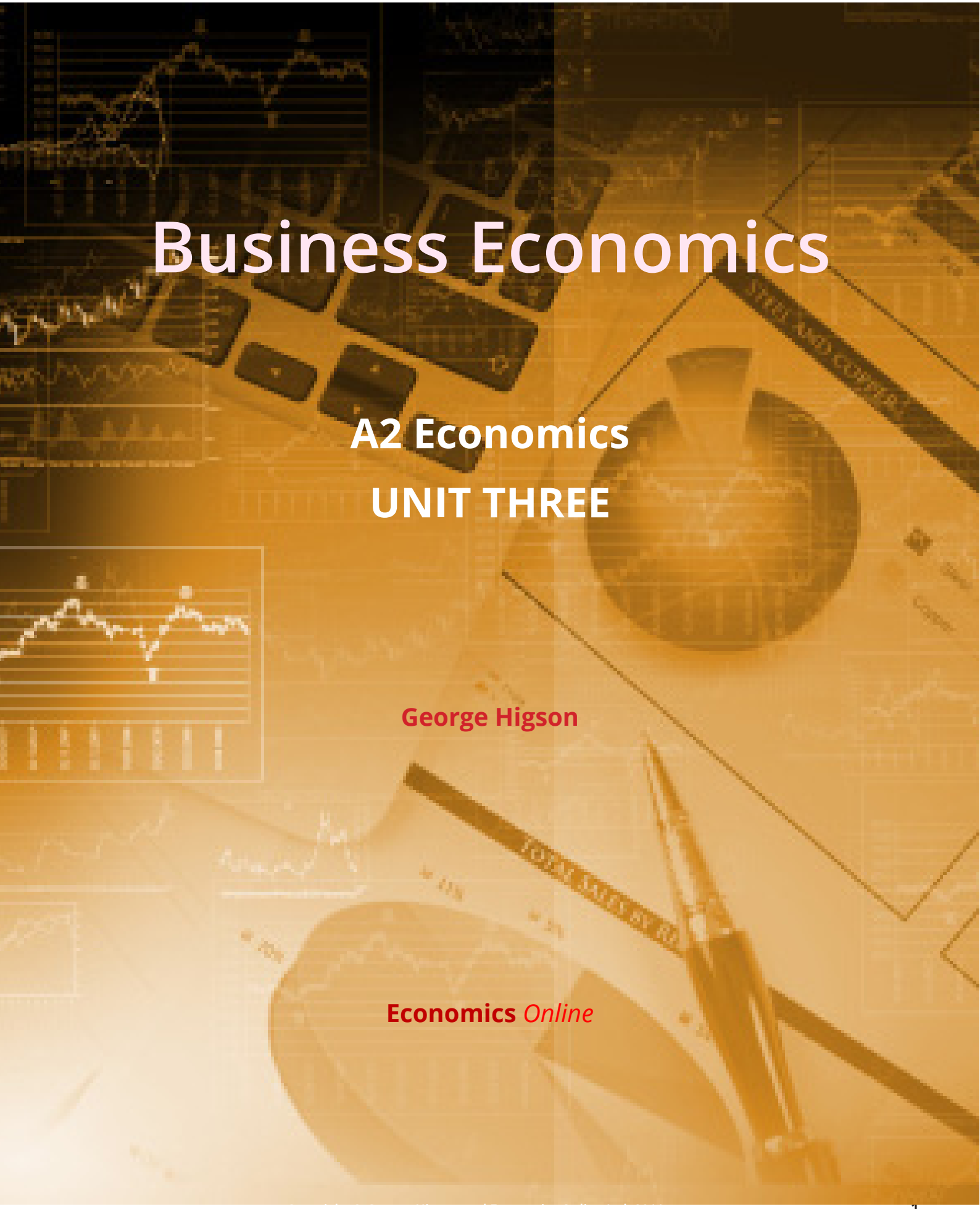


Business Economics

A2 Economics

UNIT THREE

George Higson

Economics *Online*

Contents

Enterprise

The origins of enterprise

Firms start when entrepreneurs organise resources and take risks in the expectation of earning a profit. More specifically, enterprises tend to be set up for one or more of the following reasons:

To solve a problem

Some firms originate to solve a problem faced by consumers, by other firms, or by government. For example, internet comparison websites solve the problem faced by consumers of having limited time to research the whole market for the best current deals.

To exploit an idea

Many firms start in order to exploit an original idea or an invention. If the invention can be turned into a product that adds value, it can command a price, and earn a profit. For example, *Dyson plc* was established by the inventor, James Dyson, to exploit his inventions and designs created he was in his early twenties.

To fill a gap

Some firms start because the entrepreneur identifies a gap in an existing or emerging market, such as online delivery businesses, like Amazon.

Because it can produce at lower cost

Many firms enter a market because to produce an existing product more cheaply, or more effectively, than existing firms in the market. For example, Tesco plc started in 1919 when co-founder Jack Cohen sold cheap groceries from a single stall in London's East End.

To exploit knowledge

Many firms exploit information that is not readily available, such as estate agents and travel agents.

In all cases, entrepreneurs anticipate that they will be successful and earn themselves a profit for their personal risk-taking and entrepreneurial skill. Private firms can only survive if they satisfy consumer demand effectively.

Financing enterprise

Entrepreneurs need finance to test, produce, and distribute their products. Finance can be obtained from a number of sources, including:

- The entrepreneur's own funds, called private equity
- Selling shares in their business, called share capital
- Borrowing from individuals, banks via loans and mortgages, or from other firms
- Credit from suppliers, which is similar to a loan in its effect
- Revenue from customers, which is the most important source of long term finance

Types of enterprise

There are several types of enterprise, each one distinguished by its legal ownership, including:

Private enterprise

Sole traders

Sole traders are the life-blood of a market economy. Sole traders are common in retailing and local services like plumbing and catering. With local services demand is limited, and there are many suppliers competing, so the scope for expansion is also limited. A sole trader owns the assets of the business, makes all the business decisions, bears all the risks, and, of course, retains all the profits.

Partnerships

Partnerships are owned, and usually managed, by a small number of partners, each of whom can specialise in a particular aspect of the business. Decisions will be jointly arrived at, and the risks and rewards will be spread between the partners. In certain types of partnership, not all partners bear equal risks, and some partners may have a limited liability for debts incurred by the businesses. Partnerships are common in professional and financial services such as solicitors, accountants, and estate agents.

Private Limited Companies (Ltd)

Limited companies are legally *incorporated* firms, which means that they have their own legal identity, and are owned by shareholders who have 'limited liability' for the firm's debts. Unlike sole traders and ordinary partnerships, limited, or joint stock companies, are established to take advantage of the principle of limited liability.

The rapid development of limited companies in the 18th Century provided a stimulus to the growth of private enterprise and free market capitalism. Capitalism spread because limited liability encouraged ordinary individuals to part with their savings without the risk of losing any more than the initial outlay. Today, private limited companies are common in all areas of economic activity in all sectors of the economy; from screenwriters and film producers, to restaurants and hotels.

In a limited company, shareholders appoint directors to take the key business decisions, though often the directors are also shareholders. Directors make decisions collectively as members of the Board of Directors.

Most of the significant risk taking is made by the Board of Directors, though day-to-day decision-making is commonly devolved to professional managers

Public Limited Companies (plc)

Like private limited companies, *public* limited companies are also legally incorporated and are owned by shareholders who have limited liability for the firm's debts, the difference being that public companies are allowed to sell shares to the general public. To enable them to 'go public' and 'list' their shares on the stock exchange, they must satisfy strict criteria laid down by law covering the liquidity of the business, publication of financial accounts, and the number of years the business has traded.

The regulations governing public limited companies in the UK are increasingly complex and this partly explains the recent trend towards 'de-listing' and returning to private limited status. The main advantage of being a 'plc' is that it is much easier to raise funds because shares can be offered for sale to any member of the public. Shares can also be re-sold to other members of the public via stock exchanges, so it is easy for investors to regain their liquidity. Despite tough regulations, most large firms prefer to remain 'plc's, or their equivalent in other countries.

Public enterprise

Public Corporations

Public corporations, such as the British Broadcasting Corporation (BBC), are organisations

owned by the state. They are funded in a number of ways, including:

1. Government grants and subsidies
2. License fees
3. Charges for supplying their service

A Board of Governors rather than a Board of Directors control public corporations. If their income is greater than their costs they make a 'surplus' rather than a profit.

Widespread privatisation during the 1980s and 1990s led to a reduction in the number of public corporations. Despite being owned by the state, public corporations are frequently managed along 'commercial' lines, as in the case of the BBC.

Not-for-profit organisations

As well as public corporations, many other organisations do not aim to make profits, though they may earn revenue and also be operated along commercial lines. Examples of not-for-profit organisations include Network Rail, charities like Oxfam, universities, and government sponsored organisations like the Office of National Statistics (ONS).

Do firms need entrepreneurs?

According to some early economists, the need for entrepreneurs relates to the degree of competition that exists in a market. In highly competitive markets, all participants have perfect knowledge and economic transactions will have no risk attached - hence less need for risk takers. However, in the real world of *imperfect* markets, greater risks exist and entrepreneurs can exploit this by earning higher profits.

Questions

1. Distinguish between a private limited company and a public limited company.

2. Why is 'limited liability' important to the development of enterprise?

3. Using the internet, create a short profile of three successful entrepreneurs, and the businesses they established . Create a timeline from the establishment of the business until today. Describe how the business has changed over its lifetime to date.

Motives of firms and stakeholders

Firms are organisations which may involve thousands of people directly, and millions indirectly. Those having an interest in firms are called *stakeholders*, and include owners, employees, customers and the local community.

Not all stakeholders share the same goals, or gain the same benefits from the success of a firm. For example, while entrepreneurs take business risks and expect a profit from their entrepreneurial skill and effort, managers, who are appointed by owners to make decisions, do not bear the same level of risk, and are not rewarded with profits. Managers receive a salary and other benefits, such as a company car, and may be motivated to maximise revenue, out of which salaries and other benefits are paid.

Employees and the unions they belong to will hope the business survives so that they retain their employment in the long run and receive a decent wage. The local community will hope the business provides jobs, without generating excessive external costs. The government will also hope that firms survive, prosper, and grow as their tax revenues depend on this happening.

Motives

Economists identify the following possible motives for firms:

Profit maximisation

Maximising profits means achieving the highest possible profit for the owners of a business. Profits are achieved when a firm's revenue is greater than its production costs. Profit maximisation has long been assumed to be the dominant goal of private enterprise - a view that dates back to the neo-classical economists of the late 19th Century.

Economists distinguish different types of profit, including *normal profit*, which is a reward just sufficient to keep the entrepreneur supplying their enterprise. In contrast, *super-normal*, or abnormal, profit is profit in excess of normal profit. Earning normal profit is also said to occur when the single entrepreneur or firm just covers opportunity cost and chooses to keep supplying to the market.

Sales volume maximisation

To maximise sales volume means to sell as many products as possible without making a loss. This means the firm must produce an output where the total revenue generated from sales just covers the total costs of production.

Sales revenue maximisation

To maximise total revenue means to gain the maximum possible revenue from selling a product. A firm's total revenue (TR) is the income it receives from selling a given quantity of products (Q) for particular prices (P), or in simple terms TR = P x Q.

Economic theory suggest that a price can be identified which achieves this goal. Sales revenue, or sales turnover, maximisation is associated with 'managerial' theories of business motives, which stress the influence of professional managers in the decision-making process in modern organisations.

Market share

Some firms may wish to increase their share of a market. This motive is significant for firms operating in markets with a few large competitors, called oligopolies, and where winning market share from rivals is less risky and costly than trying to win completely new customers.

Survival

Some firms may take a short-term view and simply want to remain in business - to survive.

Survival is significant for new firms and those in highly competitive markets. It is also common when there is a downturn or recession in the macro-economy, meaning that consumer spending falls across the whole economy. For example, High Street giant Woolworth went to liquidation in November 2008, as the recession took hold, after struggling to cope with debts of over £350m.

Shareholder value

To maximise shareholder value means to maximise the asset value of the business. Shareholder value is defined as the remaining (or residual) value of the business once all debts have been paid. Maximising shreholder value has become an important motive for many decision makers in large enterprises.

Ethical and environmental goals

Increasingly, firms are introducing ethical goals such as those associated with the environment and carbon emissions and with 'fair' trade.

Satisficing

Satisficing was a term first used by Herbert Simon in 1957, and means attempting to take into account a number of different and competing objectives, without attempting to 'maximise' any single one. For example, managers may first try to ensure that shareholder's get a reasonable rate of return, and then seek to reward themselves once shareholders have received their return.

Which motive?

The dominance of a goal depends upon a number of criteria, including:

Who owns the firm

Owners of a business often have different objectives that those appointed to manage the firm's operations. For example, sole traders may try to maximise profits, whereas public limited companies (plcs) may try to increase shareholder value. In contrast, not-for-profit firms may simply wish to maximise sales volume, or another, non-commercial objective.

Who manages the firm

Firms that are run by their owners, such as sole traders, may try to maximise profits, whereas firms run by professional managers may look to maximise sales revenue, given that they are usually paid a salary from revenue rather than from profit.

How large the firm is

Small firms may simply hope to survive, whereas larger firms may expect to develop market share. Motives can change over the life-cycle of the business.

What competitors are doing

If one firm, perhaps the dominant firm in a given market, introduces a new strategy this may be shared by all firms in the industry. The concern with ethical and environmental issues gained momentum as, one-by-one, large plcs introduced an ethical and environmental dimension to their operations. This is often called a 'band-wagon' effect.

The time period

In the short run, basic goals such as survival may dominate, while in the long run more challenging goals may dominate, such as maximising shareholder value, or introducing environmental goals. Goals change to reflect changing conditions and circumstances of particular firms at different points in their evolution.

Example - Tesco

When Tesco started just after the end of the 1st World War in 1919, its co-founder, Jack Cohen, was probably more concerned as a sole trader with day to day survival that anything else. Five years later, he formed a partnership with T E Stockwell to create Tesco and set up his first store. In 1932, Tesco became a private limited company and by 1947, Tesco had floated on the stock exchange and become a public limited company. This enabled capital to be raised to finance a massive expansion programme during the 1950s and 1960s. During the 1970s, Tesco built a national store network to rival its main competitor, J Sainsbury. During the 1990s, and in an attempt to win market share from its main rivals, Tesco revolutionised supermarket shopping with the introduction of the Tesco Clubcard. In 2000, it launched Tesco.com, and widened its range of products to include electrical goods and clothing.

By 2011, Tesco had stores in 13 countries and a dominant market share in the UK grocery market of around 32%, with group profits of £3.8b on revenue of £67.6b. At Tesco has evolved and expanded, and changed its legal structure, its motives, and goals have also changed. *(Source: Tescoplc.com)*

Conflicts between motives

It should be clear that there are likely to be conflicts between competing goals.

For example, the desire to maximise profits may be in conflict with a number of other goals, including sales maximisation, sales revenue maximisation and ethical goals.

These conflicts will become clearer once the main goals are analysed in subsequent chapters.

Questions

1. Why do different firms have different goals?
2. Using examples, explain why the goals of a firm may be in conflict with each other.

Revenue

The measurement of revenue

Revenue is the income a firm retains from selling its products once it has paid indirect tax, such as VAT. Revenue provides the income which a firm needs to enable it to cover its costs of production, and from which it can derive a profit. Profit can be distributed to the owners, or shareholders, or retained in the business to purchase new capital assets or upgrade the firm's technology. Revenue is measured in three ways:

Total revenue

Total revenue (TR), is the total flow of income to a firm from selling a given quantity of output at a given price, less tax going to the government. The value of TR is found by multiplying price of the product by the quantity sold.

Average revenue

Average revenue (AR), is revenue 'per unit', and is found by dividing TR by the quantity sold, Q. AR is equivalent to the price of the product, where P x Q/Q = P, hence AR is also 'price'.

Marginal revenue

Marginal revenue (MR) is the revenue generated from selling one extra unit of a good or service. It can be found by finding the change in TR following an increase in output of one unit. MR can be both positive and negative.

Example

Revenue for a hypothetical firm

Price	Quantity	TR (P x Q)	MR (Change in TR from increasing output by one unit)
10,000	1	10000	
9,000	2	18000	8000
8,000	3	24000	6000
7,000	4	28000	4000
6,000	5	30000	2000
5,000	6	30000	0
4,000	7	28000	- 2000
3,000	8	24000	- 4000

Revenue curves

Total revenue

Initially as output increases total revenue (TR) also increases, but at a decreasing rate. It eventually reaches a maximum and then decreases with further output[1].

Average revenue

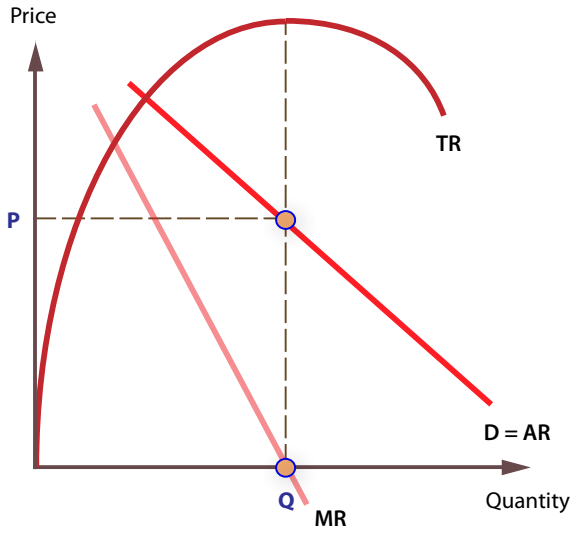
However, as output increases the average revenue (AR) curve slopes downwards. The AR

1 This assumes that firms are free to change their price. Under conditions of 'perfect' competition, the price is set and changes in revenue are at a constant rate, and the TR curve is linear.

curve is also the firm's demand curve.

Marginal revenue

The marginal revenue (MR) curve also slopes downwards, but at twice the rate of AR. This means that when MR is 0, TR will be at its maximum. Increases in output beyond the point where MR = 0 will lead to a negative MR.



Questions

Consider the example of a firm that produces and sells studio recording equipment.

Revenue for the firm, per week			
Price (AR)	**Quantity**	**TR**	**MR**
7,000	1		
6,500	2		
6,000	3		
5,500	4		
5,000	5		
4,500	6		
4,000	7		
3,500	8		
3,000	9		

1) Calculate total and marginal revenue.

   a) Plot total, average, and marginal revenue.

   b) When TR is at a maximum, what is MR?

2) If MR for a firm is always constant, what happen to AR?

The principles of production

Resource inputs

In order to produce goods and services which can be sold, and generate revenue and profits, a firm must purchase or hire scarce factor inputs. These factors of production can be fixed or variable.

Fixed factors

Fixed factors are those that do not change as output is increased or decreased, and typically include premises such as offices and factories, and capital equipment such as machinery and computer systems.

Variable factors

Variable factors are those that do change with output, which means more are employed when production increases, and less when production decreases. Typical variable factors include labour, energy, and raw materials directly used in production.

Time periods for a firm

The fundamental principles of production relate closely to the concept of time periods, of which there are four:

The very short run

A firm is said to be in its *very short run* when the only way to increase supply to the market is by using up existing stocks of inputs.

The short run

A firm is said to be in its *short run* when it can increase its output by using more variable factors, such as by hiring more workers, but not by increasing its fixed factors. In the short run firms do not use extra fixed factors, such moving to new premises, to increase output. Therefore, in the short run at least one factor of production is fixed.

The long run

A firm enters its *long run* when it increases its scale of operations. Increasing scale means that no factor of production is fixed, and all are variable. Typically, this means that a firm expands by building or renting larger premises, purchasing or leasing new machinery and employing more workers.

The very long run

A whole industry enters the very long run when there is a significant change in the use of technology. For example, the widespread use of the internet to book holidays has drastically altered how the holiday industry is structured.

Economic analysis tends to focus only on the short and long run, and largely ignores the very short and very long run.

Time periods for a market

A whole market or industry can also be considered in terms of the short and long run.

The industry short run

An industry is in its short run when its capacity is fixed. This usually means that the number of

firms in the industry is fixed, with no new firms entering or leaving the market.

The long run

This exists when there is an increase, or decrease, in the capacity of the industry to produce, and this often means that the number of firms in a given market increases, or decreases.

The law of diminishing returns

The law of diminishing marginal returns comes into play whenever a firm tries to increase output by applying additional variable inputs to a fixed factor. Production requires the combination of both fixed and variable factors to create an output. Economic theory predicts that if firms increase the number of variable factors they use, such as labour, while keeping one factor fixed, such as machinery, the extra output or returns from each additional, marginal unit of the variable factor must eventually diminish.

Diminishing marginal returns forms part of a larger principle, called the principle of variable proportions. This states that, assuming one factor is fixed, the marginal returns generated from adding new variable factors will not be constant. In fact, returns will rise at first, reach a turning point, and then eventually diminish. The law of diminishing marginal returns simply refers to the last phase of this wider principle.

Consider the following example:

Assuming that one factor remains fixed, as more workers are added, total product will increase up to a point.

