

# Economies & diseconomies of scale

This chapter introduces the concept of economies of scale and examines the difference between internal economies of scale and external economies of scale. Diseconomies of scale and their reasons are explained. The chapter concludes by studying the relationship between economies of scale and diseconomies of scale and their impact on the shape of the long-run average cost curve.

## The difference between internal and external economies of scale

### Key note

In order to understand economies of scale, you need to distinguish between the firm and the industry. A firm is a single business, such as Nissan; the industry is the collection of all firms providing that good or service – in this case, the car or motor industry.

### Key terms

**Internal economies of scale** are the advantages that an organisation gains due to an increase in its size. These advantages cause an increase in productive efficiency and thus a decrease in the average (total) cost of production.

**External economies of scale** are the advantages that an organisation gains due to growth in the size of the industry within which it operates.

**Internal economies of scale** mean that larger firms can benefit from lower average costs, which will improve their ability to compete in their marketplace.

**External economies of scale** mean that larger firms can benefit from lower average costs too, but these benefits should also be available to competitors that operate within the same industry. External economies of scale apply specifically to situations in which industries are localised, with many firms in the same area. For example, London's reputation for finance means that many support industries and services, such as consultancies and training providers, are located within easy access.

### Author tip

Remember that both small and large firms benefit from external economies of scale – it is the scale of the industry as a whole that creates the external economies. However, for internal economies of scale it is the scale of the firm that matters.

## Internal economies of scale

The main examples of internal economies of scale are:

### ■ Technical economies of scale

These focus on equipment and resources. Larger firms can benefit from:

- *expensive capital equipment*, which can lead to more efficient production, thus reducing average costs, and higher quality products, which should lead to an increase in demand.
- *economies of increased dimension*. Larger firms can use large-scale equipment

more effectively. For example, doubling the size of containers (for transportation) and doubling the size of warehouses and factories will not lead to a doubling of costs, and so average costs will fall. Smaller firms cannot reap these benefits because they are not producing and transporting enough goods to use these large-scale items efficiently.

- *specialisation and division of labour*. Larger firms can more fully exploit specialisation because they produce enough products for it to be cost effective to divide production into many specialised tasks.
- *efficient use of capital*. Specialisation ensures that each item of capital equipment is used throughout the day, rather than only being used when a worker reaches a certain stage of the production process.
- *research and development*. Large firms can afford the high costs of R&D and are thus more likely to introduce new products or processes. If these are patented, this can restrict competition and thus help firms to increase demand and benefit from high prices.
- *economies of indivisibility*. Many goods need to be produced on a large scale in order to maximise efficiency. Smaller firms may be forced to underuse machines (or use a smaller scale of production that is less efficient) and so small firms will have higher average costs than larger firms.

#### ■ Financial economies of scale

It is cheaper for larger firms to access finance because they present a lower risk than small firms. For this reason they will find it easier to get loans and overdrafts from banks. Furthermore, the terms of the loan, such as the interest rate to be paid and the length of time allowed before repayment, will be more favourable. In effect, the cost of borrowing will be cheaper for larger firms. In addition, larger firms tend to find it easier to sell shares, which are the main form of finance when a limited company is set up. Usually, the most important source of finance is retained profits. Shareholders are entitled to the profits made by limited companies but usually agree that the company can retain a significant percentage of its profits in order to buy items such as buildings and new ICT systems and machinery. Shareholders agree to this because the firm can use the retained profit to make more money in the future and so it helps to boost the share price. This allows shareholders to make money when they decide to sell their shares. In effect, retained profits are a zero cost to firms because there is no interest to be paid (although there is the opportunity cost of what they might have received if they had saved it in a bank instead of using it to buy new capital).

#### ■ Purchasing economies of scale

Greater output means that materials can be bought in bulk at lower cost. Thus large firms can buy materials much more cheaply than small firms. Large firms can also use their size to bargain with suppliers, since many of these suppliers will rely on these large firms continuing to buy the items that they make. Large firms, such as supermarkets, have been able to use this bargaining power to keep down prices of products such as milk.

