determined by the demand for and supply of labour. Thus, for any given firm, no causal relationship exists between profit and wage level. In addition, concluding that a wage is 'too low' involves making a value judgement. In a private-enterprise economy, resources will be allocated in accordance with consumers' wants only if firms' production decisions are responsive to consumers' expenditure decisions. An increase in demand for a good will at first be reflected in higher prices of that good and in higher profits from producing it. Firms will try to take advantage of this higher profitability by producing more of that good, but only if there is competition will firms fully accommodate consumers' preferences by expanding output to the point where consumers' valuation just matches the cost (including normal profit) of production.

6.10 The correct answer is B. When a market is in equilibrium, buyers can buy all they wish to at the going price and sellers can sell all they wish to at the going price; in other words excess demand and excess supply are non-existent. This does not imply that sellers would not be willing to supply more at higher prices.

6.11 The correct answer is B. At a price of \$2, excess demand exists; to clear the market, price must rise. At \$3, however, excess supply exists. Thus the equilibrium price must lie between \$2 and \$3. At a price of \$4, excess supply exists and those suppliers willing to offer services at lower prices will do so to obtain sales, i.e. there will be a movement along the supply curve. The demand curve given by the first two columns will not change as prices change.

Case Study 6.1: The Economics of Army Conscription

1

1. The output of the armed services should be treated like any other good or service. That is, the last pound spent on the army should yield the same level of social benefit as the last pound spent on any other good or service. The fact that the marginal social benefit of military service is difficult to measure is no reason not to attempt to do so. Once the estimates have been made, output should be planned to equate the marginal benefit divided by the marginal cost for the army with the corresponding ratio for civilian goods and services.
   The real cost to society of having an army is not the monetary payments made for the factors of production used in the army but the civilian output forgone, i.e. the opportunity cost. The real cost to society of adding one more soldier to the army is what that person could have produced in civilian employment, which is reflected by the corresponding civilian wage. Under the conscription system, the cost of a conscript is not the £2000 or whatever the conscript receives as army pay, but the salary the conscript would have earned in civilian life.
   For example, consider the situation in which Joe Goal, the football star, is conscripted into the army and is paid £2000 per annum. Does this mean that Joe Goal's services in the army are costing Britain £2000 each year? The answer is 'no': you have to consider what Joe Goal's opportunity cost is. If the value of his marginal product with the Heriotian Football Club is, say, £50 000 per annum, then the real cost to society of conscripting Joe Goal is the £50 000 of football services he no longer produces. Unless the benefit of having Joe Goal in the army exceeds £50 000, it is a misallocation of resources to conscript him.
2. Figure A2.10 shows the demand curve for and supply curve of soldiers. The demand curve reflects the social benefits of an additional soldier. Its downward slope is due to the fact that the benefit to national defence of adding extra soldiers decreases with the

size of the army. The supply curve reflects the social cost of an additional soldier. Its upward slope is indicative of the fact that at higher pay, more individuals would find military employment more rewarding than their civilian alternatives. At the point where the curves intersect, the marginal social benefit equals the marginal social cost; this shows the optimal army size and the equilibrium wage. It is clear that under the volunteer system, Joe Goal would not join at the equilibrium wage rate, but workers who had civilian alternatives that paid less than the army at that wage would join.

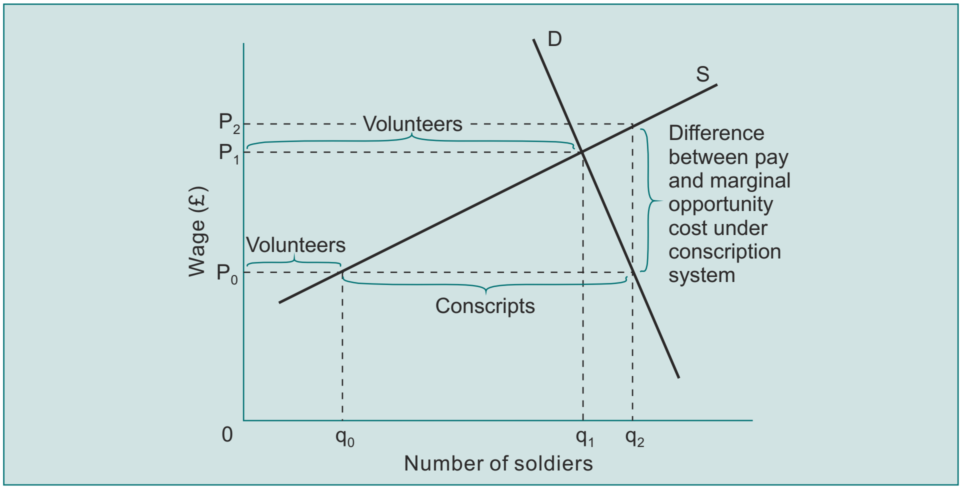


Figure A2.10 Demand for and supply of soldiers

3. The equilibrium under the volunteer system features both higher pay and a smaller army than under the conscription system. The higher pay comes about because the government is forced by the market to pay soldiers their opportunity cost. The smaller size results because government decision makers (and ultimately, taxpayers) have to take into account the full cost of labour to be used by the army. As can be seen from Figure A2.10, under the conscript system the size of the army is $0q_2$ and the pay is $0P_0$. At this rate of pay, only $0q_0$ volunteer and q_0q_2 are conscripted. In this situation, the real cost to society of conscripting the soldier with the highest opportunity cost is $0P_2$, but the benefit is only $0P_0$.
4. The conscript system causes conscripts to bear a disproportionate share of the costs of national defence. If Joe Goal, with potential civilian earnings of £50 000 per annum were conscripted and paid only £2000 in the army, he would be paying an implicit tax rate of 96 per cent. At the same time, a worker willing to serve if he were paid $0P_1$ is also worse off under the conscript system, where he is paid $0P_0$. The people who are better off under conscription are those who pay lower taxes to meet the £2000 conscription wage, i.e. do not pay the full cost of the conscripted army.
5. The operation of this nineteenth-century system means that Joe Goal and a worker earning $0P_1$ who was not conscripted could do a deal: Joe Goal could pay that worker £10 000 to take his place and they would both be better off. If the rest of society is not affected by whether the job of being a soldier is carried out by either Joe Goal or that worker, the welfare of society would be increased.
6. The volunteer system is more economically efficient than conscription; this is because, under the volunteer system, the government will hire those people who can contribute

more to society in the army than in civilian life. For example, assuming Joe Goal and a worker who volunteers would make equally good soldiers, total output would be higher under the volunteer system. Joe Goal could produce more in civilian life than in the army, and our worker more in the army than in civilian life. In other words, the conscript system, even though it costs less in military pay, is more expensive in terms of output forgone.

7. One potential drawback of the volunteer system could arise in time of war. In terms of Figure A2.10, we could expect society's demand curve to shift far to the right. The possibility could occur that the demand and supply curves would not intersect. That is, an insufficient number of men would volunteer at a price that society could pay, because of the enormous danger to soldiers. In such a situation, a conscript system may be in society's interest, but not necessarily in the interests of those who are conscripted. It is possible, of course, for the supply curve to shift to the right during wartime. This would occur if enough people were stirred by 'patriotism' to join the army, independent of the wage rate, as happened in the early stages of World War I.

Module 7

Review Questions

Multiple Choice Questions

7.1 The correct answer is C. By definition, marginal revenue equals the additional revenue that would be earned by selling one additional unit of output. The price at which a monopolist can sell an additional unit is given by the demand curve (or average revenue curve), as it is the difference in price of the last unit sold and the next unit sold. The price that attracts competition will affect a monopolist's demand curve but not its marginal revenue.

7.2 The correct answer is B. It will always pay a monopolist to add to output as long as each additional unit produced adds more to total revenue than it does to total cost, i.e. as long as marginal revenue exceeds marginal cost. Profit will be maximised at the level where marginal revenue equals marginal cost.

7.3 The correct answer is B. The demand curve facing a pure competitor is a horizontal line, i.e. it can sell as much as it wishes at the going price, over which it has no control. For example, if the going price of pencils is 10 cents, it could sell 1 or 1000 without affecting price. The competitor's average revenue would be 10 cents, as would its marginal revenue, and so P = AR = MR. For the monopolist, facing a downward-sloping demand curve, average revenue and marginal revenue differ: for the monopolist, price exceeds marginal revenue. Both a monopolist and competitor could have equal assets, both could be in a constant-cost industry, and neither need be regulated by government.

7.4 The correct answer is C. In monopolistic competition, above-normal profits will attract new firms into the industry, causing the demand curve facing the original firm to shift to the left. This process will continue until only normal profits remain, and this will prevent above-normal profits being earned by successful imitators and will prevent monopoly arising. Monopolistically competitive firms, by definition, cannot erect barriers to entry.

7.5 The correct answer is C. The marginal cost curve intersects the average variable cost curve at its minimum point, i.e. at the current output level, and at a price of \$16.50. However, at this output level, marginal revenue is only \$12. Thus, marginal cost exceeds marginal revenue, and to maximise profit the firm must contract its output. However, at a lower output level the market clearing price must be higher on a downward-sloping demand curve.

7.6 The correct answer is B. A profit-maximising monopolist produces up to the point at which marginal revenue equals marginal cost. At this output level, the price that clears the market exceeds marginal cost. Thus, the price the consumer pays for a good produced by a monopolist does not reflect the marginal cost to society of producing that good.

7.7 The correct answer is D. A profit-maximising monopolist produces up to the point at which marginal revenue equals marginal cost. For the price to equal marginal cost, output would have to be expanded in the monopolistic sector and coincidentally reduced in the competitive sector. When prices equal marginal cost in both sectors, economic efficiency prevails.

7.8 The correct answer is A. The existence of the monopoly in one of the two industries implies that, in that industry, price is higher than marginal cost. If that industry were to become competitive, then price would be bid down and output would be increased in the long run since the existence of excess profits would result in new firms entering the industry. As the output of the previously monopolised industry expanded, resources would be bid away from the competitive industry, which would result in lower output in that industry.

7.9 The correct answer is B. By definition a monopolist's demand curve is his average revenue curve; it tells what quantities can be sold at different prices.

7.10 The correct answer is D. A monopolist's profit-maximising output is determined by the intersection of the MR and MC curve. This occurs in Figure 7.17 at the output level F. The corresponding profit-maximising price, given by the demand (average revenue) curve, is E.

7.11 The correct answer is D. As long as average revenue exceeds average total cost, the monopolist will make a profit (above-normal returns). Those two curves intersect in Figure 7.17 at output point J, beyond which the monopolist will incur losses.

7.12 The correct answer is B. The marginal revenue curve indicates additions to total revenue from the sale of each additional unit of output. When marginal revenue becomes zero, additional production will detract from total revenue. Thus when marginal revenue is zero, total revenue is maximum.

Case Study 7.1: The Price of Antiques

1

1. To simplify the analysis, assume the goods in question are identical antique grandfather clocks. Since the quantity of goods coming to auction is independent of price, the supply of grandfather clocks coming to the monthly auction will be represented by a vertical supply curve. The position of the supply curve will be determined by factors other than price. The impact of the 'ring', admitted by the lawyer, will be to lower prices paid at auctions. This will come about because the demand curve shifts to the left, i.e. some of the potentially high-bidding antique dealers will not in fact bid. The demand and supply curves are shown in Figure A2.11.