Output for the firm, per week			
Workers	**Total product**	**Average product**	**Marginal product**
1	6	6	
2	14	7	8
3	28	8	14
4	36	9	8
5	40	8	4
6	42	7	2
7	42	6	0
8	40	5	-2
9	36	4	-4

Observations

What happens to productivity?

Marginal productivity is relatively low when only a few workers are employed. However, marginal productivity rises quickly as each extra worker contributes more than the previous one. Eventually marginal productivity begins to decline, in this case, with the employment of the fourth worker. With the employment of seven workers marginal product is zero, and total product is at a maximum. This means that marginal productivity is low at the extremes of output – at high and low levels.

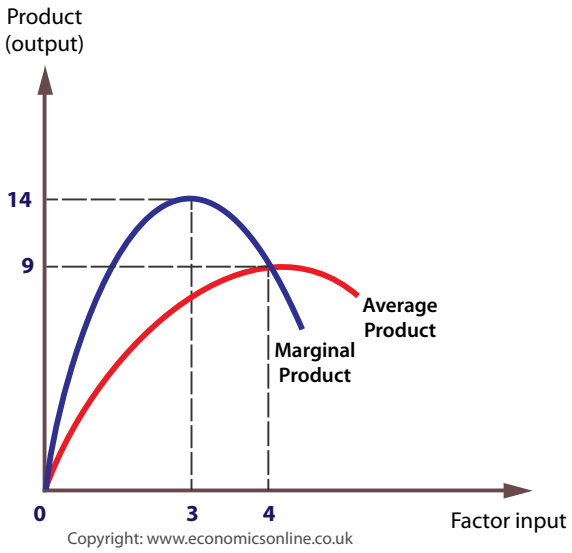
The product curves

It can be observed that, at first, the marginal returns curve increases and then decreases. The marginal returns curve cuts the average returns curve when average returns are at their peak.

How is this pattern explained?

With a small number of workers, output is low and a division of labour cannot be employed, - workers cannot specialise or develop new skills. However, marginal returns increase quickly as specialisation occurs and efficiency increases. This creates the opportunity for labour to develop skills and become more productive.

Specialisation can create its own problems and inefficiencies, especially in terms of the efficiency and productivity of labour. Eventually, marginal returns diminish as the effects of specialisation and new skills wear off. This pattern has a considerable impact on the firm's short-run cost curves.



Indeed, productivity may eventually decline due to over-specialisation. In this case workers may become bored with repetitive and monotonous work, and this may lead to labour (staff) turnover, and increased training costs. Production levels may fall during training, as workers are on traning schemes are temporarily unproductive.

Questions

A local pizza delivery business employs drivers to deliver pizzas. Each driver is responsible for a single area of the town, and the street in that area. Each driver is paid the same wage of £300 per week, though they can earn more than this with tips. The business estimates that, on a typical day it needs 5 drivers to achieve maximum productivity. In one particular week, 3 drivers are sick, and only 2 are available. The managers contacted the job centre to ask for 3 drivers to start as soon as possible. The following day, the 3 sick drivers return to work, and 3 new drivers arrived from the job centre, making a total of 8 drivers, 3 more than is really necessary.

1. Explain why, when only 2 drivers are available, marginal productivity is likely to be relatively low.

2. Explain why, when 8 drivers are available, marginal productivity is also likely to be relatively low.

Short run costs of production

Fixed, variable and total costs

Fixed costs are those that do not vary with output and typically include rents, insurance, depreciation, set-up costs, and normal profit. They are also called overheads.

Variable costs are costs that do vary with output, and they are also called direct costs. Examples of typical variable costs include fuel, raw materials, and some labour costs.

An example

Consider the following hypothetical example of a boat building firm. The total fixed costs, TCF, include premises, machinery and equipment needed to construct boats, and are £100,000, irrespective of how many boats are produced. Total variable costs (TVC) will increase as output increases.

Costs of production

Output	Total fixed costs (£000)	Total variable costs (£000)	Total costs (£000)
1	100	50	150
2	100	80	180
3	100	100	200
4	100	110	210
5	100	150	250
6	100	220	320
7	100	350	450
8	100	640	740

If we plot these, we get the following cost curves:

Total fixed costs

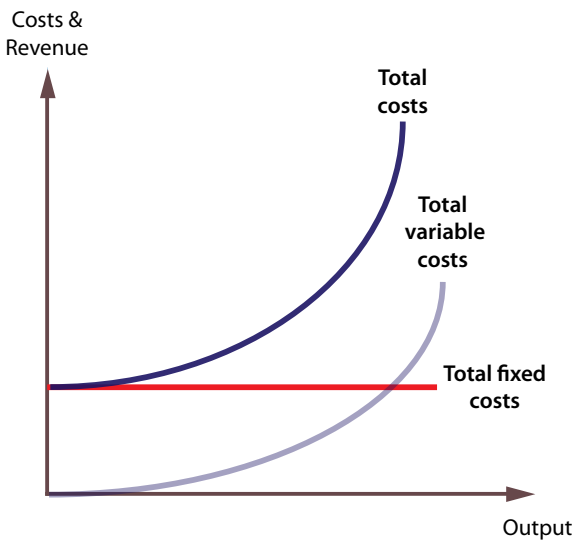
Given that total fixed costs (TFC) are constant as output increases, the curve is a horizontal line on the cost graph.

Total variable costs

The total variable cost (TVC) curve slopes up at an accelerating rate, reflecting the law of diminishing marginal returns.

Total costs

The total cost (TC) curve is found by adding total fixed and total variable costs. Its position reflects the amount of fixed costs, and its gradient reflects variable costs.

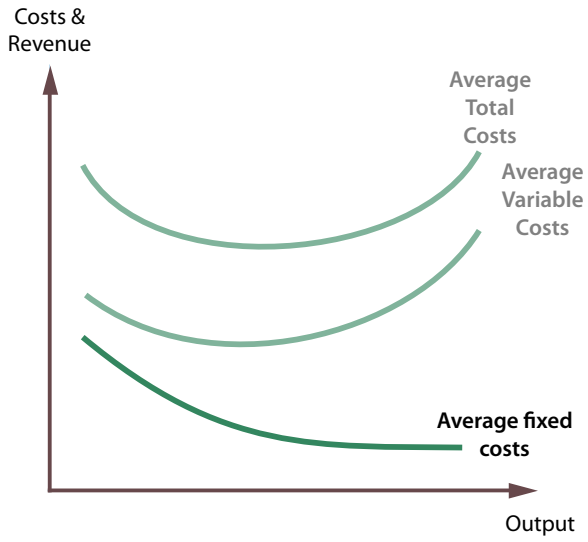


Average fixed costs

Costs of production		
Output	Total fixed costs (£000)	Average fixed costs (£000)
1	100	100
2	100	50
3	100	33.3
4	100	25
5	100	20
6	100	16.6
7	100	14.3
8	100	12.5

Average fixed costs are found by dividing total fixed costs by output. As fixed cost is divided by an increasing output, average fixed costs will continue to fall.

Graph to show average fixed costs



The average fixed cost (AFC) curve will slope down continuously, from left to right.

Average variable costs

To find average variable costs, total variable cost is divided by output.

Costs of production		
Output	**Total variable costs (£000)**	**Average variable costs (£000)**
1	50	50
2	80	40
3	100	33.3
4	110	27.5
5	150	30
6	220	36.7
7	350	50
8	640	80

The average variable cost (AVC) curve will at first slope down from left to right, then reach a minimum point, and rise again. AVC is 'U' shaped because of the *principle of variable proportions*, which explains the three phases of the curve - *increasing returns* to the variable factors, which cause average costs to fall, followed by *constant returns*, followed by *diminishing returns*, which cause costs to rise.

Average total cost

Average total cost 'ATC' is also called average cost or unit cost. Average total costs are a key cost in the theory of the firm because they indicate how efficiently scarce resources are being used. To find average total costs we must either divide total costs by output, or add AVC and AFC.

Costs of production			
Output	**Average variable costs (£000)**	**Average fixed costs (£000)**	**Average total costs (£000)**
1	50	100	150
2	40	50	90
3	33.3	33.3	67
4	27.5	25	53
5	30	20	50
6	36.7	16.6	53
7	50	14.3	64
8	80	12.5	92

Graph to show average total costs

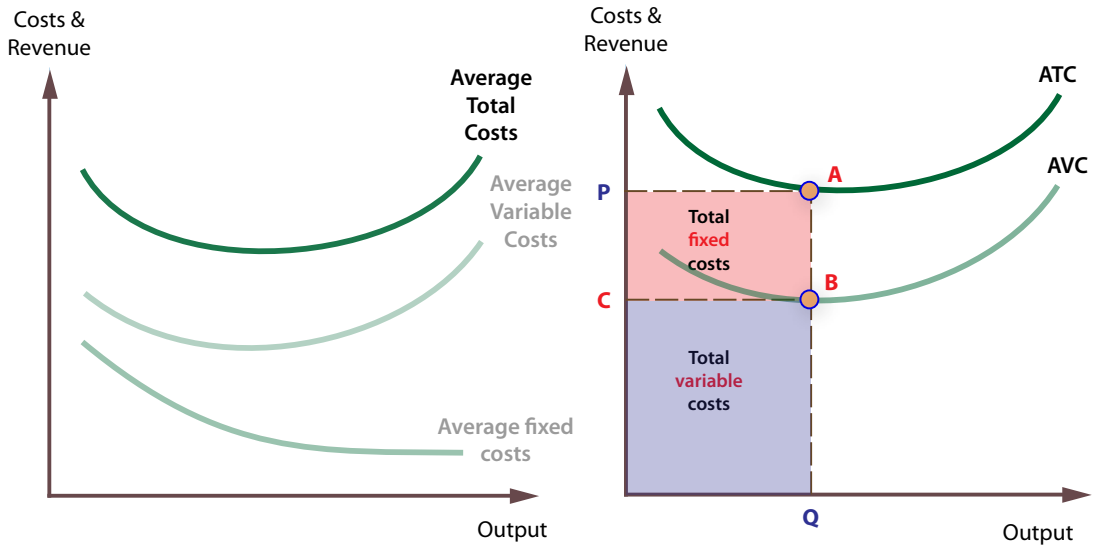
Average total cost (ATC) can be found by adding average fixed costs (AFC) and average variable costs (AVC).

The ATC curve is also 'U' shaped because it takes its shape from the AVC curve, with the upturn

reflecting the onset of diminishing returns to the variable factor.

Areas for total costs

If we take a given level of output, Q, we can see the areas representing total variable costs and total fixed costs. The two areas added together represent total costs.



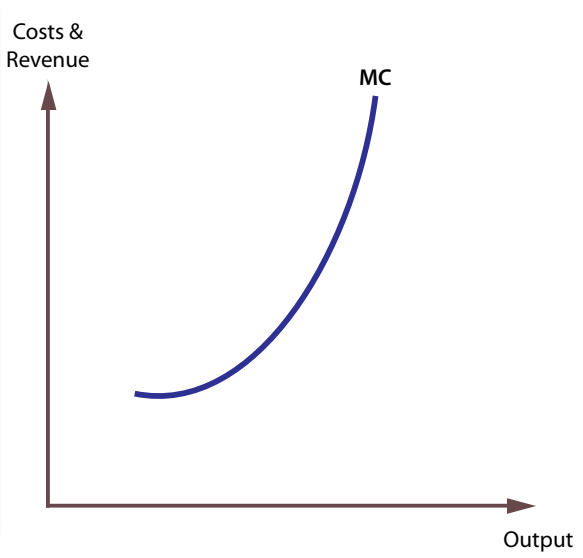
Marginal costs

Marginal cost is the cost of producing one extra unit of output. It can be found by calculating the change in total cost when output is increased by one unit.

It is important to note that marginal cost is derived solely from variable costs, and not fixed costs.

Example of marginal cost

Marginal costs		
Output	**Total costs (£000)**	**Marginal costs (£000)**
1	150	
2	180	30
3	200	20
4	210	10
5	250	40
6	320	70
7	420	100
8	570	150



Graph to show marginal costs

The marginal cost curve falls briefly at first, and then rises. Marginal costs are derived from variable costs and are subject to the ***principle of variable proportions.***

The significance of marginal cost

The marginal cost curve is significant in the theory of the firm for two reasons:

- It is the leading cost curve, because changes in total and average costs are derived from changes in marginal cost.
- The *lowest price* a firm is prepared to supply at is the price that just covers marginal cost.

Average and marginal costs

Average total and marginal cost are connected because they are derived from the same basic numerical cost data.

There are three general rules governing the relationship, which are:

1. Marginal cost will always cut average total cost from below.

2. When marginal cost is below average total cost, average total cost will be falling, and when marginal cost is above average total cost, average total cost will be rising.

3. A firm is most productively efficient at the lowest average total cost, which is also where average total cost (ATC) = marginal cost (MC).

Sunk costs

Sunk costs are those that cannot be recovered if a firm goes out of business. Examples of sunk costs include spending on advertising and marketing, specialist machines that have no scrap value, and stocks which cannot be sold off. Sunk costs are a considerable barrier to entry and exit.

Marginal costs

Costs

Average costs

Questions

1) Consider the above data for a firm (*UK Airparts Plc*) making complex electronic navigation systems for aircraft.

   a) Complete missing figures.

   b) Plot ATC, AVC and MC.

Costs of production							
Quantity	**Price (£000)**	**Fixed costs (£000)**	**Variable costs (£000)**	**Total costs (£000)**	**Average variable costs (AVC) (£000)**	**Average total costs (ATC) (£000)**	**Marginal costs (MC) (£000)**
0	100		0	100			
1	90		40	140			
2	80		60	160			
3	70		70	170			
4	60		90	190			
5	50		120	220			
6	40		190	290			
7	30		290	390			
8	20		460	560			

Economies and diseconomies of scale

Increases in scale in the long run

A firm's efficiency is affected by its size. Large firms are often more efficient than small ones because they can gain from economies of scale, but firms can become too large and suffer from diseconomies of scale. As a firm expands its scale of operations, it is said to move into its long run. The benefits arising from expansion depend upon the effect of expansion on productive efficiency, which can be assessed by looking at changes in average costs at each stage of production.

Methods of expansion

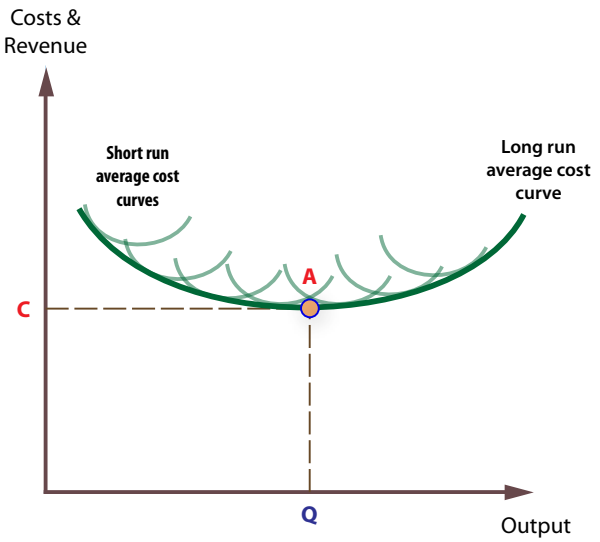
A firm can increase its scale of operations in two ways:

1. ***Internal growth***, also called organic growth

2. ***External growth***, also called integration - by merging with other firms, or by acquiring them.

By growing, a firm can expect to reduce average costs and become more competitive.

The firm's long run average cost shows what is happening to average cost when the firm expands, and is at a tangent to the series of short run average cost curves. Each short run average cost curve relates to a separate stage or phase of expansion. The specific reductions in cost associated with expansion are called economies of scale.



However, economic theory suggests that average costs will eventually rise because of diseconomies of scale. The long run cost curve for most firms is assumed to be 'U' shaped because of the impact of both economies and diseconomies of scale.

Economies of scale

There are a number of types of economy of scale, including:

Technical economies

Technical economies are the cost savings a firm makes as it grows larger, and arise from the increased use of large scale mechanical processes and machinery. For example, a mass producer of motor vehicles can benefit from technical economies because it can employ mass production techniques and benefit from specialisation and a division of labour.

Purchasing economies

Purchasing economies are gained when larger firms buy in bulk and achieve purchasing discounts. For example, a large supermarket chain can buy its fresh fruit in much greater quantities than a small fruit and vegetable supplier.

Administrative savings

Administrative savings can arise when large firms spread their administrative and manage-

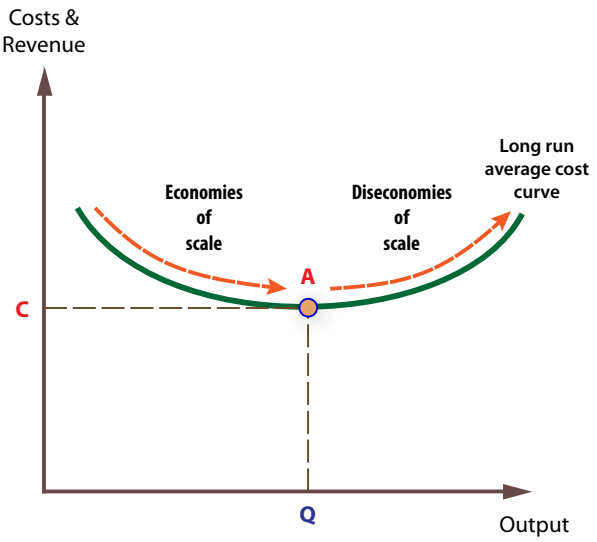
ment costs across all their plants, departments, divisions, or subsidiaries. For example, a large multi-national can employ one set of financial accountants for all its separate businesses.

Financial savings

Large firms can gain financial savings because they can usually borrow money more cheaply than small firms. This is because they usually have more valuable assets which can be used as security (collateral), and are seen to be a lower risk, especially in comparison with new businesses. In fact, many new businesses fail within their first few years because of cash-flow inadequacies. For example, for having a bank overdraft facility, a supermarket may be charged 2 or 3 % less than a small independent retailer.

Risk bearing economies

Risk bearing economies are often derived by large firms who can bear business risks more effectively than smaller firms. For example, a large record company can more easily bear the risk of a 'flop' than a smaller record label.



Diseconomies of scale

Economic theory also predicts that a firm can become less efficient if it becomes too large. The additional costs of becoming too large are called diseconomies of scale.

Examples of diseconomies include:

Poor communication

Larger firms often suffer poor communication because they find it difficult to maintain an effective flow of information between departments, divisions or between head office and subsidiaries. Time lags in the flow of information can also create problems in terms of the speed of response to changing market conditions. For example, a large supermarket chain may be less responsive to changing tastes and fashions than a much smaller, 'local' retailer.

Co-ordination problems

Co-ordination problems also affect large firms with many departments and divisions, and may find it much harder to co-ordinate its operations than a smaller firm. For example, a small manufacturer can more easily co-ordinate the activities of its small number of staff than a large manufacturer employing tens of thousands.

'X' inefficiency

'X' inefficiency is the loss of management efficiency that occurs when firms become large and operate in uncompetitive markets. Such loses of efficiency include over paying for resources, such as paying managers salaries higher than needed to secure their services, and excessive waste of resources. 'X' inefficiency means that average costs are higher than would be experienced by firms in more competitive markets.

Low motivation

Low motivation of workers in large firms is a potential diseconomy of scale that results in lower productivity, as measured by output per worker.

The 'principal-agent' problem

Large firms may experience inefficiencies related to the principal-agent problem. This prob-

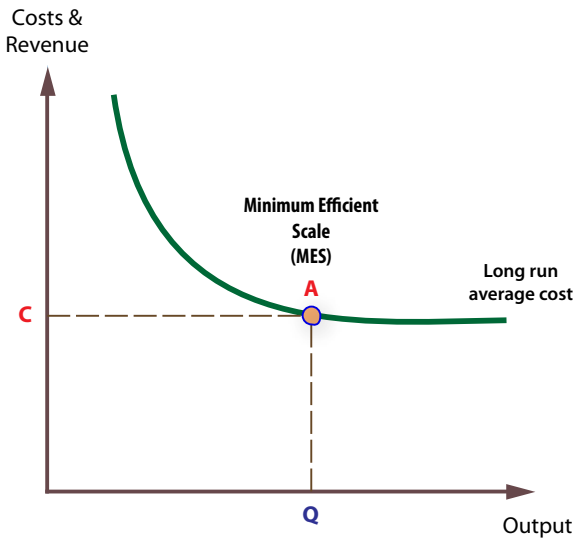
lem is caused because the size and complexity of most large firms means that their owners often have to delegate decision making to appointed managers, which can lead to inefficiencies. For example, the owners of a large chain of clothes retailers will have to employ managers for each store, and delegate some of the jobs to managers but they may not necessarily make decisions in the best interest of the owners. For example, a store manager may employ the most attractive sales assistant rather than the most productive one.

Minimum Efficient Scale (MES)

A firm's minimum efficient scale (MES) is the lowest scale necessary to achieve the economies of scale required to operate efficiently and competitively in its industry. No further significant economies of scale can be achieved beyond this scale.

Minimum efficient scale affects the number of firms that can operate in a market, and the structure of markets. When minimum efficient scale is low, relative to the size of the whole industry, a large number of firms can operate efficiently, as in the case of most retail businesses, like corner shops and restaurants.

