# Perfect Competition and Pure Monopoly: The Limiting Cases of Market Structure

*An industry is the set of all firms making the same product. The output of an industry is the sum of the outputs of its individual firms. Yet different industries have very different numbers of firms. Almost all electricity in the UK is produced by the Central Electricity Generating Board which we call a nationalized industry because it is owned and run by the state. However, some sole suppliers are private firms. Until recently almost all car windscreens in the UK were made by the Pilkington glass company. Until its patent expired, IBM was the sole supplier of golfball typewriters. In contrast, the UK began the 1980s with over 250 000 farms, 40 000 grocers, and 14 000 furniture retailers.*

How can we analyse how price and output are determined for an industry as a whole? Some hints were offered in Chapter 3, where we discussed the interaction of market demand and supply. Since then we have refined our analysis of demand in Chapters 4 and 5 and our analysis of the output supply decision of the individual firm in Chapters 6 and 7. We now combine the supply decisions of the individual firms to derive the industry supply curve and examine its interaction with the market demand curve to determine price and output for the industry as a whole.

How is this analysis affected by the size and number of firms in an industry? Why indeed do some industries have many firms but others only one? These are questions about *market structure*.

> The structure of a market is a description of the behaviour of buyers and sellers in that market.

In the next chapter we develop a general theory of market structure, showing how demand and cost conditions together determine the number of firms and their behaviour. First it is useful to establish two benchmark cases, the opposite extremes between which all other types of market structure must lie. These limiting cases are *perfect competition* on the one hand and *monopoly* or *monopsony* on the other hand.

> A *perfectly competitive* market is one in which both buyers and sellers believe that their own buying or selling decisions have no effect on the market price. A *monopolist* is the only seller or potential seller of the good in that industry. A *monopsonist* is the only buyer or potential buyer of the good in that industry.

In this and the following chapter we are interested primarily in the relationship between the number of sellers and the behaviour of sellers. For the moment we neglect the possibility of monopsony. We assume that there are many buyers whose individual downward-sloping demand curves can be aggregated to yield the market demand curve. We take up the possibility of monopsony in Chapter 10 when discussing the market not for outputs but for inputs such as labour. At present we assume that the demand side of the market is competitive, and we contrast the limiting cases on the supply side.



The economist's definition of perfect competition is different from the meaning of competition in everyday usage. The economist means that each individual, recognizing that his or her own quantities supplied or demanded are trivial relative to the market as a whole, acts on the assumption that his or her actions will have no effect on the market price. This assumption was built into our model of consumer choice in Chapter 5. Each consumer constructed a budget line on the assumption that market prices were *given* and unaffected by the quantities the consumer then chose. Changes in *market* conditions, applying to all firms or all consumers, would change the equilibrium price and hence individual quantities demanded, but each individual could neglect any feedback from his or her own actions to market price.

This concept of competition, which we now extend to firms, differs from everyday usage. Ford and Vauxhall are fighting each other vigorously for the UK motor car market, but an economist would not call them perfectly competitive. Each commands such a large share of the total market that changes in their quantities supplied will affect the market price. Each must take account of this in deciding how much to supply. They cannot regard themselves as *pricetakers*. Only under perfect competition can individuals make decisions that treat the price as independent of their own actions.

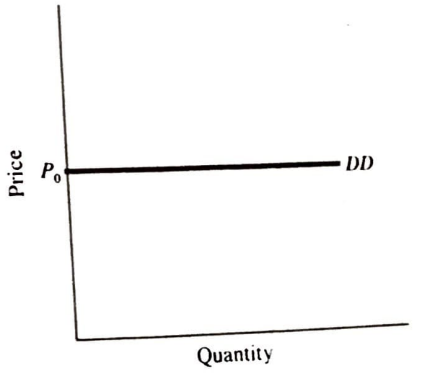
## 8-1 PERFECT COMPETITION

A perfectly competitive industry, in which all firms and consumers believe that their own actions have no effect on market price, must have many buyers and many sellers. Agricultural markets are a good example. In London the New Covent Garden fruit market confronts many buyers with many sellers. Neither buyers nor sellers believe their own actions affect the market price.

Firms in a perfectly competitive industry face a flat or horizontal demand curve as shown in Figure 8-1. No matter how much the firm sells it gets exactly the market price. If it tries to charge a price in excess of $P_0$ it will not sell any output: buyers will go to one of the other firms whose product is just as good. Since the firm can sell as much as it wants at $P_0$, there is no point contemplating a price lower than $P_0$. The individual firm's demand curve is $DD$.

This horizontal demand curve for its product is the crucial feature of a perfectly competitive firm. For this to be a plausible description of the demand curve facing the firm, we really need to have in mind an industry with four characteristics. First, there must be a large number of firms in the industry so that each is trivial relative to the industry as a whole. Second, the firms must be making a reasonably standard product, such as wheat or potatoes. Even if the car industry had a large number of firms it would not be sensible to view it as a competitive industry. A Ford Sierra is not a perfect substitute for a Vauxhall Cavalier. The more imperfect they are as substitutes, the more it will make sense to view Ford as the sole

**FIGURE 8-1 THE COMPETITIVE FIRM'S DEMAND CURVE.** A competitive firm can sell as much as it wants at the market price $P_0$. Its demand curve $DD$ is horizontal at this price.




**The competitive firm's demand curve**

supplier of Sierras and Vauxhall as the sole supplier of Cavaliers. Each producer will then cease to be trivial relative to the relevant market and will no longer be able to act as a price-taker.

This example alerts us to the problem of which goods can be grouped together within the same market or industry. We return to this issue in the next chapter. For the moment we can evade this issue. In a perfectly competitive industry all firms must be making essentially the same product, *for which they must all charge the same price.*

Even if all firms in an industry made *homogeneous* or identical goods each firm might be able to have some discretion over the price it charged if buyers have imperfect information about the quality or characteristics of the products of different firms in the industry. If you don't know much about cars you may think that a 1970 Ford Cortina being sold for £1000 must be in a better condition than a 1970 Ford Cortina being sold for £500. Hence, if we are to regard each firm in a competitive industry as being unable to affect the price for which it sells its output, it is not sufficient that all firms are selling a homogeneous product. We must also assume that buyers have almost perfect information about the characteristics of the products being sold so that they know the products of different firms in a competitive industry really are identical.

Even so, why don't all the firms in the industry do what OPEC did in 1973–74? If existing firms collectively restrict supply, they can increase the price of their output by moving the industry up its market demand curve. If the analysis of price-taking perfectly competitive firms is to have any relevance we must explain why such collective action is impossible.

One answer is that, with so many firms in the industry, the costs of organizing themselves into a cohesive group might be prohibitive. Think of all the committee meetings that would be needed. Managers might spend more time negotiating with other firms than organizing production. Nevertheless, if the market demand curve is very inelastic, the potential increase in revenue from such co-operation could be enormous, as OPEC discovered. We need a more profound answer to rule out co-operation.

Thus the fourth crucial characteristic of a perfectly competitive industry is *free entry and exit.* Even if existing firms could organize themselves to restrict total supply and drive up the market price, the consequent increase in revenues and profits would simply attract new firms into the industry, thereby increasing total supply again and driving the price back down. Conversely, as we shall shortly see, when firms in a competitive industry are losing money, some firms will close down and, by reducing the number of firms remaining in the industry, reduce the total supply and drive the price up, thereby allowing the remaining firms to survive.

To sum up, each firm in a competitive industry faces a horizontal demand curve for its product at the going market price. To be a reasonable description of the demand conditions facing a firm, the industry must have four characteristics: (1) many firms, each trivial relative to the industry as a whole; (2) a standardized or homogeneous product, so that it is legitimate to examine the industry as a whole rather than a series of sub-industries each with many fewer firms; (3) perfect customer information about product quality so that buyers recognize that the identical products of different firms really are the same; and (4) free entry and exit so there is no incentive for existing firms to collude.[1]

## 8-2 THE FIRM'S SUPPLY DECISION UNDER PERFECT COMPETITION

In Chapter 7 we developed a general theory of the supply decision of the individual firm in the short run and in the long run. First the firm uses the marginal condition ($MC = MR$) to find the best positive level of output; then it uses the average condition to check whether the price for which this output could be sold covers the relevant measure of average cost.