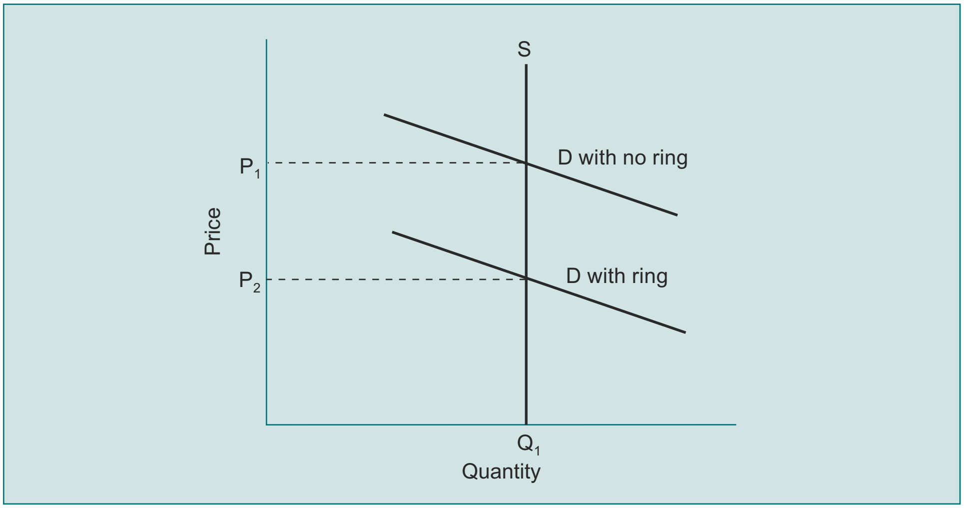


Figure A2.11 Demand and supply curves at auction sales

The impact of a 'ring' will be to make the price of grandfather clocks lower than it otherwise would have been, i.e. P_2 rather than P_1; the quantity bought and sold will not change, but will remain at Q_1. Thus the 'ring' will not affect the efficiency of the market.

2. The claims that the auctioneers 'make substantial profits anyway' is a value judgement. If substantial profits did exist and there was competition in the auctioneering business, other auctioneering firms would have set up in business in the area, thus bidding down the profits of the original firms. Only if it could be shown that there was no competition in the industry would it be possible to make out a case for 'excessive' profits. But even then, the existence of high profits on the part of the auctioneers would be no reason for these profits to be distributed to antique dealers.
3. Assuming that the antique dealers are profit-maximisers, the general public will not be better off because of the 'ring'.

   In any given time period (ignoring past stocks of grandfather clocks), the antique dealers collectively will have Q clocks for sale – all polished up and in working order. Figure A2.12 shows the equilibrium price of these clocks in the antique shops: the price that will clear the market is P. The supply curve S is vertical.

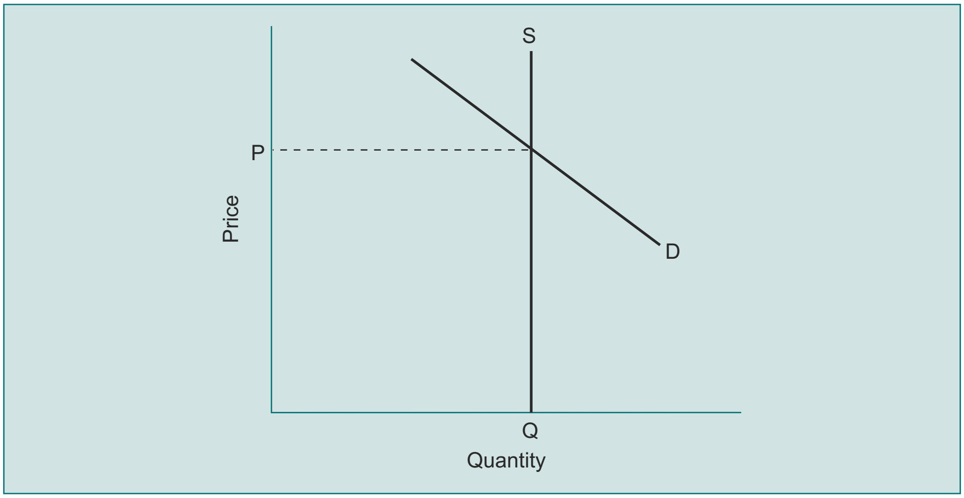


Figure A2.12 Equilibrium price

If the antique dealers are not profit maximisers but operate on a cost-plus-mark-up basis it is possible for clocks to sell at a price lower than P, but this means that the market will not clear – there will be excess demand, and some people prepared to pay a price P for a clock will not be able to obtain one and consequently will suffer a loss of satisfaction.

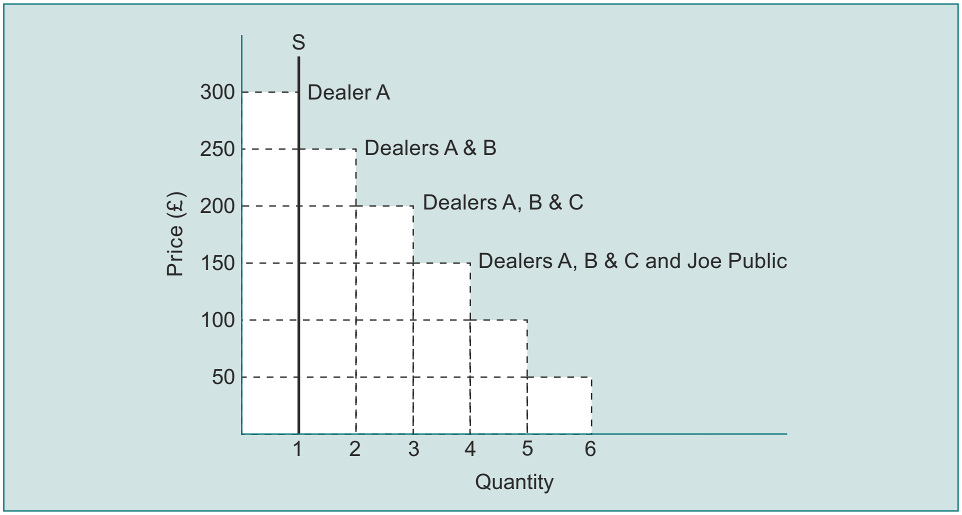


Figure A2.13 Potential loss by sellers at auction through a 'ring'

The lawyer has ignored the loss to the person whose goods are being sold at the auction. Figure A2.13 shows the potential loss to the sellers of a good in the presence of a 'ring'. Figure A2.13 shows one grandfather clock for sale, i.e. supply S: at each and every price, only one clock is offered on the market. When the bidding is opened at £50, six people are willing to buy the clock. When the bidding goes up to £100, only five people are left who are willing to buy the clock; one of the original six was willing to pay at least £50 but not as much as £100. At £150, the four people left are three dealers plus one member of the public, Joe Public. At £200 Joe Public drops out, i.e. he was willing to pay

£150 but not £200. At this stage the bidding stops because the dealers do not bid against each other.

However, if they had not made this agreement, the bidding would have gone on to £300, at which point only Dealer A would have been left willing to pay that amount for the clock. (These are all points on the demand curve.) Thus, instead of receiving £300 (less commission) for the clock, the seller only receives £200 (less commission). At the subsequent 'ring' auction, Dealer A would bid £300 for the clock, and the additional £100 would then be distributed equally among the 'ring' members. The seller's £100 has thus found its way into the pockets of the 'ring' members.

Module 8

Review Questions

Multiple Choice Questions

8.1 The correct answer is C. The disincentives facing an individual contemplating the purchase of a public good is that if anyone else buys it, the individual can enjoy it without having to pay for it and, vice versa, if an individual buys it that individual cannot prevent other people from enjoying it without paying for it. An example is national defence. This logic applies to both rich and poor.

8.2 The correct answer is B. An externality exists when one person's (creature's) actions have a positive effect (bees pollinating fruit) or negative effect (firms polluting rivers) on other members of society. When pollution occurs, the polluters ignore the extra cost that their actions impose on other people. If people were indifferent as to whether or not streams and rivers were polluted, there would be no externality (and incidentally, by definition, no pollution).

8.3 The correct answer is D. The paper mill, in polluting the river, is not taking into account the extra costs incurred by society, i.e. the pollution, in the production of paper. Thus the cost of paper will not reflect the true cost of production, and this means that paper users are being subsidised by other members of society – in this case the city taxpayers.

8.4 The correct answer is D. For efficiency to prevail in an economy, prices of goods must reflect their marginal cost to society. When pollution occurs, the firms responsible ignore the additional costs to society of the pollution; they are concerned only with the costs directly incurred by them, e.g. labour, materials, etc. The idea of a tax on polluters is to force them to pay the full societal costs, i.e. their direct costs plus the cost of pollution. Put another way, the imposed tax should equal total costs (social costs) minus private costs.

8.5 The correct answer is D. Same answer as Question 8.4.

8.6 The correct answer is C. Same answer as Question 8.4.

8.7 The correct answer is B. The river is a source of only two benefits: waste disposal and commercial fish. The greater the amount of waste deposited in the river, the greater the number of fish killed and consequently the smaller the commercial fishing catch. Because the benefits and costs to society from both activities – dumping waste and fishing commercially – are not enumerated, we do not know the best mix of these two uses of the river. What we *do* know is that the chemical producer, in deciding to dump wastes into the river, is considering the private cost of waste disposal (the cost to himself) but not the social cost. He ignores the cost that his actions impose on the commercial fishermen, and as a result the marginal social benefit from this waste disposal is less than the marginal social cost. This implies that efficiency would be increased if less waste were dumped into the river and more fish caught, even though it would mean higher-priced chemicals.

8.8 The correct answer is A. Each consumer, in voting honestly, would indicate how much of good Y that consumer wished. But by its nature other people could enjoy the public good purchased by the consumer, who in turn would enjoy more of the same good purchased by all other consumers voting honestly and independently. For instance, then, each consumer might vote for his own policeman and society would finish up with far too many policemen.

8.9 The correct answer is B. The problem of pollution is better understood in terms of a failure of the price system rather than a defect in people's moral fibre. In some instances, markets fail to provide adequate information to decision makers and as a consequence they fail to take into account some undesirable effects of their economic activity. In these cases the private cost of production is less than the total cost (including the social cost), and consumers respond by consuming too much of the goods in question because they are not paying their full cost. In order to reduce pollution, we must either reduce the external effects by consuming less of the goods whose production creates the pollution or we must use other resources to offset the external effects (e.g. anti-smog devices on cars). Either way, to obtain a pollution-reduced environment will require paying the cost of having less of other things.

Case Study 8.1: Pollution and the Metal Castings Industry

1 The dirt poured out from the iron foundries on the surrounding neighbourhood is a real cost to the people who live there, but it is not a cost borne by the iron foundry itself. Thus, in the absence of government, or some other collective action, the price of iron casting will only reflect private cost, that is, will only cover the cost of producing iron castings, and will not include the cost of cleaning up the windowsills, clothes and lungs of the people who live near the foundry, or of preventing this pollution.

To maximise profit, each firm will produce up to the point at which the firm's marginal costs equal price. Users of iron castings will purchase until the ratio of marginal utility to price for castings is equal to the corresponding ratio for any other good. However, because price now equals private – not social – marginal cost, price does not cover the full social marginal cost of production. To achieve allocative efficiency, the cost of producing castings should include the full social cost rather than merely the economic, or accounting, cost experienced by the individual firm. The marginal equivalency conditions should be met using marginal *social* costs.

This, like other externalities, is a case that requires collective action if society is to have the full cost of producing iron castings reflected in the price charged to the purchaser of them. Without some collective action to incorporate the pollution costs into the price of the finished castings, buyers of castings get them too cheap, and too many castings are produced.