However, if minimum efficient scale can only be achieved at very high levels of output relative to the whole industry, the number of firms in the industry will be small. This is case with natural monopolies, such as water, gas, and electricity supply.



Questions

1) Describe the likely economies and diseconomies of scale for:

   a) A large supermarket chain, like Tesco

   b) A large computer games manufacturer operating in several countries, like Sony

2) What is MES, and how is it likely to be different for a firm of solicitors and a computer chip manufacturer?

3) What is the significance of MES for the number of firms in an industry?

Profits

Profit has a number of meanings in economics. At its most basic level, profit is the reward gained by risk taking entrepreneurs when the revenue earned from selling a given amount of output exceeds the total costs of producing that output. This simple statement is often expressed as the profit identity, which states that:

Total profits = total revenue (TR) – total costs (TC)

However, the concept of profit needs clarification because there is no standard definition of what counts as a cost.

Normal profit

In markets which are perfectly competitive, the profit available to a single firm in the long run is called normal profit. This exists when total revenue, TR, equals total cost, TC. Normal profit is defined as the minimum reward that is just sufficient to keep the entrepreneur supplying their enterprise. In other words, the reward is just covering opportunity cost - that is, just better than the next best alternative.

The accounting definition of profits is rather different because the calculation of profits is based on a straightforward numerical calculation of past monetary costs and revenues, and makes no reference to the concept of opportunity cost. Accounting profit occurs when revenues are greater than costs, and not equal, as in the case of normal profit. To the economist, however normal profit is a cost, and is included in total costs of production.

Super-normal (economic) profit

If a firm makes more than normal profit it is called super-normal profit. Supernormal profit is also called economic profit, and abnormal profit, and is earned when total revenue is greater than the total costs. Total costs include a reward to all the factors, including normal profit. This means that, when total revenue equals total cost, the entrepreneur is earning normal profit, which is the minimum reward that keeps the entrepreneur providing their skill, and taking risks. The level of super-normal profits available to a firm is largely determined by the level of competition in a market – the more competition the less chance there is to earn super-normal profits.

Super-normal profit can be derived in three general cases:

1. By firms in perfectly ***competitive markets*** in the short run, before new entrants have eroded their profits down to a normal level.

2. By firms in less than perfectly competitive markets, such as firms operating under ***monopolistic competition*** and ***competitive oligopolies*** when innovating or reducing costs, and earning ***head start profits***. These will eventually be eroded away, providing further incentive to innovate and become more cost efficient.

3. By firms in highly ***uncompetitive markets***, like collusive oligopolies and monopolies, who can erect barriers to entry protect themselves from competition in the long run, earning persistent above-normal profits.

Marginal profits

Marginal profit is the additional profit from selling one extra unit. A profit per unit will be achieved when marginal revenue (MR) is greater than marginal cost (MC). At profit maximisation, marginal profit is zero because MC = MR.

Profit maximisation

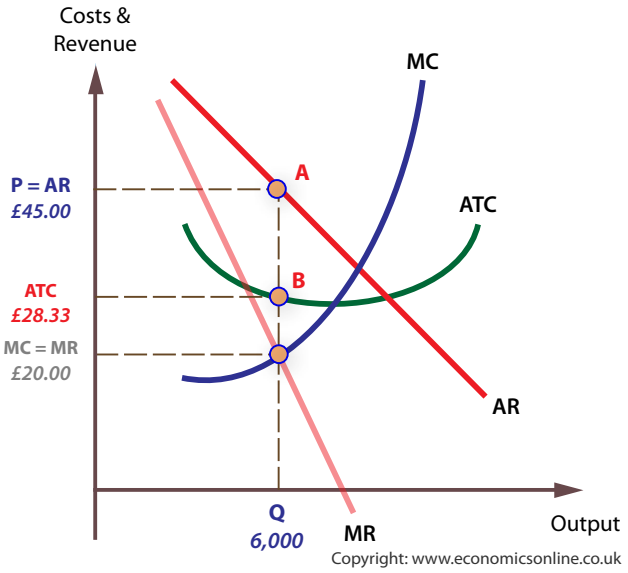
Firms achieve maximum profits when marginal revenue (MR) is equal to marginal cost (MC), that is when the cost of producing one more unit of a good or service is exactly equal to the revenue derived from selling one extra unit[2]. Consider the following example of a firm producing paper for printing. The 'units' are boxes of 10 packs of printing paper.

Costs, revenue, and profits

Price (£)	Quantity (000)	Total Fixed costs (£000)	Total Variable costs (£000)	Total costs (£000)	Average costs (ATC) (£)	Marginal costs (MC) (£)	Total Revenue (£000)	Marginal Revenue (£)	Profit (TR - TC) (£000)
60	3	50	65	115	38.3		180		65
55	4	50	85	135	33.8	20	220	40	85
50	5	50	100	150	30	15	250	30	100
45	6	50	120	170	28.3	20	270	20	100
40	7	50	145	195	27.9	25	280	10	85
35	8	50	175	225	28.1	30	280	0	55
30	9	50	190	260	28.9	35	270	- 10	10
25	10	50	250	300	30	40	250	- 20	- 50
20	11	50	300	350	31.8	50	220	- 30	- 130

Profits are clearly maximised at a price of £45 per box, where the firm sells 6,000 boxes. This is also where the marginal cost per box (at £20) exactly equals the marginal revenue per box. Actual profits are £100,000 (which is TR at 6,000 boxes - £270,000 - less TC - £170,000). This value is also the difference between AR (P) and ATC, times the amount sold - i.e. Profits = AR (P) at £45 less £28.333 x 6000 [which is £100,000].

This can be shown diagrammatically. Profit is maximised at output Q in the diagram below.



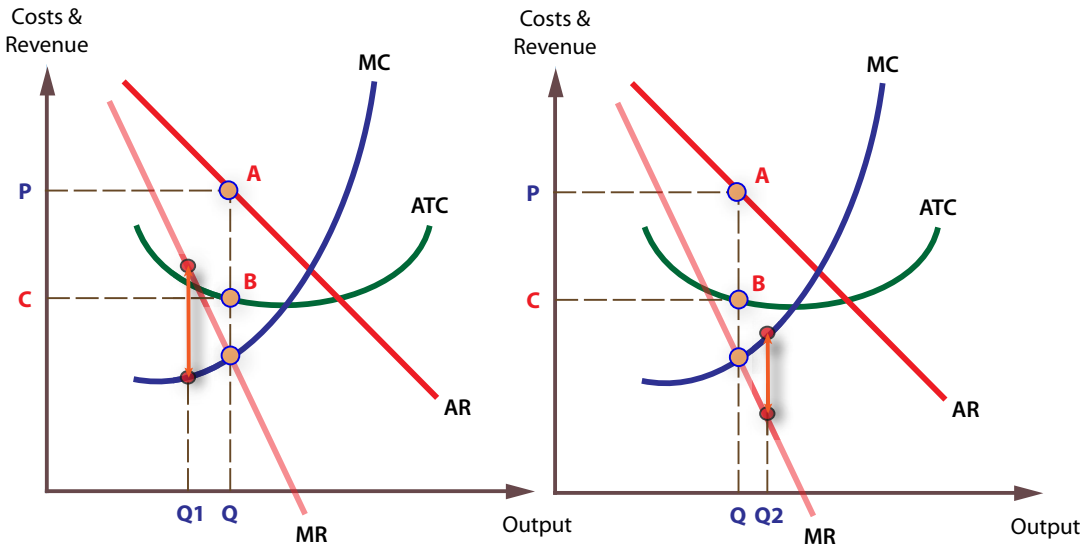
The value of super-normal profit is the profit per unit (verical distance A - B, times the quantity Q. As an area on the graph, it is P=AR, A, B, ATC.

If the firm stops short of producing Q, (at Q1 below) then MR is great than MC, and marginal profit is still greater than zero. Hence, the firm should increase output. If the firm produces greater than Q, (at Q2, below) MC is greater than MR, and marginal profit is negative. Hence, the firm should reduce its output. Only when MR = MC, at Q, will total profits be maximised.

2 Marginal revenue (MR) is the additional revenue from selling one extra unit. Given that the figures for output are in thousands, the change in total revenue must be divided by 1000. Similarly, the change in total cost is divided by 1000 to get the marginal cost per unit (MC).

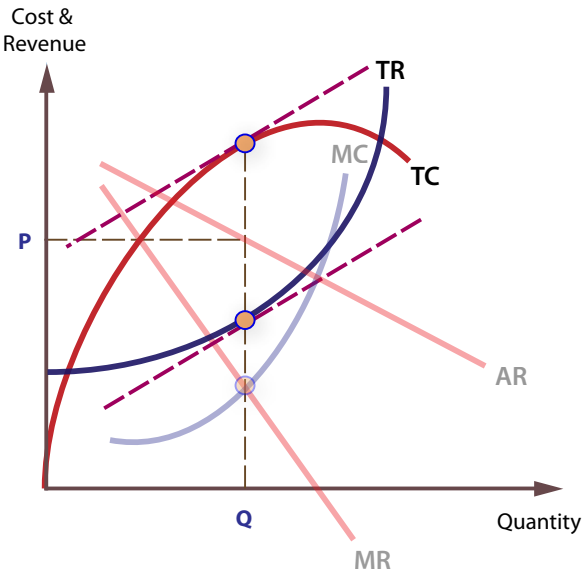
Conflicts of objectives

As can be seen from this example, profit max, revenue max and sales max all occur at different levels of output (6, 8 and 9 respectively).



Showing profit maximisation using total cost and revenue curves

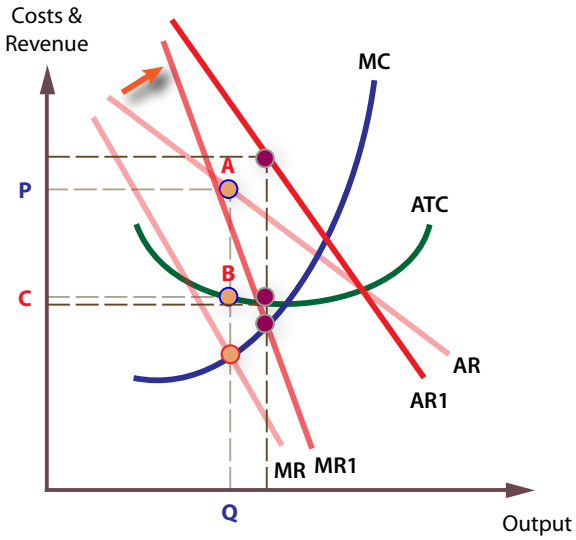
Profit maximisation is at Q because total revenue (TR) is the greatest it can be above total costs (TC). At this point, the gradient of the cost and revenue curves will be identical.



Not all firms are profit maximisers. Profit maximisation is the most likely objective for a firm whose owners are involved in day-to-day decision making, such as with small and medium sized enterprises (SMEs).

Competitiveness and profits

The fewer the firms in a market, the less competitive it is likely to be, and the greater the potential for super-normal profits.



The number of firms and profits

Assuming firms are selling substitute products, the effect of fewer firms is less competition, and this will reduce elasticity of demand. This creates steeper AR and MR curves, and increases the level of supernormal profits for each firm.

Distributed and retained profits

When profits are generated, they can be retained by the firm, or distributed to its owners. A government may provide tax incentives for those firms that retain their profits, and use them for investment. Distributed profits are generally subject to corporation tax.

Questions

1) Consider the following total costs and average revenue (price) for a small firm making boats.

Costs, revenue, and profits									
Q	Price (£000)	Fixed costs (£000)	Variable costs (£000)	Total costs (£000)	Average costs (ATC) (£000)	Marginal costs (MC) (£000)	Total Revenue (£000)	Marginal Revenue (£000)	Profit (TR - TC)
0	20			10					
1	18			16					
2	16			20					
3	14			28					
4	12			38					
5	10			50					
6	8			64					
7	6			88					
8	4			106					

a) Complete all the missing figures.

b) Plot ATC, MC, AR and MR.

c) What is the profit maximising output?

d) Show the area representing super-normal profits.

e) What is sales revenue maximising output?

f) Explain what happens to efficiency if the firm chooses to operate at profit maximisation.

2) Consider the following data for a small firm, which manufactures steel pressing machines for the motor vehicle industry.

Costs, revenue, and profits									
Q	**Price (£000)**	**Fixed costs (£000)**	**Variable costs (£000)**	**Total costs (£000)**	**Average costs (ATC) (£000)**	**Marginal costs (MC) (£000)**	**Total Revenue (£000)**	**Marginal Revenue (£000)**	**Profit (TR - TC)**
0	100		0	100					
1	90		40	140					
2	80		60	160					
3	70		70	170					
4	60		90	190					
5	50		120	220					
6	40		190	290					
7	30		260	360					
8	20		460	560					

a) Complete the missing figures. (10)

b) Plot ATC, MC, AR and MR. (10)

c) What is profit maximising output? (4)

d) Show the area representing super-normal profits on your graph. (4)

e) What is sales revenue maximising output? (4)

f) What is the most productively efficient output? (2)

g) Assume the firm chooses to operate at profit maximisation – evaluate this decision in terms of efficiency. (10)

h) It has recently been under threat from new entrants into the market. What could the firm do in response? (8)

i) What would be the impact of a 50% rise in variable costs on the firm's profitability? (4)

j) Show the impact of these changes on your graph. (4)

The firm's short run supply curve

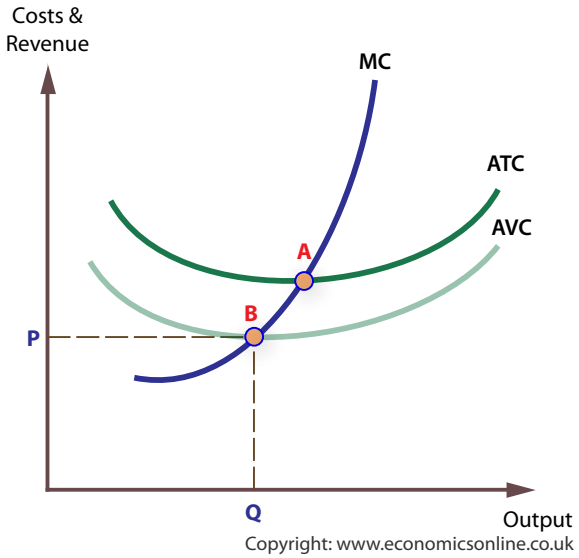
Understanding the nature of a firm's supply curve helps explain how price, output, revenue, and profits are determined. Neo-classical economic theory suggests that a firm's decision to supply in the long run is determined by whether it can cover all of its production and distribution costs. If a firm cannot cover all its costs in the long run it will, clearly, go out of business, often referred to as shutting down.

The importance of variable costs

However, economic theory also indicates that, in the short run, the firm does not need to cover all of its costs to carry on supplying. In the short run the firm needs only to cover its variable costs, at Q (B) in the diagram. This is largely because covering variable cost ensures than an output can be produced in the future. If variable costs cannot be covered then no further output can be made. In addition, fixed costs have already been paid for prior to any marginal decision to supply, so will not enter into the firm's short run calculations.

Given that the fixed costs are historic, the entrepreneur will be prepared to forgo a contribution to these costs in an attempt to keep the firm running. However, this cannot continue indefinitely, and unless all costs are covered, and the firm at least breaks-even, the firm will eventually shut down.

The firm will break-even (where total revenue = total cost) when price (AR) = ATC (at point A in the diagram).



The short run supply curve

The firm's short run supply curve is its marginal cost (MC) curve above its average variable cost (AVC).

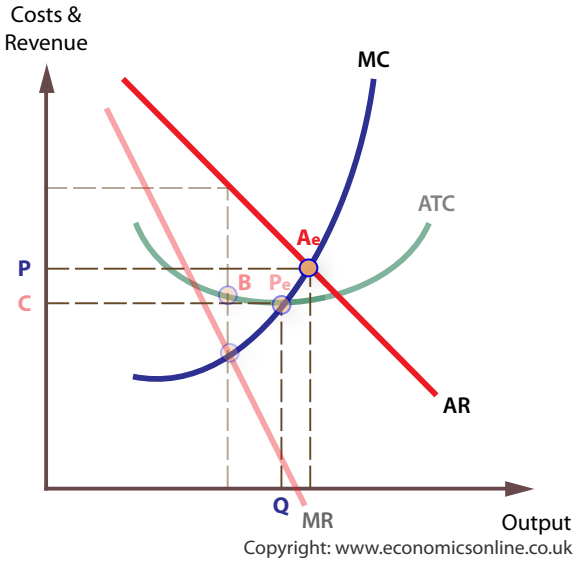
Efficiency

The importance of efficiency

Assessing the efficiency of firms is a powerful means of evaluating performance of firms, and the performance of markets and whole economies. There are several types of efficiency, including allocative and productive efficiency, technical efficiency, 'X' efficiency, dynamic efficiency and social efficiency

Allocative efficiency

Allocative efficiency occurs when consumers pay a market price that reflects the private marginal cost of production. The 'condition' for allocative efficiency for a firm is to produce an output where marginal cost, MC, just equals price, P.

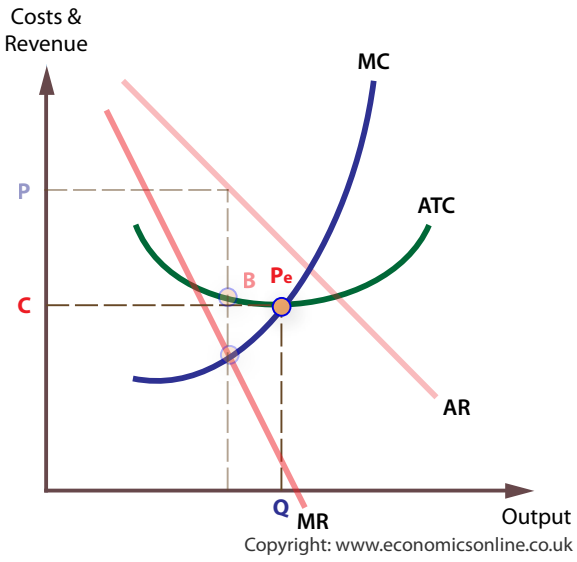


Productive efficiency

Productive efficiency occurs when a firm combines resources in such a way as to produce a given output at the lowest possible average total cost. Costs will be minimised at the lowest point on a firm's short run average total cost curve. This also means that ATC = MC, because MC always cuts ATC at the lowest point on the ATC curve.

Technical efficiency

Technical efficiency relates to how much output can be obtained from a given input, such as a worker or a machine, or a specific combination of inputs. Maximum technical efficiency occurs when output is maximised from a given quantity of inputs.



The simplest way to differentiate productive and technical efficiency is to think of productive efficiency in terms of cost minimisation by adjusting the 'mix' of inputs, whereas technical efficiency is output maximisation from a given mix of inputs.

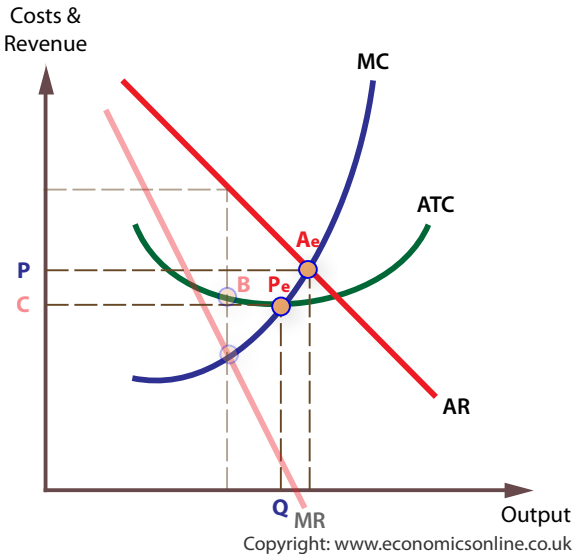
Identifying allocative and productive efficiency points

To identify which output a firm would produce, and how efficient it is, we need to combine data on both costs and revenue.

We can assume that most real firms face a downward sloping demand (AR) curve, and MR falls at twice the rate. Diagrammatically, productive efficiency occurs where ATC is at its lowest, and is equal to MC.

'X' efficiency

X' efficiency is a concept that was originally applied to management efficiencies by Harvey Leibenstein in the 1960s. The concept can be applied specifically to situations where there is more or less motivation of management to maximise output, or not.



Copyright: www.economicsonline.co.uk

'X efficiency occurs when the output of firms, from a given amount of input, is the greatest it can be. It is likely to arise when firms operate in highly competitive markets where managers are motivated to produce as much as possible.

When markets are less than perfectly competitive, as in the case of oligopolies and monopolies, there is likely to be a loss of 'X' efficiency, with output not being maximised due to a lack of managerial motivation.

Dynamic efficiency

The concept of dynamic efficiency is commonly associated with the Austrian Economist Joseph Schumpeter and means technological progressiveness and innovation.

Neo-classical economic theory suggests that when existing firms in an industry, the incumbents, are highly protected by barriers to entry they will tend to be inefficient. Schumpeter argued that this is not necessarily the case; indeed, firms that are highly protected are more likely to undertake risky innovation, and generate dynamic efficiency.

Firms can benefit from two types of innovation:

Process innovation

Process innovation occurs when new production techniques are applied to an existing product. For example, this is common in the production of motor vehicles where firms are continually looking to develop new methods and production processes.