#### ■ Marketing economies of scale

Costs such as advertising can be spread over more units of output if a firm produces on a large scale and so advertising campaigns tend to be more cost effective for large firms. Smaller firms cannot afford to use expensive forms of media, such as

television advertising, and therefore are less likely to be successful. The internet and online advertising has reduced many costs of advertising and so this economy is less prominent than it was when TV advertising dominated advertising and promotions. However, sophisticated databases still enable larger firms to be able to target their marketing more effectively than most small firms.

■ **Risk-bearing economies of scale**

Large firms often diversify into different products and different markets. They can also ensure that they have a choice of suppliers when ordering supplies. These actions help to protect such firms from sudden changes such as a fall in demand, a decline in a particular market or the liquidation of a supplier. The firm is less at risk from the detrimental effects of these changes because it has alternative products to sell or suppliers from which to buy materials. It should be noted that risk-bearing economies may protect a firm from the risk of overspecialisation, but this can mean that the firm is operating in so many markets that it may not fully exploit the other internal economies of scale.

■ **Managerial economies of scale**

Large firms can employ specialist managers, thus benefiting from division of labour amongst their management staff as well as their production line employees. Large firms can also afford to employ the most successful senior managers. Good decision making by the chief executive, for example, can mean huge increases in profits, and so large firms can usually offer much higher salaries in order to attract the best managers.

**Author tip**

The classification of different internal economies of scale varies between sources. This classification is often a matter of personal preference. However, categorising internal and external economies precisely is important because internal economies result from the firm's scale, whereas external economies of scale result from the scale and localisation of the industry.

**External economies of scale**

If a firm operates within a large industry, this can help it to reduce its average costs regardless of its size, especially if it is located in a place where that particular industry is concentrated.

Examples of external economies of scale are:

■ **Specialist firms and infrastructure**

In areas where an industry is located, specialist firms tend to locate too. These firms may be suppliers, such as suppliers of parts for car manufacturers in the area around Sunderland or close to Oxford and Swindon. Alternatively, they may specialise in distribution, such as refrigerated transport of agricultural products in Lincolnshire. Firms of all sizes will benefit from having easy access to these specialist firms. The local infrastructure, such as the transport system and facilities provided by the local council, might be adapted to suit the needs of firms within industries that dominate that locality. This concentration of firms can also greatly reduce transport costs and so cut average costs.

### ■ Training and education

Local colleges and universities will offer courses suited to the local community, thus helping all firms involved in the main industries in that area. Training firms will also provide facilities and courses geared towards the needs of the local community. In the UK, many business parks have developed in university cities in order to take advantage of the skilled graduates available.

### ■ Reputation

Sometimes a city or area gains an excellent reputation for the provision of a certain good or service. This can greatly assist firms that are located in that city or area, as they can use its reputation to boost sales of their goods. For example, any financial firm operating in the City of London will benefit from its reputation for financial services. Similarly, Milan has a reputation for fashion items.

#### Key note

Both internal and external economies help to cut costs. For internal economies of scale, these benefits apply mainly to large firms; for external economies of scale, they apply to all sizes of firms, but only if they are located close to other firms in the industry.

Developments in ICT have tended to reduce some of the benefits of large-scale production because small firms can often match the ICT resources of larger firms. Similarly, physical locations have become less important in many service industries. As a consequence, ICT has reduced the effect of both internal and external economies of scale.

## Reasons for diseconomies of scale

Diseconomies of scale occur because as the firm (or industry) gets larger, problems begin to appear. These problems lead to higher average costs of production.

### Internal diseconomies of scale

As output increases, some firms find that their average costs of production start to increase rather than decrease. In most instances this tends to occur at the highest levels of output, when the benefits of internal economies of scale seem to have been exhausted or negated.