Here the principle of marginalism applies. If we were to eliminate all pollution from iron foundries (zero dirt emission from the cupolas), we would probably end up raising the price of castings so high that few, if any, would be bought. As a practical matter, this standard might well force the production of metal blocks by other devices, such as machining or new electric processes for moulding metals. Costs could be very high, and the shift would not be necessarily desirable socially. Just how far we should force the iron foundries to go in eliminating emissions from their cupolas should be determined by balancing off the marginal costs of installing and operating the emission control equipment against the marginal benefits derived (the cleaner air).

There is no general way of saying just when this equation of marginal costs with marginal benefits will be reached in any given case – it depends on the particular circumstances involved. In the Midlands case, for example, you may or may not feel that going from the 83 per cent emission control level to the 86.3 per cent level is worth the added cost.

1. Whether students picket or not should depend on their evaluation of the marginal costs to themselves of using scarce time and the marginal benefits of their action. The decision must take into account:
   (a) the effect on economic efficiency; and
   (b) the effect on the distribution of income.
   The available information indicates that forcing the foundries to eliminate pollution would have considerable effects on output and employment in the metal castings industry, while contributing only a relatively small improvement to pollution control; many people would be put out of work, thus reducing their incomes. At the same time, Midlands residents showed no enthusiasm for doing anything about car emissions – the major air pollutant.
2. If the Midlands authorities impose rigorous emission-control standards on its iron foundries, their costs will increase and they will have to raise their prices. However, if iron foundries elsewhere do not have to meet the same emission standards, many Midlands foundries will go out of business at the higher prices; customers will simply take their orders where they can buy iron castings cheaper. An alternative would be for the Midlands' citizens to pay for the cleaner air they want by using tax receipts to subsidise foundries to end pollution. Thus, the people of the Midlands must decide whether the benefits of additional clean air through foundry legislation are worth the extra cost. In particular, the argument is about the 83 per cent versus the 86.3 per cent difference: a small marginal benefit for a large marginal cost.
3. If the Midlands decides to live with its air pollution, and if all such pollution were confined to the Midlands, government legislation that imposed pollution control standards more severe than the Midlands citizens wanted would be inefficient. Why? Because the Midlands citizens would be forced to use scarce resources to produce good clean air that they did not believe was worth the cost. If, however, the effects of Midlands found-

ry pollution were found in areas outside the Midlands, a wider view of benefits must be taken to set against the costs.

4. The principle of marginalism also throws light on the question of whether a flat rate is more efficient than a tax (charge) per unit of pollution emitted into the air. The absolute limit (say 83 per cent emission control) falls very unevenly on differently sized firms. The cost of installation is about the same for a very small firm (one that uses its furnace only for an hour or two a day) as for a very large firm (which may run its furnace for many hours a day and hence emit far more pollution into the air). The principle of marginalism says efficiency would be achieved by charging a tax, or charge, for so much per pound of particles emitted into the air. This would mean that the larger firm would have to pay a much larger tax than would the smaller firm because it emits so much more dirt into the air per day. The larger tax burden on the larger producer would put a much heavier pressure on that producer to reduce emissions than it would on the smaller producer, whose tax would be much lighter as a result of the lesser emissions of dirt. Thus the tax would put financial pressure where it would do the most to clean up pollution.
5. It is possible to identify the impact of different methods of taxation and subsidy on economic efficiency. For example, the tax per pound, which falls both on producers and consumers, will affect output decisions and consequently the prices paid by consumers, by changing the relationship between marginal cost and marginal revenue. However, if a subsidy is paid to the firms in order to make it possible to install and maintain the anti-pollution devices, there will be no effect on output and price, since marginal cost will not be affected.

   However, the different types of income distribution that would result from different policies involve considerations of equity rather than efficiency, and they cannot be judged on the basis of economic analysis.

Module 9

Review Questions

Multiple Choice Questions

9.1 The correct answer is B. The actions of any one firm in a competitive labour market affect neither the going wage rate nor the equilibrium level of employment. Were a firm operating at an output level at which the wage rate exceeded labour's marginal revenue product, profit-maximising behaviour dictates that the firm reduce its level of employment until the contribution of the last worker hired (marginal revenue product) just equals the wage rate.

9.2 The correct answer is C. The profit-maximising monopsonistic firm will hire labour up to the point at which the benefit to the firm of hiring one more worker (the value of the marginal product of labour) is equal to the cost to the firm of hiring that worker (the marginal cost of labour). Consequently the monopsonist will not hire labour up to the point at which the value of the marginal product of labour equals the wage rate because, given the upward-sloping supply curve of labour, the marginal cost of labour will exceed the wage rate. This is what economists mean by 'the exploitation of labour' – Marxists have a different definition.

9.3 The correct answer is D. Profit-maximising firms hire factors of production up to the point at which their contribution to output (the value of their marginal product) equals the going rate. In competitive markets, wages are determined for managing directors and unskilled workers by demand and supply conditions. The reason the wages of the former typically exceed those of the latter is because of relative demand and supply conditions.

9.4 The correct answer is B. The fact that employers were forced to pay a higher-than-competitive equilibrium wage rate does not imply that they did not equate wage and marginal product. Indeed, the laying-off of workers indicates that they moved in that direction. Thus, there is no reason to suppose that response A is correct. Since total wage income rose, however, not only were the employed coalminers better off than before but they were in a position to pay the laid-off workers their original wage income, themselves their original wage income, and still have some income left over.

9.5 The correct answer is C. Equilibrium exists in a competitive economic system when no incentive exists for any factor of production to seek alternative employment. This neither implies equality of income among people with similar abilities – they could have different preferences for work versus leisure – nor does it imply no one will starve, i.e. that someone with no marketable skills will be unemployed. Also, an unskilled labourer could work very hard but have a much lower income than a skilled football player who may work very little.

9.6 The correct answer is A. Since the wage of \$9 per day is constant, the marginal cost of labour is \$9. A profit-maximising firm will hire labour up to the point at which the value of the marginal product of labour just equals the wage rate. This equality is never actually reached in the table. In hiring a sixth worker, the addition to output is 10 units ($70 - 60$) worth \$10 (\$1 × 10). This exceeds the wage rate of \$9, and this worker should therefore be hired. However, the value of the marginal product of the seventh worker is only \$8 and therefore he should not be hired.

9.7 The correct answer is B. A profit-maximising firm will hire an additional worker if that worker contributes more to revenue (value of his marginal product) than to cost (going wage rate). In the table the value of the marginal product of labour steadily declines as workers are hired. The third worker adds \$1.40 (\$6.20 – \$4.80) and costs \$1.20, and therefore should be hired. However, the fourth worker adds \$1 but costs \$1.20 and therefore should not be hired.

9.8 The correct answer is C. Prior to the \$1.60 minimum-wage legislation, the farmer was hiring three pickers. However, profit-maximising behaviour dictates that the farmer now hires only two pickers because the third picker's marginal product of \$1.40 (\$6.20 − \$4.80) is less than the wage rate of \$1.60. Thus that picker will be fired and become worse off, although the other two workers will benefit. Output per worker (the value of the average product of labour) will rise, but total output will fall.

Case Study 9.1: Does the Eurozone Need a Minimum-Wage Law?

1 The effect of the minimum-wage law on firms and the portion of the labour market that would be affected is illustrated in Figure A2.14.

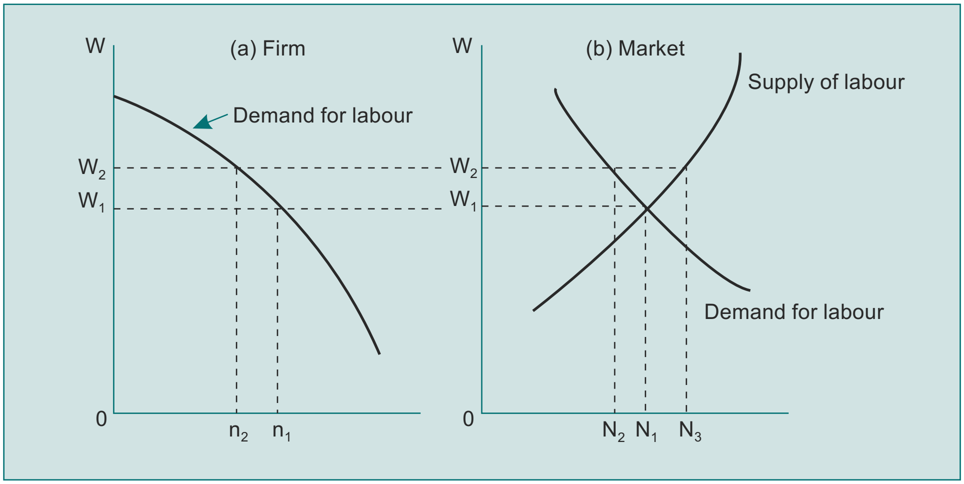


Figure A2.14 Minimum wage and the demand and supply of labour

1. A firm's demand for labour is based on the marginal revenue product of labour. As more labour is hired, the marginal product of labour falls; hence marginal revenue product declines as more labour is hired. As a result, the demand curve for labour by the firm is downward-sloping, as shown in Figure A2.14a. The effect of increasing the wage rate from $0W_1$ to $0W_2$ is to reduce the demand for labour from $0n_1$ to $0n_2$.
2. Since the market demand for labour is the sum of the demands of all firms, it is also downward-sloping, as shown in Figure A2.14b. The effect of increasing the wage rate from $0W_1$ to $0W_2$ is to reduce the demand for labour from $0N_1$ to $0N_2$.
3. The supply of labour is upward-sloping, as shown in Figure A2.14b. The increase in wage rate from $0W_1$ to $0W_2$ will increase the number of people willing to work from $0N_1$ to $0N_3$.
4. The labour market is initially in equilibrium at a wage of $0W_1$ and a labour force of $0N_1$. When the government imposes a minimum wage of $0W_2$, the quantity of labour demanded fails by N_1N_2. However, the quantity of labour supplied increases at the same time by N_1N_3 due to the higher wages offered. This causes unemployment of N_2N_3.
5. The benefits of the minimum wages are confined to those workers who remain employed at the new higher wages, i.e. the $0N_2$ workers who are now paid $0W_2$ instead of $0W_1$.
6. The workers who are made worse off are those who are fired because of the minimum-wage law, i.e. the N_2N_1 workers who were willing to work at wage rate $0W_1$ but who are now prevented by legislation from doing so.
7. Many entrants to the labour force serve apprenticeships, i.e. they learn skills and tasks that qualify them for more-responsible and higher-paying jobs. During such apprenticeships, some wage rates are zero, and certainly a substantial number are below the proposed minimum wage; this is because the marginal revenue product of young entrants is very low or zero. Opponents of minimum wages, therefore, fear that such

training schemes would be curtailed and leave groups of young people with few labour-force skills and little opportunity of entering the labour market.

8. The increase in labour costs would have the effect of increasing firms' marginal cost; hence the price of final goods would be increased for profit-maximising firms. At the same, time the reduction in total employment would lead to a reduction in the total output of goods and services, making the total output of goods and services lower than it would have been in the absence of the minimum wage.
9. The profit-maximising rule for a firm of not hiring any factor input whose marginal product is less than the going price dictates that the higher the price the fewer units of that factor that will be employed. Thus the imposition of a minimum wage, while helping some workers, is likely to harm the least productive lowest-paid workers, i.e. those whom it was allegedly designed to help. If society's goal is to make sure every family has at least a minimum standard of living, the imposition of a minimum wage does not guarantee the achievement of such a goal.