Product innovation

Product innovation occurs when firms generate new or improved products. For example, this is common in many consumer product markets, including electronics and communications.

Social efficiency

Social efficiency exists when all the private and external costs and benefits are taken into account when producing an extra unit. Private firms only have an incentive to consider external costs if they are forced to internalise them through taxation or through the purchase of permit to pollute.

Knowledge and efficiency

Information failure is a type of inefficiency that can affect markets and firms in certain circumstances. There are various types of information failure.

The principal-agent problem

The principal-agent problem is associated with large firms, where ownership and control are in the hands of different people. The principal-agent problem can occur whenever owners of a firm appoint managers to make key decisions. The owners are the principals, and those appointed to run and manage the business are the agents. This separation causes asymmetric information, where the agents know more than the owners do, and this creates the need for owners to construct mechanisms to monitor and check the performance of agents. The problem develops because the owners and managers usually have different objectives, so the owners cannot trust the managers to act on their behalf, creating the need for constant checking. This leads to inefficiencies in terms of the need to employ checkers and complex monitoring systems.

The principal-agent problem can also occur in the public sector, where the government (as principals) appoint managers to undertake the day-to-day operations of publicly owned enterprises. Conflicts between agents and principals can frequently occur. For example, managers of the railway network may want to generate maximum revenue, whereas the government may want a safer railway system.

Solutions to the principal-agent problem

A firm can adopt a number of strategies to resolve the principal-agent problem, including:

1. Allocating shares to managers of a firm, so that they understand the shareholders' objectives, and are more likely to consider their view when making day-to-day decisions.

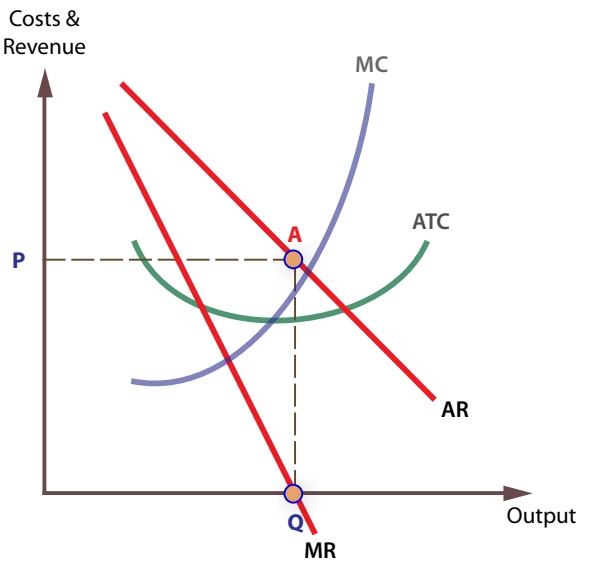
2. Using incentives tied to profits, such as with performance related pay.

Revenue and sales maximisation

Revenue maximisation

Maximising sales revenue is an alternative to profit maximisation and occurs when the marginal revenue, MR, from selling an extra unit is zero. The condition for revenue maximisation is, therefore, to produce up to the point where MR = 0.

This is also at the same level of output where PED = 1, namely at the mid-point of the average revenue/demand curve (point A in the diagram).

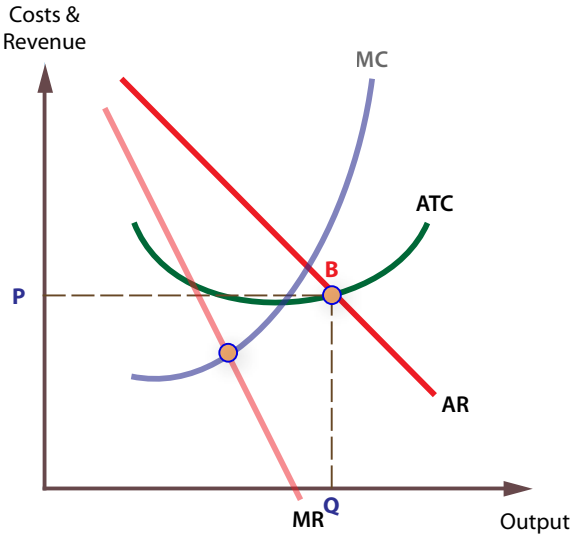


Sales maximisation

Sales maximisation is another goal and occurs when the firm sells as much as possible without making a loss. This occurs when average revenue (AR) = average revenue (ATC), at point B in the diagram.

Not-for-profit organisations may choose to operate at this level of output, as may profit making firms faced with certain situations, or employing certain strategies.

An example of this would be predatory pricing where, so long as costs are covered, a firm may reduce price to drive rivals out of the market. To the right of Q, the firm will make a loss, and to the left of Q sales are not maximised.



Questions

1) Using diagrams, contrast the sales revenue maximisation position for a firm with a downward sloping AR/D curve with the sales maximisation and profit maximisation positions.

2) Using a diagram, explain why, when MR = 0, TR must be maximised.

Competition and market structures

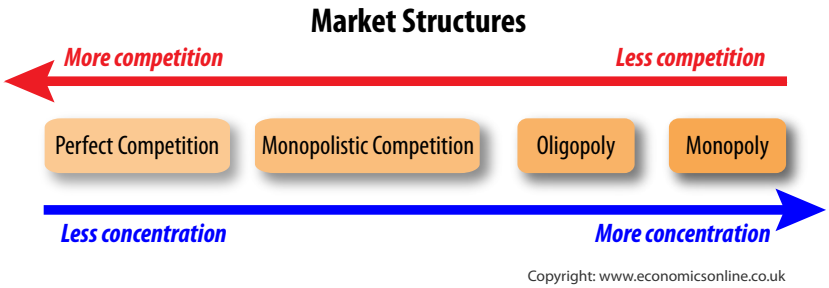
There are several market structures in which firms can operate. The type of structure influences the firm's behaviour, whether it is efficient, and the level of profits it can generate. Neo-classical theory of the firm distinguishes a number of market structures, each with its own characteristics and assumptions. The structure of a market refers to the number of firms in the market, their market shares, and other features that affect the level of competition in the market. Market structures are distinguished mainly by the level of competition that exists between the firms in the particular market.

Competitive structure vs competitive behaviour

As well as considering market structures, modern theory also looks at the behaviour, or conduct of firms, their performance, and the level of 'contestability' in the market. A market might have an uncompetitive structure, with only a small number of firms competing, but the behaviour of firms might be highly competitive, as is the case in the UK with the supermarket sector.

Market structures

In neo-Classical theory, market structures are classified in term of the presence or absence of competitors. When there are few competitors, the market is said to be uncompetitive, with market power 'concentrated' in the hands of the few firms that exist. There is a spectrum of structure, from 'perfect' competition to 'pure' monopoly.



Perfect competition

A perfectly competitive market is a hypothetical market where competition is at its highest possible level. Neo-classical economists argued that perfect competition would produce the 'best' possible outcomes for consumers and society.

Key characteristics

Perfectly competitive markets exhibit the following characteristics:

Perfect knowledge

There is perfect knowledge, with no information failure or time lags. Knowledge is freely available to all participants, which means that risk-taking is minimal and the role of the entrepreneur is limited.

No barriers

There are no barriers to entry into or exit out of the market.

Homogeneous products

Firms produce homogeneous, identical, units of output that are not branded.

Homogeneous inputs

Each unit of input, such as units of labour, are also homogeneous.

The firm is a price taker

No single firm can influence the market price, or market conditions. The single firm is said to be a price taker, taking its price from the whole industry.

Large number of firms

There are a very large numbers of firms in the market.

No need for government

There is no need for government regulation, except to make markets more competitive.

No externalities

There are assumed to be no externalities, that is, no external costs or benefits.

Normal profits in the long run

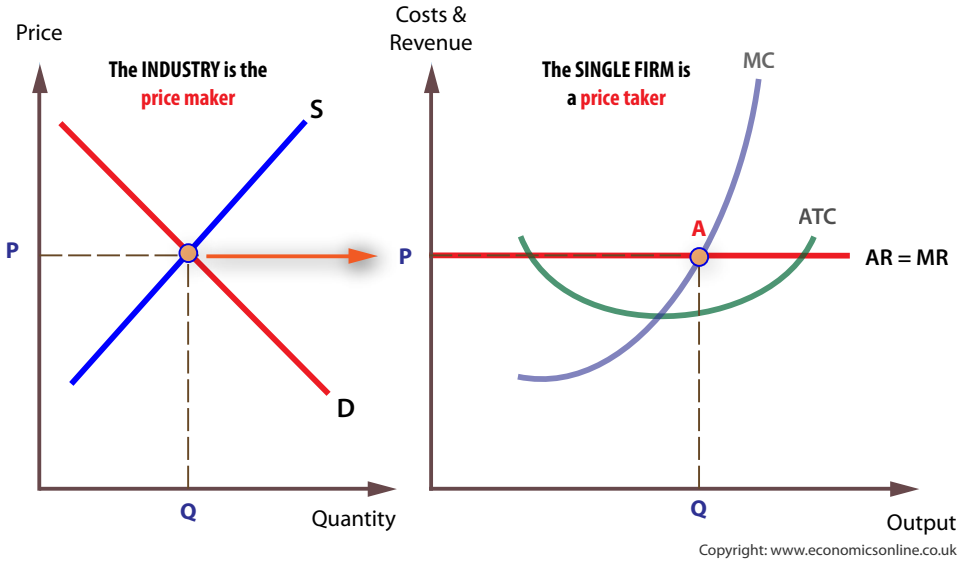
Firms can only make 'normal' profits in the long run, but they can make abnormal profits in the short run.

Leaving does not affect the market

If a firm leave the market they do not take important knowledge or technology with them.

The firm as price taker

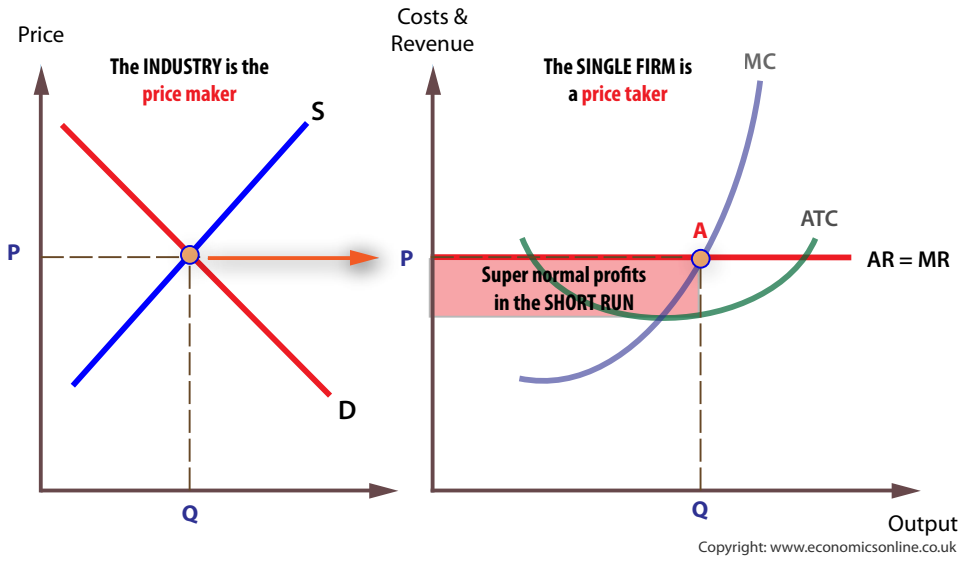
The single firm takes its price from the industry and is, consequently, referred to as a *price taker*. The industry is composed of all firms in the industry and the market price is where market demand is equal to market supply. Each single firm must charge this price and cannot diverge from it.



Equilibrium in perfect competition

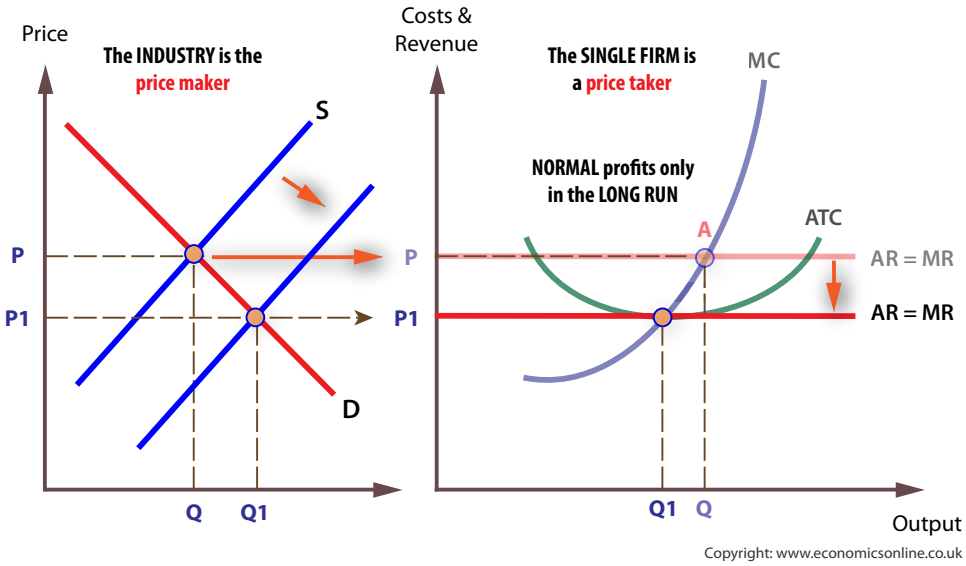
In the short run

Under perfect competition, firms can make super-normal profits or losses.



In the long run

However, in the long run firms are attracted into the industry if the incumbent firms are making supernormal profits. This is because there are no barriers to entry and because there is perfect knowledge. The effect of this entry into the industry is to shift the industry supply curve to the right, which drives down price until the point where all supernormal profits are gone. If firms are making losses, they will leave the market as there are no exit barriers, and this will shift the industry supply to the left, which raises price and enables those left in the market to derive normal profits.



The super-normal profit derived by the firm in the short run acts as an incentive for new firms to enter the market, which increases industry supply and market price falls for all firms until only normal profit is made.

The advantages of perfect competition

It can be argued that perfect competition will yield the following benefits:

No information failure

Because there is perfect knowledge, there is no information failure and knowledge is shared evenly between all participants.

No monopoly power

There are no barriers to entry, so existing firms cannot derive any monopoly power.

Opportunity cost is covered

Only normal profits made, so producers just cover their opportunity cost.

No need for advertising

There is no need to spend money on advertising, because there is perfect knowledge and firms can sell all they can produce. In addition, selling unbranded goods makes it hard to construct an effective advertising campaign.

Maximum welfare

There is maximum possible consumer surplus and economic welfare.

Maximum efficiency

There is maximum allocative and productive efficiency, where P = MC and MC = ATC.

Maximum choice

There is maximum choice for consumers.

How realistic is the model?

Very few markets or industries in the real world are perfectly competitive. For example, how homogeneous is the output of real firms, given that even the smallest of firms working in manufacturing or services try to differentiate their product. Most of the other assumptions, such as perfect knowledge and no barriers to entry are equally unrealistic, as is the assumption of no externalities.

Although unrealistic, it is still a useful model in two respects. Firstly, many primary and commodity markets, such as coffee and tea, exhibit many of the characteristics of perfect competition, such as the number of individual producers that exist, and their inability to influence market price. Secondly, for other markets in manufacturing and services, the model is a useful yardstick by which economists and regulators can evaluate levels of competition that exist in real markets.

Questions

1. If all firms operated in 'perfectly competitive markets' what benefits would be generated? Use a diagram to support your answer.

2. 'A firm under perfect competition is faced with a perfectly elastic demand curve'. Do you agree? Justify your answer.

3. Mythica is an economy that used to operate under a military dictatorship, but has recently reformed and encouraged 'free market capitalism' and free trade between itself and the rest of the world. Assume that the global market for MP3 players is currently perfectly competitive, with the global price being exactly $100. A new firm starts up in Mythica selling MP3 players, and, so far, it is the only supplier. It currently sells its MP3 players for $150.

4. Using diagrams, explain what is likely to happen to the market for MP3 players in Mythica, including a comment on prices, output and profits over the next 5 years

Monopolistic competition

The model of monopolistic competition describes a common market structure in which firms have many competitors, but each one sells a slightly different product. Monopolistic competition as a market structure was first identified in the 1930s, by American economist Edward Chamberlin, and English economist Joan Robinson.

Many small businesses operate under conditions of monopolistic competition, including independently owned and operated high-street stores and restaurants. In the case of restaurants, each one offers something different and possesses an element of uniqueness, but all are essentially competing for the same customers.

Key characteristics

Monopolistically competitive markets exhibit the following characteristics:

Independent decision making

Each firm makes independent decisions about price and output, based on its product, its market, and its costs of production.

Large numbers of competitors

There are usually a large number of independent firms competing in the market.

Good knowledge

Knowledge is widely spread between participants, but it is unlikely to be perfect. For example, diners can review all the menus available from restaurants in a town, before they make their choice. Once inside the restaurant, they can view the menu again, before ordering. However, they cannot fully appreciate the restaurant or the meal until after they have dined.

Important role for entrepreneur

The entrepreneur has a more significant role than in firms that are perfectly competitive because of the increased risks associated with decision-making.

Low barriers to entry and exit

There is freedom to enter or leave the market, as there are no major barriers to entry or exit.

Product differentiation

A central feature of monopolistic competition is that products are differentiated. There are four main types of differentiation:

1. ***Physical product differentiation***, where firms use size, design, colour, shape, performance, and features to make their products different. For example, consumer electronics can easily be physically differentiated.

2. ***Marketing differentiation***, where firms try to differentiate their product by distinctive packaging and other promotional techniques. For example, breakfast cereals can easily be differentiated through packaging.

3. ***Human capital differentiation***, where the firm creates differences through the skill of its employees, the level of training received, distinctive uniforms, and so on.

4. ***Differentiation through distribution***, including distribution via mail order or through internet shopping, such as Amazon.com, which differentiates itself from traditional bookstores by selling online.

Firms are price makers

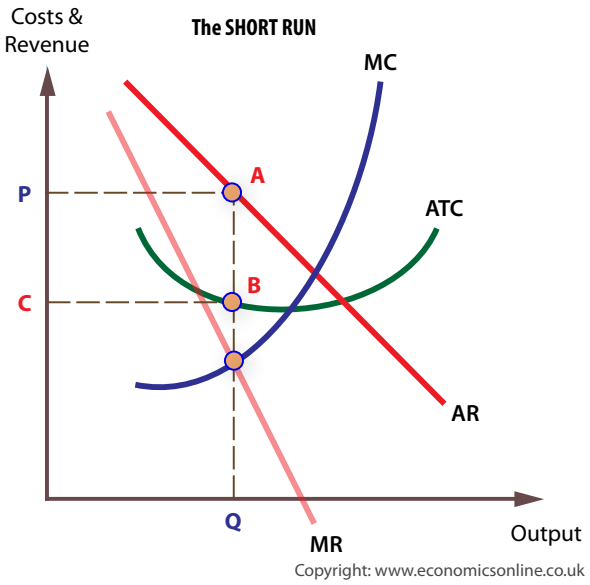
Firms are price makers and are faced with a downward sloping demand curve. Because each firm makes a unique product, it can charge a higher or lower price than its rivals. The firm can set its own price and it does not have to 'take' it from the industry as a whole, though the industry price may be a guideline, or act as a constraint. This also means that the demand curve will slope downwards.

Advertising

Firms operating under monopolistic competition usually have to engage in advertising. Firms are often in fierce competition with other (local) firms offering a similar product or service, and may need to advertise on a local basis, to inform customers of their differences. Common methods of advertising are local press and radio, local cinema, posters, leaflets and special promotions.

Profit maximisation

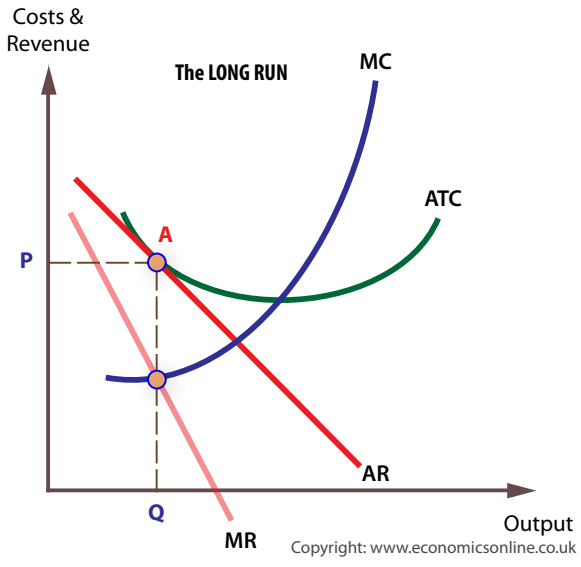
Monopolistically competitive firms are assumed to be profit maximisers because firms tend to be small with entrepreneurs actively involved in managing the business.



Equilibrium under monopolistic competition

In the short run, supernormal profits are possible, but in the long run new firms are attracted into the industry, because of low barriers to entry, good knowledge and an opportunity to differentiate.

As new firms enter the market, demand for the existing firm's products becomes more elastic and the demand curve shifts to the left, driving down price. Eventually, all super-normal profits are eroded away. Clearly, the firm benefits most when it is in its short run and will try to stay in the short run by innovating and further product differentiation.