This general theory must hold for the special case of perfectly competitive firms. *The special feature of perfect competition is the relationship between marginal revenue and price.* The competitive firm faces a horizontal demand curve as in Figure 8-1. Unlike the more general case, in which the firm faces a downward-sloping demand curve, the competitive firm does *not* bid down the price as it sells more units of output. Since there is no effect on the revenue from existing output, the marginal revenue from an additional unit of output is simply the price received.

This special feature of a perfectly competitive firm has far-reaching consequences. It is so important we show this feature as equation (1):

$$\text{(Marginal revenue) } MR = P \text{ (price)} \qquad (1)$$

### The Firm's Short-run Supply Curve

Figure 8-2 shows again the short-run cost curves – marginal cost *SMC*, average total cost *SATC*, and average variable cost *SAVC* – shown in Figure 7-10. From equation (1) the marginal condition for the best level of positive output now implies

$$SMC = MR = P \qquad (2)$$

Suppose the firm faces a horizontal demand curve at the price $P_4$ in Figure 8-2. Equation (2) implies that the firm chooses the output level $Q_4$ to reach the point *D*, at which price equals marginal cost.

Next, the firm checks whether it would rather shut down in the short run. From Chapter 7 we know that it will shut down only if the price $P_4$ at which output can be sold fails to cover short-run variable costs of producing this output. In Figure 8-2 $P_4$ exceeds *SAVC* at the output level $Q_4$. Not only does the firm wish to produce this output, it also makes profits in the short run. The point *D* lies above the point *G*, the short-run average total cost (including overheads) of producing $Q_4$.

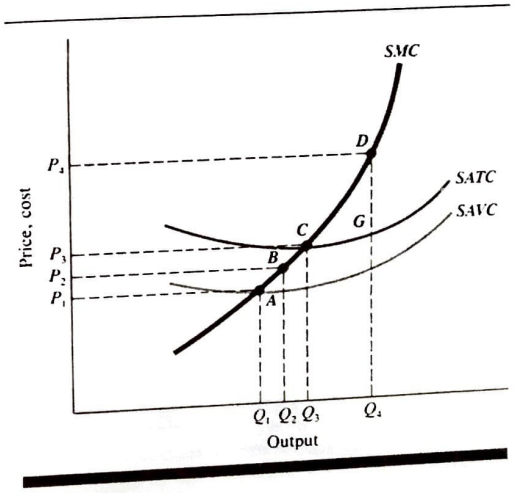



FIGURE 8-2 SHORT-RUN SUPPLY DECISIONS OF THE PERFECTLY COMPETITIVE FIRM. The perfectly competitive firm produces at that level of output at which price is equal to marginal cost, provided it makes more profit by producing some output than none at all. The firm's short-run supply curve is the *SMC* curve above the point *A*, the shutdown point below which the firm cannot cover average variable costs *SAVC* in the short run.

[1] Many factors may inhibit entry and exit. IBM's patent prevented other firms entering the golfball typewriter industry. Until 1980 de Beers controlled virtually all diamond mines in the non-communist world, preventing new firms entering the diamond industry. Some economists believe that in many countries doctors and lawyers, acting through their professional bodies, have restricted entry to the medical and legal professions. Conversely, the UK government has sometimes put political pressure on the National Coal Board not to exit from loss-making activities in South Wales where the NCB is a major employer of local labour.

Suppose the firm had faced a different price. In the short run the firm should produce positive output for any price above $P_1$. Any price below $P_1$ lies below the minimum point on the *SAVC* curve and the firm cannot find an output at which price covers *SAVC*. Given any price such as $P_2$, above $P_1$, the firm produces $Q_2$, the output at which price equals marginal cost.

> The curve showing the quantity the firm wants to produce at each price is the firm's *supply curve*.

The short-run supply curve is thus the *SMC* curve above point *A*, the point at which the *SMC* curve crosses the lowest point on the *SAVC* curve.

Between points *A* and C (prices $P_1$ and $P_3$) the firm will be making short-run losses, since price is less than average total cost. But it will be recouping some of its overheads. At any price above $P_3$, the point at which the *SMC* curve crosses the lowest point on the *SATC* curve, the firm is making short-run profits. For example, at the price $P_4$ the profit per unit of output is the distance *DG*, the difference between price and average total cost per unit of output. Remember that these profits are economic or supernormal profits after allowing for the economic costs, including the opportunity costs of the owners' financial capital and work effort, summarized in the *SAVC* and *SATC* curves.

The price $P_1$ is called the *shutdown price*, the price below which the firm reduces its losses by choosing not to produce at all.

### The Firm's Long-run Supply Curve

The same principles apply in deriving the long-run supply curve of the perfectly competitive firm. Figure 8-3 shows the firm's average and marginal costs in the long run. Remember that the long-run marginal cost curve *LMC* will be flatter than the *SMC* curve since the firm can freely adjust all factors of production only in the long run.

Facing a price $P_4$, the marginal condition leads the firm to choose the long-run output level $Q_4$ at the point *D*. Again we must check whether it is better to shut down than to produce this output. In the long run, shutting down means leaving the industry altogether.

In the long run the firm exits from the industry only if price fails to cover long-run average cost *LAC* at the best positive output level. At the price

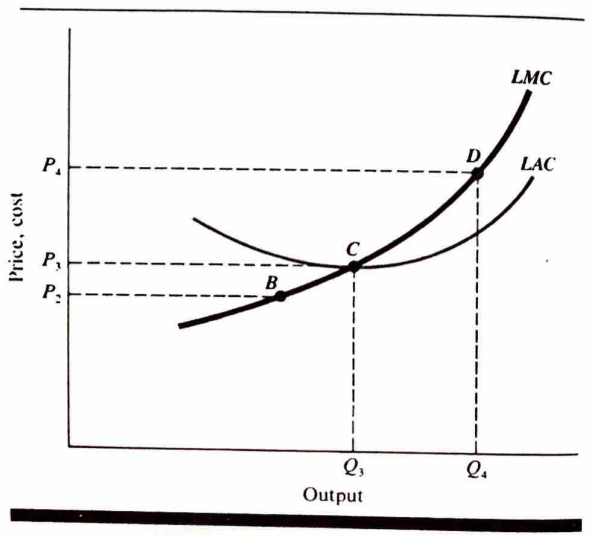



**FIGURE 8-3 LONG-RUN SUPPLY DECISIONS OF THE PERFECTLY COMPETITIVE FIRM.** The perfectly competitive firm produces at that level of output at which *P* is equal to marginal cost, provided it makes more profit by producing some output than none at all. It therefore chooses points on the *LMC* curve. At any price above $P_3$ the firm makes profits because price is above long-run average cost (*LAC*). At any price below $P_3$, such as $P_2$, the firm makes losses because price is below long-run average cost. It therefore will not produce any output at prices below $P_3$. The long-run supply curve is the *LMC* curve above point *C*.



$P_2$ the marginal condition leads to the point *B* in Figure 8-3, but the firm is losing money and should leave the industry in the long run.[2]

Thus the firm's *long-run supply curve*, the schedule relating output supplied to price in the long run, is the portion of the *LMC* curve to the right of point C corresponding to the price $P_3$. At any price below $P_3$ the firm can find no positive output at which price covers *LAC*. At the price $P_3$ the firm would produce $Q_3$ and just break even after paying all its economic costs. The firm would be making only normal profits.

> When economic profits are zero we say the firm is making *normal profits*. Its accounting profits just pay the opportunity cost of the owners' money and time.

### Entry and Exit

The price $P_3$ corresponding to the minimum point on the *LAC* curve is called the *entry* or *exit price*. Firms are making only normal profits. There is no incentive to enter or leave the industry. The resources tied up in the firm are earning just as much as their opportunity costs, what they could earn elsewhere.

Any price less than $P_3$ will induce the firm to leave, or exit from, the industry in the long run. At any price above $P_3$ the firm can find a long-run output level, such as $Q_4$ in Figure 8-3, that yields supernormal profits. $P_3$ is the minimum price required to keep the firm in the industry.