Module 10

Review Questions

Multiple Choice Questions

10.1 The correct answer is A. Compared with country Y, country X can produce a unit of wheat and also a unit of cloth with few resources; thus country X has an absolute advantage in the production of both commodities. Country X, however, requires three times more man-days to produce cloth compared with wheat, while country Y requires only twice the number of man-days to produce cloth compared with wheat. Therefore, since country X has a comparative advantage in wheat production and country Y has a comparative advantage in cloth production, country X will export wheat and import cloth.

10.2 The correct answer is B. A higher world standard of living means a higher per capita gross world product. This could occur in a variety of ways – for example, the same world output and a lower population or a lower world output with a disproportionately lower population. Assuming a particular world population, however, a higher output would be required to increase living standards. Higher output comes about with an increase in the quantity and/or quality of the factors of production or by a more efficient allocation of these factors. Although an increase in a protective tariff could benefit one country, a universal increase in protective tariffs would lead to a less efficient allocation of world resources and, consequently, a decrease in world output. Both an increase in worker skills and an increase in capital goods would increase output.

10.3 The correct answer is B. The effect of opening up trade with Russia would be to increase per capita income in both countries because it would expand the opportunities for consumers in each country to consume combinations of goods that are more in accord with their preferences, and it would also permit each country to specialise in the production of those goods for which it has a comparative advantage. Although there would definitely be an increase in US national income, not all workers and employers would share evenly in this increase. It is likely, at least in the short run, that those in import-competing industries would face a decline in income. This decline would not be as great, however, as the increase in incomes in export industries.

10.4 The correct answer is C. The increase in world trade, by permitting more efficient resource allocation, would lead to higher world income and a higher average living standards. It would even be possible, as a result, to make everyone better off if appropriate income-redistribution schemes were adopted. But even though the average would rise, some people could experience a decline in income and a lowering of living standards.

10.5 The correct answer is A. The fact that the exchange value of exports and imports is equal does not mean that those who are trading are not better off. Indeed, if both sides did not benefit from trading, trade would not take place. Each side typically would be willing to give up more than it has to for what it gets. The amount that each country has to give up for what it gets is determined by market forces. The main idea in this question is perhaps most easily expressed by a demand–supply curve figure showing consumer and supplier surpluses.

10.6 The correct answer is C. In the absence of any offsetting changes, appreciation of the pound sterling means that more of any other currency will be exchanged for one pound sterling, i.e. one pound will now buy more euro and, consequently, a holiday in Europe would be less expensive for a holidaymaker from the UK. The pound, by exchanging for more euro, will be able to command more imports, i.e. imported goods will become cheaper.

10.7 The correct answer is B. A Swiss franc still buys the same basket of goods in Switzerland. One pound sterling, however, now exchanges for 2.8 Swiss francs instead of 2.5 Swiss francs, that is, one pound sterling can now buy more Swiss goods than before. Swiss goods have become less expensive for British consumers, just as British goods have become more expensive for Swiss consumers.

10.8 The correct answer is D. Under fixed exchange rates, a trade deficit – unless offset by net capital inflows – will lead to an outflow of foreign currency holdings. In this example, both the deficit on the trade account and the net outflow of capital will have to be matched by an outflow of foreign currency holdings.

10.9 The correct answer is B. If a country had a persistent trade deficit and no net capital inflow, for the balance of payments accounts to balance would require a net outflow of foreign currency holdings. When such holdings of foreign currency were exhausted, the currency would have to depreciate. However, a net trade deficit could be sustained if it were matched by net capital inflows or a combination of net capital inflows and outflows of foreign currency holdings. An appreciation of a currency in the presence of a trade deficit would only be possible over time if net capital inflows exceed the trade deficit.

Case Study 10.1: A Tale of Cigarettes and Whisky

1

1. When two individuals trade with each other, they exchange goods with a relatively low marginal utility for goods with a relatively high marginal utility. This has the effect of making them both better off as a result of exchange. This benefit can be regarded as a surplus from exchange that will somehow be shared between them. However, a third party may become involved whose only role is to bring two individuals together who have the potential to benefit from exchange. Such an individual performs a useful role and can keep some of the surplus for himself; this does not remove the incentive for the other parties to trade since they are still made better off. In this case, Joe was the middle man and he made it possible for the Americans and the islanders to increase their utility by providing the means for exchanging goods with a low marginal utility for goods with a high marginal utility. What neither the Americans nor the islanders realised was that Joe was keeping some of the surplus for himself. Thus instead of a two-way split of the surplus, there was a three-way split. So long as there was a difference in the relative value placed on cigarettes and whisky by the Americans and the islanders, there was scope for achieving benefits from exchange, and for Joe to keep some of the surplus for himself.

   The reason the Americans were so happy with Joe's deal was that they actually gave up two packets of cigarettes for each bottle of whisky, although they would have been willing to give up to three packets. Similarly, the distillery workers gave up only one bottle of whisky for each packet of cigarettes, when they would have been willing to give up three bottles. Both sides felt they had made a favourable deal because neither was forced to give up the maximum it would be prepared to relinquish in order to obtain the other good.
2. Prior to the intervention of Joe Bloggs, the American flyers' marginal utility of malt whisky was relatively high – they did not have any malt whisky – and the distillery workers' marginal utility of American cigarettes was relatively high – they did not have any American cigarettes. For the US flyers, one miniature bottle of whisky was worth three packets of American cigarettes; for the distillery workers, the reverse ratio held – one packet of American cigarettes was worth three miniature bottles of whisky.

   A going rate for cigarettes, in terms of whisky, would quickly be established. Any US flyer offering more cigarettes than the going rate would be swamped by offers of whisky; he would have no incentive to do so because he would be able to obtain whisky at the going rate. Similarly, if the flyer were to offer fewer cigarettes than the going rate, he would not be offered any whisky because the distillery workers could do better at the going rate.

   As American cigarettes became available to the distillery workers, their marginal utility of cigarettes would fall, and their marginal utility of malt whisky would rise. Similarly, for the American flyers their marginal utility of malt whisky would fall and their marginal utility of cigarettes would rise. Without knowing how the relative marginal utilities would change, it is impossible to know what the exchange rate would be. What is known is that, in the absence of changes in tastes, preferences, and other variables, it would fall between the extreme ratios stated, i.e. one packet of cigarettes would command not more than three miniature bottles of whisky and not less than one-third of a miniature bottle.

3. Joe Bloggs's privileged position existed only because the American flyers and the distillery workers could not communicate directly. When fraternisation, and consequently free exchange of offers and bids for cigarettes and whisky, was allowed, Joe's usefulness came to an end and he was unable to obtain his share of the surplus. His standard of living therefore fell.

Module 12

Review Questions

Multiple Choice Questions

12.1 The correct answer is C. The quantity and quality of resources in a nation together set an upper limit on what a nation can produce; together they determine potential output. The extent to which resources are utilised determines what is actually produced, i.e. actual output. Average well-being, which is total output divided by the number of people in a nation, depends therefore on both its stock of resources and how fully they are utilised.

12.2 The correct answer is B. When an economy is operating at any point on its production possibilities frontier, it is producing maximum output from its factors of production. It follows that producing more of any good necessitates withdrawing some resources from the production of other goods. The production possibilities frontier implies nothing about the effects of population growth on productive capacity, nor does it imply that an economy will automatically produce on the frontier. Finally, the production possibilities frontier says nothing about the optimal combination of outputs, which is dependent on people's preferences.

12.3 The correct answer is C. Point L lies beyond the production possibilities frontier CA shown in Figure 12.12. It therefore represents a level of output that is unattainable given the current stock of factors of production and the current state of technical knowledge. A point such as L may be achieved at some period in the future given an increase in the supply of the factors of production (land, labour, capital, and enterprise) and/or an increase in technical knowledge, but is beyond the current productive capacity of the economy. Points within the frontier would indicate unemployment and an inefficient use of resources. A leftward-shifting frontier would indicate declining productive capacity.

12.4 The correct answer is A. At point P in Figure 12.12, more resources are allocated to the production of capital goods than at point N; consequently potential output will be higher in the following year, which would allow the economy to achieve a higher growth rate in actual output. Point P is feasible because it lies on the production frontier.

12.5 The correct answer is B. The shift in the production possibilities frontier from CA to CB shown in Figure 12.12 denotes increased productive potential, which may be due to an increase in the supply of the factors of production and/or improved technical knowledge. In the example illustrated, the economy's productive potential to manufacture capital goods has improved but its productive potential to manufacture consumption goods has remained unchanged. This suggests an improvement in capital goods technology but not in consumption goods technology.

12.6 The correct answer is A. Since point N lies on the production possibilities frontier CA in Figure 12.12, no unemployment exists while CA is the production possibilities frontier. However, when CB becomes the production possibilities frontier, unemployment exists when the economy is operating at any point within the frontier, e.g. at point P. The shift in the production possibilities frontier from CA to CB increases the economy's potential output. The movement from point N to point P increases output of capital goods at the expense of consumption goods; actual output does not increase.

12.7 The correct answer is A. Potential output can increase if any of the following occur: an increase in labour, an increase in capital, or technological progress. However, if in any year there is a reallocation of resources to the production of consumption goods away from capital goods, potential output will be lower in the future than it otherwise would have been.

12.8 The correct answer is A. When acting in an advisory capacity to government, an economist's role is to give professional opinions about the expected outcomes of a policy or policies. Such information is crucial to rational policy-making. For example, a tax cut plus an increase in the money supply may lower the unemployment rate, lower interest rates, and increase investment expenditure; but an advising economist can point out that the tax cut may only have little effect on the unemployment rate but may *increase* interest rates and hence may reduce investment expenditure. Most, if not all, economists have opinions about what society's goals should be but those are value judgements and should not influence the economist in his advisory capacity on the consequences of alternative policies.

12.9 The correct answer is B. The potential GNP of an economy will rise if the supply of any factor of production increases or if there is an advance in technical knowledge. When the proportion of the population available for employment increases, the labour supply increases (for instance an increase in the number of married women looking for work) and potential GNP increases. A decrease in the rate of unemployment would imply an increase in *actual* GNP but not an increase in potential GNP because the supply of labour available for employment has not changed; the only change is in the utilisation of the given labour force. A decrease in output per worker would not increase potential GNP. Although increased government expenditure on civil servants' pay would result in an increase in measured GNP, potential GNP would not have changed. For potential GNP to increase, the productivity of civil servants would need to rise.

12.10 The correct answer is A. Potential output is the level of output that would be achieved at full capacity utilisation, i.e. it is the level of output that would be achieved if all factors of production were fully and efficiently employed. In practice, unemployment cannot be zero because there is always some frictional or structural unemployment arising from the fact that the economy is always undergoing change. Change involves some period of unemployment as workers move from one job to another in the same industry or from one industry to another. At full capacity output, therefore, there will be some frictional and structural unemployment, but there will be no unemployment due to deficient demand. Where demand-deficient unemployment exists, it indicates that the level of demand is not sufficient to provide jobs for all people willing and able to work at going wage rates and conditions of employment, so that full capacity utilisation is not achieved. Potential output can differ from actual output, which is the level of output achieved in the current period.

12.11 The correct answer is B. Actual output is the level of output realised in a given period by currently employed labour and capital, and includes both consumption goods and capital goods. The output that would be realised if all resources were fully employed is the potential output of the economy. Only at full employment is actual output equal to potential output. At all levels of employment below full employment, actual output is less than potential output.