Examples of monopolistic competition

Examples of monopolistic competition can be found in every high street. Monopolistically competitive firms are most common in industries where differentiation is possible, such as in the restaurant business, hotels and pubs, general specialist retailing and consumer services, such as hairdressing. Hairdressing salons can differentiate their service in several ways, including opening hours, male/female/unisex, ability to book ahead and, of course, the different personalities, skills and abilities of their individual hairdressers.

The survival of small firms

The existence of monopolistic competition partly explains the survival of small firms in modern economies. The majority of small firms in the real world operate in markets that could be said to be monopolistically competitive.

The advantages of monopolistic competition

Monopolistic competition can bring the following advantages:

Contestability

There are no significant barriers to entry, therefore markets are relatively contestable and firms find it easy to enter and leave the market.

Choice

Differentiation creates diversity, choice, and utility. For example, a typical high street in any town will have a number of different restaurants from which to choose.

Efficiency

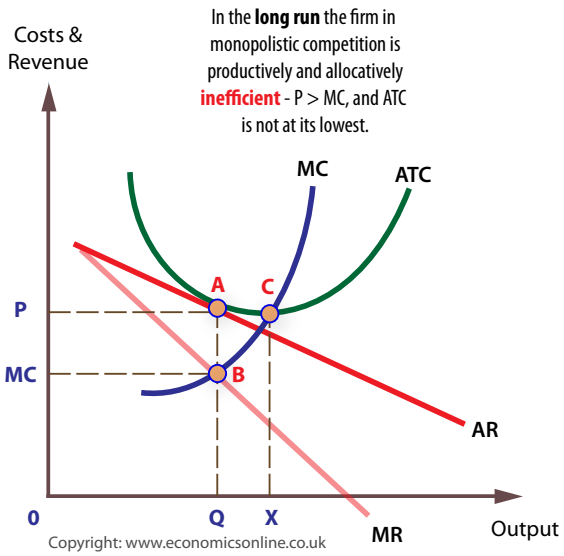
The market is more efficient than monopoly but less efficient than perfect competition - less allocatively and less productively efficient. However, they may be dynamically efficient, innovative in terms of new production processes or new products. For example, retailers often constantly have to develop new ways to attract and retain local custom.

The disadvantages of monopolistic competition

There are several potential disadvantages associated with monopolistic competition, including:

Waste

Some differentiation does not create utility but generates unnecessary waste, such as excess packaging. In addition, advertising can be considered wasteful, though most is 'informative' rather than 'persuasive'.



Allocative inefficiency

As the diagram illustrates, there is allocative inefficiency in both the long and short run, assuming profit maximisation, because price is above marginal cost in both cases. In the long run the firm is less allocatively inefficient, but it is still inefficient.

Excess capacity and productive inefficiency

There is a tendency for excess capacity because firms can never fully exploit their fixed factors, as mass production is difficult. This means that they are productively inefficient in both the long and short run. However, this may be outweighed by the advantages of diversity and choice.

Questions

1. To what extent are hairdressers a good example of firms operating under monopolistic competition?

2. Which of the following markets/industries are closest to the economist's definition of monopolistic competition - you may select more than one:

a. Airlines

b. Petrol retailers

c. Plumbers

d. Independent electrical retailers

e. A coffee grower in Brazil

f. An estate agent

g. Electricity generation

h. Retail banks

3. Draw separate diagrams to show a firm under monopolistic competition making supernormal profits in the short run and normal profits in the long run.

4. Evaluate monopolistic competition as a market structure.

5. A firm in long run equilibrium under monopolistic competition will exhibit (select just *one* correct answer)

a. Allocative but not productive efficiency

b. Productive but not allocative efficiency

c. Neither productive nor allocative efficiency

d. Super-normal profits

e. Both allocative and productive efficiency

Oligopoly

Defining and measuring oligopoly

An oligopoly is a market structure in which a few firms dominate. When a market is shared between a few firms, it is said to be highly concentrated. Although only a few firms dominate, it is possible that many small firms may also operate in the market. For example, major airlines like British Airways (BA) and Air France operate their routes with only a few close competitors, but there are also many small airlines catering for the holidaymaker or offering specialist services.

Concentration ratios

Oligopolies may be identified using concentration ratios, which measure the proportion of total market share controlled by a given number of firms. When there is a high concentration ratio in an industry, economists tend to identify the industry as an oligopoly.

Example of a hypothetical concentration ratio

The following are the annual sales, in £m, of the six firms in a hypothetical market: A = 56, B = 43, C = 22, D = 12, E = 3, and F = 1. In this hypothetical case, the 3-firm concentration ratio is 88.3%. That is 56 + 43 + 22/137 x 100, which is 121/137 x 100.

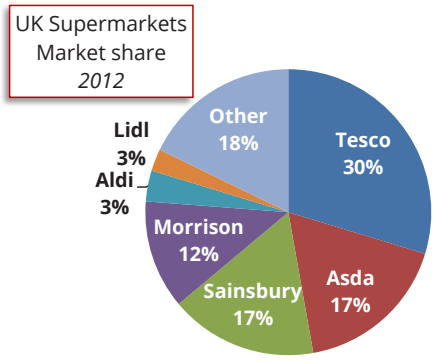
Examples

Fixed line rental

While there are around 170 telephone service suppliers in the UK[3] the fixed-line market is dominated by two main suppliers, BT and Virgin Media, with a 3-firm concentration ratio for fixed-line telephone supply of 89% in 2006.

UK supermarkets

The UK supermarket sector is dominated by just four firms:



The Herfindahl – Hirschman Index (H-H Index)

This is an alternative method of measuring concentration, and for tracking changes in the level of concentration following mergers. The formula is (H–H) = X2 + Y2 + Z2; where X, Y and Z is the percentage of the top three firm's market shares.

If the index is below 1000, the market is not considered concentrated, while an index above 2000 indicates a concentrated market or industry – the higher the figure the greater the concentration. Mergers between oligopolists increase concentration and monopoly power and are likely to be the subject of regulation.

3 Source: simplyswitch.com,

Key characteristics

The main characteristics of firms operating in a market with few close rivals include:

Interdependence

Firms operating under conditions of oligopolistic competition are said to be ***interdependent*** , which means they cannot act independently of each other. A firm operating in a market with just a few competitors must take the potential reaction of its closest rivals into account when making its own decisions. For example, if a petrol retailer like Texaco wishes to increase its market share by reducing price, it must take into account the possibility that close rivals, such as Shell and BP, may reduce their price in retaliation. An understanding of game theory and the prisoner's dilemma helps appreciate the concept of interdependence.

Strategy

Strategy is extremely important to firms that are interdependent. Because firms cannot act independently, they must anticipate the likely response of a rival to any given change in their price, or their non-price activity. In other words, they need to plan, and work out a range of possible options based on how they think rivals might react.

Oligopolists have to make critical strategic decisions, such as:

1. Whether to compete with rivals, or collude with them.

2. Whether to raise or lower price, or keep price constant.

3. Whether to be the first firm to implement a new strategy, or whether to wait and see what rivals do. The advantages of 'going first' or 'going second' are called '1st-mover' and '2nd-mover' advantage. Sometimes it pays to go first because a firm can generate ***head-start profits***. 2nd mover advantage occurs when it pays to wait and see what new strategies are launched by rivals, and then try to improve on them or find ways to undermine them.

Barriers to entry

Oligopolies and monopolies frequently maintain their position of dominance in a market because it is too costly or difficult for potential rivals to enter the market. These difficulties are called barriers to entry. Barriers can be erected deliberately by the incumbent(s), or they can exploit barriers that naturally exist in the market.

Natural entry barriers include:

Economies of large-scale production.

If a market has significant economies of scale that have already been exploited by the incumbents, new entrants are deterred.

Ownership or control of a key scarce resource.

Owning scarce resources that other firms would like to use creates a considerable barrier to entry, such as an airline controlling access to an airport.

High set-up costs.

High set-up costs deter initial market entry, because they increase break-even output, and delay the possibility of making profits. Many of these costs are sunk costs, which are costs that cannot be recovered when a firm leaves a market, and include marketing and advertising costs and other fixed costs.

High R&D costs (research and development)

When firms spend money on research and development (R & D), it is often a signal to potential

entrants that they have large financial reserves. In order to compete, new entrants will have to match, or exceed, this level of spending in order to compete in the future. This deters entry and is widely found in oligopolistic markets, such as pharmaceuticals and chemicals.

Artificial barriers include:

Predatory pricing.

A firm may deliberately lower price to try to force rivals out of the market.

Limit pricing.

Limit pricing means the incumbent firm sets a low price, and a high output, so than entrants cannot make a profit at that price. This is best achieved by selling at a price just below the average total costs (ATC) of potential entrants. This signals to potential entrants that profits are impossible to make. The incumbent is exploiting its superior knowledge of the market, and production costs, for its own advantage.

Superior knowledge

An incumbent may, over time, have built up a superior level of knowledge of the market, its customers, and its production costs. This superior knowledge can deter entrants into the market.

Predatory acquisition

Predatory acquisition involves taking-over a potential rival by purchasing sufficient shares to gain a controlling interest, or by a complete buy-out. As with other deliberate barriers, regulators, like the Competition Commission, may prevent this because it is likely to reduce competition.

Advertising

Advertising is another sunk cost - the more that is spent by incumbent firms the greater the deterrent to new entrants.

A strong brand

A strong brand creates loyalty, 'locks in' existing customers, and deters entry.

Loyalty schemes

Schemes such as Tesco's Club Card, help oligopolists retain customer loyalty and deter entrants who need to gain market share.

Exclusive contracts, patents and licences

Contracts, patents, and licences make entry difficult as they protect existing firms who have won the contract, or who own the license or hold the patent. For example, contracts between specific suppliers and retailers can exclude other retailers from entering the market.

Vertical integration

Vertical integration can 'tie up' the supply chain and make life difficult for potential entrants, such as an electronics manufacturer like Sony having its own retail outlets (Sony Centres), and a brewer like Heineken owning its own chain of UK pubs, which it acquired from the brewers Scottish and Newcastle in 2008.

Collusive oligopolies

Another key feature of oligopolistic markets is that firms may attempt to collude, rather than compete. If colluding, participants act like a monopoly and can enjoy the benefits of higher profits over the long term.

Types of collusion

Overt

Overt collusion occurs when there is no attempt to hide agreements, such as the when firms form trade associations like the Association of Petrol Retailers.

Covert

Covert collusion occurs when firms try to hide the results of their collusion, usually to avoid detection by regulators, such as when fixing prices.

Tacit

Tacit collusion arises when firms act together, called acting in concert, but where there is no formal or even informal agreement. For example, it may be accepted that a particular firm is the price leader in an industry, and other firms simply follow the lead of this firm. All firms may 'understand' this, but no agreement or record exists to prove it. If firms do collude, and their behaviour can be proven to result in reduced competition, they are likely to be subject to regulation. In many cases, tacit collusion is difficult or impossible to prove, though regulators are becoming increasingly sophisticated in developing new methods of detection.

Competitive oligopolies

When competing, oligopolists prefer non-price competition in order to avoid price wars. A price reduction may achieve strategic benefits, such as gaining market share, or deterring entry, but the danger is that rivals will simply reduce their prices in response. This leads to little or no gain, but can lead to falling revenues and profits. Hence, a far more beneficial strategy may be to undertake non-price competition.

Pricing strategies of oligopolies

Oligopolies may pursue the following pricing strategies:

Predatory pricing

Oligopolists may use predatory pricing to force rivals out of the market. This means keeping price artificially low, and often below the full cost of production.

Limit pricing

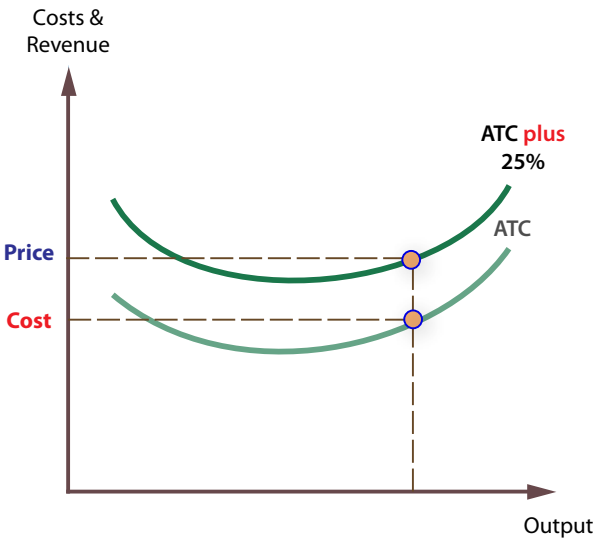
They may also operate a limit-pricing strategy to deter entrants, which is also called entry forestalling price.

Colluding over price

Oligopolists may collude with rivals and raise price together, but this may attract new entrants.

Cost-plus pricing

Cost-plus pricing is a straightforward pricing method, where a firm sets a price by calculating average production costs and then adding a fixed mark-up to achieve a desired profit level. Cost-plus pricing is also called rule of thumb pricing.



There are different versions of cost-pus pricing, including full cost pricing, where all costs - that is, fixed and vari-

able costs - are calculated, plus a mark up for profits, and contribution pricing, where only variable costs are calculated with precision and the mark-up is a contribution to both fixed costs and profits.

Cost-plus pricing is very useful for firms that produce a number of different products, or where uncertainty exists. It has been suggested that cost-plus pricing is common because a precise calculation of marginal cost and marginal revenue is difficult for many oligopolists. Hence, it can be regarded as a response to information failure. Cost-plus pricing is also common in oligopoly markets because it is likely that the few firms that dominate may often share similar costs, as in the case of petrol retailers.

However, there is a risk with such a rigid pricing strategy as rivals could adopt a more flexible discounting strategy to gain market share.

Non-price strategies

Non-price competition is the favoured strategy for oligopolists because price competition can lead to destructive price wars, with examples including:

Quality

Trying to improve quality and after sales servicing, such as offering extended guarantees, is a key option for oligopolists.

Advertising

Spending on advertising, sponsorship and product placement - also called hidden advertising – is very significant to many oligopolists. The UK's football Premiership has long been sponsored by firms in oligopolies, including Barclays Bank and Carling.

Promotions

Sales promotion, such as buy-one-get-one-free (BOGOF), is associated with the large supermarkets, which is a highly oligopolistic market, dominated by three or four large chains.

Schemes

Loyalty schemes, which are also common in the supermarket sector, such as Sainsbury's Nectar Card and Tesco's Club Card.

Each strategy can be evaluated in terms of how successful is it likely to be, will rivals be able to copy the strategy, will the firms get a '1st - mover' advantage, and how expensive is it to introduce the strategy. If the cost of implementation is greater than the pay-off, clearly it will be rejected. In addition, it is important for oligopolists to assess how long it will take to work. A strategy that takes five years to generate a pay-off may be rejected in favour of a strategy with a quicker pay-off.

Price stickiness

The theory of oligopoly suggests that, once a price has been determined, it will 'stick' at this price. This is largely because firms cannot pursue independent strategies. For example, if an airline raises the price of its tickets from London to New York, rivals will not follow suit and the airline will lose revenue: the demand curve for the price increase is relatively elastic. Rivals have no need to follow suit because it is to their competitive advantage to keep their prices as they are.

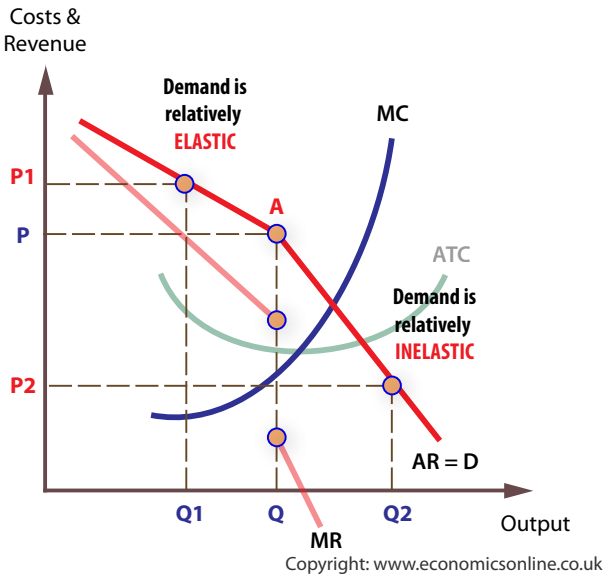
However, if the airline lowers its price, rivals would be forced to follow suit and drop their prices in response. Again, the airline will lose sales revenue and market share. The demand curve is relatively inelastic in this context.

The kinked demand curve

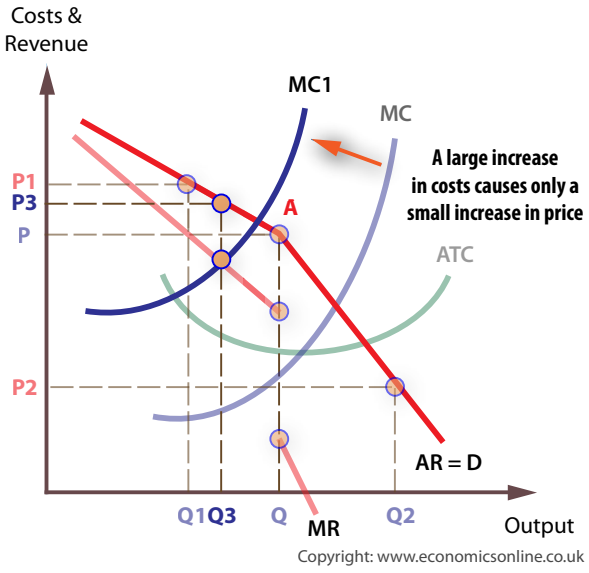
The reaction of rivals to a price change depends on whether price is raised or lowered. The elasticity of demand, and hence the gradient of the demand curve, will be different. The demand curve will be 'kinked', at the 'current' price, **P**.

Even when costs rise, price does not rise in line given the elastic nature of demand.

If marginal revenue and marginal costs are added to the diagram, it is possible to show that profits will also be maximised at price P. Profits will always be maximised when MC = MR, and so long as MC cuts MR in its vertical portion then profit maximisation is still at P. Furthermore, if MC changes in the vertical portion, of the MR curve, price still sticks at P. Even when MC moves out of the vertical portion, the effect on price is minimal, and consumers will not gain the benefit of any cost reduction.



Copyright: www.economicsonline.co.uk



Copyright: www.economicsonline.co.uk

A game theory approach

Pricing strategies can also be looked at in terms of game theory; that is, in terms of strategies and payoffs. There are three possible price strategies, with different pay-offs and risks:

- Raise price
- Lower price
- Keep price constant

The choice of strategy will depend upon the pay-offs, which depends upon the actions of competitors. Raising price or lowering price could lead to a beneficial pay-off, but both strategies can lead to losses, which could be potentially disastrous. In short, changing price is too risky to undertake. Therefore, although keeping price constant will not lead to the single best outcome, it may be the least risky strategy for an oligopolist.

Game theory also predicts that there is a tendency for cartels to form because co-operation is likely to be highly rewarding. Co-operation reduces the uncertainty associated with the mutual interdependence of rivals in an oligopolistic market. While cartels are 'unlawful' in most countries, they may still operate, with members concealing their unlawful behaviour.

Game theory evolved during the 20th Century to try to understand all kinds of games, from chess to war scenarios. Game theory has been widely applied to the behaviour of producers with few or only one competitor.

What is a game?

All games have the following:

1. Rules, which govern conduct of the players
2. Pay-offs, such as win, lose or draw
3. Strategies, which influence the decision making process.

In applying game theory to the behaviour of firms, we can suggest that firms face a number of strategic choices which govern their ability to achieve a desired pay-off, including decisions on price and output, such as whether to raise lower or hold prices. There are also likely to be decisions on products, such as whether to keep existing products or develop new ones.

In terms of pay-offs arising from the choice of strategy, there are a number of alternatives, including more profits for shareholders, greater market share, improved chances of survival, squeezing out a rival , and avoiding regulation.

The Prisoner's Dilemma

The Prisoner's Dilemma is a classic game that illustrates the choices facing oligopolies. As you read the scenarios, you can play the role of one of the prisoners.

The scenario

Robin and Tom are prisoners and they have been arrested for a petty crime, of which there is good evidence of their guilt. If found guilty they will each receive a 2 year prison sentence. During the interview the police officer becomes suspicious that the two prisoners are also guilty of a serious crime, but is not sure he has any evidence. Robin and Tom are placed in separate rooms and cannot communicate with each other. The police officer tries to get them to confess to the serious crime by offering them some options, with possible pay-offs.

Evaluation The options

Each is told that if they both confess to the serious crime they will receive a sentence of 3 years. However, each is also told that if he confesses and his partner does not, then he will get a short prison sentence of 1 year, and his partner will get a long sentence of 10 years. They know that if they both deny the serious offence they are certain to be found guilty of the lesser offence, and will get a 2-year sentence.

The pay-off matrix

Robin \ Tom	Confess	Deny
Confess	3 yr / 3 yr	1 yr / 10 yr
Deny	10 yr / 1 yr	2 y / 2 y

What would you do if you were one of them? Give an answer before you read on.