The main examples of diseconomies of scale are as follows:

#### ■ Coordination difficulties

- There may be a loss of control by management as a firm grows, particularly if the firm has expanded internationally. Growth may also lead to increasing workloads for managers, who thus lose some degree of control of their subordinates.
- Individuals are less likely to follow organisational policies if the level of control is reduced. This may mean that individuals do not follow the company approach and may make poor decisions without senior managers being aware of the problems this might cause.
- In order to help senior managers keep control, large firms often have rigid and inflexible policies, which are imposed to limit the loss of control described above. Greater control can improve efficiency, but by limiting the power given to subordinates, who are likely to be the people dealing directly with customers,

#### Key term

**Internal diseconomies of scale** are the disadvantages that an organisation experiences due to an increase in size. These cause a decrease in productive efficiency and therefore an increase in the average (total) costs of production.

a firm will reduce its ability to respond to customer needs. This is likely to lead to the alienation of customers.

■ **Communication problems**

- In large firms, the effectiveness of communication is reduced. Messages can be distorted and it is possible that communications do not reach everyone. Sometimes the time taken for a message to reach an individual can cause problems for the firm.
- In very large firms, managers may find it more difficult to meet their subordinates and so there may be a lack of understanding of particular issues.
- In large firms, large-scale approaches to communication, such as the circulation of standard emails, are more common. Employees may feel undervalued and demotivated.

■ **Technical diseconomies**

These occur where production on a very large scale becomes extremely difficult to organise as efficiently as smaller-scale production, and so average costs start to increase.

■ **Excessive bureaucracy**

As firms get larger, the number of levels of management increases and this may slow down decision making and add to the costs of production.

■ **Industrial relations problems**

The poor quality of communication within the firms may lead to more staff leaving and higher absences. In extreme cases it may cause industrial disputes, such as a strike. There is a strong correlation between the size of a firm and the number of industrial disputes occurring.

■ **Less flexibility**

In a time of rapid change, flexibility is vital because customer needs are constantly changing. The structure of many large firms makes it difficult for them to adapt to change as quickly as smaller firms.

**External diseconomies of scale**

When an industry becomes very large, this can cause difficulties, especially within the geographical area in which firms in that industry are concentrated. These difficulties can lead to an increase in average costs of production for firms.

■ **Pollution and traffic congestion**

Pollution and traffic congestion are common external diseconomies of scale because facilities often lag behind the growth of an area. Pollution can add costs or lower quality of output, whereas transport problems affect the speed and reliability of transport and thus increase average costs of production.

■ **Competition**

Competition can be fierce because many competitors will be located nearby in order to gain the external economies of the area. This may hinder a firm's ability to attract customers and may lead to lower prices and lower profits.

■ **Higher costs of factors of production**

Due to the high numbers of firms operating in the area, there may be shortages of factors of production, especially land. As a result, average costs may rise.

**Key term**

**External diseconomies of scale** are the disadvantages that an organisation experiences due to growth in the size of the industry within which it operates.

## Internal economies and diseconomies of scale and the shape of the long-run average cost curve

Figure 15.1 shows a typical long-run average cost curve and its relationship to short-run average costs.  $SRAC_1$  shows the short-run cost curve of the firm when it is operating one factory,  $SRAC_2$  shows the short-run cost curve of the firm when it is operating two factories, and so on, up to  $SRAC_7$ , which shows the short-run cost curve of the firm when it is operating seven factories.