However, we can also interpret Figure 8-3 as the decision facing a potential entrant to the industry. The cost curves now describe the post-entry costs, which may be higher than the costs of existing firms in the industry. For example, if existing firms have all the best locations, new entrants may have to build factories further away from the market and incur higher transport costs. Nevertheless, $P_3$, the price that just covers the lowest average cost at which the entrant could produce, is the critical point at which entry becomes attractive. Any price above $P_3$ yields supernormal profits and means that the return on the owners' time and money will be higher than their opportunity costs, the highest return that these resources could earn elsewhere in the economy.

### The Long-run and Short-run Supply Decisions of the Competitive Firm

Figure 8-4 summarizes the preceding discussion. For each level of fixed factors there exists a different *SMC* curve and short-run supply curve (*SRSS*). The long-run supply curve (*LRSS*) is flatter than *SRSS* because extra factor flexibility in the long run makes the *LMC* curve flatter than the *SMC* curve. The *SRSS* curve starts from a lower shutdown price because in the short run the firm will produce if it can cover average variable costs. In the long run all costs are variable and must be covered if the firm is to remain in the industry. In either case, the competitive firm's supply curve is the part of the marginal cost curve above the point at which it is better to produce no output at all. Table 8-1 sets out this principle.

## 8-3 THE INDUSTRY SUPPLY CURVES

A competitive industry comprises many firms. In the short run two things are fixed: the quantity of fixed factors employed by each firm in the industry, and the number of firms in the industry. In the long run, each firm can vary all its factors of production, but the number of firms can also change through entry and exit from the industry.

### The Short-run Industry Supply Curve

Just as we can add individual demand curves by buyers to obtain the market demand curve, we

[2] The firm's behaviour in the short run and in the long run is rather like the behaviour of a good poker player. In the short run the poker player is dealt a particular hand and plays the hand if it is likely to be profitable. If not, the player temporarily shuts down by throwing in the hand. Over time the player gets new hands just as a firm can gradually rearrange its factors of production. When the poker player realizes that, whatever the cards, the long-run outlook is bad because other players are better, the player should leave the game altogether. Similarly, if a firm realizes that, however it adjusts its factors of production, it is going to make losses in the long run, it should leave the industry.

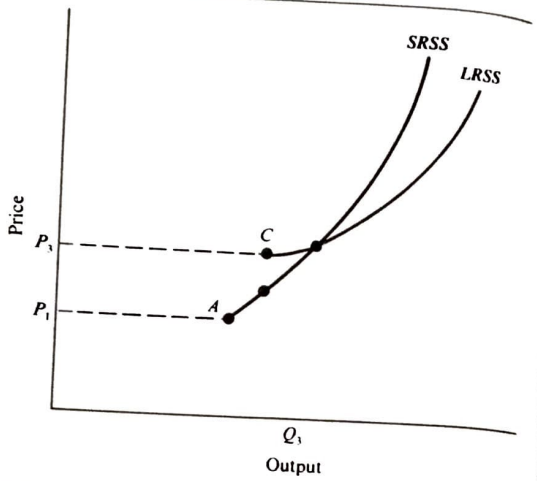



**FIGURE 8-4 SHORT- AND LONG-RUN SUPPLY CURVES OF THE COMPETITIVE FIRM.** Taken from the two previous figures, the short-run supply curve is the firm's *SMC* curve above *A* and the long-run supply curve *LRSS* is the firm's *LMC* curve above *C*. $P_1$ is the shutdown price in the short run and $P_3$ the entry and exit price in the long run. If the firm happens to begin with the stock of fixed factors it would choose at the lowest point on its *LAC* curve, then *C* will actually lie on the *SRSS* curve.

can add the individual supply curves of firms to obtain the industry supply curve. Figure 8-5 shows how. It does not specify whether we are discussing the short or the long run. Either way, at each price we add together the quantities supplied by each firm to obtain the total quantity supplied at that price.

In the short run the number of firms in the industry is given. Suppose there are two firms, A and B. Each firm's short-run supply curve is the part of its *SMC* curve above its shutdown price. Figure 8-5 assumes that firm A has a lower shutdown price than firm B. Firm A has a lower *SAVC* curve, perhaps because of a more favourable geographical location or superior technical know-how. Each firm's supply curve is horizontal at the shutdown price. At a lower price, no output is supplied.

At each price, the industry supply $Q$ is the sum of $Q^A$, the supply of firm A, and $Q^B$, the supply of firm B. Thus at the price $P_3$, $Q_3 = Q_3^A + Q_3^B$. The industry supply curve is the horizontal sum of the separate supply curves. Notice the industry supply curve is discontinuous at the price $P_2$. Between $P_1$ and $P_2$ only the lower-cost firm A is producing. At $P_2$ suddenly firm B starts to produce as well.

When there are many firms, each with a different shutdown price, there are a large number of very small discontinuities as we move up the industry supply curve. In fact, since each firm in a competitive industry is trivial relative to the total, the industry supply curve is effectively smooth.

**TABLE 8-1**
**THE SUPPLY DECISION OF THE PERFECTLY COMPETITIVE FIRM**

| MARGINAL CONDITION | AVERAGE CONDITION | |
|---|---|---|
| | SHORT-RUN | LONG-RUN |
| Produce output where $P = MC$ | If $P < SAVC$ shut down temporarily | If $P < LAC$ leave industry |

## Comparing Short and Long-run Industry Supply Curves

Figure 8-5 may also be used to derive the long-run industry supply curve. For each firm the individual supply curve is the portion of the *LMC* curve above the firm's entry and exit price. However, unlike the short run, the number of firms in the industry is no longer fixed. Not only can existing firms leave the industry, but also new firms can enter. Instead of horizontally aggregating at each price the quantities supplied by the existing firms in the industry, we must horizontally aggregate the quantities supplied by existing firms *and firms that might potentially enter the industry.*

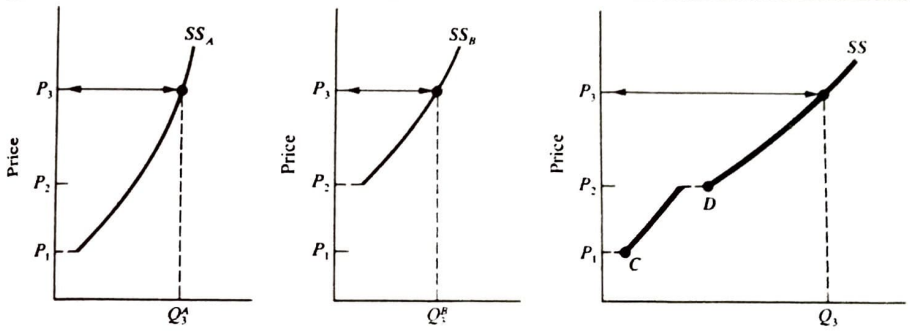
At a price below $P_2$ in Figure 8-5 firm B will not be in the industry in the long run. As we contemplate prices above $P_2$ we must recognize that firm B *will* wish to enter the industry in the long run. As the market price rises, the total industry supply rises in the long run for two distinct reasons: each existing firm will move up its long-run supply curve, and new firms will find it profitable to enter the industry.

Conversely, at lower prices, the higher-cost firms will begin to lose money and will decide to leave the industry. Entry and exit in the long run play a role analogous to shutdown in the short run. In the long run, entry and exit affect the number of producing firms whose output must be horizontally aggregated to obtain the industry supply. In the short run, although the number of firms in the industry is given, the fraction that is producing rather than being temporarily shut down is not. Again, the industry supply curve is the horizontal sum of the outputs of those actually producing at the given market price.

Figure 8-6 illustrates what these arguments imply about the relation between short-and long-run industry supply curves. The long-run supply curve is flatter for two reasons: each firm can vary its factors more appropriately in the long run and



**FIGURE 8-5 DERIVING THE INDUSTRY SUPPLY CURVE.** The industry supply curve *SS* shows the total quantity supplied at each price by all the firms in the industry. It is obtained by adding at each price the quantity supplied by each firm in the industry. With only two firms A and B the figure shows how at a price such as $P_3$ we add $Q_3^A$ and $Q_3^B$ to obtain the output $Q_3$ on the industry supply curve. Since firms can have different shutdown prices or entry and exit prices, the industry supply curve can have step jumps at points such as *C* and *D* where an extra firm starts production. However, with many firms in the industry, each trivial relative to the industry as a whole, the step jumps in the industry supply curve when another firm starts production are so small that we can effectively think of the upward sloping industry supply curve as smooth.



has a flatter supply curve (Figure 8-4); and higher prices attract additional firms into the industry, causing industry output to rise by more than the additional output supplied by the firms previously in the industry.