12.12 The correct answer is D. Actual output is the level of output realised in the current period. Potential output is the level of output that would be realised if all factors of production were fully and efficiently employed. A higher level of actual output can be achieved without any increase in potential output, provided that potential output is greater than actual output initially. This occurs whenever there are unemployed factors of production in the economy. Given the level of potential output, the higher is actual output the more fully are the factors of production utilised and, therefore, the lower is the unemployment rate.

12.13 The correct answer is C. Actual output is the level of output realised; potential output is the level of output that would be realised if all resources were fully employed. Classifying all resources under the headings capital and labour, it follows that if potential output exceeds actual output, at least one of these two sets of resources is not fully employed.

12.14 The correct answer is A. When potential and actual output are equal, all resources are fully employed. Full employment of labour, however, is defined to exclude that unemployment occurring as people change jobs, as old industries close and new industries begin, and as industry relocates. Unemployment associated with such changes is known as frictional and structural; it is never zero, and is an integral part of the market system. People willing and able to work at going wage rates and conditions for whom no jobs exist are unemployed because of demand deficiency. When such unemployment appears, actual will be less than potential output.

12.15 The correct answer is C. The unemployment rate is determined by the gap between potential and actual output. Thus, if both potential output and the unemployment rate increase, actual output is growing more slowly than potential output. One possible cause of this growing gap is that certain groups of the population, such as married women, are joining the labour force (and increasing potential output) faster than they are being employed.

12.16 The correct answer is C. Structural unemployment occurs as an economy's infrastructure changes, i.e. as some industries decline and new ones emerge; this is a continuous process and thus structural unemployment is always present in a market economy. The substitution of oil for coal, which increases employment opportunities in the oil industry but decreases those in the coal industry, causes structural unemployment. The building worker seeking employment with another company comes under the classification 'frictional unemployment'. Unemployment caused by lack of demand through a credit squeeze would be 'demand-deficient' unemployment. The bricklayer who retired would not be considered a member of the labour force.

12.17 The correct answer is B. Frictional unemployment arises when an individual is not working because he/she is in the process of changing jobs within the same geographical region, occupation and industry. Such a process normally takes time because of imperfect information on the part of both employers and employees. Structural unemployment can be conceived as frictional unemployment with a long time lag, but job changing in this case requires some change of region and/or occupation, and/or industry. Demand-deficient unemployment arises because of insufficient demand for labour and could be attributable to macroeconomic policy or inflexible market wages.

12.18 The correct answer is A. The presence of high unemployment in an economy means that actual output is below potential output. Average output per worker employed equals total output divided by the number of workers actually employed. At the end of the year, total output was higher but so was the number of workers employed. Insufficient information is given to determine whether average output per worker employed increased, remained constant, or decreased. If the least productive workers are hired last, then average output per worker would fall as full employment was approached.

12.19 The correct answer is A. In economy Z, more guns and more butter were produced in $t + 1$ compared with Year t; therefore economic growth occurred. In economy Y, the output of guns remained constant and the output of butter fell; thus there was a decline in total output. In economy X, the output of guns declined and the output of butter rose; however, since no information is given about the relative worth of guns in relation to butter (e.g. their prices), no conclusions can be drawn about total output because the units (number of guns and tons of butter) are not additive.

12.20 The correct answer is D. Since actual GNP increased from 100 to 150 between time periods t and $t + 2$, and since population remained constant, GNP per capita increased. Thus I is incorrect. Increases in potential GNP come about from a net increase in technological change, the labour force and the capital stock. Thus either or both of the labour force and technological change could have increased between t and $t + 1$ but could have been offset by a reduction in the capital stock, resulting in a constant potential GNP – for example, the labour force could have increased due to a rise in the participation rate of teenagers, but potential output could have remained constant due to declining technology and/or a decrease in the capital stock. Thus neither II nor III must be correct.

12.21 The correct answer is D. The most appropriate policy depends upon the opinion of the people it affects. Whether the government should emphasise lowering the inflation rate or lowering the unemployment rate will depend upon which evil – inflation or unemployment – society considers to be more burdensome. Similarly, if society places a high value on low inflation rates in the future and believes that this can be attained by suffering high unemployment today, III might be chosen as the best short-run policy. Thus the most appropriate policy depends entirely on the subjective opinions of society.

12.22 The correct answer is D. A shift to the left of the Phillips Curve indicates only a more favourable trade-off between inflation and unemployment; it does not indicate where the economy will be on the shifted Phillips Curve. For example, the economy might originally be at point B in Figure 12.13 (zero inflation and an eight per cent unemployment rate); policies might take the economy to a position of say, a one per cent inflation rate and a five per cent unemployment rate on the shifted Phillips Curve; thus the inflation rate will have risen despite the shift on the curve. Similarly, if the economy had started out at point A, policies could have led to an increase in the unemployment rate on the shifted curve. Thus, neither I nor II is necessarily true.

Module 13

Review Questions

Multiple Choice Questions

13.1 The correct answer is A. Given that the production possibilities frontier has not shifted, the economy is capable of producing the same amount of goods and services (national output) in year 2 as in year 1. However, since national output is lower, the economy is no longer on its production possibilities frontier and some resources are not being fully utilised. Thus, unemployment of labour and/or capital exists in the economy. It is possible that the output of both consumption goods and capital goods has decreased, but not necessarily. If the output of consumption goods decreases more than the amount by which the output of capital goods increases, the effect will be a fall in total output, and so II is not necessarily correct.

13.2 The correct answer is B. In the circular flow of income, households serve a dual purpose: from firms they buy goods and services, and to firms they sell the services of their labour, land and capital. The decision to invest in plant and equipment is made by firms, not by households.

13.3 The correct answer is D. Returns to factors of production equal the value of final goods and services produced, i.e. national income equals national output; thus I is false. Total output minus the value of goods and services consumed determines what resources are available for investment; thus II is false.

13.4 The correct answer is D. National income is a flow that can be measured in three ways: by value added in each activity (output), by the incomes accruing to factors of production (income), and by expenditure on all final goods and services (expenditure). The three methods of measuring national income must, by definition, yield the same result, errors of arithmetic apart. Hence the money value of the flow of goods and services (£200 billion) must equal the money value of expenditures on goods and services (£200 billion).

13.5 The correct answer is C. National income is a flow that can be measured in three ways: by value added in each activity (output), by the incomes accruing to factors of production (income), and by expenditure on all final goods and services (expenditure). The three methods yield identical results provided there are no arithmetical errors.

13.6 The correct answer is C. The simple two-sector model of the economy demonstrates the circular flow of income and expenditure, in which the flow of income to households from producers for the use of their resources must be equal to the flow of expenditure from households to producers for goods and services purchased. The simple model does not demonstrate any necessary conflict between households and firms. The distribution of expenditure between goods and services will depend on tastes, and only by chance would lead to equality of expenditure on the two categories. Finally, wages and salaries are only one receipt of households, which also receive other income from land and capital.

13.7 The correct answer is C. The purchase of machine tools and the accumulation of goods produced by a company in inventory are both included as investment expenditure. In purchasing shares, claims to assets are exchanged for cash and such a transaction creates no new goods or services.

13.8 The correct answer is C. Gross National Product (GNP) consists of the flow of all final goods and services produced in a given period. It therefore consists of consumption and investment goods but excludes intermediate goods, which are used only to produce final consumption or investment goods. Bread is a consumption good; hence the value of bread produced is part of GNP. However, flour and wheat are intermediate goods whose values are reflected in the price of bread. To include in the GNP accounts the value of bread, flour and wheat produced would be to count the flour twice and the wheat three times.

13.9 The correct answer is B. National income is the sum of income accruing to factors of production. Thus for a receipt to be included as part of national income, it must be in return for the use of a resource in the production of current output. It could represent a payment on borrowing (i.e. return to capital) but need not. Payment for a good or service is part of Gross National Expenditure. It cannot represent an increase in the recipient's resources as it is a return for services rendered from existing resources; the recipient could add to his stock of resources in the future by, for example, using the income from his labour services to purchase more capital goods.

13.10 The correct answer is C. National income can be measured by summing the expenditure on all final goods and services, by summing the values added by each producer and by summing the incomes earned by each factor of production. Arithmetical errors apart, each of the three methods of calculating national income will yield the same result. Hence summing expenditures on final goods and services is one method of measuring national income, and the result must be equal to the result obtained by summing values added at each stage of production. The income paid out as wages and salaries does not amount to national income because it excludes payment to other factors of production. Disposable income is a receipt, while investment is an expenditure. The sum of money in circulation and bank deposits is always substantially less than expenditures on final goods and services in a normal accounting period, since the same unit of money can be spent many times over within a single accounting period.

13.11 The correct answer is C. By definition, GNI equals GNE. Thus an inequality in practice must be due to errors and omissions. Since GNI is found by adding together all factor incomes, and GNE by adding together all final expenditures, it is inevitable in the real world that discrepancies will occur. A 'residual error' item is included in national income accounts to ensure equality. This residual error is sometimes called the 'statistical discrepancy'.

13.12 The correct answer is A. Since only sales of final goods and services are summed to yield GNP, the further addition of sales of intermediate goods and the sales of assets, i.e. goods produced in past periods, would yield a total greatly in excess of GNP.

13.13 The correct answer is C. Replacement investment is that investment which is necessary to make good any depreciation and maintain the productive capacity of the existing stock of capital goods. Net investment is the difference between gross and replacement investment, i.e. gross investment − replacement investment = net investment. Net investment can be positive, zero or negative depending on whether gross investment exceeds, is equal to or is less than replacement investment. When replacement investment is less than gross investment, then net investment is positive. This means that gross investment is sufficient to make good any depreciation and add to the productive capacity of the existing capital.

13.14 The correct answer is D. National income by definition equals national output. If more consumption goods are produced than people wish to purchase, the excess will be added to stocks, causing investment to increase. This process ensures that actual saving must always equal actual investment. It does not guarantee, however, that planned saving will always equal planned investment.

13.15 The correct answer is C. A 'closed economy' is one in which there is no international sector, i.e. one in which there are no imports and no exports.

13.16 The correct answer is C. If people buy fewer goods and services than businesses expect them to, unsold goods will be added to investment expenditure in the form of unintended stocks. Thus actual investment expenditure will be greater than planned investment expenditure. Since saving is that part of income not spent, and since the amount of goods sold equals the amount of goods bought, actual saving must equal actual investment.

13.17 The correct answer is B. Same explanation as 13.16.

13.18 The correct answer is B. Increased private sales of any item produced in previous time periods will not affect the equilibrium level of national income because such transactions are not included in the national accounts. An increase in the rate of savings will be matched by a decrease in consumer expenditure, thus reducing the equilibrium level of national income.

13.19 The correct answer is C. Saving, in the national income accounting sense, is defined as the part of income that is not consumed, i.e. $Y - C = S$ where Y is income, C is consumption and S is savings. Savings are that part of income that is not currently spent and are thus a leakage from the flow of income. Investment expenditure, which is made possible by savings, is an injection to the flow.

13.20 The correct answer is C. Equilibrium is a situation from which there is no tendency to depart. Macroeconomic equilibrium occurs where national income is not changing and where there are no forces tending to change it.

Module 14

Review Questions

Multiple Choice Questions

14.1 The correct answer is D. The marginal propensity to consume (MPC) is calculated by dividing the change in consumption by the accompanying change in disposable income, i.e. $\Delta C/\Delta Y$. As disposable income increases, the accompanying changes in consumption and the marginal propensity to consume are as follows:

Disposable income ($)	Increase in income ($)	Consumption expenditure ($)	Increase in consumption ($)	MPC ($\Delta C/\Delta Y$)
200	25	205	20	
225	25	225	20	0.8
250	25	245	20	0.8
275	25	265	20	0.8
300	25	285	20	0.8

14.2 The correct answer is C. The average propensity to consume of disposable income is found by dividing consumption expenditure by disposable income.