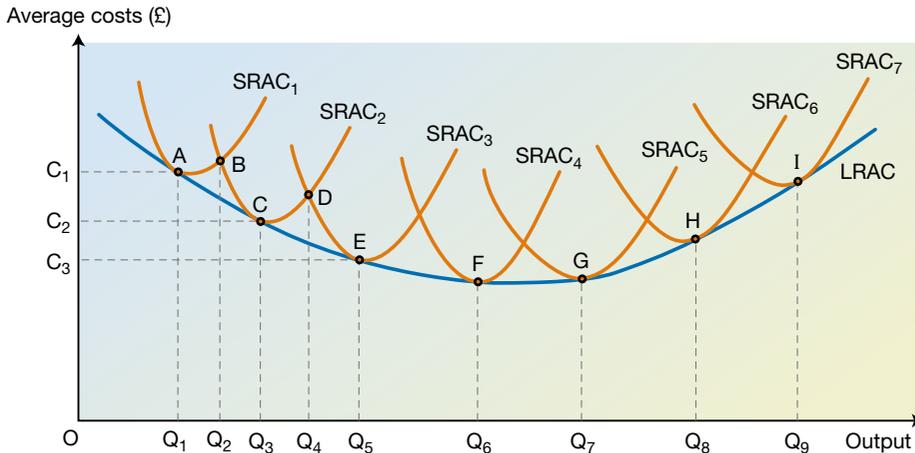


Figure 15.1 The long-run average cost curve

In the short run the firm cannot change its capacity, represented in this diagram by the number of factories it uses. When the firm's demand is low, it will therefore only have a low output of goods. If its output is low, it will only require one factory and so its short-run average cost curve is the line  $SRAC_1$ . If its output is  $OQ_1$ , then its average cost is shown by Point A. Point A shows a high level of productive efficiency as it is close to the lowest point of  $SRAC_1$ . If output increases beyond  $OQ_1$ , average cost starts to rise (as shown by the move from A to B).

In the long run a firm can change its capacity. The firm can see that if output is below  $OQ_2$ , it is cheaper to produce using one factory. However, Point B shows the output ( $OQ_2$ ) at which average costs of production are the same for both one factory or two factories. In the case of one factory, it is approaching maximum capacity and so average costs are rising (as shown by the move from A to B). In the case of two factories, Point B is where average costs are falling, but they are still quite high because fixed costs are much higher if there are two factories.

If output increases beyond  $OQ_2$ , the firm will find its average costs are lower if it operates two factories. In the long-run, firms can add new factories and so, if the firm has anticipated the increasing demand for its goods, it will have planned to have completed a second factory by the time demand (and thus output) reaches  $OQ_2$  units.

The same logic applies when Point C is reached. It is worth keeping the second factory until Point D is reached (at output  $OQ_4$ ). Beyond this level of output it becomes cheaper to operate with a third factory.

Based on the assumption that the firm can only increase its output by building new factories, its long-run average cost curve would follow the sequence shown above and therefore be a line joining points A, B, C, D and E.

In practice, firms can increase output without the need for a new factory. The existing

factory may be extended or additional capital equipment introduced, so that small increases in capacity can be planned. However, once output reaches  $OQ_3$  it would be cheaper to use a second factory rather than use additional capital in a single factory. For this reason, firms can avoid the rise in average costs shown by the move from A to B. The introduction of new capital allows average costs to keep falling so that the long-run average costs line falls along the line from A to C.

Similarly, once the firm is operating with two factories, new capital equipment can be introduced into the second factory as output rises from  $OQ_3$  to  $OQ_4$  to prevent average costs rising from C to D. The introduction of new capital into the second factory allows average costs to keep falling so that the long-run average costs line falls along the line from C to E.

The long-run average cost curve is often described as an ‘envelope’ curve because it envelops the SRAC curves. In Figure 15.1 it is shown by the line ACEFGHI. Why is it shaped like a shallow ‘U’?

The explanation is provided by internal economies and diseconomies of scale. Up until Point F, the LRAC curve slopes downwards, showing a steady fall in average costs of production. This fall in LRAC continues until output reaches  $OQ_6$ . This fall in LRAC is caused by internal economies of scale, which are improving productive efficiency and thus lowering average costs.

Between points A and F there are **internal economies of scale because LRAC is falling**.

Between points F and G the LRAC line is horizontal. This means that average costs are staying at the same level because any internal economies of scale are being cancelled out by internal diseconomies of scale.