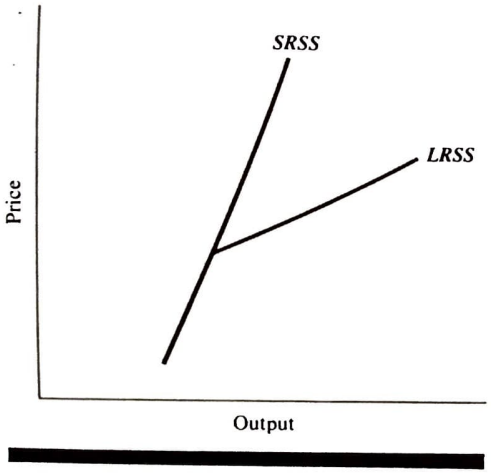
Conversely, when the price falls firms initially move down their (relatively steep) short-run supply curves. Provided short-run average variable costs are covered firms will continue to produce and may not reduce output very much. In the long run each firm will reduce output further since all factors of production can now be varied. In addition some firms will leave the industry since they are no longer covering long-run average costs. Thus, in response to a price reduction, industry output will fall by more in the long run than it does in the short run.

The Marginal Firm Suppose there is a large number of firms, each making the same product for sale at the same price but having slightly different cost curves. Figure 8-7 shows the cost curves for two firms, a low-cost firm A and a high-cost firm B. Some firms have costs lying between those of A and B and others have even higher costs than B.

The long run is the period in which all adjustment – both in factors and in number of firms – has been completed. There is no further entry and exit. Suppose the long-run price is $P^*$ in Figure 8-7. The low-cost firm A is producing $Q_A$ and making healthy profits, since $P^*$ exceeds $LAC$ at the output $Q_A$. Slightly higher-cost firms are making slightly less profit. Firm B is the last firm that can survive in the industry. It is just breaking even producing $Q_B$. It is the *marginal firm* in the industry. A slight price fall would force it to leave the industry.

All firms with higher costs than firm B cannot compete in the industry if the long-run price is $P^*$. Suppose one potential entrant has an $LAC$ curve whose lowest point is only slightly above $P^*$. It is the *marginal firm* waiting to enter the industry. If anything causes $P^*$ to rise a little, this marginal firm will enter the industry.

**FIGURE 8-6 SHORT- AND LONG-RUN INDUSTRY SUPPLY CURVES.** The long-run industry supply curve *LRSS* is flatter than the short-run industry supply curve *SRSS*. Each firm has a flatter supply curve in the long run because inputs can be varied more appropriately than in the short run. The *LRSS* curve also reflects changes in the number of firms in the long run as firms enter or exit from the industry.

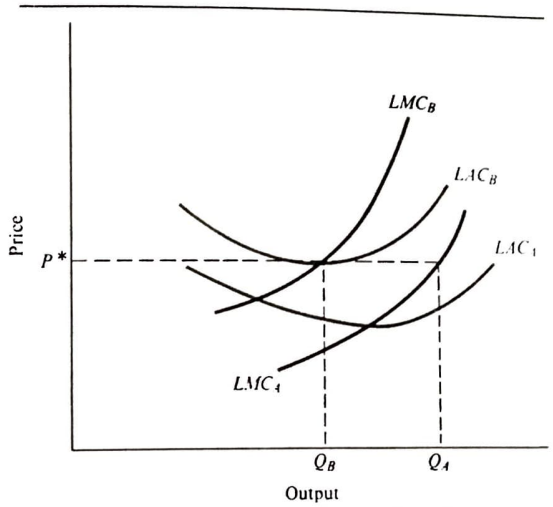



## The Horizontal Long-run Industry Supply Curve

Each firm has a rising $LMC$ curve and hence a rising long-run supply curve. The industry supply curve is somewhat flatter. Higher prices not merely induce existing firms to produce more; they also induce new firms to enter the industry. In the extreme case the industry long-run supply curve is horizontal. This case occurs when all existing firms and potential entrants have *identical cost* curves. This is illustrated in Figure 8-8. Each firm has the same $LAC$ curve and will supply along the part of its $LMC$ curve that is not below $C$. Any supply curve shows the minimum price required to elicit a certain quantity of output. Below $P^*$ no firm will wish to supply. Although it takes a price above $P^*$ to persuade each individual firm to produce more than $Q_1$, no higher price than $P^*$ is required to expand industry output.



**FIGURE 8-7 THE MARGINAL FIRM IN THE INDUSTRY.** Suppose firms have different cost curves. Firm A, the lowest-cost firm in the industry, has long-run average costs $LAC_A$ and marginal costs $LMC_A$. Firm B faces much higher costs $LAC_B$ and $LMC_B$. Other firms have intermediate costs. At the price $P^*$ firm A produces $Q_A$ and makes profits. Firm B produces $Q_B$ and just breaks even. Firm B is the marginal firm in the industry, the highest-cost producer that can remain in the industry in the long run.



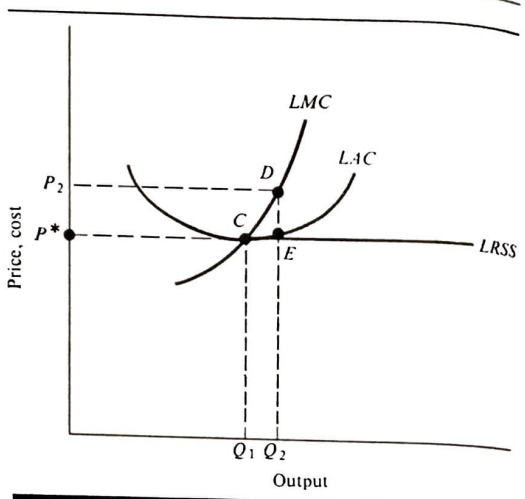

Consider any price such as $P_2$ above $P^*$. Each firm produces $Q_2$ and makes supernormal profits since point $D$ lies above point $E$. Since potential entrants face the same cost curves, there would be a flood of new firms entering the industry. In fact, we would say that the output of the industry would be infinite. There would be an infinite number of firms each producing $Q_2$. A similar argument applies for any price in excess of $P^*$, the price that just covers minimum average costs at the output $Q_1$.

Hence, for any finite output, the industry supply curve is horizontal in the long run at the price $P^*$. It is not necessary to offer a higher price to bribe existing firms to move up their individual supply curves. Industry output can be expanded by the entry of new firms alone. At any price below $P^*$ no output will be produced. In Figure 8-8 we show the long-run industry supply curve $LRSS$ as a horizontal line at the price $P^*$. Moving along this line, we are simply adding more and more firms each producing $Q_1$.

There are two reasons why the general case of a rising long-run industry supply curve is much more likely than the special case of a horizontal long-run supply curve for a competitive industry. First, it is unlikely that every firm and potential firm in the industry has identical cost curves. For example, some firms are likely to have better managers or a more favourable location for producing the industry's product. When firms have different costs there cannot be unlimited expansion of industry output at a constant price. Existing firms face rising marginal costs, and high-cost firms, presently excluded from the industry, require higher prices before it is profitable to enter the industry and contribute to supply.

Second, even if all firms face the same cost curves, the industry long-run supply curve may not be horizontal. We draw a cost curve for given technology *and* given input prices. Although each firm is small relative to the total and can affect neither output prices nor input prices when it acts alone, the collective expansion of output by all firms may bid up input prices. If so, it requires a higher output price to allow an increase in industry output that will bid up input prices and shift the cost curves for each individual firm upwards. Without a rise in the output price firms cannot survive in the industry at the higher level of

**FIGURE 8-8 THE HORIZONTAL LONG-RUN INDUSTRY SUPPLY CURVE.** When all existing firms and potential entrants have identical costs, industry output can be expanded without offering a price higher than $P^*$ to induce firms to move up their *LMC* curves. $P^*$ is the price at which entrants can survive in the industry in the long run. The long-run industry supply curve is the horizontal line *LRSS* at $P^*$. Industry output can be indefinitely expanded at this price by increasing the number of firms that each produce $Q_1$.




industry output and individual costs. Thus in general we expect the long-run supply curve of the industry to be rising. It requires a higher price to call forth a higher total output.

