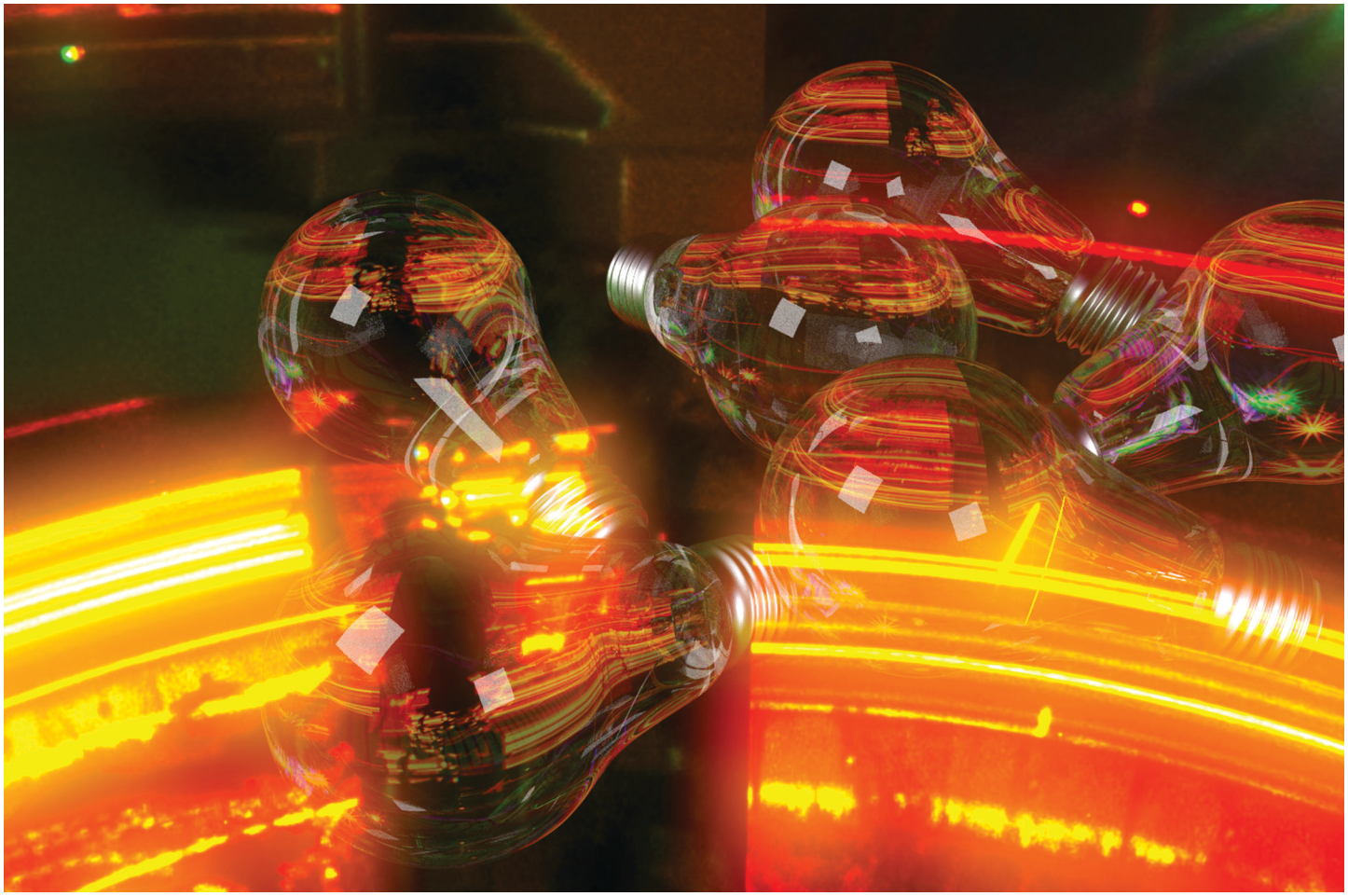
OpenLearn

The Open University

Innovation, markets and industrial change



About this free course

This free course provides a sample of Level 1 study in Sociology www.open.ac.uk/courses/find/social-sciences.

This version of the content may include video, images and interactive content that may not be optimised for your device.

You can experience this free course as it was originally designed on OpenLearn, the home of free learning from The Open University – www.open.edu/openlearn/people-politics-law/politics-policy-people/sociology/innovation-markets-and-industrial-change/content-section-0

There you'll also be able to track your progress via your activity record, which you can use to demonstrate your learning.

Head of Intellectual Property, The Open University
The Open University

Contents

Introduction

This course looks at the role of innovation in the development of industries and considers how production costs change as sales increase and as new technology is introduced into the production process. It looks at the relation between consumer demand for a good and that good's price, and at how the relation between output and production costs in different markets can dramatically affect industry structure. In describing these issues, the course introduces the range of activities that constitutes economics: formulating theories, modelling, debate and persuasion, analysis of data, understanding the behaviour of economic institutions such as companies and households, and analysing economic processes. It also seeks to show how all of these economic techniques can be used to build up a rich understanding of innovation and economic change.

This OpenLearn course provides a sample of Level 1 study in Sociology.

Learning Outcomes

After studying this course, you should be able to:

- appreciate the importance of technological change, costs of production and consumer preferences to the changing organisation of production
- understand the relation between the quantity demanded of a good and its price as represented by the demand curve
- understand economic models of the relation between firms' costs and output
- analyse the role of technology and costs in influencing industry structure over the life cycle.

1 Technological change, demand and costs

The new economy

Over the past 40 years global computing power has increased a billionfold. Number-crunching tasks that once took a week can now be done in seconds. Today a Ford Taurus car contains more computing power than the multimillion-dollar mainframe computers used in the Apollo space programme. Cheaper processing allows computers to be used for more and more purposes. In 1985, it cost Ford $60,000 each time it crashed a car into a wall to find out what would happen in an accident. Now a collision can be simulated by computer for around $100. BP Amoco uses 3D seismic-exploration technology to prospect for oil, cutting the cost of finding oil from nearly $10 a barrel in 1991 to only $1 today …

Thanks to rapidly falling prices, computers and the Internet are being adopted more quickly than previous general-purpose technologies, such as steam and electricity. It took more than a century after its invention before steam became the dominant source of power in Britain. Electricity achieved a 50% share of the power used by America's manufacturing industry 90 years after the discovery of electromagnetic induction, and 40 years after the first power station was built. By contrast, half of all Americans already use a personal computer, 50 years after the invention of computers and only 30 years after the microprocessor was invented. The Internet is approaching 50% penetration in America 30 years after it was invented and only seven years since it was launched commercially in 1993.