The dilemma is that their own 'pay-off' is wholly dependent on the behaviour of the other

prisoner. To avoid the worse-case scenario (10 years), the safest option is to confess and get 3 years. If collusion is possible they can both agree to deny (and get 2 years), but there is a very strong incentive to cheat because, if one denies and the other confesses, the best outcome of all is possible - that is 1 year. Fearing that the other may cheat, the safest option is to confess.

Maximin and Maximax

There are two basic strategies that could be followed by prisoners in this situation, or by firms operating in an uncertain oligopolistic market – the pessimistic ***maximin*** strategy, where the player makes a choice which will give the 'best of the worst' pay-offs, and the optimistic ***maxi-max*** strategy, which will give the 'best of the best' payoff.

In the case of the Prisoner's Dilemma, the worst pay-off for Robin if he confesses is 3 years, and the worst for denying is 10 years, hence the maximin strategy is to confess. The best pay-off for confessing is 1 year, and for denying is 2 years, hence the maximax strategy is also to confess. In this case confession is a dominant strategy because both maximin and maximax point to the same option. Confessing is also ***Nash equilibrium*** because once it is chosen any change will make the prisoner worse off. Nash equilibrium is the best available option taking the other player's likely decision into account.

Nash equilibrium

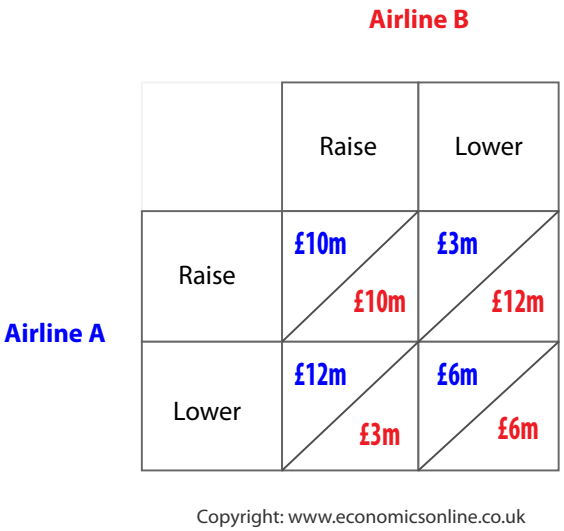
Nash equilibrium, named after Nobel winnng economist, John Nash, is a solution to a game involving two or more players who want the best outcome for themsleves and must take the actions of others into account. When Nash equilibrium is reached, players cannot improve their payoff by independently changing their strategy. This means that it is the best strategy assuming the other has chosen a strategy and will not change it.

Example - Airlines

The following example can be used to illustrate maximin and maximax strategies for firms operating under oligopolistic conditions. The pay-offs are in terms of profits.

In the case of airline A, the worst payoff for raising price is £3m (£10m and £3m), and the worst payoff from ***lowering price*** is £6m (£12m and £6m), so lowering is the *best of the worst* (**£6m** for lowering and £3m for raising).

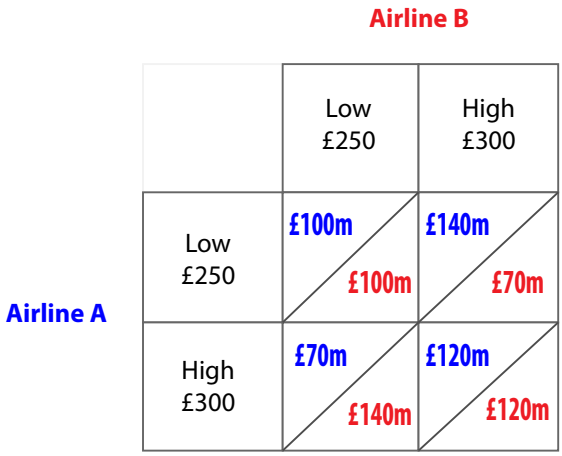
For the optimistic maximax strategy, the best pay-off for raising is £10m, and from lowering is £12m, hence the *best of the best* is also achieved by ***lowering price***, at **£12m**. In this case, lowering is both maximin and maximax, and it is a ***dominant strategy*** and is likely to be chosen by both airlines.



It is also Nash equilibrium because once it is chosen any change will make the firm worse off. Of course, lowering is not the best single option for both, as they could each get a higher pay-off of £10m by raising price together. This would only work through collusion, and, as it is regarded as anti-competitive behaviour, it runs the risk of an investigation by the OFT or Competition Commission. Even collusion might not work as there is the risk of the other firm(s) ratting on the agreement. If one party elects to be a whistle blower they may become exempt from any penalties imposed for cartel-like behaviour.

Consider a second hypothetical example of the two Airlines, and return ticket prices to New York.

In this case, the maximax strategy for airline A would be operate as a low cost-low price airline. At a low price the best of the best is £140m. The maximin strategy also to lower price (the best of the worst is £100m). The dominant strategy is to reduce price, and this is also Nash equilibrium.



Implications of game theory

Game theory provides many insights into the behaviour of oligopolists. For example, it indicates that generating rules for behaviour may take many of the risks out of competition, such as:

1. Employing a simple ***cost-plus pricing*** method that is shared by all participants. This would work well in situations where oligopolists share similar or identical costs, such as with petrol retailing.

2. Implicitly agreeing a ***price leader*** with other firms as followers. In the Airline example, firm A may lead and raise price, with B passively following suit. In this case, both would generate revenues of £120.

3. Supermarkets implicitly agreeing some lines where ***price cutting*** will take place, such as bread or baked beans, but keeping price constant for most lines.

4. Generally keeping prices ***stable*** to avoid price retaliation.

Cartels

A cartel is a grouping of producers that work together to protect their interests. Cartels are created when a few large producers decide to co-operate with respect to aspects of their market. Once formed, cartels can fix prices for members, so that competition on price is avoided. In this case cartels are also called price rings. They can also restrict output released onto the market, such as with OPEC and oil production quotas, and set rules governing other aspects of the behaviour of members. Setting rules is especially important in oligopolistic markets, as predicted in game theory. A significant attraction of cartels to producers is that they set rules that members follow, thus reducing risks that would exist without the cartel.

The negative effects on consumers include:

1. ***Higher prices*** - cartel members can all raise prices together, which reduces the elasticity of demand for any single member.

2. ***Lack of transparency*** - members may agree to hide prices or withhold information, such as the hidden charges in credit card transactions.

3. ***Restricted output*** - members may agree to limit output onto the market, as with OPEC and its oil quotas.

4. ***Carving up a market*** - cartel members may collectively agree to break up a market into regions or territories and not compete in each other's territory.

When are cartels most powerful?

They are at their most powerful when there are high barriers to entry into the market or industry, and when all members can be 'policed' by a dominant member.

Cartel-like behaviour

Some firms may act as though there is a cartel, and undertake cartel-like behaviour, even though there is no formal cartel, and this may be subject to investigation by the regulators.

Complex monopolies

When firms collude and act as though they are a single firm, but where collusion is tacit, a situation of complex monopoly is said to exist.

Example - The Siemens led electronic equipment cartel

In January 2007, the European Commission imposed a record fine of €750m on 11 European power equipment firms, led by the German firm Siemens. The Commission argued that Siemens, along with 10 other firms, had 'carved-up' the European power equipment market between 1988 and 2004. The market had been carved-up along geographical lines and through a quota system. One of the cartel members, ABB, had escaped a fine because it has been a 'whistle blower' and provided important evidence to the Commission. Source: Reuters, 2007.

Examples of oligopoly

Oligopolies are common in the airline industry, banking, brewing, soft-drinks, supermarkets and music. For example, the manufacture, distribution and publication of music products in the UK, as in the EU and USA, is highly concentrated, with a 4-firm concentration ratio of around 75%, and is usually identified as an oligopoly.

The disadvantages of oligopolies

Oligopolies are significant because they generate a considerable share of the UK's national income, and they dominate many sectors of the UK economy, but they are also criticised on a number of obvious grounds, including:

Reduced choice

High market concentration reduces consumer choice.

Reduced competition

Cartel-like behaviour reduces competition, and can lead to higher prices, and reduced output.

Barriers to entry

Firms can be prevented from entering a market because of deliberate barriers to entry.

Loss of economic welfare

There is a potential loss of economic welfare, and a loss of allocative and productive efficiency.

The advantages of oligopolies

However, oligopolies may provide the following advantages:

Competitive oligopolies

Oligopolies may adopt a highly competitive strategy, in which case they can generate similar benefits to more competitive market structures, such as lower prices. In this case, even though there are a few firms their behaviour may be highly competitive.

Dynamic efficiency

Oligopolists may be dynamically efficient in terms of innovation and new product and process development. The super-normal profits they generate may be used to innovate, in which case the consumer may gain.

Price stabilisation

Price stability may bring advantages to consumers and the macro-economy because it helps consumers plan ahead and stabilises their expenditure, which may help stabilise the trade cycle.

Monopsonists

The special case of the monopsonist – the single buyer of labour – is an important one. A monopsonist is a single buyer of labour, such as De Beers, the diamond producer, and the major employer of diamond workers in South Africa. Monopsonists are common in some small towns, where only one large firm provides the majority of employment.

Because of their ***buying power***, monopsonists are able to influence the price they pay compared with buyers in more competitive markets. Pure monopsonists are rare because suppliers normally have alternative outlets for their good or service. However, ***monopsony power*** is significant in certain sectors of the economy.

In the case of ***supermarkets***, as with other dominant buyers, the price paid to suppliers is often forced down so that the supermarkets can reduce costs and generate higher profits. Alternatively they can reduce their prices, assuming they operate a cost plus pricing strategy. In turn this can threaten rival suppliers, so increasing the monopsony power of the major supermarkets. In an increasingly globalised world the supermarkets are free to source supplies from around the world, thus making it difficult for smaller suppliers to compete.

Questions

Questions 1 and 2 are multiple choice.

1. Fixing a price just below the average cost of potential entrants into a market is called:

a. Price discrimination

b. Predatory pricing

c. Cost-plus pricing

d. Limit pricing

e. Break-even pricing

2. If an oligopolist raises price it will:

a. Increase its profits

b. Reduce its profits

c. Expect a big change in its revenue

d. Make entry into the market less likely

e. Make entry into the market more likely

3. There is clear evidence that the UK supermarket sector is increasingly dominated by a few firms, led by Tesco, Sainsbury's and ASDA.

a. How do supermarkets compete with each other? (6)

b. Explain why UK supermarkets are increasingly dominated by a few large firms. (6)

c. What are the likely effects of this domination on:

d. Prices (2)

e. Profits (2)

f. Efficiency (2)

g. Economic welfare (2)

h. What could the regulators do to limit the power of supermarkets? (4)

i. Why would major airlines operating in Europe prefer non-price competition? Support your answer with reference to the kinked demand curve and game theory. (16)

Monopoly

A pure monopoly is a single supplier in a market. For the purposes of regulation, monopoly power exists when a single firm controls 25% or more of a particular market. A monopoly defined by such legislation is called a legal monopoly.

Formation of monopolies

Monopolies can form for a variety of reasons, including:

Exclusive ownership of a resource

If a firm has exclusive ownership of a scarce resource, such as Microsoft owning the Windows operating system brand, it has monopoly power over this resource and is the only firm that can exploit it.

Government legislation

Governments may grant a firm monopoly status, such as with the Post Office, which was given monopoly status by Oliver Cromwell in 1654. The Royal Mail Group finally lost its monopoly status in 2006, when the market was opened up to competition.

Patents and copyright

Producers may have acquired patents over designs, or copyright over ideas, characters, images, sounds or names, giving them exclusive rights to sell a good or service, such as a song writer having a monopoly over their own music and lyrics.

Mergers

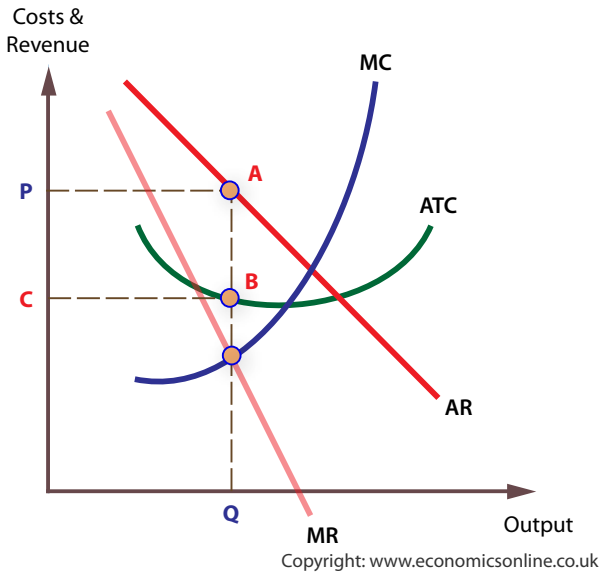
A monopoly could be created following the merger of two or more firms. Given that this will reduce competition, such mergers are subject to close regulation and may be prevented if the two firms gain a combined market share of 25% or more.

Key characteristics

Super-normal profits in the long run

Monopolies can maintain super-normal profits in the long run. As with all firms, profits are maximised when MC = MR. In general, the level of profit depends upon the degree of competition in the market, which for a pure monopoly is zero. With no close substitutes, the monopolist can derive super-normal profits, area PABC.

A monopolist with no substitutes would be able to derive the greatest monopoly power.



The advantages of monopoly

Monopolies can be defended on the following grounds:

Economies of scale

They can benefit from economies of scale, and may be natural monopolies, so it may be argued that it is best for them to remain monopolies to avoid the wasteful duplication of infrastructure that would happen if new firms were encouraged to build their own infrastructure.

Dynamic efficiency

It has been consistently argued by some economists that monopoly power is required to generate dynamic efficiency, that is, technological progressiveness. This is because innovation is more likely in large enterprises and this can lead to low-cost production. It is argued that a firm needs a dominant position to bear the risks associated with innovation. Firms need to be able to protect their intellectual property by establishing barriers to entry; otherwise, there will be a free rider problem. Why spend large sums on R&D if new ideas or designs are instantly copied by rivals who have not allocated funds to R&D? However, monopolies are protected from competition by barriers to entry and this will generate high levels of supernormal profits.

If some of these profits are invested in new technology, costs are reduced via process innovation. This makes the monopolist's supply curve to the right of the industry supply curve. The result is lower price and higher output in the long run. In addition, Austrian economist Joseph Schumpeter argued that inefficient firms, including monopolies, would eventually be replaced by more efficient and effective firms through a process called creative destruction.

Exports

Domestic monopolies can become dominant in their own territory and then penetrate overseas markets, earning a country valuable export revenues. This is certainly the case with Microsoft.

The disadvantages of monopoly

Monopolies can be criticised because of their potential negative effects on the consumer, including:

Low output

Monopolies can restrict output onto the market.

Restricting choice for consumers.

A market dominated by one firm inevitably means that the consumer has less choice.

High prices

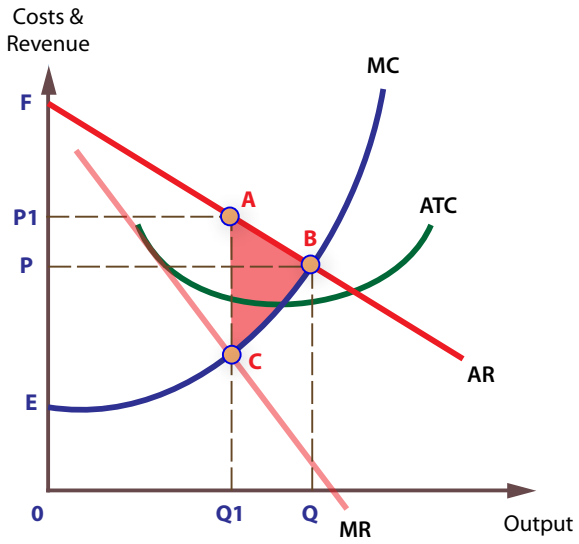
The traditional view of monopoly stresses the costs to society associated with higher prices. Because of the lack of competition the monopolist can charge a higher price (P1) than in a more competitive market (at P).

Reduced consumer surplus and welfare

The area of economic welfare under perfect competition is E, F, B. The loss of consumer surplus if the market is taken over by a monopoly is P P1 A B. The new area of producer surplus, at the higher price P1, is E, P1, A, C. Thus, the overall (net) loss of economic welfare is area A B C. This is also called deadweight loss.

The area of deadweight loss for a monopolist can also be shown in a more simple form, comparing perfect competition with monopoly. This diagram assumes that average cost is constant, and equal to marginal cost (ATC = MC).

Under perfect competition, equilibrium price and output is at P and Q. If the market is controlled by a single firm, equilibrium for the firm is where MC = MR, at P1 and Q1. Under perfect competition, the area representing economic welfare is F, B, E, but under monopoly the area of welfare is F, A, C, E. Therefore, the deadweight loss is the area A, B, C.



The wider and external costs of monopolies

Widespread dominance by monopolies can also lead to a less competitive economy in the global market place, and a less efficient economy, with less productive and allocative efficiency.

The economy is also likely to suffer from 'X' inefficiency, which is the loss of management efficiency associated with markets where competition is limited or absent.

Finally, it is possible that fewer are employed in economy, as higher prices and lower output may mean a reduced demand for labour.

Natural monopolies

A natural monopoly is a distinct type of monopoly that may arise when there are extremely high fixed costs of distribution, such as exist when large-scale infrastructure is required to ensure supply. Examples of infrastructure include cables and grids for electricity supply, pipelines for gas and water supply, and networks for rail and underground. These costs are also sunk costs, and they deter entry and exit.

In the case of natural monopolies, trying to increase competition by encouraging new entrants into the market creates a potential loss of efficiency. The efficiency loss to society would exist if the new entrant had to duplicate all the fixed factors - that is, the infrastructure. It may be more efficient to allow only one firm to supply to the market because allowing competition would mean a wasteful duplication of resources.

Economies of scale

With natural monopolies, economies of scale are very significant so that minimum efficient scale is not reached until the firm has become very large in relation to the total size of the market. Minimum efficient scale (MES) is the lowest level of output at which all scale economies are exploited. If MES is only achieved when output is relatively high, it is likely that few firms will be able to compete in the market. When MES can only be achieved when one firm has exploited the majority of economies of scale available, then no more firms can enter the market.

Public utilities

Natural monopolies are common in markets for 'essential services' that require an expensive infrastructure to deliver the good or service, such as in the cases of water supply, electricity, and gas, and other industries known as public utilities. Because there is the potential to exploit monopoly power, governments tend to nationalise or heavily regulate them.

Regulation

If public utilities are privately owned, as in the UK, since privatisation during the 1980s, they usually have their own special regulator to ensure that they do not exploit their monopoly status. Examples of regulators include OFGEM, the energy regulator, and OFCOM, the telecoms and media regulator.

Railways as a natural monopoly

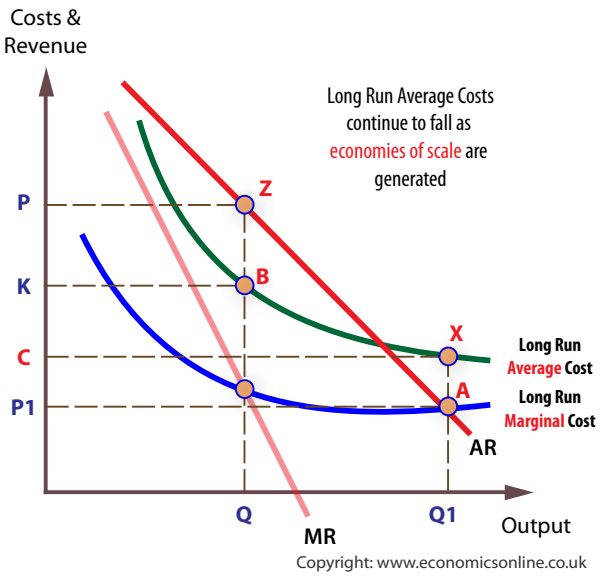
Railways are a typical example of a natural monopoly. The very high costs of laying track and building a network, as well as the costs of buying or leasing the trains, would prohibit, or deter, the entry of a competitor. To society, the costs associated with building and running a rival network would be wasteful.

Avoiding wasteful duplication

The best way to ensure competition, without the need to duplicate the infrastructure, is to allow new train operators to use the existing track; hence, competition has been introduced, without duplication of costs. This is called opening-up the infrastructure, and it is an approach which is frequently adopted to deal with the problem of privatising natural monopolies and encouraging more competition, as in the case of Telecoms, where the network is provided by BT, and gas, where the network is provided by National Grid (previously Transco.)

The natural monopoly diagram

With a natural monopoly, average total costs (ATC) keep falling because of continuous economies of scale. In this case, marginal cost (MC) is always below average variable cost (ATC) over the whole range of possible output.



Profits

In order to maximise profits the natural monopolist would charge Q, and make super-normal profits of P, Z, B, K. If unregulated, and privately owned, the profits are likely to be excessive. In addition, the natural monopolist is likely to be allocatively and productively inefficient.

Losses

To achieve allocative efficiency, the regulator will have to impose an excessive price-cap (at P1). The output needed to be allocatively efficient, at Q1, is so high that the natural monopolist is forced to make losses, given that ATC is above AR at Q1. Allocative efficiency is achieved when price (AR) = marginal cost (MC), at A, but at this price, the natural monopolist makes a loss.