## 8-4 COMPARATIVE STATICS FOR A COMPETITIVE INDUSTRY

Having discussed the industry supply curve in the short run and the long run, we can now examine how supply and demand interact to determine equilibrium price in the short run and the long run.

> In *short-run equilibrium* the market price equates the quantity demanded to the total quantity supplied by the *given* number of firms in the industry when each firm produces on its short-run supply curve.
>
> In *long-run equilibrium* the market price equates the quantity demanded to the total quantity supplied by the number of firms in the industry when each firm produces on its long-run supply curve. Since firms can freely enter or exit from the industry, the marginal firm must make only normal profits so that there is no further incentive for entry or exit.

We now examine equilibrium in a competitive industry and apply the method of comparative static analysis introduced in Chapter 3.

> *Comparative statics* examines how equilibrium changes when there is a change, for example, in demand or cost conditions.

### The Effect of an Increase in Costs

First we discuss the effect of an increase in costs that hits all firms. Perhaps there has been an increase in the price of a raw material, or in the wage rate which must be paid to workers in the industry. For simplicity, we discuss the case in which all firms face the same costs and the long-run supply curve of the industry is horizontal. The same general principles carry over to the case in which the industry supply curve slopes upwards in the long run.

Figure 8-9 summarizes the implications of our analysis of a competitive industry. The industry faces the downward-sloping demand curve *DD*. Initially, the long-run supply curve is $LRSS_1$ and the market clears at the price $P_1^*$ and the total output $Q_1^*$. The short-run industry supply curve is $SRSS_1$. The market is in both short-run and long-run equilibrium.

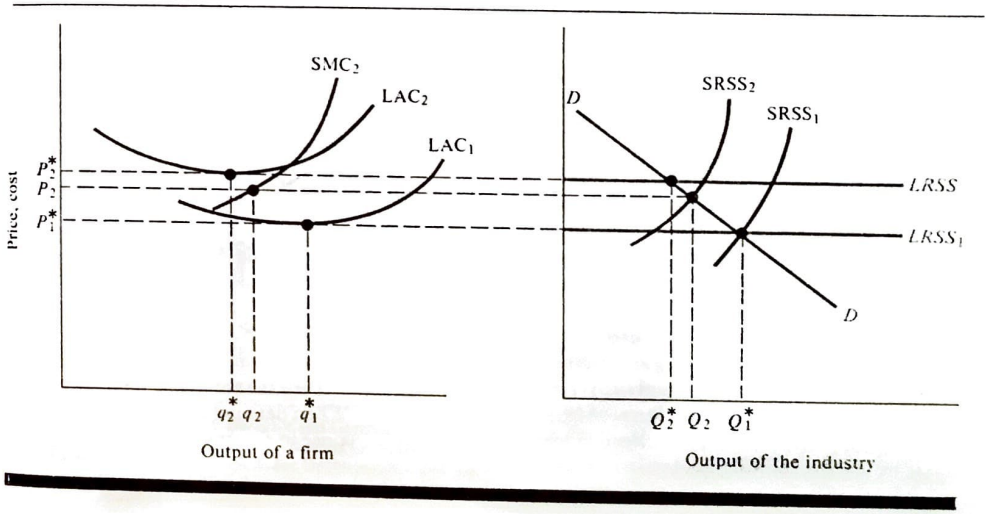
The left-hand figure shows what is going on at the level of the firm. Each firm is producing $q_1^*$ at the lowest point on its average cost curve $LAC_1$. This point must also be the lowest point on its *SATC* curve and hence also lies on its *SMC* curve, though the initial position of these two curves is not shown in Figure 8-9. If there are $N_1$ firms in the industry, total output $Q_1^*$ is $N_1$ times the individual firm's output $q_1^*$.

Now suppose there is an increase in input prices that raises costs for all firms. $LAC_2$ is the new long-run average cost curve for a firm. In the short run the firm has some fixed factors. $SATC_2$ and $SAVC_2$ depict average total and average variable costs at this level of fixed factors. Short-run marginal costs $SMC_2$ pass through the lowest point of both these curves. The part of $SMC_2$ lying above $SAVC_2$ is the firm's short-run supply curve. In the short run the number of firms in the industry remains fixed.

Horizontally adding these short-run supply curves for the given number of firms, we obtain the new industry short-run supply curve $SRSS_2$. The new short-run equilibrium occurs at $P_2$, where $SRSS_2$ crosses the demand curve. Each firm sets $P_2$ equal to $SMC_2$ and produces an output $q_2$. Together the $N_1$ firms produce $Q_2$. In Figure 8-9 firms are now covering their variable costs but not their fixed costs at the price $P_2$. They are losing money.

As time elapses two things happen: fixed factors can be varied, and firms can leave the industry. Long-run equilibrium occurs at the price $P_2^*$ since the new long-run industry supply curve $LRSS_2$ is horizontal at $P_2^*$, which just covers minimum long-run average costs. Each firm produces $q_2^*$ and the

**FIGURE 8-9 THE EFFECT OF A COST INCREASE ON A COMPETITIVE INDUSTRY.** The industry begins in long-run equilibrium producing $Q_1^*$ at a price $P_1^*$. Each identical firm produces $q_1^*$ at the lowest point on $LAC_1$. The long-run supply curve $LRSS_1$ is horizontal at $P_1^*$. When costs increase, firms have fixed factors and the number of firms is given in the short run. Each firm produces $q_2$ where the short-run equilibrium price $P_2$ equals $SMC_2$. Together these firms produce $Q_2$. Since firms are losing money, in the long run some firms leave the industry. The new long-run supply curve $LRSS_2$ for the industry is horizontal at $P_2^*$, the minimum point on each firm's new long-run average cost curve $LAC_2$. In the long run each firm produces $q_2^*$. Industry output is $Q_2^*$.



number of firms contracts to $N_2$ such that $Q_2^*$ equals $q_2^*$ times $N_2$.

Figure 8-9 makes two points about the change in the long-run equilibrium position. First, the rise in average costs is eventually passed on to the consumer in higher prices. In long-run equilibrium the marginal firm (here, all firms, since they are identical) must make only normal profits to prevent an incentive for further entry or exit. To allow normal profits, prices must rise to cover the increase in minimum average costs.

Second, since higher prices reduce the total quantity demanded, industry output must fall. Unless the rise in costs takes a strange form which greatly reduces the minimum average cost output for each individual firm (so that $q_2^*$ lies well to the left of $q_1^*$), the reduction in total industry output will be achieved in part by a reduction in the number of firms that can survive in the industry in the long run.

### A Shift in the Market Demand Curve: An Example from the Coal Industry

Figure 8-10 illustrates the effect of an outward shift in the market demand curve from *DD* to *D'D'*. We show only the effects at the industry level. You should try to draw your own diagram showing what is happening for the individual firm, as we did in Figure 8-9.

The industry begins in long-run equilibrium at the point *A*. Overnight, each firm has fixed factors and the number of firms is fixed. Horizontally adding their short-run supply curves, we obtain the industry supply curve *SRSS*. The new short-run equilibrium occurs at the point *A'*. When demand first increases it requires a large price rise to persuade individual firms to move up their steep short-run supply curves with given fixed factors.

In the long run, firms can adjust all factors and move on to their flatter long-run supply curves. In addition, supernormal profits attract extra firms into the industry. Figure 8-10 assumes that the long-run industry supply curve is rising. Either it takes higher prices to attract higher-cost firms into the industry, or the collective expansion bids up some input prices, or both. The new long-run equilibrium occurs at *A"*. Relative to short-run equilibrium at *A'* there is a further expansion of total output but a more appropriate choice of factors of production and the entry of new firms combine to increase supply and reduce the market-clearing price.

We know demand curves depend on tastes, incomes, and prices of related goods. Since tastes and incomes change only slowly, the most spectacular example of a demand curve shift is probably provided by the oil price shock in 1973–74 when oil prices tripled. Since oil and coal are substitutes as energy sources, we should expect a large outward shift in the demand for coal.

In many European countries the coal industry is at least partially regulated by the government. In the UK the coal industry is actually nationalized. The best illustration of the working of a freely competitive coal industry is the case of the United States.