Disposable income ($)	Consumption ($)	APC (C/Y)
200	205	1.02
225	225	1.00
250	245	0.98
275	265	0.96
300	285	0.95

At the lowest level, the value of APC exceeds 1 but declines to a value of 0.95 as income increases.

14.3 The correct answer is C. Figure A2.15 represents the consumption function described and the savings function derived from the consumption function. Consumption increases at a lower rate than income; as a result the average propensity to consume, C/Y, falls as income rises. Thus, responses A and B are incorrect. For income levels less than y_1, consumption exceeds income, so that dissaving takes place. Thus the average propensity to save is negative for income levels less than y_1, is equal to zero at y_1, and is positive thereafter. Thus response D is incorrect. The marginal propensity to save, S/Y, is given by the slope of the savings function and equals 1 − MPC, since MPC + MPS = 1. Like MPC, MPS is constant and positive for all levels of income.

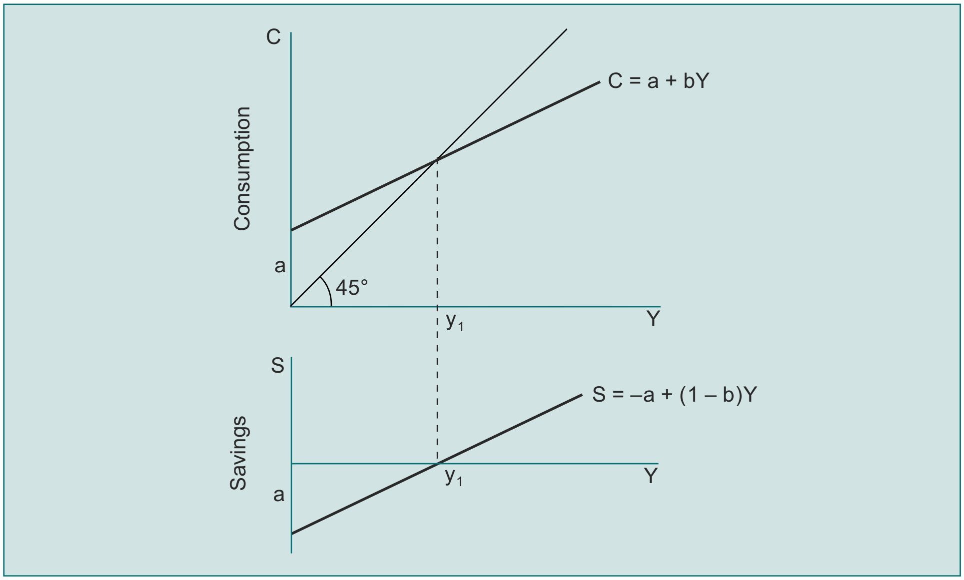


Figure A2.15 Consumption and savings functions

14.4 The correct answer is B. If $C = 15 + 0.75Y_d$, then when $Y_d = 100$

$$C = 15 + (0.75 \times 100) = 90$$

Since S (saving) = $Y_d - C$, S equals 10 when $Y_d = 100$.

If, alternatively, $Y_d = 300$, then

$$C = 15 + (0.75 \times 300) = 240$$

Hence, both I and II are correct. However, as consumption is 15 at zero income and as consumption increases by 75 for every increase in income of 100, it follows that consumption must always be greater than saving. Thus III is incorrect.

14.5 The correct answer is B. In an economy operating below full employment, i.e. with unemployed factors of production, increased investment creates employment and income for previously unemployed factors of production. Factor incomes rise and this, in turn, will result in higher consumption. Higher consumption will create employment and income in a further range of activities, and lead to another round of additional consumption, and so on. The eventual increase in aggregate demand will be some multiple of the initial increase in demand.

14.6 The correct answer is A. If the marginal propensity to consume were zero, then the multiplier would be 1. Since there would be no induced expenditure effect, any decline in investment expenditure would therefore not reduce consumption. The fall in aggregate demand would be restricted to the initial decline in investment demand without any secondary effects.

14.7 The correct answer is D. The consumption function is a schedule showing the level of planned consumption at each possible level of income.

14.8 The correct answer is B. In economic models the values of some variables are assumed to be determined by external events, i.e. external to the model. These variables are known as exogenous variables. Endogenous variables are affected by exogenous variables but the reverse is not true.

14.9 The correct answer is A. The value of C depends upon Y, and therefore C is an endogenous variable. The value of Y depends upon both C and I, and therefore Y is an endogenous variable. The value of I is always 100, independent of the values of C and Y, therefore I is an exogenous variable.

14.10 The correct answer is D. Solving for Y, we have

$$\begin{aligned} \mathrm{Y} &= \mathrm{C} + \mathrm{I} \\ &= 50 + 0.5\mathrm{Y} + 100 \\ \therefore \mathrm{Y} - 0.5\mathrm{Y} &= 150 \\ \therefore 0.5\mathrm{Y} &= 150 \\ \therefore \mathrm{Y} &= 300 \end{aligned}$$

14.11 The correct answer is C. The average propensity to consume (APC) is equal to consumption (C) divided by income (Y). Thus

$$\begin{aligned} \mathrm{APC} &= \frac{\mathrm{C}}{\mathrm{Y}} \\ &= \frac{50+0.5\mathrm{Y}}{\mathrm{Y}} \\ &= \frac{50}{\mathrm{Y}} + 0.5 \end{aligned}$$

14.12 The correct answer is B. Solving for Y given I = 200, we have

$$\begin{aligned} Y &= C + I \\ &= 50 + 0.5Y + 200 \\ \therefore 0.5Y &= 250 \\ \therefore Y &= 500 \end{aligned}$$

From this we can deduce that

$$\begin{aligned} C &= 50 + 0.5Y \\ &= 300 \end{aligned}$$

14.13 The correct answer is A. The value of the investment multiplier is given by the formula

$$\text{Investment Multiplier} = \frac{1}{1-\text{MPC}}$$

and

$$\text{MPC} = \frac{\Delta C}{\Delta Y}$$

Since

$$\begin{aligned} C &= 50 + 0.5Y \\ \frac{\Delta C}{\Delta Y} &= 0.5 \end{aligned}$$

and so the value of the investment multiplier is

$$\frac{1}{1-\text{MPC}} = \frac{1}{1-0.5} = 2$$

This value is independent of the level of I.

14.14 The correct answer is D. Solving for Y with I = 0, we have

$$\begin{aligned} Y &= C + I \\ Y &= C + 0 \\ \therefore Y &= C \end{aligned}$$

Now solving for C:

$$\begin{aligned} C = 50 + 0.5Y &= 50 + 0.5C \\ \therefore 0.5C &= 50 \\ \therefore C &= 100 \end{aligned}$$

14.15 The correct answer is B. The value of MPC is given by the slope of the consumption function. Since C is a straight line in Figure 14.8, the MPC must be constant. The APC is found by dividing C by Y_d. Since C has a positive intercept equal to 2000, C/Y_d must decline as Y increases.

Where C intersects the 45° line, $C = Y_d$ and consequently savings $(Y_d - C) = 0$. To the left of the intersection point $S < 0$ (dissaving); to the right of the intersection $C < Y$ and therefore $S > 0$ (positive saving).

Module 15

Review Questions

Multiple Choice Questions

15.1 The correct answer is A. The value of national output, which equals national income (Y), is determined by the sum of consumption expenditure, investment expenditure, and government expenditure, i.e. $Y \equiv C + I + G$.

15.2 The correct answer is C. Equilibrium income is established when planned injections equal planned withdrawals. If the consumption function is $C = 200 + 0.8(Y - T)$, the savings function is

$$
\begin{aligned}
\text{S} &= -200 + 0.2(\text{Y} - \text{T}) \\
\text{Planned injections} &= \text{I} + \text{G} = 125 + 250 = 375
\end{aligned}
$$

Planned withdrawals are given by S + T, and so

$$
\begin{aligned}
\text{Planned withdrawals} &= -200 + 0.2(\text{Y} - 25) + 25 \\
&= -200 + 0.2\text{Y} - 5 + 25 \\
&= -180 + 0.2\text{Y} \\
\text{In equilibrium, I} + \text{G} &= \text{S} + \text{T, and so} \\
375 &= -180 + 0.2\text{Y} \\
\text{Y} &= 2775
\end{aligned}
$$

15.3 The correct answer is A. The basic proposition of the income–expenditure model is that the level of national income, and therefore the level of employment, is determined by the level of aggregate demand. Hence unemployment arises because of insufficient aggregate demand.

15.4 The correct answer is B. A \$1000 million cut in government expenditure would, through the multiplier effect, decrease GNP by some multiple amount, i.e. $-G \times$ multiplier $= -$GNP. The same multiplier process works for an increase in tax, i.e. $T \times$ multiplier $= -$GNP. If $-G$

and T were of the same magnitude, e.g. $1000 million, the resultant decreases in GNP would differ by $1000 million; this is because the tax multiplier process would only begin with the decrease in consumption that results from the increase in tax, e.g. $750 million if MPC were 0.75.

15.5 The correct answer is D. The accelerator principle assumes a fixed relationship between the capital stock and national output, which equals national income. Thus, a change in national income will produce a change in net investment expenditure.

15.6 The correct answer is D. The upward shift of the investment function shown in Figure 15.16 has come about because of an increase in the return on investment. The rate of interest is merely the 'price' of investment; hence changes in the rate of interest affect the position on the curve, not the curve itself. Increases in business tax rates and wage rates will reduce the return on investment, but labour-saving innovations will increase the return on investment.

15.7 The correct answer is B. The investment curve shows how much firms are willing to invest given the rate of interest. An increase in the rate of interest from 15 per cent to 20 per cent between time periods 1 and 2 would have compensated for the increased willingness of firms to invest in time period 2. Replacement investment is included in the investment curve, as is the timing of investment expenditure. The investment curve relates to the intentions of firms in the country; hence imports are irrelevant.

15.8 The correct answer is A. At a 15 per cent rate of interest, firms would be willing to invest $0i_1$ in time period 1. After the investment curve had shifted in period 2, firms would be willing to invest $0i_3$ at 10 per cent rate of interest.

15.9 The correct answer is D. A shift in the investment expenditure curve (the investment curve) occurs as a result of a change in some factor that affects the return on investment. Technological advance, and/or reduced wage costs and/or a reduction in corporation tax would cause the investment curve to shift upwards as, other things being equal, they would increase the rate of return expected from an investment. An increase in corporation tax would shift the investment curve downwards – that is, other things being equal, an increase in corporation tax would reduce the rate of return expected from an investment.

15.10 The correct answer is C. The simple accelerator model assumes that a fixed capital output ratio holds in all situations. In consequence, a given change in income will produce the same change in investment, whatever the rate of the economy. The simple accelerator model, therefore, views investment as an automatic process uninfluenced by expectations. Further, it assumes that there is never any excess capacity that would allow a firm to meet increased demand without new investment.

15.11 The correct answer is A. The policy increases income through the balanced budget multiplier process. An increase in government expenditure accompanied by the same increase in taxation will raise the level of aggregate demand. In the first round, the increase in government expenditure increases the aggregate demand by the amount of the increase; subsequent multiplier effects are exactly offset by the multiplier effects from increased taxes.