Between points F and G there are **constant returns to scale** because LRAC remains the same. Between  $OQ_6$  and  $OQ_7$  units of output, average costs are at their lowest level. This lowest level is first reached at Point F (output level  $OQ_6$ ). This output level ( $OQ_6$ ) is referred to as the **minimum efficient scale (MES)**. If output is lower than the MES, then LRAC is higher. If output is higher than the MES, then LRAC is either equal to or higher than it is at the MES.

From Point G onwards, the diseconomies of scale outweigh the economies of scale and so LRAC slopes upwards.

Between points G and H, and H and I, there are **internal diseconomies of scale because LRAC is rising**.

#### Key note

The LRAC line shown in Figure 15.1 is a very common LRAC curve. However, it is feasible for a firm never to reach an output at which the internal diseconomies of scale outweigh the internal economies of scale. In this case the LRAC line will continue to fall. (In effect it is the LRAC between points A and F in Figure 15.1.)

It is also possible (but less likely) for a firm always to experience rising costs as output increases because the diseconomies outweigh the economies at low levels of output. In this case the LRAC line will continue to rise. (In effect it is the LRAC between points G and I in Figure 15.1.)

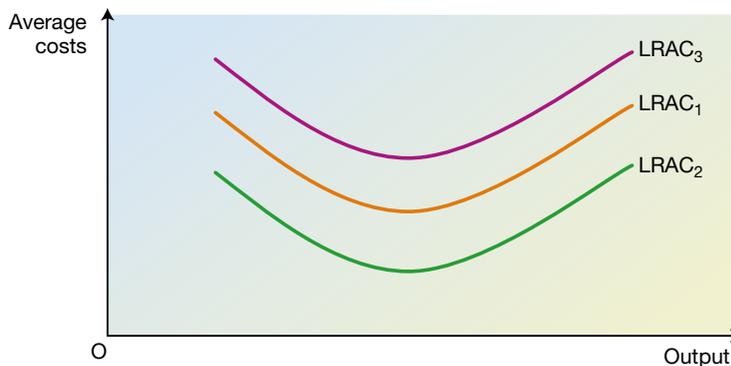
It is possible, but very unlikely, that constant returns to scale may occur. In this case the LRAC line will be horizontal. (In effect it is the LRAC between points F and G in Figure 15.1.)

## External economies and diseconomies of scale and the shape of the long-run average cost curve

Figure 15.1 shows the impact of **internal** economies and diseconomies of scale on the LRAC. This diagram relates average costs of production to the level of output and so the LRAC shows how internal economies reduce LRAC as output rises, but when output is very high the diseconomies are more likely to have an impact.

**External economies and diseconomies of scale** are *not* related to the size/output of the firm; instead they relate to the size of the industry. For this reason their impact is different, as shown in Figure 15.2.

In Figure 15.2,  $LRAC_1$  is the original long-run average cost curve of the firm. External factors, such as improvements in the local infrastructure or local training courses being provided for employees, lead to greater efficiency. This lowers the average cost curve for the firm as, regardless of its level of output, these improvements enable it to produce goods at a lower average costs. As a result of these external economies of scale, the long-run average cost curve falls from  $LRAC_1$  to  $LRAC_2$ .



**Figure 15.2** The impact of external economies and diseconomies of scale on the LRAC curve

If the firm experiences external diseconomies of scale, perhaps through traffic congestion or more expensive raw materials in the local area, then its average costs will rise. As a result of these external diseconomies of scale, the long-run average cost curve rises from  $LRAC_1$  to  $LRAC_3$ .

### REALWORLD ECONOMICS 15.1

## Economies of scale in shipping

In the 1950s, the 'container' was introduced into trade between nations. These standardised metal storage units are based on a standard size (known as a TEU) of 20 feet by 8 feet (or 2 TEUs – 40 feet by 8 feet). These dimensions allow them to fit neatly into rail, road and sea transportation systems.

For long-haul trade, the 1950s

container ships could carry between 500 and 800 TEUs. In 2013, shipping giant Maersk launched its latest container ship – able to carry 18,000 TEUs.