(The Economist, 23 September 2000, pp. 5, 10)

That is how *The Economist* discussed the issue of information technology as a general purpose technology.

Question

How would you summarise the argument about prices, costs and information technology put forward in *The Economist*?

It seems to us that the argument is that improvements in information technology caused a decline in the costs that firms face in producing goods. Falling prices have in turn encouraged a rapid take-up by consumers of products embodying information technology. Economists place costs on the supply side of markets, where firms produce and sell goods such as cars, oil, personal computers and Internet access. Take-up by consumers is on the demand side of the market. The quotation suggests that new information technology *causes* a decline in costs and hence in prices that enables large numbers of consumers to buy safer cars, personal computers and so on. Technological change, a supply-side phenomenon, is seen as a prime cause of economic change.

In D202_2, the author looked at the impact of newly introduced technologies in the early phase of the US auto industry and the PC industry, highlighting the similarities between what we are observing today in the IT-based industries and what we observed 100 years ago in an industry we now consider mature. As we write, technological change continues to be very rapid, and two particular technological developments will be used in this chapter to illustrate our discussion of the forces shaping the organisation of industrial production. The dilemma faced by car manufacturers in adjusting to the hydrogen economy' of the future will be discussed in Section 4. We also discuss digital technology, which has revolutionised sound reproduction as well as the capture and transmission of visual images in DVD, digital television and digital cameras. Car manufacturers are also exploiting its potential. For example, the Citroën C5 'bristles with the latest digital technology to improve your motoring experience' (see Figure 1). An advertisement for the C5 alludes to the origins of digital technology, showing instructions as sequences of 'switches' in a computer as being 'on' or 'off', or set to '1' or '0'.

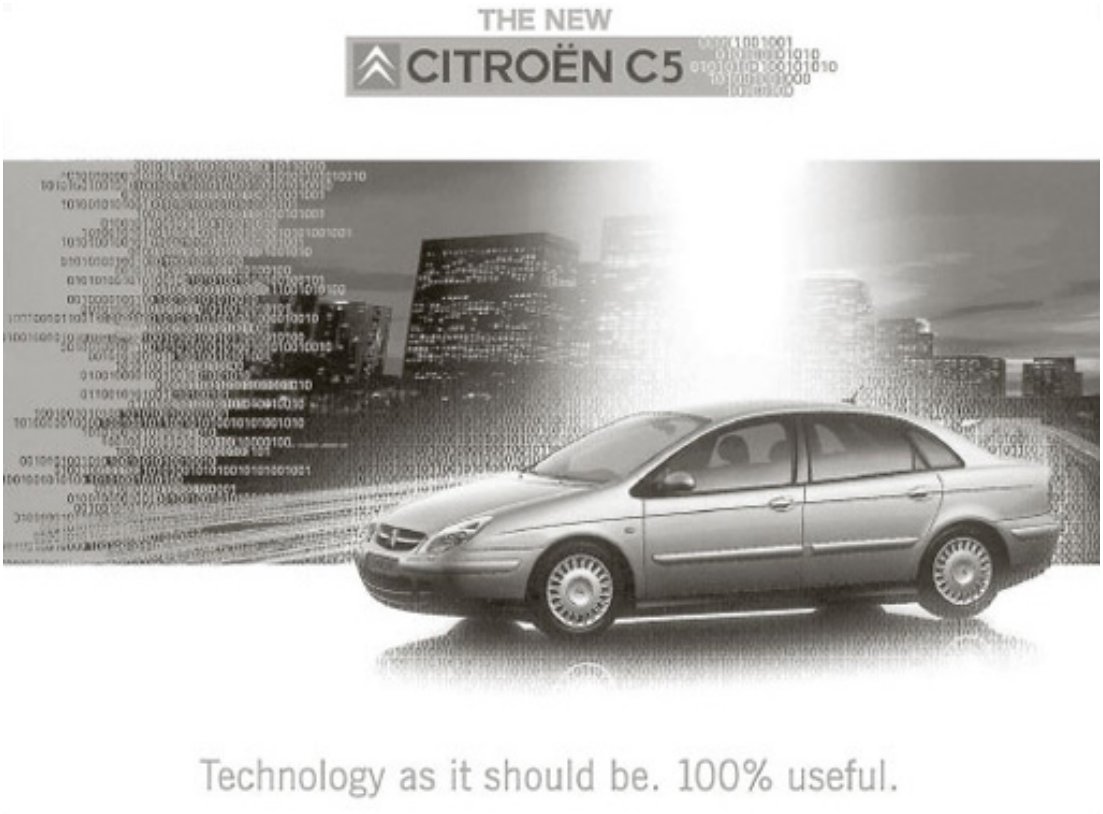


Figure 1: An advertisement for the Citroën C5

Copyright © Citroën UK Ltd

In this course we will introduce the use of economic models of markets, firms and industries to examine the relationships among consumer demand, technological change and costs. Section 2 develops the idea expressed in *The Economist* extract that falling prices were responsible for the rapid take-up of new products by consumers. This is an example of a widely observed relation between the quantity demanded of a good by consumers and the price of the good: the lower the price, the greater the quantity demanded. However, there are other influences on market demand. For example, consumers are likely to buy more goods if their incomes increase. So Section 2 explains

how economists analyse the interaction between price and other influences. This analysis is the first part of the theory of consumer demand.

The rest of the course focuses mainly on the supply side of the market, exploring the role of costs and technological change in the organisation of production. The objective is to understand the process by which a firm – initially one among many similar firms jostling for position – emerges ahead of the pack to achieve an advantage over its competitors, as Ford did in the US automobile industry. What was so special about Ford? More generally, how can we account for the change in structure that so many industries seem to undergo? Why do most of the many small firms so common in the early years of new industries disappear to leave an established industry dominated by a few large firms? Why does the heterogeneity or extreme variety of those small firms in new industries give way to a much greater degree of similarity, indeed standardisation, among the few survivors?