A public utility's losses could be dealt with in a number of ways, including subsidies from the government, and price discrimination, where revenue could be increased by charging some consumers a different price than others.

Questions

Questions 1 – 5 are multiple choice.

1) A distinguishing feature of a natural monopoly is that:

   a. It is the only supplier in a given market

   b. It will be nationalised

   c. It will always make losses

   d. Average costs rise continuously with output

   e. Average costs fall continuously with output

2) If a monopolist switches from profit maximisation to sales maximisation it will plan to:

   a. Reduce price

   b. Increase price

   c. Reduce output

   d. Increase MR

   e. Increase super-normal profits

3) Price discrimination by a monopolist can only be beneficial if:

   a. Advertising costs do not rise

   b. Price elasticities of demand are the same in both markets

   c. It creates a barrier to entry

   d. Consumers can move freely from one market to another

   e. There is no seepage by consumers between markets

4) Cartels are least likely to be formed when:

   a. There are no barriers to entry

   b. The industry is highly concentrated

   c. There is a weak regulatory regime

   d. The industry is dominated by a few firms

   e. Collusion is easy

5) The profit maximising monopolist will always:

   a. Make profits

b. Derive economies of scale

c. Produce at the lowest average total cost

d. Produce in the elastic portion of the AR curve

e. Produce up to the point where the extra costs of production are less than the extra sales revenue

6) Using a diagram, and with reference to an example, explain that is meant by 'a natural monopoly'.

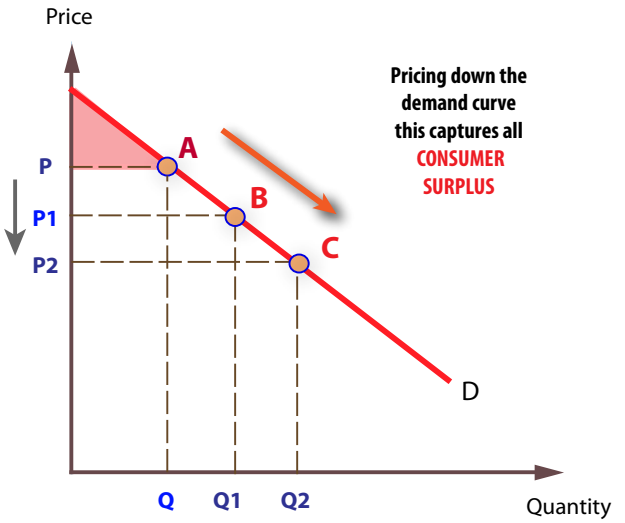
7) To what extent should natural monopolies be regulated?

Price discrimination

Price discrimination is the practice of charging a different price to different consumers, for the same good or service. There are three of types of price discrimination – first-degree, second-degree, and third-degree price discrimination.

First degree

First-degree discrimination, alternatively known as perfect price discrimination, occurs when a firm charges a different price for every unit consumed. The firm is able to charge the maximum possible price for each unit which enables the firm to capture all available consumer surplus for itself. In practice, first-degree discrimination is rare.



Second degree

Second-degree price discrimination means charging a different price for different quantities, such as quantity discounts for bulk purchases.

Third degree

Third-degree price discrimination means charging a different price to different consumer groups. For example, rail and tube travellers can be subdivided into commuter and casual travellers, and cinema goers can be subdivide into adults and children. Splitting the market into peak and off peak use is very common and occurs with gas, electricity, and telephone supply, as well as gym membership and parking charges. Third-degree discrimination is the commonest type.

Necessary conditions for successful discrimination

Price discrimination can only occur if certain conditions are met:

1. The firm must be able to identify different market segments, such as domestic users and industrial users.

2. Different segments must have different price elasticities (PEDs).

3. Markets must be kept separate, either by time, physical distance and nature of use, such as Microsoft Office 'Schools' edition which is only available to educational institutions, at a lower price.

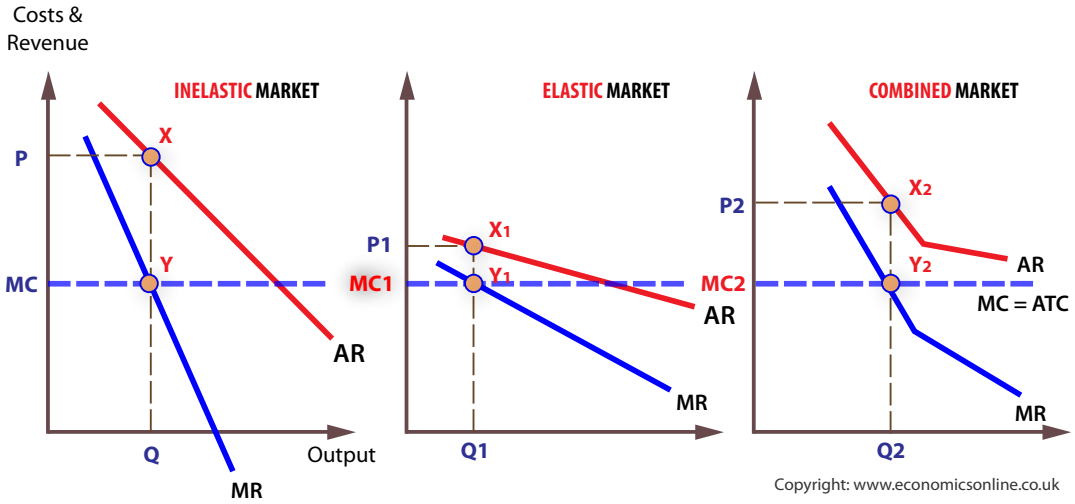
4. There must be no seepage between the two markets, which means that a consumer cannot purchase at the low price in the elastic sub-market, and then re-sell to other consumers in the inelastic sub-market, at a higher price.

5. The firm must have some degree of monopoly power.

Diagram for third degree price discrimination

If we assume marginal cost (MC) is constant across all markets, whether or not the market is divided, it will equal average total cost (ATC). Profit maximisation will occur at the price and

output where MC = MR. If the market can be separated, the price and output in the inelastic sub-market will be P and Q and P1 and Q1 in the elastic sub-market.



Profits will be the area MC, P, X, Y + MC1, P1, X1, Y1. If the market cannot be separated, and the two submarkets are combined, profits will be the area MC2, P2, X2, Y2. If the profit from separating the sub-markets is greater than for combining the sub-markets, then the rational profit maximizing monopolist will price discriminate.

Market separation and elasticity

Discrimination is only worth undertaking if the profit from separating the markets is greater than from keeping the markets combined, and this will depend upon the elasticities of demand in the sub-markets. Consumers in the inelastic sub-market will be charged the higher price, and those in the elastic sub-market will be charged the lower price.

Questions

1. Using an appropriate diagram, explain why a major airline charges different prices for the same London-Heathrow to New York-JFK flight.

2. Virgin Rail, which is currently trying to achieve maximum sales revenue, is considering buying the franchise to offer a new rail service from Birmingham to Manchester. If it goes ahead it will be a local monopoly. It has undertaken market research and believes there are two distinct markets. The commuter market, catering for those who live in Birmingham, but work in Manchester, and the 'casual traveller' market, catering for those on trips, on visits to friends and university interviews, tourists and business travellers going to exhibitions.

Its market research shows that, at various ticket prices, demand from the two markets will be:

Reurn ticket price (£)	Estimated sales to commuters per year (000)	Estimated sales to casual travellers per year (000)
120	14	
110	16	
100	18	
90	20	
80	22	
70	24	4
60	26	8
50	28	12
40	30	16
30	32	20
20	34	24
10	36	20

a. Give advice to Virgin about its pricing policy, in particular:

b. Should it have one price for its tickets irrespective of the type of customer, or two prices – one for each market?

You must state the price or prices and you must justify your advice.

Also, support your advice with a graph.

If you argue that two prices are best, give advice about how best to operate this system.

Growth, mergers and acquisitions

The long run

The long run for a single firm is entered when it uses more fixed and variable factors to increase its scale of production.

Growth

Firms grow in order to achieve their objectives, including increasing sales, maximising profits or increasing market share. Firms grow in two ways; by internal expansion and through integration.

Internal expansion

To grow organically, a firm will need to retain sufficient profits to enable it to purchase new assets, including new technology. Over time, the total value of a firm's assets will rise, which provides collateral to enable it to borrow to fund further expansion.

The importance of branding

One of the most common strategies for internal growth is to build the firm's brand, which provides the firm with many advantages. Once a brand is established, less advertising is required to launch new products. Internal growth often provides a low risk alternative to integration, although the results are often slow to arrive.

External expansion

The second route to achieve growth is to integrate with other firms. Firms integrate through mergers, where there is a mutual agreement, or through acquisitions, where one firm purchases shares in another firm, with or without agreement. There are a number of different types of integration, including:

Vertical integration

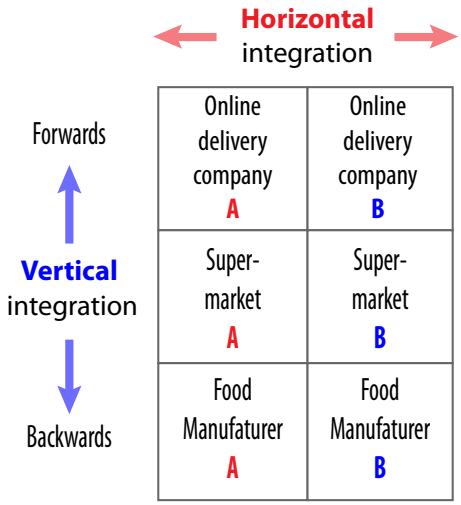
Vertical integration occurs when firms merge at different stages of production. There are two types of vertical integration, backwards and forwards.

Backwards

Backward vertical integration occurs when a firm merges with another firm that is nearer to the source of the product, such as a car producer buying a steel manufacturer.

Forwards

Forwards vertical integration occurs when a firm merges to move nearer to the consumer, such as a car producer buying a chain of car showrooms.



Horizontal integration

Horizontal integration occurs when firms merge at the same stage of production, such as a merger between two or more supermarkets, food manufcturers or car producers. Horizontal integration is also called ***lateral integration***.

Conglomerate integration

Conglomerate, or diversified integration, occurs when firms operating in completely different markets, merge - such as a car producer merging with a travel agency. In this case, firms tend to retain their original names, and are owned by a 'holding' company. Conglomerate mergers can involve firms in unrelated markets, or where complementary goods are involved.

Multi-nationals

Many firms grow by integrating with foreign firms, which is increasingly common in the globalised world economy, and is a key part of the globalisation process. Cross-border mergers contribute to inward investment between countries. The UK is a major global investor, and in 2005 topped the world league table for receiving FDI (Foreign Direct Investment), with inward investment of $220b[4]. Mergers and acquisitions account for a large share of FDI.

Bank mergers

Like all firms, banks can derive considerable benefits from merging, including economies of scale. In addition, there are considerable benefits to financial institutions from merging rather than expanding organically. Over time, banks will have built up a range of low, medium, and high-risk borrowers. To expand organically, a bank may have to take on higher risk customers. However, if a bank acquires another bank it will not need to increase its average risk because it will acquire a range of customers of all risks. Banks can also merge to help secure extra liquidity.

The advantages of mergers

Mergers can bring about the following advantages:

Economies of scale

Firms that merge can benefit from a range of economies of scale, such as cost savings associated with marketing and technology.

Lower costs

In the case of vertical integration there are savings in terms of not paying '3rd party' profits. For example, if a tour operator owns its own hotels it will not need to pay profits to the hotel, and will be able to keeps costs and prices down.

Economies of scope

Economies of scope are also available to firms that merger, and are benefits associated with using the fixed assets of one firm to produce output for the other firm.

Unexpected synergies

Unexpected synergies are unpredicted benefits that arise when firms merge or undertake a joint venture, such as when two pharmaceutical companies merge, and create a new drug.

Rationalisation

Rationalisation is the process of eliminating parts of a business that are inefficient or unprofitable, and is a possible consequence of two or more firms merging.

Sharing knowledge

When firms merge, they can share knowledge with each firm benefitting from the knowledge and experience acquired by the other. With vertical integration, information asymmetries can be reduced or removed.

4 Source: UNCTAD, 2005.

Protection from take-over

The merger of two firms may send out a signal to other firms not to attempt a take-over bid.

Spending on R&D

Firms that merge may be able to allocate more funds to Research and Development (R&D) and generate new products as a consequence. This may increase their competitiveness and profitability in the long run.

The disadvantages of mergers

Mergers can generate s number of disadvantages, including:

Less competition

Increased market concentration and less competition are obvious disadvantages of a merger between two large firms.

Diseconomies of scale

Firms that merge may experience diseconomies of scale, such as difficulties with co-ordination and control. This will increase average cost in the long run, and reduce profitability.

Higher prices

Higher prices are a likely consequence of a merger because, with less competition, demand is more inelastic and raising price will raise revenue.

Restricted output

There may be less output from the merged firm, compared with combined output of the two firms.

Loss of jobs

Rationalisation is likely to lead to lost jobs as the merged firms attempt to increase profitability. For example, two advertising agencies that merge could dispense with two design departments, and share one.

Reduced choice

Consumers are likely to suffer from reduced choice following a merger of two close competitors. This is a common criticism of banking and supermarket mergers, and one reason why they are the subject of scrutiny.

Increased barriers

The economies of scale and scope derived from a merger may increase barriers to entry and make the market less contestable. In the case of forward vertical integration, new entrants may be denied access to outlets for its products. With backwards vertical integration, new entrants may find it difficult to secure a source of supply of materials or products.

Questions

1) In 2005, Virgin Atlantic attempted a takeover of BMI (British Midland International Airlines).

   a) What type of integration is this merger be an example of? (2)

   b) Outline the likely benefits of this merger. (6)

   c) Under what circumstances might the Competition Commission or BERR become involved in an airline merger? (8)

   d) What options are available to the Competition Commission in terms of regulating this merger? (4)

   e) Evaluate alternative pricing strategies for a typical airline. (10)

   f) To what extent is the UK air travel market a contestable one? (10)

2) European Cinema Screens (2004)

Odeon	608
UCI	443
Vue	406
UCG	396
Show Case	323
TOTAL	*2653*

Extract 1

'...In locations where both circuits operate directly competing cinemas, aligned distributors normally supply films to their aligned circuit but not the other. This practice, by reducing competition for screens among aligned distributors and reducing pressure on the two circuits to compete for films on merit, makes the market less responsive to consumer preferences. We condemned it in our 1983 report, but stopped short of making a recommendation because of the then parlous plight of the industry. The industry is now stronger and we make recommendations intended to bring this practice to an end.

The second practice which we find to be against the public interest concerns minimum exhibition periods. Distributors sometimes insist on lengthy minimum exhibition periods-perhaps four weeks or longer-as a condition of supplying exhibitors with prints of popular films. This practice creates problems particularly for single-screen cinemas or those with few screens, reducing their freedom to respond to consumer demand, and adds to the difficulties faced by independent distributors in getting their films shown. We recommend that minimum exhibition periods should be restricted to a maximum of two weeks on first release and one week subsequently.

Another issue is vertical integration, to which critics of the film industry frequently point as distorting the market and creating barriers to entry, particularly for British films. All the Hollywood studios rely upon their respective affiliates to distribute their films in the UK. Given the state of competition among the studios and in the distribution market generally, we do

not object to this practice, which is common worldwide. Four of the seven Hollywood studios also have ownership links with UK exhibitors. We have examined whether these links result in dealings between distributors and exhibitors being other than at arm's length. Our analysis shows a slight degree of preference between vertically linked parties at the margin, but the evidence does not warrant an adverse public interest finding. We suggest nevertheless that certain indicators should be monitored by the Office of Fair Trading so that the matter is kept under review.....'

Monopolies and Mergers Commission 1994 (which was replaced by the Competition Commission)

Extract 2

The UK cinema sector witnessed its biggest shake-up for over a decade on Friday when private equity firm Terra Firma acquired both the Odeon and UCI circuits in one fell swoop.

The resulting estate, valued at approximately £580m and which is estimated to represent 35% of the UK's box office revenues, will almost certainly fall foul of the regulations concerning market share. Previous Commission reports on the exhibition sector set the limit of control of box office share by any single operator at between 25% - 27%.

Odeon investor Robert Tchenguiz and US private equity group Blackstone are also reported to have made the final Odeon bid shortlist, while VUE Entertainment and BC Partners withdrew from the auction for Odeon in the penultimate round of bidding. Source Pearl & Dean (2004)

a) With reference to the data, describe the type of market structure in this industry. (3)

b) What type or types of integration has taken place in this industry? (3)

c) Explain why the practice of price discrimination is likely to be common in cinema ticket pricing? (8)

d) Examine the likely benefits to Terra Firma of acquiring Odeon and UCI. (12)

e) To what extent is such an acquisition likely to be against the public interest? (14)

Multiple choice - select the correct option

3) The regulator agrees to set the pricing of Water Companies as RPI – X. This means that:

a) The Regulator expects there to be inflation

b) Water companies will have to reduce their prices

c) The Regulator expects efficiency gains to the value of X

d) X represents expected increases in competition

Contestable markets

Key characteristics

The theory of contestable markets is associated with the American economist William Baumol. In essence, a contestable market is one with zero entry and exit costs. This means there are no barriers to entry and no barriers to exit, such as sunk costs and contractual agreements.

The existence, or absence, of sunk costs and economies of scale are the two most important determinants of contestability. Based on these criteria, natural monopolies are the least contestable markets. ***Asymmetric information*** is also a key barrier to entry. Incumbents are likely to know much more about their industry than potential entrants. With no barriers to entry into a market, it can be argued that the threat of entry is enough to keep incumbents 'on their toes'. This means that even if there are a few firms, or a single firm, as with oligopolistic and monopolistic markets, a market with no barriers will resemble a highly competitive one.

Potential entrants can operate a hit and run strategy, which means that they can 'hit' the market, given there are no or low barriers to entry, make profits, and then 'run', given there are no or low barriers to exit.

The implications of contestability

If we assume there are only a few firms in a market, and there are few barriers to entry and exit, then we can state that potential entrants can freely enter and leave the market and could, if they wished, operate a 'hit and run' strategy. Indeed, just the threat of entry is enough to 'keep firms on their toes', to the extent that existing firms behave 'as if' the market has a highly competitive market structure.

Evaluation

The theory of contestable markets is often seen as an alternative to the traditional, Neo-classical, theory of the firm. Perfectly contestable markets can deliver the theoretical benefits of perfect competition, but without the need for a large number of firms. Firms are forced to keep excess profits to a minimum, and move towards sales maximisation rather than profit maximisation.

Questions

Typically, how contestable are the following markets?

1. The market for soft drinks
2. Hairdressing
3. Major television broadcasting
4. Internet browsers

Regulation

Neo-classical theory

As Adam Smith noted in the late 18th Century, '..people of the same trade seldom meet together...without the conversation ending in a conspiracy against the public, or in some contrivance to raise prices.' (Wealth of Nations, 1776). This view dominated Classical and Neo-Classical theory for 150 years. The Neo-Classical analysis of firms is deeply rooted in the belief that monopolies are inherently harmful, and that a merger between competitive firms will reduce competition and increase monopoly power. The Neo-Classical view was that monopolies would cause a misallocation of scarce resources, with prices rising well above competitive prices. In short, regulatory authorities should be suspicious of the motives behind meetings of firms, alliances and formal mergers, and closely monitor and control the anti-competitive behaviour of monopolies.

The modern approach

The modern view is more pragmatic, and recognises that monopolies and mergers should be judged on a case by case basis, and it should not be assumed that they are against the public's interest. The modern approach accepts that monopolies can create economic benefits as well as costs, including the benefits of economies of scale, innovation and dynamic efficiency, and export earnings.

Legislation

The Competition Act 1998

The Competition Act 1998 prohibits a number of activities by firms, including:

1. The formation and operation of cartels.

2. The abuse of a firm's dominant position on a national or local level.

3. Concerted practice, such as firms colluding instead of competing. For example:

4. Fixing price, such as a number of book publishers fixing the minimum resale price of books sold by separate book stores, or raising price together, or fixing output.

5. Fixing terms of business, such as agreeing to the same delivery times or terms of payment.

6. Carving up a market, which means that firms agree to split up a market and not compete in the different sectors of the market.

The Enterprise Act 2002

The Enterprise Act, 2002, amended the Competition Act and strengthened the power of the regulators, especially in terms of detecting and punishing abuse of market dominance and cartel-like behaviour. The main provisions of the Act were:

1. Assessment of mergers to be less influenced by politicians and more independent.

2. New powers for regulators to investigate markets, such as the power to use covert surveillance.

3. Criminalisation of cartels, with the UK regulators becoming tougher than those in the EU.

4. Disqualification of directors for breach of the competition rules.

5. Consumer groups can complain about uncompetitive practices.

6. There was a shift of emphasis from considering the public interest criteria to a more narrow concern regarding the effect of behaviour on competition.

Regulatory structure in the UK

In the UK, the regulation of firms and promotion of competition is undertaken by the Department for Business, Innovation and Skills (BIS), the Office of Fair Trading (OFT), and the Competition Commission.