Larger ships lead to three main internal economies of scale. The capital costs of these massive ships (over 400 metres in length) are high, but not in relation to the amount

of cargo that they can carry. The design of these vessels means that a crew of only 13 people is needed per journey, instead of the 23 people needed on much smaller ships. Fuel costs are also lower. The combination of these factors has meant that the cost of transporting a container from China to Europe is \$218 for an 18,000 TEU container

ship compared to \$333 for a 13,000 TEU container ship. Other ship operating costs, such as insurance and administration, are also lower. The efficiency of these ships is so great that the cost of transporting T-shirts from China to Europe amounts to less than

1.5 pence per T-shirt.

However, to date only the world's largest shipping line – the Danish firm, Maersk – has been able to justify the expenditure needed to order these new ships from their South Korean shipyards. Many countries, such as New Zealand, do

not have ports that can cope with such large ships, and some trade routes, such as those using the Panama Canal, are unsuitable for these ships.

There is also the question of capacity. On average, Maersk's ships operate at 85% capacity between Asia and Europe, but they are only 55% full on journeys from Europe to Asia. The slowing growth rate in China and the recession and its after-effects have led to a situation in which there is too much capacity in the shipping industry. Maersk is relying on economies of scale to push smaller competitors out of the market, since they are unable to afford these huge container ships.



The latest container ship – Maersk exploits economies of scale to gain competitive advantage

**Discussion point**

Identify the main internal and external economies and diseconomies of scale featured in the article. Do you think Maersk is taking too big a risk?

## Review questions

Total: 25 marks

- 1 A firm is able to reduce its average costs because it is located in close proximity to other firms within the industry in which it operates. This is an example of an:
  - A External economy of scale
  - B External diseconomy of scale
  - C Internal economy of scale
  - D Internal diseconomy of scale

(1 mark)
  
- 2 Internal economies of scale must exist when:
  - A Average cost rises as the firm's output rises
  - B Total cost rises as the firm's output rises
  - C Average cost falls as the industry's output grows
  - D Average cost falls as the firm's output rises

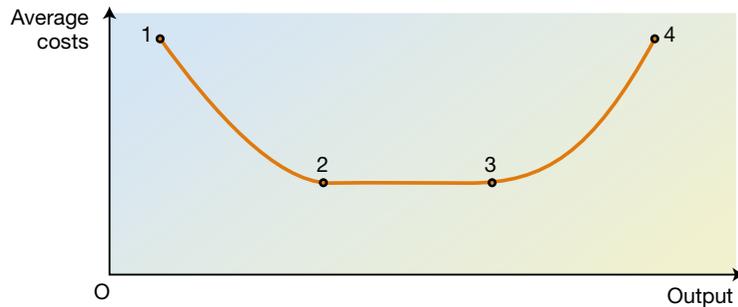
(1 mark)
  
- 3 Which *one* of the following is most likely to be an internal economy of scale?
  - A Bulk buying
  - B Industrial disputes
  - C Reputation of an area
  - D Specialised transport firms in the region

(1 mark)

- 4 Which *one* of the following is most likely to be an external economy of scale?
- A Ease of acquiring finance
  - B Good transportation networks
  - C Research and development skills
  - D Indivisibility of production

(1 mark)

Figure 15.3 applies to questions 5, 6 and 7.



**Figure 15.3**

- 5 Diseconomies of scale occur between points:
- A 1 and 2
  - B 2 and 3
  - C 2 and 4
  - D 3 and 4
- (1 mark)
- 6 The minimum efficient scale (MES) is at point:
- A 1
  - B 2
  - C 3
  - D 4
- (1 mark)
- 7 Productive efficiency is at its greatest between points:
- A 1 and 2
  - B 1 and 3
  - C 2 and 3
  - D 3 and 4
- (1 mark)
- 8 Explain one possible cause of external economies of scale. (4 marks)
- 9 Distinguish between internal economies of scale and external diseconomies of scale. (6 marks)
- 10 Explain two possible causes of internal economies of scale. (8 marks)