In exploring these questions, economists use models of the relation between the output, technology and costs of firms. In Section 3, we define technology in an economic model of a firm, and use it to explore the link between technology and costs. We begin from a model in which firms take technology as given and their ability to change their costs is severely constrained, and then build in progressively more decision-making flexibility over the adoption and use of technology.

Section 4 links this analysis of technology and costs to the model of the industry 'life cycle' introduced in DD202_2. The life-cycle model represents an industry as if it were a biological organism going through the stages of birth, growth, maturity and decline, and is used to consider the interaction of demand and technology in shaping industrial structure. This section further extends the analysis of firms, technology and costs to include models where firms' freedom to learn and create is one driver of technological change itself. This helps us to understand how a particular firm can become the 'leader of the pack' through innovation and how it can then gain an advantage over competitors by reducing its costs through large-scale production.

2 Market demand

2.1 Industry and markets: what do we mean?

Case study: Digital outsells film

Sales of digital cameras have overtaken traditional 35 mm cameras for the first time. According to monthly figures collated by national electric and photo retailer Dixons, digital camera sales outstripped 35 mm cameras during the month of April. 'This is a sea change in consumer photography,' said Dixons marketing director Ian Ditcham. 'As a leading photographic retailer, Dixons is a clear barometer of consumer trends,' he said. The main reasons for the popularity of digital cameras are falling prices, the growth of home PC and internet usage and the instant delivery of images without the need for processing.

(Adapted from Outdoor Photography, August 2001, no. 15, p. 4)

Questions

1. What do you think are the main similarities between the arguments in this case study and the second paragraph of 'The new economy' quotation discussed in Section 1?
2. Figure 2 illustrates a number of different makes of digital camera. What other products do you normally associate with the manufacturers of these cameras?

Figure 2: A selection of digital cameras from the buyer's guide of *What Camera?* Magazine

Copyright © IPC Media Ltd

The quotation and the case study describe the take-up by consumers of new products embodying innovations in information technology. Both link the popularity of new technology among consumers with falling prices and draw attention to the rapidity of consumer take-up. There is something new for Dixons and other retailers to sell and for consumers to buy. A new product, the digital camera, has created a new market. Figure 2 shows that this market is supplied by a number of manufacturers normally associated with the production of a range of other consumer goods. Nikon and Olympus make traditional cameras. Epson is a well-known brand of personal computers and printers. Sony is probably best known for the Walkman, the first personal audio-cassette player. Casio puts its name on watches and calculators. Fuji is particularly interesting as a manufacturer not only of cameras but also of film, the medium under challenge from digital.

It seems natural to us to say that these firms from different *industries*, including computers, electronics and optical equipment, have together created a new *market*, that is, a market for a new product, the digital camera. (An industry is a group of firms producing a broadly related range of goods using similar technologies. A market is constituted by the buying and selling of goods or services.) In other words, industries can be thought of as being about the production of a range of broadly related goods, and markets are about the sale and purchase of a more narrowly defined set of goods. For example, the optical equipment industry produces cameras, photocopiers, microscopes, telescopes and so on in order to supply different markets. There are consumer markets for disposable cameras, entry-level compact cameras or cameras for the serious amateur and so on. There are industrial markets for cameras for professionals or microscopes for specialist use in medical and scientific research. So markets can be identified in terms of their consumers and the purposes for which those consumers are buying goods, as well as in the more familiar terms of the goods themselves. Phrases such as ‘the market for medium format cameras’ are common. But a medium format camera might appeal to the serious, and affluent, amateur as well as the professional. It is not easy to say where one market ends and another begins.

Precisely where we draw the line between ‘industry’ and ‘market’ depends on which aspect of economic activity we want to analyse. Producing and selling are both stages in a single complex process of making profits from the *supply* of goods by turning inputs into outputs that consumers are willing and able to buy. It is helpful to use the term ‘industry’ if we want to direct attention towards suppliers (as producers) and the technologies they are using. Most countries use a system known as the Standard Industrial Classification to assign organisations to industries. Under this system, each organisation is assigned to an industrial category according to the principal goods or services it produces. The system is essential to the collection of statistics that enable us to estimate the relative contributions of the different industries to national income, and to detect which parts of the economy are growing or contracting over time.

To speak of a market directs attention towards the relations between suppliers (as sellers) and buyers of a particular good. It is usually taken for granted that buyers and sellers are engaged in *voluntary* exchange, of money for goods, and that buyers are able to exercise *choice* (to buy or not to buy, to buy this good rather than that one). ‘Market’ is therefore a politically charged term, unlike ‘industry’, shown, for example, by the prevalence in political debate of the expression ‘free market’.

The focus in Section 2 will be on consumer demand for products and in particular the price at which they are offered for sale. It is therefore appropriate to use the term ‘market’ in this section, reserving ‘industry’ for later sections where our attention is directed towards the costs incurred by firms in producing goods.

2.2 Market demand and price

This subsection will explore the widely observed relationship between the quantity demanded of a good by consumers and the price of the good: the lower the price, the greater the quantity demanded. This relationship underlies the way in which falling prices are responsible for the rapid take-up of new products by consumers, as reported in the quotations above. We focus on the market demand curve, which represents the demand of all the consumers in a given market. However, as well as the price of the good, there are other influences on market demand, discussion of which will be postponed until Section 2.3. This makes it possible to take a step-by-step approach and to begin by considering the influence of price alone.

The relationship between demand and price can be represented in different ways: in words, in a diagram or by using algebra. We expressed it in words in the preceding paragraph: the lower the price, the greater the quantity demanded. This relationship can also be shown in a diagram, known as a demand curve (always by convention a 'curve' though it may be drawn as a straight line). Figure 3 shows a demand curve, and we look at it in detail in a moment. Note first that as part of the step-by-step approach, the demand curve is drawn on the assumption that the price of the product is the only relevant variable influencing demand for the product. A 'variable' is a precisely defined aspect of the economy, such as the price of a good, that can take a range of values (such as £1, £2, £3,…).