How did higher oil prices affect the freely competitive coal industry in the United States? Table 8-2 presents some statistics for the 1970s which confirm the prediction of Figure 8-10. In 1974-77, immediately following the oil price shock, there was a 52 per cent rise in the real price of coal but only a modest 12 per cent rise in production of coal. This matches the move from *A* to *A'* in Figure 8-10.

For the period 1978–80 output rises a lot but the real price falls back, as the move from *A'* to *A"* predicts. Table 8-2 confirms that firms were being attracted into the industry as the theory implies. These additional firms provide a substantial amount of the increase in total output. Many of these new coal mines were quite small relative to the large mines previously in operation. We see from Table 8-2 that the addition of these new smaller mines actually reduced average output per mine. Only at the higher prices could these small and almost certainly higher-cost mines survive in the coal industry.

Thus the messages of Figure 8-10 are confirmed. When demand increases there must eventually be a rise in the price. This price rise has three effects



**TABLE 8-2**
**THE COAL INDUSTRY IN THE UNITED STATES**

| YEARS | REAL PRICE OF COAL* | COAL OUTPUT* |
|---|---|---|
| 1970–73 | 100 | 100 |
| 1974–77 | 152 | 112 |
| 1978–80 | 145 | 129 |

| YEAR | REAL PRICE | NO. OF ESTABLISHMENTS | EMPLOYEES |
|---|---|---|---|
| 1972 | 102 | 3365 | 155 000 |
| 1977 | 147 | 5275 | 242 000 |

* Index for 1970–73 = 100
*Sources: Survey of Current Business* (various issues), and *Statistical Abstract of the United States*, 1981.

which act to restore long-run equilibrium. First, by moving consumers up the demand curve, the price rise partly mitigates the increase in quantity demanded. Second, the price rise induces existing firms to expand along their long-run supply curves and produce more output. Finally, the price rise entices new firms into the industry.

In the short run the price *overshoots* its long-run position. In Figure 8-10 the point *A'* lies above the point *A"*. Consumers may well complain about the large price increase in the short run, especially since firms in the industry are temporarily making large profits. But these profits fulfil an important role in the adjustment process, for they act as the signal to potential entrants that this is an industry that can profitably be entered. Entry helps increase long-run supply and mitigate the initial price increase. As entry takes place and existing firms also manage to adjust their previously fixed factors, the industry gradually moves from *A'* to *A"* in Figure 8-10. Eventually, the extra output competes away the supernormal profits by bidding the price down, and the industry comes to rest at *A"*, its new long-run equilibrium position.

## 8-5 COMPETITION IN WORLD MARKETS

Changes in conditions in domestic markets are often the result of events in other countries. The

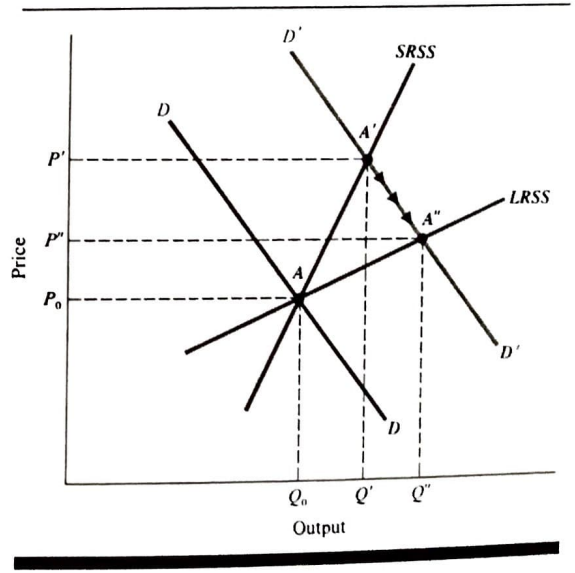



**FIGURE 8-10 A SHIFT IN DEMAND IN A COMPETITIVE INDUSTRY.** The industry begins in long-run equilibrium at *A*. When the demand curve shifts from *DD* to *D'D'* the new short-run equilibrium occurs at *A'*. As fixed factors are gradually adjusted and new firms enter the industry, equilibrium gradually moves from *A'* towards *A"*, the new long-run equilibrium.

BOX 8-1

## THE ELASTICITY OF SUPPLY

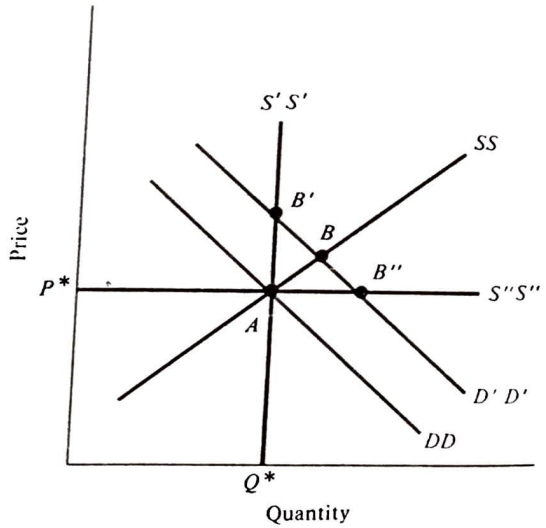
The elasticity of supply measures the responsiveness of the quantity supplied to a change in the price of that commodity. As with elasticity of demand, we use percentage changes to adjust for the absolute level of price and the absolute scale of output.

$$\text{Elasticity of supply} = \frac{\text{percentage change in quantity supplied}}{\text{percentage change in price}}$$

Because supply curves slope upward, the elasticity of supply is *positive*. As we move along a supply curve, positive price changes are associated with positive output changes. The more elastic is supply the larger the percentage increase in quantity supplied in response to a given percentage change in price. Thus, elastic supply curves are relatively flat and inelastic supply curves relatively steep.

Unlike the special case of unit-elastic demand, the case in which a price change leaves revenue unchanged, a supply elasticity of unity has no special significance. Since price and quantity move in the same direction as we move along a supply

SUPPLY ELASTICITIES Along the supply curve $SS$ the supply elasticity is positive. Higher price is associated with higher output. The vertical supply curve $S'S'$ has a zero supply elasticity. Higher prices lead to a zero change in output. The supply curve $S''S''$ has an infinite supply elasticity. Any price increase above $P^*$ leads to an infinite increase in quantity supplied. Beginning from equilibrium at $A$, a demand shift from $DD$ to $D'D'$ leads to a new equilibrium at $B'$, $B$, or $B''$ depending on the elasticity of supply. The more inelastic is supply the more the demand increase leads to higher prices rather than higher quantities. In the extreme cases, the move from $A$ to $B'$ reflects only a price increase and the move from $A$ to $B''$ reflects only a quantity increase.






curve, higher prices are always associated with higher revenue whatever the supply elasticity.

The diagram shows a typical supply curve $SS$ with a positive supply elasticity and also shows the two limiting cases. The vertical supply curve $S'S'$ has a zero supply elasticity. A given percentage change in price is associated with a zero percentage change in quantity supplied. The horizontal supply curve $S''S''$ has an infinite supply elasticity. Any price increase above the price $P^*$ would lead to an infinite increase in quantity supplied.

The elasticity of supply is crucial in telling us how much the equilibrium price and quantity will change when there is a shift in demand. The diagram shows that a demand shift from $DD$ to $D'D'$ leads to higher price rises and lower quantity rises the more inelastic is supply.

decision by OPEC to raise oil prices in 1979–80 quickly led British producers of North Sea oil to follow suit. Wool prices in Britain and the rest of the European Community change when there is a drought in Australia, one of the world's largest wool suppliers. We now discuss how competitive markets in different countries are linked together and show why shifts in foreign supply or demand curves affect domestic markets.

When a commodity is internationally traded, its price in one country cannot be independent of its price in another country. In the extreme case, the 'Law of One Price' will hold.

> If there were no obstacles to trade and no transport costs, the *Law of One Price* implies that the price of a given commodity will be the same all over the world.