The Department for Business, Innovation and Skills (BIS)

The Department for Business, Innovation and Skills (BIS) was created in 2009 with the merger of the Department for Business, Enterprise and Regulatory Reform (BERR), and the Department for Innovation, Universities and Skills (DIUS).

The main objectives of BIS are:

- To promote free and fair markets, with increased competition
- To increase productivity and improve skills
- To promote science and innovation, and promote the commercial exploitation of knowledge
- To create the right conditions for business success
- To improve economic performance of the UK regions, and to reduce the gap in growth rates between the regions

The Office of Fair Trading (OFT)

The OFT is an independent body whose main role is to try to ensure that markets work effectively. As its name suggests, it looks at 'unfair' and uncompetitive trading. It has separate divisions (offices) that regulate the privatised utilities, including OFGEM, OFWAT, and OFCOM. It is the main 'referring body', referring cases to the Competition Commission.

The OFT's main objectives are:

- To identify and put right trading practices which are against the consumer's interests.
- To regulate the provision of consumer credit.
- To investigate anti-competitive practices, including restrictive practices, such as manufacturers forcing retailers to fix a minimum price.
- To investigate abuse of market power, when a firm has a dominant position, and cartel-like behaviour.
- To help promote market structures which encourage competitive behaviour.

It can impose fines of up to 10% of turnover (for 5 years) when necessary, and in 2004 the OFT gained new powers to use covert surveillance to investigate anti-competitive practices.

The sub-offices

The OFT's sub-offices were established to regulate the privatised utilities and encourage competition.

The Competition Commission

The Competition Commission (CC) is also an independent public body which conducts in-depth inquiries into markets, mergers and the regulation of the major regulated industries. It undertakes inquiries following referrals gto it by other authorities, most notably the OFT.

In terms of mergers, the Commission must assess whether a merger will reduce competition. After investigating, it may recommend that the merger goes ahead, is prohibited, or is allowed to go ahead with modifications. In deciding which option to implement, the Commission will consider whether, after the merger, competition is maintained.

Regulation of privatised utilities

Nationalisation

Between the 1930s and 1950s, most of the public utilities in the UK, such as gas, water, and electricity, were taken over by the state via a process called nationalisation. The main motive for nationalisation during this period was to ensure a coordinated approach to production and supply to ensure economic survival and efficiency in the face of war, and post-war reconstruction.

However, by the late 1970s it was becoming apparent that these utilities were suffering because:

- They were being managed ineffectively and inefficiently. The principal-agent problem is highly relevant to public sector activities. The managers of the utilities were not required to meet any efficiency objectives of the state, and there was growing criticism because these industries were protected from competition, and hence had become increasingly 'X' inefficient.
- The nationalised industries had no power to raise capital on the open market. If they needed funds for investment, they would have to compete with other government spending departments, like education, health, and defence. This is one reason why there was considerable under-investment in these industries.
- To resolve these problems and to generate revenue for the government, many State-owned industries during the 1980s were sold off to the private sector through a process called privatisation.

Privatisation and regulation

The major privatisations in the UK were gas, electricity, oil, telecoms, and coal. However, many of the privatised utilities are also natural monopolies. With a natural monopoly, the role of the regulator is to act as a surrogate competitor to the privatised, natural monopoly. In doing this the regulator can make up for the missing contestability found with natural monopolies.

Regulatory options

Regulators have a number of options, including:

Price controls

Regulators can set price controls and formulae, often called price capping. This means forcing the monopolist to charge a price below profit maximising price. For example, in the UK the RPI – 'X' formula has been widely used to regulate the prices of the privatised utilities. In the formula, the RPI (Retail Price Index) represents the current inflation rate. 'X' is a figure which is set at the expected efficiency gain which the regulator believes would have existed had the firm operated in a competitive market. However, there is a dilemma with price controls – price-capping results in lower prices, but lower prices also deter entry into the market. Regulators may remove price caps if they judge that competition in the market has increased sufficiently, as in the case of OFCOM who removed BT's price cap in 2006.

Rate-of-return regulation

An alternative to price-cap regulation is rate-of-return regulation. Rate of return regulation, which was developed in the USA, is a method of regulating the average price of private or privatised public utilities, such as water, electricity and gas supply. The system, which employs accounting rules for the calculation of operating costs, allows firms to cover these costs, and earn a 'fair' rate of return on capital invested. The 'fair' rate is based on typical rates of return which might be expected in a competitive market.

However, rate-of-return regulation is often criticised because, unlike in an actual competitive market, a reduction in costs will not improve its situation, and hence there is little incentive to control costs. In fact, it will be to the advantage of the monopolist to allow costs to inflate because prices will then be allowed to rise. This would not happen in a competitive market because demand would form a constraint against such price rises.

A further general weakness is that regulators are unlikely to have perfect knowledge about the costs of production of the monopolist, and cannot make an effective judgement about whether the costs are being controlled effectively, or not.

Windfall taxes

Regulators could impose a windfall tax on excessive profits, which this would encourage the firm to reinvest its profits, rather than distribute them to shareholders. This tax would not alter the output of the firm; hence consumers would not suffer from falling output.

Prohibiting mergers

Regulators can prevent mergers or acquisitions, or set conditions for successful mergers.

Breaking up a monopoly

The regulators might insist on the break-up of the monopoly. For example, in 2004 the UK telecoms regulator Ofcom recommended that BT be split into two businesses: 'retail' and 'wholesale'.

Nationalisation

A less popular option would be to bring the monopoly under public control, in other words to nationalise it.

Forcing un-bundling

Effective regulation may also involve bringing down barriers to entry, such as forcing the incumbent to allow potential rivals to have access their network or infrastructure. This is referred to as opening-up or unbundling their infrastructure.

Yardstick competition

Regulators can introduce yardstick competition, such as setting punctuality targets for British Rail based on the best-performing European rail systems. It is also possible to split up a service into regional sections to compare the performance of one region against another. This is applied in the UK to both water and rail.

Licensing

The regulator or relevant government department can create a licensing system, such as with the train operating companies (TOCs) and Royal Mail for letter post, which can be extended or withdrawn, subject to the performance of the licence operator.

Self-regulation

In some industries, the regulator might allow self-regulation. Certain industries may be allowed to self-regulate by establishing a 'code of conduct' by which industry members agree to abide. In 2002, the main UK supermarkets established a voluntary code of conduct following

criticism by the Competition Commission in 2000. Critics argue that self-regulation is unlikely to provide sufficient incentive for firms to behave responsibly.

Questions

Evaluate alternative ways to regulate the following:

1) UK banks

2) European airlines

3) UK telecommunications

Nationalisation

Most of the UK's major strategic heavy industries and public utilities were nationalised between 1946 and the early 1950s, only to be returned to the private sector between 1979 and 1990.

Examples of nationalisation

1. 1946 - The Bank of England was the first organisation to be nationalised by the new Labour government of Clement Atlee.

2. 1947 - The Coal industry was nationalised in 1947 when over 800 coalmines were taken under public ownership and a National Coal Board (NCB) was established to manage the industry on commercial lines. The NCB became the British Coal Corporation in 1987, and this was wound up in 1997 as the industry was privatised. *(Source: National Archives)*.

3. 1948 - Railways were nationalised to help rebuild the network infrastructure and re-equip the rolling stock after the destructive effects of the Second World War.

4. 1949 - Steel was first nationalised in 1949, and privatised a year later by the new Conservative government. It was re-nationalised in 1967 when over 90 of steel capacity was put under the control of the British Steel Corporation (BSC). Steel was returned to the private sector once more in 1988.

5. 2008/9 - A number of key UK banks became subject to full or part-nationalisation from early 2008 as a response to the financial crisis and banking collapse. The first bank to become nationalised was the Northern Rock in February 2008, and by March 2009, the UK Treasury had taken a 65% stake in the Lloyds Banking Group and a 68% stake in the Royal Bank of Scotland (RBS).

The advantages of nationalisation

- The main motive for nationalisation during the post-war period was to ensure a co-ordinated approach to production and supply to ensure economic survival and efficiency in the face of war, and post-war reconstruction. For example, the advantage of a nationalised rail network, as with other natural monopolies, was that central planning could help create a more organised and co-ordinated service. This argument was applied widely to the so-called commanding heights of the economy.

- It can also be argued that much infrastructure provides a considerable external benefit to individuals and firms. For example, a nationally and centrally funded and efficient rail network helps keep road traffic down and hence reduces pollution and congestion. It may also help reduce business costs, which may be passed on to other businesses.

- Another advantage of national ownership is that economies of large scale can be gained that would not be available to smaller, privately owned enterprises. For example, a nationalised rail service could purchase materials, rail track, and rolling stock on a large scale, thereby reducing average costs and supplying more efficiently than smaller operators.

- In more recent times the failure of major banks has highlighted the fact that, under national ownership and control, failing banks can be funded more quickly and for larger amounts than under private ownership. This enables the banking 'infrastructure' to be rebuilt, as well as ensure the closer regulation of banks in the future.

The disadvantages of nationalisation

- By the late 1970s it became increasingly apparent that many of the industries nationalised between 1945 and 1951 were running into difficulties. The major problems that these industries faced were:

- They were being managed ineffectively and inefficiently. The principal-agent problem is highly relevant to public sector activities given that the managers of the utilities were generally not required to meet any efficiency objectives set by the state. There was growing criticism that, because these industries were protected from competition, they had become increasingly 'X' inefficient.

- Nationalised industries were also prone to suffer from moral hazard, which occurs whenever individuals or organisations are insured against the negative consequences of their own inefficient behaviour. For example, if a particular nationalised industry made operating losses, the government would simply cover those loses with subsidies. Knowing that the taxpayer would come to the rescue meant that the inefficient behaviour could continue. This is, perhaps, the most significant criticism of the recent 'bail out' of failing banks. Given that they know the taxpayer will bail them out this may be an encouragement to continue with their inefficient and highly risky lending activities.

- In addition, the nationalised industries had limited scope to raise capital for long-term investment and modernisation because they would have to compete with other government spending departments, like education, health, and defence. The result was a prolonged period of under-investment in these industries.

- By the late 1970s, and throughout the 1980s, most UK's major State owned industries were sold off to the private sector through privatisation. The intention was that, back in the free market, these industries would become more efficient and would be able to modernise by having greater access to the capital markets, and by employing more modern and dynamic management. Privatisation also generated huge revenues for the UK Treasury as well as allowing tax cuts and creating an environment where other supply-side reforms could be implemented. Following the banking collapse of 2009, nationalisation was put firmly back on the agenda, if only in terms of the financial system.

Questions

1) To what extent should the following be nationalised?

   a) UK banks

   b) UK railways

2) Should the UK Post Office be privatised?

Case 1 – Railways

Reorganisation

Rail services in the UK have been reorganised twice in the last 12 years. Rail was first privatised in 1993 with the creation of Railtrack as the monopoly owner of the infrastructure. In the mid 1990s, train operating franchises were sold off to private firms such as Virgin and Stagecoach. In 2000, the Labour government created the Strategic Rail Authority. Because of the poor safety record, in 2002 Network Rail was created to replace Railtrack – Network Rail is a not-for-profit company. By 2005, there were 26 separate train companies, with franchises running from 5 to 8 years. Network Rail requires large subsidies from central government, averaging £4b per year (2005). The Office of Rail Regulation is the rail regulator.

Evaluation

Critics argue that the structure is 'too fragmented' – fragmentation is one way of resolving the natural monopoly dilemma - separate owners of infrastructure and suppliers, but it creates co-ordination and control issues, such as:

- ***Diseconomies of scale*** – difficult to co-ordinate operations across such a large structure.
- ***Information failure*** – information does not flow smoothly between parts of the structure.
- ***The principal-agent problem***, in terms of profits vs safety – should the railways operate in the public interest or commercial interest?

Would re-nationalisation help?

Many feel that bringing back railways into the public sector would create a net benefit, especially because:

- Safety would not need to be compromised.
- The rail network can be co-ordinated better.
- More funds can be allocated to investment in infrastructure.
- More account can be taken of the positive externalities associated with rail travel.

However:

- Governments also fail in terms of information failure, over-interference, and excessive bureaucracy.
- Nationalisation does not guarantee efficiency.
- Is it fair on the general taxpayer to subsidise rail transport which some may never use, despite the external benefits.
- There may be less, rather than more, access to capital, as rail has to compete with other demands on the Exchequer[5].

5 Source: The Times, 10/01/04

Case 2 – Telecoms

BT loses its monopoly status

British Telecom lost its protected monopoly status when it was privatised in 1984. Between 1984 and 2005 BT's market share of for fixed-line calls fell from 100% to less than 65%.

The UK Broadband Market		
Supplier	**Customers**	**Market share**
BT	15,000,000	62.5
NTL/Telewest	4,000,000	16.7
Talk Talk	2,400,000	10.0
Homecall	600,000	2.5
Sky Talk	500,000	2.1
Tiscali	250,000	1.0
Post Office	200,000	0.8

Source: Uswitch, 2005

The de-regulation of the telecoms market, which followed privatisation, opened up the market to rival suppliers, and with the application of new technology, barriers to entry fell appreciably. The network now can be used by other firms through a process called Wholesale Line Rental (WLR). After being threatened with brake-up BT created its own independent division, Openreach, in 2004 to try to ensure fair access to its network for rival operators.

Unbundling

In 2005, Ofcom set an RPI - X pricing formula for BT's charges for the use of its network to encourage rivals to enter the market. By 2006 over 300 independent firms had access to the network infrastructure, making the UK telecoms market one of the most competitive in the world. Even so, OFCOM has requested that BT speeds up its unbundling process, and has focussed a large part of its investigation on BT's broadband pricing. Today, the market can best be described as an oligopoly, with a three-firm concentration ratio of 89%.

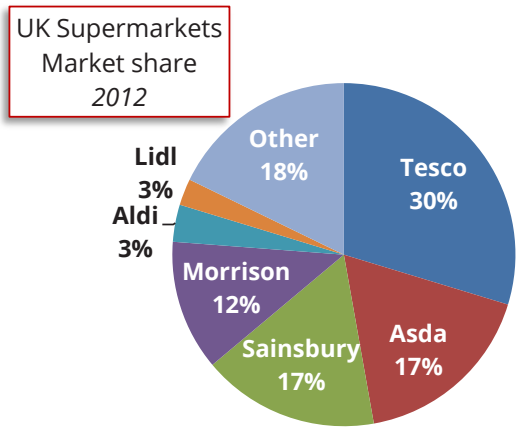
Case 3 – Supermarkets

Increasing concentration

There is clear evidence that the UK supermarket sector is increasingly dominated by a few firms, led by Tesco, Sainsbury, and ASDA. The market is clearly oligopolistic and the pricing strategy of the supermarkets can be understood using a game theory approach.

It has been noted that farmers and growers claim they are suffering in the face of the increasing monopsony power of the major supermarkets. The pace of concentration appears to have accelerated over the last five years.

Tesco was prevented by the OFT from purchasing Safeway in 2002, but was allowed to buy 1200 'convenience' stores. By 2005, Tesco had acquired nearly 6% of the convenience food market[6].



The decline of independent retailers

The rise of the large supermarket chains has, unsurprisingly, coincided with the decline in independent high street retailers. Between 2005 and 2010, the number of independent retailers has fallen by a third, from 35,000 to 25,000.

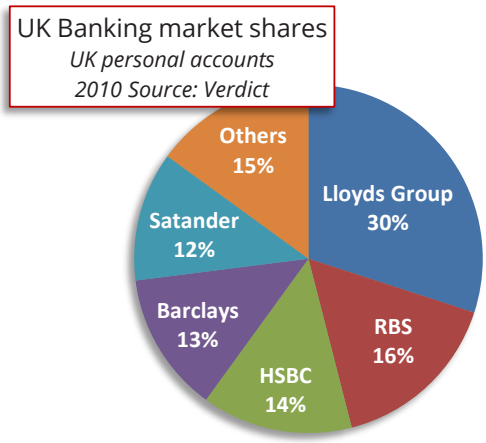
6 Source: Verdict

Case 4 - UK banks

Oligopoly

The UK banking sector is dominated by a few very large banks, including Barclays, The Royal Bank of Scotland (RBS), and HBOS. In term of the value of assets, the market is clearly oligopolistic.

In response to the credit crunch, the UK banking sector has accelerated the process of integration. In January 2009, the Halifax Bank of Scotland (HBOS) merged with Lloyds TBS to create the Lloyds Group.



Complex monopoly

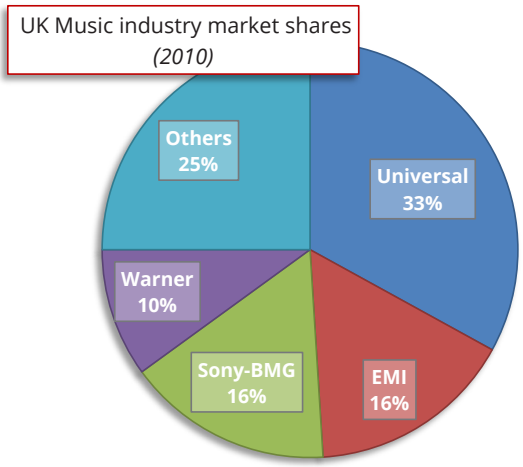
In 2002, the Competition Commission concluded that a number of the largest banks operated a 'complex monopoly' in the supply of services to small and medium sized enterprises (SMEs) which resulted in reduced competition to the detriment of the customers. For example, customers were reluctant to switch banks because they all offered very similar benefits.

Case 5 - The music industry

Music mergers

In 2004, Sony Music merged with BMG, the German based music giant, in a deal worth £3.4b, reducing the number of music 'majors' from five to four.

In an attempt to block this merger it was reported that Impala, a trade organisation representing over 2000 independent music labels, demanded that the EU Commission for Competition block this merger because it would weaken music diversity and limit independent producer's access to retailers and to the radio. The EU Commission did look at the vertical integration in the EU music industry, but even after widespread protests from independent labels and consumer groups,, the merger was allowed to go ahead on the grounds that the growth of internet piracy was sufficient justification for the merger[7].



By 2011, the UK market shares were:

Price fixing of downloaded music

In 2006 the four leading US music producers, Sony BMG, Universal, Warner, and EMI were accused of price fixing downloaded music tracks.

7 Source – The Times, January 27th and 30th, 2004

Case 6 – Microsoft

The EU vs Microsoft

There has been a long running battle between the EU Competition Commission and the US giant, Microsoft, over alleged abuses of its dominant position. The EU started to investigate the Microsoft Corporation in 1999 after a complaint from Sun Microsystems that Microsoft had withheld vital information about its operating system. After 5 years of investigation it made its final ruling.

Under this ruling, in 2004 Microsoft was forced to provide its rivals with technical details of its operating system, a type of forced 'unbundling' of the infrastructure upon which computer programmes run. Microsoft was also fined €497m (£340m) for continued abuse of its near monopoly position, and forced to offer its operating system to computer manufacturers without Media Player bundled into the package. The EU Commission based its rulings on the view that competition would eventually be eliminated in the Media Player market, with consumers having less choice, and with innovation stifled. (Source: eurunion.org – 2005).

By 2008 Microsoft's market shares in Europe were:

- Browser market share: 76%, Mozilla was second with 17%
- Office software share: 95%

Source: Net Applications, 2008

Case 7 - Soap powder

The European Chemical Cartel

In 2006, seven European chemical companies were fined a total of £275m for operating an illegal cartel that fixed chemical prices in continental Europe and the UK. The particular chemical in question was hydrogen peroxide, and agent for bleaching textiles and an ingredient in many domestic soap powders. The seven companies were found guilty by the European Competition regulator, the European Competition Commission, of two particular practices: price fixing of chemicals and 'carving-up' the chemical market across Europe.

The cartel came to light in 2002 when one company involved became a 'whistle blower' and gained immunity from prosecution. The fines were the third largest imposed by the European Commission, mainly reflecting the fact that the participants of the cartel were repeat offenders[8].

8 Source: The Times, May 2006.

Evaluation of competition policy

Increasing powers

Since the Competition Act (1998) and the Enterprise Act (2002) regulators have had more power to act against abuses of monopoly power.

However:

- Very few mergers are actually investigated each year. Since the Enterprise Act (2002), not a single major cartel has been investigated by the OFT (2008).
- Despite heavy fines, covert collusion is difficult to prove. New powers have been given to the regulators to undertake covert surveillance of firms to establish whether collusion is taking place.
- Tacit collusion is almost impossible to prove. Statistical techniques could be used to establish correlations between price movements in 'theory' and in practice.
- The problem of 'regulatory capture' where the regulator begins to 'take sides' with the firm or industry it is regulating, that is, there is a potential 'principal-agent' problem.
- The problem of 'cheating' or 'finding loopholes', such as getting round the regulations by moving into an adjacent market. For example, a large grocery retailer moving into the 'convenience' store market could be seen as a way of circumventing competition policy.
- A major criticism is that 'single' markets are inadequately defined. For example, in 2004 the OFT allowed Tesco to purchase Adminstore (the owners of Cullens, Europa and Harts) because it regarded 'convenience stores' to be a separate market from 'grocery' stores. This gave Tesco 6% of the convenience food market as well as 26% of the grocery market. While Tesco argued that the two markets are quite distinct, critics, like the Forum for Private Business, argued that the markets are almost indistinguishable and that Tesco should not have been allowed to enter the convenience food market.