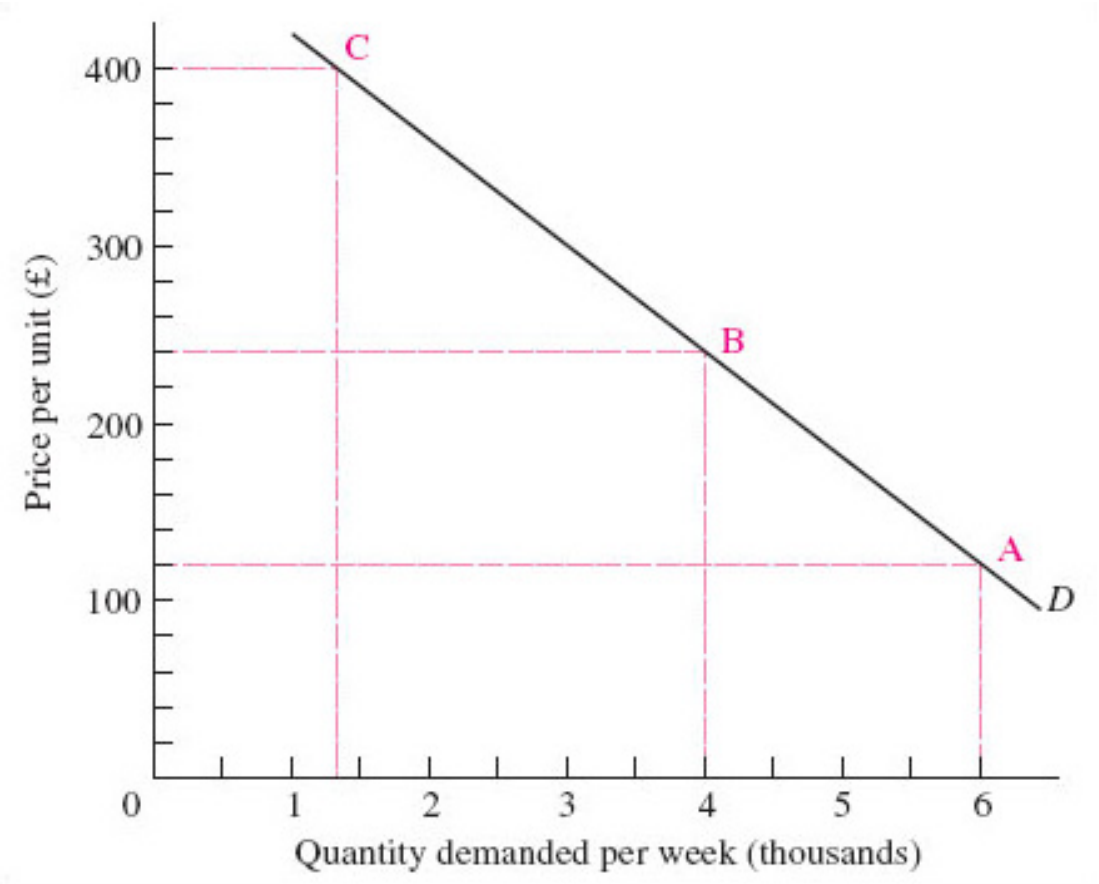


Figure 3: A market demand curve for electronic personal organisers

All the other influences on market demand are held constant while we look at the relationship between demand and price. This procedure is usually known by the Latin

phrase *ceteris paribus*, which means 'other things being equal'. It is the foundation for constructing economic models, which abstract from the complexities of real economic life to concentrate on one or two variables that seem to be important. Once we have understood how two variables are related – how a change in one affects the other, *ceteris paribus* – it is possible to move on, dropping the assumption that other things have remained equal. We can introduce other variables, gradually making the model more complex by considering the effects of changes in them.

An economic model therefore provides a systematic way of thinking about causal relationships. We can use them to formulate hypotheses about cause and effect, such as 'lower prices caused the increase in sales'. That puts us in a position to look for evidence that might lead us to accept or reject the hypothesis.

The market demand curve is a very simple economic model in that it abstracts from the many things going on in a market to focus on only two: the quantity demanded of the good and its price. Look at Figure 3, and notice that the vertical and horizontal axes are 'anchored' at the zero point, called the origin. As with all such diagrams, movements up the vertical axis, and along the horizontal axis to the right, represent higher values. Each point on the curve *D* shows the quantity demanded (measured on the horizontal axis) at a particular price (measured on the vertical axis). The market demand curve therefore shows the quantity demanded at each and every price by all the consumers in a particular market. We say 'each and every price' to draw attention to the fact that in drawing the market demand curve as a continuous line, economists are making estimates. The good may not have been offered for sale at 'each and every' price but only a small number of selected prices. Drawing the market demand curve as a continuous curve on a diagram such as Figure 3 shows estimates of what demand would be at other prices.

Figure 3 shows a hypothetical market demand curve for electronic personal organisers (hypothetical because it is not based on actual sales figures but is being used purely as an illustration of market demand curves in general). Electronic devices are particularly good at storing files, allowing them to be used in different ways, and have come to dominate the market for portable information storage. The market demand curve depicts the quantity consumers demand, depending on price. This 'quantity demanded' is not necessarily the number of electronic personal organisers that people need, but what they are willing and able to purchase at different prices.

The demand curve in Figure 3 shows that with a price of £120, the quantity demanded of electronic personal organisers is 6000 units per week (point A). For a higher price at, say, £240, *ceteris paribus*, we can find out how many units will be purchased by moving along the demand curve to point B, where the quantity demanded is only 4000. If the price was as high as £400 smaller quantities would be purchased; we move along the demand curve to point C, and see that the quantity demanded is only 1300 units. This inverse relationship between the price of a good (or service) and the quantity demanded (shown by the market demand curve sloping downwards and to the right) is known as 'the law of demand'.