15.12 The correct answer is B. Suppose the budget was balanced, i.e. G = T, and that Y was distributed among C, I and G in the desired proportions, for example 60 per cent, 20 per cent and 20 per cent; also suppose that the economy was in equilibrium but operating at less than full employment. The achievement of full employment requires fiscal policy, i.e. an increase in G, a cut in taxes, or both. Once such a policy has been enacted, there is no guarantee that the new level of G will equal the new level of taxes, i.e. that the budget will still be balanced, and no guarantee that the original distribution of C, I and G will remain; this is because different fiscal actions will affect each of C, I and G by varying amounts. It is therefore incorrect that the distribution among C, I and G cannot be affected by policy. It is possible to have a balanced budget at full employment, but there is no guarantee that this will occur. The economist's advice is therefore based on the fact that it is not possible to control all five variables, Y, T, C, I and G, simultaneously. The fact that the three objectives may have different levels of importance is no reason for dropping any one of them.

15.13 The correct answer is B. If households were to save more of their income, then in the short run they would consume less of their income. Businesses would find inventories (stocks) accumulating because their production plans would be geared to the former higher levels of consumption expenditure. They would cut back on production to eliminate the unintended inventory accumulation, thereby reducing national income and increasing unemployment.

15.14 The correct answer is D. An in-built stabiliser is an element of demand that changes automatically in an anticyclical fashion, thus helping to stabilise national income. An in-built stabiliser could be any form of taxation that increases automatically when income rises and falls automatically when income falls. A progressive tax has this property, since high marginal tax rates result in a higher proportion of income being collected as income increases, and vice versa.

15.15 The correct answer is C. An in-built stabiliser ensures that when economic activity increases, proportionately more spending power is leaked out into taxation, and when economic activity declines, proportionately less tax is leaked out. As a result, in-built stabilisers tend to reduce fluctuations in economic activity. Thus in-built stabilisers do not ensure that tax collected and government expenditure are the same, nor that the same amount of tax is collected each year.

15.16 The correct answer is D. Imports are the demand for goods and services produced in other economies and are a withdrawal from the circular flow. Exports are the demand by foreigners for goods and services produced within the domestic economy and are an injection into the circular flow. Since $Y = C + I + G + X - Z$, Y can continue to grow when $Z > X$ as long as the growth in $C + I + G$ exceeds the growth in the difference between Z and X.

15.17 The correct answer is B. The value of the multiplier depends on the leakages from the circular flow of income. The marginal propensity to import and the marginal tax rate are both leakages. The value of government expenditure does not affect the value of the multiplier.

15.18 The correct answer is A. National income = C + I + G + X − Z. Thus an increase in Z will make national income lower than it otherwise would have been; a decrease in X would have the same effect.

15.19 The correct answer is A. Calculation of the multiplier must take into account all leakages, including the marginal propensity to import, therefore I is correct. The marginal propensity to save is 1 − MPC, hence it was implicitly used to calculate the multiplier; therefore II is incorrect. Equality of exports and imports does not set a constraint on the level of national income, therefore III is incorrect.

15.20 The correct answer is C. The level of GNP can fall for a number of reasons – for example, a reduction in investment demand, a reduction in export demand, or a reduction in consumer demand. Hence an increase in exports, an increase in consumption, or an increase in investment demand, could not account for a reduction in imports. At the lower level of GNP, customers could buy less imports despite the fact that the marginal propensity to import had increased.

Module 16

Review Questions

Multiple Choice Questions

16.1 The correct answer is C. Whenever actual output is less than potential output, appropriate fiscal and monetary policies will lead to an increase in actual output. The size of the increase will depend upon the multiplier, which increases with the marginal propensity to consume, i.e. it does not have to equal one to prevent positive multiplier effects. Expansion requires unemployed resources (including labour), or technological change; it does not require labour to be a fixed percentage of the population. The ratio between capital and output, although affecting the rate of expansion, will not stop the expansion process. Expansionary aggregate demand policies will only be ineffectual when the economy has reached full employment, i.e. when productive capacity is fully utilised.

16.2 The correct answer is D. The upper limit to the level of national output in the short run is potential output, which is determined by the supply of factors of production. This fact is independent of the prices of factors of production. Aggregate demand can increase in the short run and variable factors of production can be substituted for each other in the short run.

16.3 The correct answer is C. The upper limit to the level of national output is determined by the economy's potential output. More labour added to a given capital stock will always increase an economy's potential output, as will more capital added to a given labour force. Therefore, neither the capital stock on its own, nor the labour force on its own, will determine potential output.

16.4 The correct answer is C. Potential output is defined as the output achievable when all resources are fully employed. In practice, however, resources can be over-fully employed for short periods, e.g. 50 hours per week for some members of the labour force. Because over-full employment cannot be sustained in the long run, potential output is defined with reference to a given practical rate of capacity utilisation of capital. There is no necessary connection between the level of employment and potential GNP. Full employment is defined as actual GNP equal to, not greater than, potential GNP.

16.5 The correct answer is B. If the economy had been at full employment initially, it would be impossible for the growth of actual output to exceed the growth of potential output by 1 per cent per annum for five years. Hence, the observation that the growth of actual output has exceeded the growth of potential output over that period must imply that, initially, actual output was below potential output. Further, it must imply that the gap between actual output and potential output is falling – that is, the deflationary gap is decreasing. Currently the economy need not be at full employment, but it could now be.

16.6 The correct answer is D. An increase in aggregate demand would decrease a given deflationary gap between potential and actual output. An increase in investment expenditure and a balanced increase in government expenditure and tax revenue would increase actual output, the latter policy operating by the balanced budget multiplier. An increase in imports would decrease aggregate demand because imports are a leakage from the circular flow of income.

16.7 The correct answer is D. Decreases in income tax would stimulate the economy; decreases in government expenditure would depress the economy. Had the decrease in government expenditure equalled the tax revenue increase, the net impact would have been to decrease aggregate demand by the balanced budget multiplier. It could then be inferred that the concern of the government was the inflationary gap. However, without knowing the magnitudes of the changes, it is not possible to predict the net effect and thus not possible to deduce the government's concern.

16.8 The correct answer is B. Decreasing a deflationary gap requires an expansionary fiscal and/or monetary policy. Increasing taxes, decreasing unemployment compensation, and decreasing government expenditure would all increase the deflationary gap. An increase in subsidies to exporting industries would stimulate exports, increase aggregate demand, and reduce the deflationary gap.

16.9 The correct answer is C. Decreasing an inflationary gap requires a deflationary fiscal and/or monetary policy. Reducing income taxes, reducing imports, and raising unemployment compensation would all increase aggregate demand and increase the inflationary gap. A reduction in government expenditure would decrease aggregate demand and reduce the inflationary gap.

16.10 The correct answer is C. An inflationary gap exists when the level of demand at full-employment income (Y_f in Figure 16.6) is greater than the level of demand necessary to maintain full-employment income. The gap is the excess of demand over that necessary to maintain full-employment income. In the diagram, this is given by the vertical distance between the 45° line and D, i.e. UV.

16.11 The correct answer is D. Y_f represents full employment income. While national income for short periods can exceed Y_f, for example by working 50-hour weeks, such a level of demand creates inflationary pressures, causing money national income to increase without a matching increase in real output.

16.12 The correct answer is B. Higher money incomes will be associated with higher real living standards only if rising money incomes can continuously buy more goods and services. This is possible only if the prices of goods and services are not rising as fast as money incomes. Increases in money incomes are normally associated with increases in spending. Real GNP declines when aggregate demand declines, not when money incomes increase.

16.13 The correct answer is C. Money GNP rose in Year $t + 1$ while real GNP fell; hence prices increased in Year $t + 1$. The growth in money GNP was greater than the growth in real GNP in Year $t + 2$, indicating rising prices. Real GNP and money GNP are equal in the base year chosen and can subsequently be equal if increases in output are compensated by reductions in prices, and vice versa. Money GNP can be a misleading measure of economic performance because price changes and real output changes cannot be separated.

16.14 The correct answer is C. If actual output is equal to potential output, an increase in government expenditure will cause prices to rise, but real national output will not rise because no idle resources will be available. If all of the additional income generated by the increased government expenditure were spent on imports, no multiplier process would result.

16.15 The correct answer is A. An increase in government expenditure of \$5 million will result in a matching increase in income, which in turn will lead to an increase in consumption, the size of the increase being dependent on leakages from the circular flow. This multiplier process will continue to lead to an increase in output that will be some multiple of the original increase in government expenditure and will decrease the deflationary gap. A decrease in tax by \$5 million will not be matched by an increase in national income of \$5 million because it is a transfer of income from the government to households. However, households will spend a proportion of the \$5 million on consumption expenditures, which will start off the multiplier process similarly to that outlined above with the first round of expenditure and income creation missing.

Module 17

Review Questions

Multiple Choice Questions

17.1 The correct answer is D. Because money is a medium of exchange, it passes from hand to hand in financing purchases or settling debt. As a medium of exchange, money is used in a number of transactions within a year; hence the money supply is much smaller than disposable annual income or total annual spending. Although some money is used for transactions purposes, money is also held for precautionary and speculative purposes.

17.2 The correct answer is B. The value of money is determined by its buying power. When price inflation occurs, the same amount of money buys fewer goods and services and so the value of money in terms of its buying power declines. The value of money is unrelated to the cost of its production. In modern times the money supply is not backed by gold or silver. The interest rate is determined by the demand for and supply of money in financial markets.

17.3 The correct answer is B. If the desired ratio of cash reserves to deposits (known as the cash ratio) is 1:5, each \$1 of cash reserves can support \$5 of deposit liabilities. Therefore, if cash reserves are \$35m, this will support \$175m in demand deposits. As demand deposits are initially \$100 million the banks would be able to expand the money supply by \$75m without destroying the desired cash ratio.

17.4 The correct answer is D. The banks have excess cash reserves and therefore wish to increase lending provided that they can find creditworthy customers. If demand deposits, and therefore the money supply, do not increase, this must mean either that the banks cannot find creditworthy customers or that such customers do not wish further credit.

17.5 The correct answer is A. Using monetary policy to curb a deflationary gap requires an expansion of the money supply. Increasing commercial bank reserve requirements would decrease the money supply; raising interest rates would decrease consumption and investment expenditures. Thus options I and III would not reduce, but would increase, the deflationary gap. The government, by purchasing its own securities (option II), increases the supply of money; this will lead to lower interest rates and hence higher aggregate demand.

17.6 The correct answer is D. Two factors that would diminish the lending of money by commercial banks would be:

(a) a request to hold larger amounts of any deposit as a cash reserve; and
(b) withdrawals of deposits by customers, since this would force banks to call in loans.

The raising of interest rates by the Bank of England will increase the cost to commercial banks of borrowing from the Bank of England but will not diminish their ability to lend money; indeed, a higher interest rate may increase the flow of deposits to commercial banks.

The releasing of Special Deposits by the Bank of England and the reduction of the reserve requirement will increase the ability of commercial banks to lend money.

17.7 The correct answer is C. An amount of \$100 deposited in a bank with a cash ratio of 10 per cent means that the bank must retain \$10 but can loan out the remaining \$90. If that \$90 is in turn deposited in a bank, \$9 (i.e. 10 per cent) must be retained, and \$81 can be loaned out. This process will continue, the ultimate increase in the money supply being \$100 + \$90 + \$81 + … = \$1000. The mathematical formula for the credit creation multiplier is 1/cash ratio = 1/0.1 (0.1 = 10 per cent). The total expansion equals the initial deposit × multiplier = \$100 × 10 = \$1000.