Without trade barriers and transport costs, suppliers would always wish to sell in the market with the highest price but consumers would always wish to purchase in the market with the lowest price. The commodity could be simultaneously traded in two different countries only if its price were the same in both markets.[3]

In practice, transport costs and trade restrictions such as tariffs (taxes levied only on imports) allow international differences in the price of a commodity. Nevertheless, unless these costs and restrictions are prohibitive, international competition will ensure that prices of the same good in different countries generally move together.

We now show how international trade affects competitive markets. To highlight this issue, we assume transport costs and trade restrictions are negligible. Producers and consumers throughout the world are essentially part of a unified world market for the commodity.

### Equilibrium in the Domestic Market

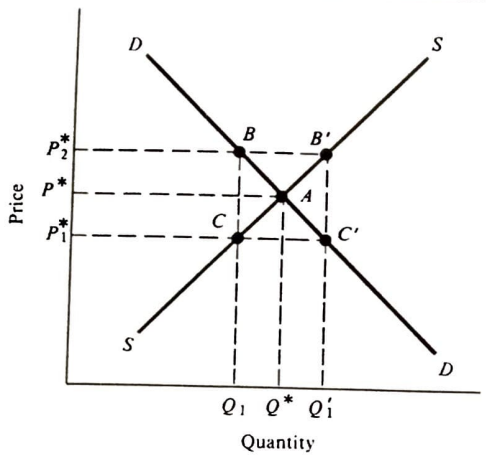
Figure 8-11 shows the domestic supply curve $SS$ and the domestic demand curve $DD$ for such a commodity. Suppose first that there is no international trade, perhaps because the domestic country has enormous tariffs in imports. The domestic market will be in equilibrium at the point $A$, at which price is $P^*$ and quantity is $Q^*$.

Now suppose tariffs are abolished and transport costs are negligible. There is a world supply curve, which horizontally aggregates the supply curve of each country, and a world demand curve, which horizontally aggregates the demand curve of each country. Together these determine a world equilibrium price for the commodity. Suppose the domestic country is small relative to the world and must take the world price as given.

One of three things can happen, and Figure 8-11 illustrates each of these cases. Suppose, first, that the world equilibrium price is $P^*$, exactly the price that would have cleared the domestic market

[3] In the early 1980s the UK price and German price of BMW cars was so different that UK consumers found it profitable to fly to Germany, buy a BMW for Deutschmarks, and drive it home. Such examples are the exception, not the rule.

**FIGURE 8-11 DOMESTIC EQUILIBRIUM AND WORLD PRICES.** *DD* and *SS* show the domestic supply and demand curves for a commodity competitively traded in world markets. In the absence of trade, domestic equilibrium occurs at *A*. When trade is possible at the world price $P^*$, equilibrium occurs at *A*. When trade is possible at the given world price $P_1^*$, domestic producers supply $Q_1$ and domestic consumers demand $Q_1'$. The excess demand (the horizontal distance between *C* and *C'*) is met from imports. Conversely, when world prices are $P_2^*$, domestic producers supply $Q_1'$, domestic consumers demand $Q_1$, and the excess supply (the horizontal distance between *B* and *B'*) is exported.




in isolation. Point *A* continues to describe equilibrium in the domestic market. The Law of One Price is satisfied. Consumers cannot buy the good more cheaply abroad and producers cannot sell the good at a higher foreign price. Domestic supply exactly caters for domestic demand and the domestic country neither imports nor exports the good.

Now suppose the given world price is $P_1^*$. If domestic suppliers attempt to charge a higher price domestic consumers will simply import the good and pay $P_1^*$. But domestic suppliers will produce at least $Q_1$ since they can always export the good if domestic consumers will not buy it. Hence the domestic market is in equilibrium at the price $P_1^*$, at which producers supply $Q_1$, consumers demand $Q_1'$, and the quantity $(Q_1' - Q_1)$, corresponding to the horizontal distance between *C* and *C'*, is imported from abroad. Conversely, if the given world price is $P_2^*$, domestic consumers will demand the quantity $Q_1$ but domestic producers will supply $Q_1'$. The quantity $(Q_1' - Q_1)$ corresponding to the horizontal distance between *B* and *B'* will now be exported to consumers abroad.

## The Effect of Changes in World Conditions on the Domestic Market

When our industries compete in world markets, a change in the world price, reflecting a shift in the world supply curve or the world demand curve, will affect the domestic market. Figure 8-11 may be used to show why.

Suppose a drought in Australia reduces the world supply of wool. The world price of wool rises. Suppose originally the world price was $P_1^*$ in Figure 8-11. British farmers were producing $Q_1$ but clothing manufacturers in Britain were additionally importing $(Q_1' - Q_1)$ since their total quantity demanded was $Q_1'$. The Australian drought raises the world price above $P_1^*$ and this has two effects. First, it reduces the quantity of wool demanded by British clothing manufacturers. Second, it allows British farmers to charge higher prices and move up their supply curve, expanding output and attracting new resources into the farming industry.

Since the domestic quantity demanded has declined but the domestic quantity supplied has increased, the higher world price of wool has led to a fall in imports. Indeed, if the world price rises sufficiently, Figure 8-11 implies that the UK would become a net exporter of wool.

## Foreign Trade as a Shock Absorber

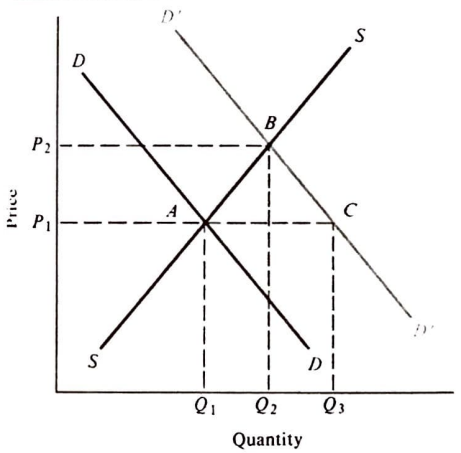
In the previous example, changes in world prices of commodities induced changes in the quantities domestically produced and consumed. The rest of the world was a source of shocks to the domestic economy. But trade is a two-way street. When the source of the shock is domestic, foreign trade may act as a shock absorber which cushions the domestic economy from the effect of changes in domestic conditions.



Figure 8-12 shows what happens if the domestic demand curve shifts to the right. In the absence of international trade, domestic equilibrium moves from *A* to *B*. To attract additional supply the price must rise. Higher prices move consumers up their new demand curve *D'D'*. Output increases from $Q_1$ to $Q_2$ but not to $Q_3$, the output that consumers would now be prepared to buy had the price not risen.

Suppose instead that the economy competes freely in world markets. If the world price is $P_1$ the original position of the economy will be at the point *A*. When the domestic demand curve shifts from *DD* to *D'D'* the new equilibrium position is now *C* not *B*. Consumers are no longer frustrated by higher prices and domestic producers produce exactly as before. Consumers obtain the extra quantity *AC* entirely from imports.

**FIGURE 8-12 FOREIGN TRADE AS A SHOCK ABSORBER.** In the absence of trade, a demand shock that shifts the domestic demand curve from *DD* to *D'D'* moves the domestic equilibrium from *A* to *B*. Producers must adjust quantity and consumers face higher prices. If the country trades at given world prices $P_1$, the demand shock has no effect on domestic producers who still produce at *A*. Prices remain fixed at $P_1$ and consumers import the amount *AC*.




Thus, if domestic demand curves shift around a lot, foreign trade acts to cushion domestic producers from wild swings in the demand for their output. Similarly, competitive foreign trade prevents domestic shocks spilling over into price changes. The domestic price is anchored by the Law of One Price.

To sum up, international trade tends to transmit shocks from the world economy to the domestic economy but tends to send domestic shocks abroad through the shock absorber of imports and exports, thereby cushioning the domestic economy.

This brief look at international trade, which we examine in greater detail in Chapter 31, also reminds us that the relevant definition of the market or the industry may be a good deal wider than that of the domestic economy. When transport costs are low and trade restrictions unimportant, it is in the *world* market that we must seek the forces that determine the equilibrium price of a good.

## 8-6 PURE MONOPOLY: THE OPPOSITE LIMITING CASE

The perfectly competitive firm is too small to worry about the effect of its own output decision on industry supply. It can sell as much as it wants at the market price. Before setting out a general theory of market structure, we discuss the opposite limiting case on the supply side, the case of pure monopoly.