So economic models can be stated in words and represented in diagrams. They can also be represented by using algebra. The claim that demand depends on – or changes in response to – price can be written as:

$$D = f(P), \text{ ceteris paribus}$$

which is read as demand (*D*) *is a function of* price (*P*), all other influences held constant. Both demand and price vary in this model; quantity demanded is the *dependent* variable, since it changes in response to price, the *independent* variable. So this algebraic

statement makes clear the causal relation proposed by the model, while the diagram, Figure 3, showed the *negative* relationship it proposes between demand and price: a higher price results in a lower quantity demanded (*ceteris paribus*).

However, there may be some exceptions to this 'law' of demand. Early in the twentieth century, Thorstein Veblen, an American institutional economist, analysed cultural influences on consumption. In *The Theory of the Leisure Class* (Veblen, 1912) he suggested that it can be important to show off your wealth by means of conspicuous consumption. The rich can demonstrate their wealth by buying goods that are widely known to be very expensive and beyond the reach of most consumers. The term *Veblen goods* is used to denote luxury items, such as exclusive jewellery, cars or designer clothes which may therefore be in greater demand at higher prices.

At the lower end of the income scale, consumers in very poor countries may actually buy less of a very basic good, such as rice, when its price falls. This is because they can use the spending power released by the fall in the price of the basic foodstuff to replace some rice with a greater variety of foods. Such goods are called *Giffen goods* because the influential economist Alfred Marshall, apparently in error, gave Sir Robert Giffen credit for discovering this exception to the general law of demand.

2.3 Other influences on market demand

What about other variables which may affect demand? Let us consider four such variables. As is often the case in economics, the first two points involve understanding some rather formal relationships between variables, in this case price and income.

1. *The price of other goods*. Two goods *x* and *y* are known as *substitutes* if the quantity demanded of good *x* increases after a *rise* in the price of good *y*. The rise in the price of good *y* causes consumers to switch to good *x*. For example, if the quantity demanded of electronic personal organisers increases after a rise in the price of (print) diaries, these goods are substitutes. On the other hand, two goods *x* and *y* are *complements* if the quantity demanded of good *x* increases after a *fall* in the price of good *y*. The fall in the price of good *y* causes consumers to buy more and also to buy more goods that are used with it, such as good *x*. For example, if the quantity demanded of electronic personal organisers increases after a fall in the price of desktop computers (to which data can be downloaded), these goods are complements.
2. *The incomes of consumers*. The amounts of income consumers have at their disposal determines the absolute level of consumption: people with high incomes tend to purchase more of most goods and services than people with low incomes. But the level of income also influences the kind of things that people buy. If a country's national income goes up, households have more total purchasing power, and more goods and services of most types will be bought. A *normal good* is one for which the quantity demanded increases when incomes rise. In countries where the average level of income is relatively high, most goods are normal goods (e.g. refrigerators, cars, haircuts, houseplants and electronic personal organisers). Studies of households in poverty show their very limited purchases. There are some products, such as basic foodstuffs, or black and white television sets, of which fewer items will be purchased as income rises. For example, if poor households in low-income countries get richer, they are able to substitute beans, meat or fish for part of their basic grain diet of rice, maize or millet. Less rice will therefore be purchased.

When demand for a commodity falls as incomes go up, the commodity is called an *inferior good*.

3. *Socio-economic influences*. J.S. Duesenberry, an American economist, suggested that demand for particular commodities, as well as consumption expenditure in general, is affected by a 'demonstration effect', where people feel social pressures to purchase what others have. J.K. Galbraith, another American economist, talks of a dependence effect, where wants are dependent on the very process by which they are satisfied, since producers use advertisements and sales people to persuade us to purchase what they are making. Many commentators suggest that for a wide range of goods 'consumption to be' has replaced 'consumption for use'. That is, the consumption of many goods and services has become an expression of identity and self-definition, with less emphasis placed on the use value of the product. As more people turn to shopping for entertainment and leisure, and as the phrase 'retail therapy' increasingly enters into everyday language in high-income countries, it becomes more difficult to understand the demand side of the economy without taking account of society's norms and values.
4. *The expected future price of the good.* Expectations about future prices may affect demand. People may buy non-perishable goods now because they think that prices will be higher in the future or delay purchases in the belief that price cuts are imminent.

We can use algebra to express very concisely what it will take us a paragraph to say in words. A demand function is an algebraic expression of the idea that the demand for a good depends on its price and on the other variables discussed above. We can now write the demand function for a commodity, x, in the following expanded form:

$$D_x = f(P_x, P_r, Y, Z, P_e)$$

This demand function states that the market demand for commodity x, D_x, is a function of, or depends on, five variables. P_x is the price of commodity x itself, while P_r stands for the price of related goods, that is, substitutes and complements. Y is the standard symbol in economics for the income of all consumers or households together. Then there are socio-economic variables, labelled Z. Finally, P_e stands for expected future prices. The number of variables on the right-hand side of the equation shows that the complete market demand function for a good is quite complex.

So far we have looked at a hypothetical market demand curve for electronic personal organisers on the assumption that while the price of the good changes other things remain equal. What happens if we relax the *ceteris paribus* assumption? In other words, supposing that commodity x is electronic personal organisers, what happens if any of the items on the right-hand side of the demand function change? A change in any of the variables other than the price of electronic personal organisers can be shown by a *shift* in the whole demand curve, because the *ceteris paribus* assumption, on which the original curve was drawn, has been dropped. The shift may be to the left or right, depending on the cause of the change.