Module 18

Review Questions

Multiple Choice Questions

18.1 The correct answer is C. If V is constant and M is increasing at 4 per cent per annum, MV will be increasing at 4 per cent per annum. Given the equality MV = PY, PY must also be increasing at 4 per cent per annum. If Y is increasing at 2 per cent per annum then P must be increasing at approximately 2 per cent per annum also.

18.2 The correct answer is C. Since Y is held constant, real national output cannot increase or decrease. In order to maintain the equality MV = PY, P, the price level, must increase as M increases.

18.3 The correct answer is C. If the economy has unemployed resources, real output (Y) may change. The level of prices can change in the short run. The efficiency with which money is used can change, hence V may change. The money supply variables can change at the same time but it is not possible to predict the value of MV or PY despite the fact that MV must equal PY.

18.4 The correct answer is D. In deciding whether to hold cash, i.e. preference for liquidity, or interest-bearing assets, individuals will take into account:

(a) the return they would obtain from interest-bearing assets; and
(b) the prospective capital gain or loss from holding interest-bearing assets given a fall or rise in interest rate.

If investors believe that interest rates are going to rise, then this is equivalent to expecting the price of assets to fall; hence they will be reluctant to acquire interest-bearing assets and will hold most of their assets in cash. The reverse applies if individuals expect the rate of interest to fall.

18.5 The correct answer is A. The price of long-term interest-bearing assets moves inversely to the rate of interest. If individuals expect that the rate of interest is likely to fall, then this is equivalent to expecting the price of long-term interest-bearing assets to rise. In these circumstances, they are likely to buy long-term interest-bearing assets. The price level will not affect this decision.

18.6 The correct answer is B. Individuals hold some cash in order to finance day-to-day transactions. The convenience derived from these cash balances outweighs the cost of holding cash, that cost being interest forgone. If individuals believe that the rate of interest is likely to rise in the future, they are likely to retain their assets in cash rather than bonds, as the price of bonds will fall if the rate of interest rises. Option III is incorrect as savings may be held in the form of assets other than cash.

18.7 The correct answer is A. Interest rates are determined by the demand for and the supply of money. In recessions, aggregate demand decreases and, as incomes fall, the demand for money for transactions purposes falls. As demand falls and supply remains unchanged, the rate of interest falls. The money supply never automatically increases; the velocity of circulation does not follow any particular pattern; savings do not affect the rate of interest directly.

18.8 The correct answer is C. While a decrease in personal income tax would stimulate aggregate demand, an increase in the rate of interest would be deflationary in that it would discourage both investment and consumption expenditures. Thus the proposals would work in opposite directions and tend to cancel each other out.

18.9 The correct answer is B. The reduction of taxes and the increase in government expenditure would both increase aggregate demand and, through the multiplier process, increase GNP and decrease unemployment. However, both would also increase the demand for money, and in the absence of an increase in the money supply would raise interest rates. Businesses would be faced with two opposing forces: increased aggregate demand stimulating higher investment expenditure, and increased interest rates leading to less investment opportunities. Since private investment expenditure has fallen, the latter forces were obviously stronger. The decrease in private investment expenditure would, through the multiplier process, decrease GNP and employment, thus offsetting the increase due to the tax reduction and increased government expenditure.

18.10 The correct answer is D. This is a difficult question. If the economy were very depressed, an increase in the money supply would not necessarily stimulate investment and consumption expenditure because of poor business expectations. An increase in government expenditure and/or a decrease in taxes with no increases in the money supply would increase the demand for money and increase interest rates; this in turn could cause investment expenditure to decrease, offsetting the positive effects and leaving unemployment unchanged. While the apparently correct policy is for fiscal and monetary expansion, if the increase in government expenditure and/or tax reduction caused interest rates to rise in spite of an increase in the money supply (i.e. the money supply increase was not large enough), the resultant increase in interest rates could cause investment and/or consumption expenditure to fall. Thus there is no guarantee of success.

Module 19

Review Questions

Multiple Choice Questions

19.1 The correct answer is B. The liquidity trap is a floor below which the interest rate does not fall. Holders of speculative balances expect the interest rate to rise and are unwilling to commit such balances to investment purposes. Thus any increase in the money supply is held in idle balances.

19.2 The correct answer is A. Option I is the definition of the LM curve. The LM curve is concerned only with equilibrium in the money market and has no relationship whatsoever with factors affecting the IS curve. Thus option II is wrong.

19.3 The correct answer is D. If the marginal propensity to consume were large, a small increase in autonomous expenditure would bring about a relatively large increase in national output, reflected in an elastic IS curve. Thus option I is wrong. If the marginal efficiency of investment curve were elastic, a change in the rate of interest would bring about a relatively large change in the volume of investment. Thus option II is wrong.

19.4 The correct answer is C. The increase in the money supply would shift the LM curve to the right. The cut in taxes would shift the IS curve to the right. Thus the new intersection point of the IS and LM curves would be to the right of the original intersection point. Thus the equilibrium level of Y would be higher than the original. However, the impact on R would be determined by the relative shifts: R could increase, remain constant or decrease.

19.5 The correct answer is D. The interest rate consistent with full employment occurs when the IS curve intersects the dotted Y_F line at point C in Figure 19.15. However, C is below the interest rate R_0, which represents the liquidity trap. Thus increase in the money supply will not produce a low enough interest rate to achieve full employment.

19.6 The correct answer is B. To reach point D in Figure 19.15 and to achieve equilibrium in both the real goods and monetary sectors, both the IS curve and LM curve would have to shift to the right. Thus answers C and D are wrong. It can be reached, however, but since point D lies to the right of Y_F, point D represents an inflationary gap situation.

19.7 The correct answer is B. To achieve all stated goals, the IS and LM curves must intersect at point E in Figure 19.15. This requires a decrease in the money supply to shift the LM curve to the left, accompanied by expansionary fiscal policy.

19.8 The correct answer is A. The increase in government expenditure in the absence of offsetting factors would increase the demand for money and, in the absence of an increase in the money supply, raise interest rates, thereby producing a negative money multiplier on investment and consumption expenditure. Thus option I is correct. If the gap between potential and actual output were not large enough, this could prevent the full value of the multiplier being realised. Thus, any increase in potential output would help rather than hinder the full multiplier effect occurring. Thus option II is incorrect.

19.9 The correct answer is D. If the IS curve is vertical and/or if the economy is operating at full employment, an increase in the money supply will not result in an increase in national income. Thus option I is incorrect. If the LM curve is horizontal, i.e. the economy is in the liquidity trap, a decrease in taxes shifting the IS curve to the right may leave the rate of interest unchanged. Thus option II is incorrect.

19.10 The correct answer is B. Since all the changes are in the real goods sector, the LM curve is not affected. Thus responses A and D are wrong. The increase in taxes would shift the IS curve to the left, and the increase in government expenditure would shift the IS curve to the right, as would the trade surplus. However, since we do not know the magnitude of the relative changes, we cannot tell where the IS curve will finish up.

Module 20

Review Questions

Multiple Choice Questions

20.1 The correct answer is B. Since real GNP increased by 4 per cent between Year *t* and Year *t* +1 and since the labour force only increased by 3 per cent, average output per worker must have increased. Since money GNP rose by 6 per cent and real output by only 4 per cent, prices must have risen by about 2 per cent. Since the labour force grew by 3 per cent but potential output remained constant, the capital stock must have decreased.

20.2 The correct answer is B. Other things being equal, the lower the price of a good or resource, the larger will be the quantity demanded. Thus, if prices were completely flexible and responded quickly to situations of excess demand and excess supply, all markets – including the labour market – would clear. Therefore full employment would always exist. In the real world, many markets are 'sticky', i.e. respond slowly to excess demand/supply. By definition, national income equals national output; the rate of productivity increase determines potential, not actual, output; to date no nation has sufficient resources to satisfy all of its citizens' wants.

20.3 The correct answer is A. Increasing transfer payments and increasing government expenditure on goods and services, while both stimulatory, would increase the relative size of the public sector. Leaving the tax structure and government expenditure at their existing levels would not help cure a recession. Only a reduction in the level of taxation would meet both goals, i.e. increasing GNP but reducing the fractions G/GNP and T/GNP.

20.4 The correct answer is C. Longer repayment periods would increase the demand for consumer durables and consequently increase inflationary pressures. Thus option I is incorrect. Increasing down-payment requirements would reduce the demand for consumer durables. Selling securities by the central bank would reduce the cash reserves of the commercial banks and would therefore result in a contraction of the money supply. Thus both options II and III would reduce inflationary pressures.

20.5 The correct answer is C. An increase in potential GNP requires an increase in the labour force, the capital stock, or technological progress. The lowering of interest rates could achieve an increase in investment expenditure and an increase in potential GNP, but this is a long-run strategy because, in the short run, an increase in investment expenditure adds only negligibly to the existing capital stock. An increase in investment expenditure leads to an increase in aggregate demand; given the economy is at over-full employment already, this will lead to even greater inflationary pressures.

20.6 The correct answer is B. Inflationary pressures would be increased by any factor that increases aggregate demand. An increase in investment expenditure, an increase in exports, or an increase in government expenditure, would each increase aggregate demand and consequently the inflationary gap. However, an increase in savings results in a decrease in consumption expenditures, which would reduce aggregate demand and the inflationary gap.

20.7 The correct answer is D. Holding the money supply constant, an increase in investment expenditure will increase the demand for money and, given no offsetting decrease in the demand for money by households or government, increase interest rates. The increase in investment expenditure will stimulate the economy, and the increase in interest rates will reduce aggregate demand. Since opposing forces will exist, it is not possible to predict whether consumption expenditure will increase, remain constant, or decrease. Likewise it is not possible to predict what will happen to imports.

20.8 The correct answer is B. Loans are normally made at some negotiated interest rate. When borrowed money is returned, its value will be inversely related to the inflation rate during the loan period; thus unanticipated inflation will benefit borrowers at the expense of lenders, since the negotiated interest rate will not take the unanticipated inflation into account. Thus income will be redistributed. Since the economy is at full employment, there will be no increase in unemployment but the average price level will have risen. No information is given regarding potential output.

20.9 The correct answer is A. As increasing tax rates and reducing government expenditure must lower aggregate demand, option II is incorrect. Option I could account for the failure of prices to respond immediately to the deflationary policies adopted, as it may take time to modify expectations and behaviour.

20.10 The correct answer is C. To reduce an inflationary gap requires a deflationary policy. An increase in the money supply is expansionary. A decrease in the money supply would cause interest rates to rise, investment and consumption expenditures to fall, and thus reduce aggregate demand and inflationary pressures.

20.11 The correct answer is B. To assist the government in decreasing the gap between actual and potential output, unanticipated changes would have to exert some positive influence on aggregate demand. An increase in foreign tariffs would reduce the demand for exports, hence reducing aggregate demand. An increase in the marginal propensity to consume would increase consumption expenditures, hence increasing aggregate demand.

20.12 The correct answer is C. Since expenditure on imports is a leakage from the circular flow of income, the effect of the tariffs and quotas will be to increase aggregate demand and hence cause Y to exceed Q. An increase in government expenditure would add further to aggregate demand, as would a decrease in taxes and a subsidy to exporters, thus increasing the inflationary gap. A decrease in the money supply, given the increase in the transactions demand for money resulting from the higher level of aggregate demand caused by the reduction in imports, would lead to higher interest rates and, consequently, lower investment expenditure by firms and lower consumption expenditure by households.

Index