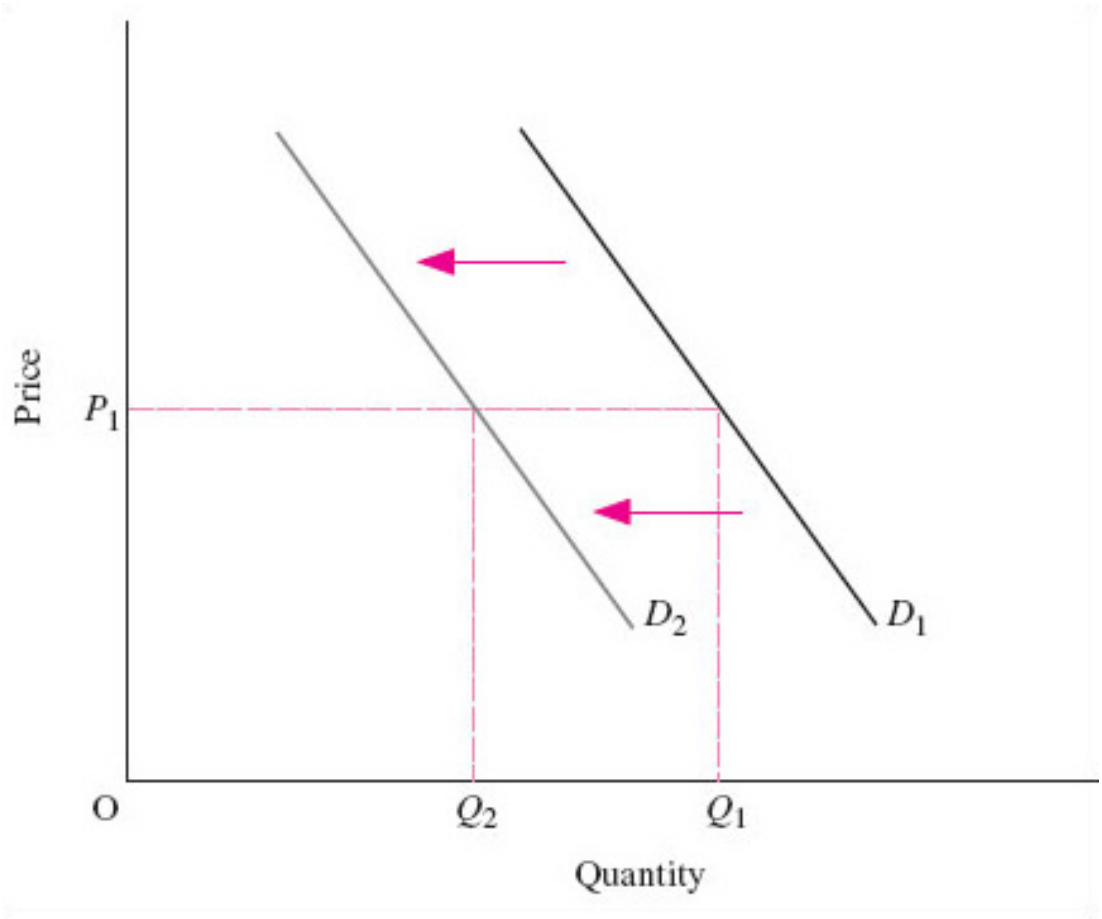


Figure 4: A shift in a market demand curve

Figure 4 shows such a shift in demand. This figure does not have a numerical scale on the axes, as in Figure 3. Instead, the axes are just labelled 'Price' (P) and 'Quantity' (Q). 'Quantity' should be understood as 'quantity per period of time', such as a day, a week or a year. Diagrams such as this allow us to focus on the direction of change of variables and their implications, rather than exact magnitudes. Let us suppose that consumer incomes fall, so that everyone has less to spend on normal goods. The quantity demanded will therefore decrease at all prices. For example, at price P_1 quantity demanded falls from Q_1 to Q_2. This means that we now have a new relationship between quantity demanded and price, and this is shown by the new market demand curve D_2 in Figure 4 lying to the left of the original one, D_1.

The key points about movements along and shifts of the market demand curve are:

- If the price of the good changes while all other variables remain the same, this is reflected in a movement *along* the market demand curve.
- If a variable other than the price of the good itself changes, the whole curve shifts, showing that after the change in the variable more (or less) is now demanded at each price.

Here is an exercise to help you think about shifts in the market demand curve.

Exercise 1

Think for a few minutes about how other things besides price may *not* remain equal in the market for electronic personal organisers. Work out what effect each will have on the demand curve. Then see if you can complete Table 1 below.

Table 1: Shifts in the market demand curve for electronic personal organisers

Change in variable	Effect on demand curve
Decrease in income	Decrease in quantity demanded at all prices. Demand curve shifts to the left
Increase in income	
Rise in price of a substitute	
Fall in price of a substitute	
Rise in price of a complementary good	
Fall in price of a complementary good	
Change in socio-economic influences in favour of electronic personal organisers	
Change in socio-economic influences away from electronic personal organisers	

Table 1 (completed): Shifts in the market demand curve for electronic personal organisers

Change in variable	Effect on demand curve
Decrease in income	Decrease in quantity demanded at all prices
	Demand curve shifts to the left
Increase in income	Increase in quantity demanded at all prices
	Demand curve shifts to the right
Rise in price of a substitute	Demand curve shifts to the right
Fall in price of a substitute	Demand curve shifts to the left
Rise in price of a complementary good	Demand curve shifts to the left
Fall in price of a complementary good	Demand curve shifts to the right
Change in socio-economic influences in favour of electronic personal organisers	Demand curve shifts to the right
Change in socio-economic influences away from electronic personal organisers	Demand curve shifts to the left

3 Firms, costs and technology

3.1 Introduction

In this section the focus turns towards the supply side of the market, towards firms and industries, exploring the importance of costs and technological change in the organisation of production. The objective is to understand some of the different kinds of change in industrial structure, namely changes in the number and size of firms in an industry. One such change saw the emergence of Ford, initially one among many similar firms jostling for position in the US automobile industry, as the industry 'leader'. What was so special about Ford? Henry Ford was the first car maker to introduce an innovative assembly-line production technology. This gave him a competitive advantage over firms using more traditional and, for that reason, more expensive processes. Since consumers were unwilling to pay higher prices for broadly similar products, Ford's rivals were forced to take up the same new methods of production if they wanted to compete.

This story raises some important questions about competition, technology and costs. How do firms respond to changing conditions in their industry? What can a 'typical' firm do to cut its costs? What are the main constraints on its behaviour? Economists use models of the relationships between technology, costs and output of firms to explore these questions. This section examines these relationships in a basic economic model of a firm. It explores the scope the firm has for cutting its costs in the short and long run, and the impact of changing technology on the firm's costs. The distinction between the short run and the long run is important because it is based on assumptions about how much room to manoeuvre the firm has in responding to changing market conditions and draws attention to the role of investment in the firm's response. It is in the short run that the firm's actions are subject to the greatest constraint. In the analysis in this section, firms take technology as given. In Section 4 we remove this constraint, restoring to firms some influence in shaping technology and hence their costs to their own advantage.

3.2 Technology and costs in the short run

Advertising leaflets are dropping through letter boxes around the UK, as we are writing this chapter, from cable suppliers trying to attract new customers for their services. They promise to provide a telephone line, a bundle of television channels, an Internet connection, home shopping and movies-on-demand, all at a 'bargain price'. These leaflets raise some interesting questions. How does expanding output of cable services by selling to new customers make it possible to offer them for sale at a lower price? What happens to the costs of producing cable services as the firm increases its output of them?

In analysing the costs, technology and output of firms, economists create a model of the typical or representative firm. A firm is an organisation that buys inputs, such as land, labour and capital, and transforms them into an output of goods and services for sale. We call these inputs to the firm's production process *factors of production*.

A firm that wants to expand output is likely to need more inputs. For example, in order to increase output the cable supplier would probably recruit more workers, such as the technicians who install cable boxes in customers' homes and the telephonists who

respond to customer enquiries. It might also need to buy in more of the equipment used by these workers, such as telephones, computers and engineering tools, and it might need to rent more office space. In other words, the firm increases its input of factors of production, namely labour (the technicians and telephonists), capital (telephones, computers and engineering tools) and land (office space). The firm is assumed to have done this by purchasing the additional inputs it needs from the relevant factor markets, for example by recruiting new workers in the labour market.

The firm can combine factors of production in various ways to create output, but is limited by the technology available to it. The best production methods available to the firm are summarised in its *production function*, which identifies the maximum output a firm can produce from each available combination of inputs. We can write it out using algebra in exactly the same way as we wrote out the demand function in Section 2:

$$Q = f(F_1, F_2, \ldots, F_n)$$

This says that the firm's output (*Q*) (the dependent variable) is a function of the various factors of production (the different *F*s, up to any number 'n' of them) such as land, different types of labour and machinery (the independent variables). This production function is the firm's *technology*. The firm is assumed to do the best it can with its inputs, without waste. If there is technological change, the firm can get more output from its inputs, that is, increase their productivity. In the models we discuss in this section all firms have access to the same technology.

The firm's objective is not to produce maximum output, but rather to make as much profit as possible. However, reducing costs is an important competitive tool in the search for profits, as the cable supplier's offer of a 'bargain price' to new customers illustrates. So how do firms' costs change as output changes? What makes it possible to supply more customers at a lower price? Are there any limitations to this process? The distinction between the short run and the long run helps to answer these questions.

Question

We referred above to the cable supplier recruiting new workers. Can you think of any employers or industries that have found this difficult?

In 2001 in the UK the health and education industries were beset with staff shortages and in some areas recruited from overseas. It is likely to take most employers longer to increase the quantity of skilled than unskilled labour. Skills may not be available in the locality and workers may have to be enticed to move from other jobs, involving periods of notice. New workers may have to be trained and the training period may well be lengthy. BBC television reported on 4 September 2001 that Arriva, a UK train and bus operator, had announced the cancellation of up to a hundred services a day because of a shortage of train drivers. Arriva had the capital equipment, the trains, to run the services but could not do so until it had increased the amount of skilled labour at its disposal by training new drivers. In this example the firm is constrained in its response to the demand for rail services because the quantity of one factor of production, labour, is fixed.

Economists describe the situation in which Arriva finds itself as **the short run**. A firm is operating in the short run when it is unable to change the quantity it uses of at least one of its factors of production. In other words, in the short run the firm can vary the quantity it

uses of some of its factors of production but there is at least one which is fixed. The term denotes this constraint on the firm's behaviour rather than a particular period of historical time. The firm's actions take place in actual or historical time and so the short run will eventually come to an end, in Arriva's case when the drivers have been trained in a matter of months. Therefore, the short run, understood as the condition of being unable to change the quantity used of at least one factor, refers to different periods of historical time, depending on the circumstances of each particular firm. Another example from transport illustrates this variety. London's Heathrow, Stansted and Gatwick airports are constrained in responding to rising demand for air travel from London and south-east England by the planning process required before new runways or terminals can be constructed. In this case the fixed factor is capital and the short run is measured in years, perhaps even decades.

> **Heathrow Terminal 5 to get go-ahead**
>
> After an inquiry that began in 1995, the government is poised to give the go-ahead for Heathrow's Terminal 5. It is due to open in 2007, subject to final approval.
>
> (Adapted from The Sunday Times, 4 November 2001, p. 2)

In the short run, therefore, the firm has one or more variable factors of production but at least one fixed factor. How does the firm's output change in the short run as it increases the amount of a variable factor? For example, let us suppose that the cable supplier recruited more labour, such as technicians and telephonists, while using unchanged quantities of both capital and land. Hiring more labour allows it to increase output. Initially, output may rise faster than the inputs of labour. Each additional worker may not need a computer all to themselves and, up to a point, extra workers can share the same office space. So initially increases in the amount of labour may generate larger ‘returns’ in the form of increases in output. At this stage in its expansion of output in the short run, therefore, the firm is experiencing increasing returns to a factor of production, in this case labour.

However, this fortunate situation has its limits. If the firm continues to expand output by increasing the input of labour with unchanged inputs of capital and land, a point will be reached after which the employment of even more workers brings successively smaller increases in output. For example, there may be so many workers that each one spends some time idle in the queue waiting to use a computer. The input of labour is rising faster than output, and the firm is now experiencing diminishing returns to a factor of production. Diminishing returns pose a serious constraint on the expansion of a firm in the short run.

These relationships between inputs and outputs in the short run influence the firm's costs. Returning to the cable supplier, it is now possible to frame a more precise question: what happens to the costs of producing these services as the firm increases its output of them in the short run? To analyse this, we need to distinguish *total costs* (*TC*) and the *average cost* (*AC*) of production. A firm's total costs are the expenses incurred in buying the inputs necessary to produce the firm's output. Average cost is the cost per unit of output. Average cost (*AC*) is therefore the total cost of production (*TC*) divided by the number of units produced (*Q*):

$$AC = \frac{TC}{Q}$$

We can draw a short-run average cost (*SRAC*) curve that models the relationship between different levels of output and average cost (*AC*) in the short run (Figure 5). Each

point on the *SRAC* curve represents the average cost in the short run (measured on the vertical axis) of producing a quantity of output (measured on the horizontal axis). Output (*Q*) is measured per time period, such as a year. The *SRAC* curve is expected to be 'U'-shaped, as shown in Figure 5. The downward-sloping section of the *SRAC* curve indicates that as output expands from a low level, average costs in the short run fall. Eventually, however, the *SRAC* curve begins to slope upwards, showing that at output levels above Q_1average costs rise as output increases.

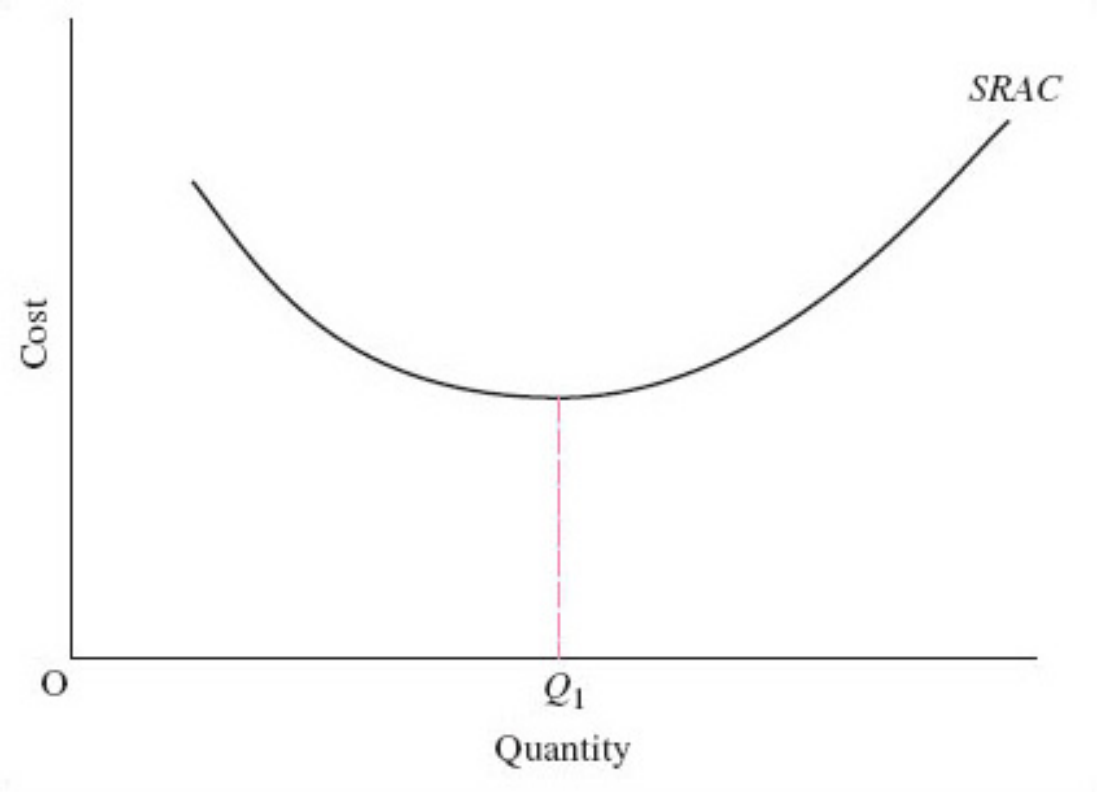


Figure 5: A short-run average cost (*SRAC*) curve

The shape of the *SRAC* curve arises from the technical constraints on short-run output expansion just described. We assume that the price of inputs is constant. As output expands the fixed cost of the fixed factors of production, such as the cable system itself, is being spread over an expanded number of customers. Increasing returns to a factor of production means that output rises faster than the variable input, so the cost of the variable factor per unit of output falls too. So initially the average cost of production (*AC*) falls as output rises.

Eventually, however, diminishing returns to the variable factor of production sets in. Output starts to rise more slowly than the variable input so the additional cost required to produce additional units of output starts to rise. Eventually total cost (fixed plus variable costs) will start to rise faster than output: that is, average costs will start to rise. On Figure 5 this happens as output rises above Q_1.

To sum up, in the short run the firm's ability to reduce costs as output rises is constrained by diminishing returns. In the long run, opportunities to invest in factors such as plant and machinery mean that the quantity of all factors of production can be varied. How does that affect the firm's costs?

3.3 Long-run costs and economies of scale

What makes it possible to offer more output for sale at a lower price? That was one of the questions with which Section 3.2 opened. Part of the answer is that the firm's cost curves, which reflect the technology it is using, may display falling average cost as output increases over a range of output levels. The other part of the answer is that market demand must be sufficient to justify successive expansions of output. A firm such as the cable supplier seeking to increase sales by offering 'bargain prices' to its new customers is making an assumption about the size of the market. This section examines the relation between output and average cost on the assumption that the size of the firm's market is sufficiently large to justify increases in its input of all factors of production.

There is greater scope for cutting costs in *the long run*. In the long run the firm can increase inputs of all factors of production: labour, capital and land. This corresponds in reality to firms making investments. Firms invest in labour by training key personnel such as train drivers, technicians and telephonists. They may invest in items of capital equipment from telephones and computers to new factories or warehouses. They may invest in land by buying more office space or, literally, perhaps by buying a 'greenfield' site for a new runway. In modelling firms in the long run, it is still assumed that they are operating with given technology available to all firms.

How exactly do economists analyse the effects of investment of this kind on the firm's costs? The long-run average cost (*LRAC*) curve models the relationship between changes in output and average cost (*AC*) in the long run (Figure 6). Each point on the *LRAC* curve represents the average cost, in the long run, of producing a given quantity of output. The shape of the *LRAC* curve varies with the technology in use by firms in a particular industry. An increase in output in the long run is described as an increase in the *scale* of production, the phrase reflecting the implications of investment in plant and equipment. In the long run, a firm can adjust the quantities of all its factors of production to produce its desired level of output at the lowest possible cost. To achieve higher levels of output, a firm will need to buy more factors of production. How will costs change?