A monopolist is the sole supplier and potential supplier of the industry's product. The firm and the industry coincide. The sole national supplier need not be a monopolist if the good or service is internationally traded. The Royal Mint is the sole supplier of UK coins and is a monopolist. British Steel, although effectively the sole UK steel supplier, is not a monopolist since it must compete with imports. Like many monopolists in the UK, British Rail is a nationalized industry. The state

makes price and output decisions and may not aim primarily to maximize profits. For example, it may decide to subsidize rural users to preserve a sense of national unity. The behaviour of nationalized industries will be discussed in Chapter 17.

Here we are concerned with the decisions of a private profit-maximizing monopolist. Bricks are so heavy that huge transport costs effectively insulate national markets from one another. Since 1968 the London Brick Company has been the sole UK supplier of fletton bricks. Based on a particular clay known as Oxford clay, these bricks enjoy a substantial cost advantage over all other bricks because their higher carbon content greatly reduces the cost of firing the clay. Other examples of *private* monopolists can be given (e.g., Rank Xerox copiers until the 1970s), but the analysis of this section has a wider significance. In many countries there is currently discussion of whether to 'privatize' state-run monopolies. The analysis in the remainder of this chapter illustrates how we might expect such industries to behave if they were restored to private ownership.

## 8-7 PROFIT-MAXIMIZING OUTPUT FOR A MONOPOLIST

In Chapter 6 we developed the general theory of supply for an individual firm. To maximize profits it chooses the output at which marginal revenue *MR* equals marginal cost *MC* (*SMC* in the short run and *LMC* in the long run). The firm then checks that it is covering average costs (*SAVC* in the short run and *LAC* in the long run).

The special feature of a competitive firm is that *MR* equals price. Selling an extra unit of output does not bid down the price and reduce the revenue earned on previous units. The price at which the extra unit is sold *is* the change in total revenue.

In contrast, the monopolist's demand curve *is* the industry demand curve, which slopes down. As we explained in Chapter 6, this implies *MR* is less than the price at which the extra unit of output is sold. The monopolist recognizes that extra output reduces revenue from previous units because price falls as we move down the demand curve.

Figure 8-13 reminds you of our previous discussion of the relationship between price, marginal revenue, and total revenue when the demand curve slopes down. The more inelastic the demand curve, the more an extra unit of output will bid down the price and reduce revenue from existing units. At any output, *MR* lies further below the demand curve the more inelastic is demand. Also, the larger the existing output, the larger the revenue loss from existing units when the price is reduced to sell another unit. For a given demand curve, *MR* falls increasingly below price the higher the output level from which we begin.

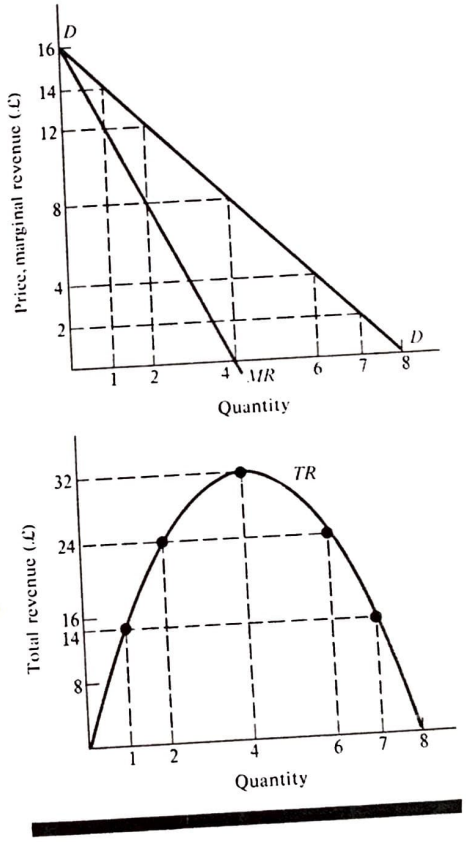
Beyond a certain output (4 units in Figure 8-13), the revenue loss on existing output exceeds the revenue gain from the extra unit itself. Marginal revenue becomes negative. Hence Figure 8-13 shows total revenue starting to fall at this output. Further expansion reduces total revenue.

On the cost side, there is only one producer, and the discussion of the cost curves for a single firm in Chapter 7 carries over directly. The monopolist has the usual cost curves, average and marginal, short-run and long-run. For simplicity we discuss only the long-run curves. The following analysis is easily supplemented by inclusion of short-run cost curves to explain how a monopolist makes the transition from one long-run equilibrium to another when demand or cost conditions alter.

There is one other crucial aspect of our definition of monopoly. Not only is a monopoly the sole existing supplier, it need take no account of new entrants to the industry. When existing suppliers take account of the threat of new firms entering the industry they are not monopolists. The behaviour of such firms forms the basis of the next chapter. Without anticipating that discussion of when entry will or will not be possible, we simply assume for the moment that the monopolist need not take any account of potential entry. Imagine that the firm is the sole legal licensee (the Royal Mint), the sole patent holder (Rank Xerox, Polaroid, IBM golfball typewriters), or simply has an enormous cost advantage over its nearest rival (the London Brick Company).



**FIGURE 8-13 DEMAND, TOTAL REVENUE AND MARGINAL REVENUE.** Total revenue (*TR*) equals price times quantity. From the demand curve *DD* we can plot the *TR* curve at each quantity. Maximum *TR* occurs at £32, when 4 units are sold for £8 each. Marginal revenue (*MR*) shows how *TR* changes when quantity is increased a small amount. *MR* lies below the demand curve *DD*. From the price received for the extra unit we must subtract the loss in revenue from existing units as the price is bid down. This effect is larger the higher is existing output and the more inelastic is the demand curve. At a particular output, the *MR* curve lies further below *DD* the larger is output and the more inelastic the demand curve. Beyond an output of 4 units, *MR* is negative and further expansion reduces total revenue.




### Profit-maximizing Output

In Chapter 6 we showed why setting *MR* equal to *MC* would lead to the profit-maximizing level of positive output. When *MR* exceeds *MC*, an additional unit of output will add more to revenue than to costs and will increase profits. When *MC* exceeds *MR*, the last unit has added more to costs than to revenue. Profits would be increased by cutting back output. When *MR* equals *MC* output is at the profit-maximizing or loss-minimizing level, given that the firm produces anything at all.

Then the monopolist must check whether at this output the price (or average revenue) covers average variable costs in the short run and average total costs in the long run. If not, the monopolist should shut down in the short run and leave the industry in the long run. In the latter case, the industry will probably cease to exist. Table 8-3 summarizes the criteria by which a profit-maximizing monopolist decides how much to produce.

Figure 8-14 shows the average cost curve *AC* with its usual U-shape. The marginal cost curve *MC* passes through the lowest point on the *AC* curve. The marginal revenue curve *MR* lies below the down-sloping demand curve *DD*. Setting $MR = MC$, the monopolist chooses the output level $Q_1$. However, to find the price for which $Q_1$ units can be sold we must look at the demand curve *DD*. The monopolist sells $Q_1$ units of output at a price $P_1$ per unit. Profit per unit is given by $P_1 - AC_1$, price minus average cost when $Q_1$ is produced. Total profits are given by the shaded area $(P_1 - AC_1) \times Q_1$.

Even though we are studying the long run, the monopolist continues to make these *supernormal* profits. They are sometimes called *monopoly* profits. They are pure profit after making all cost deductions for the opportunity cost of the owners' time and money. Unlike the competitive industry, supernormal profits of a monopolist are not eliminated in the long run by the entry of new firms. We have already established that an industry

**TABLE 8-3**
**MONOPOLIST'S CRITERIA FOR MAXIMIZING PROFITS**

| | MARGINAL CONDITION | | | AVERAGE CONDITION | | | |
|---|---|---|---|---|---|---|---|
| | | | | SHORT-RUN | | LONG-RUN | |
| DECISION | *MR > MC* Raise output | *MR = MC* Optimal output | *MR < MC* Lower output | *P ⩾ SAVC* Produce | *P < SAVC* Shut down | *P ⩾ LAC* Stay in business | *P < LAC* Exit from industry |

is a monopoly only if the sole existing supplier need take no account of the possibility of entry. By ruling out the possibility of entry, we remove the mechanism by which supernormal profits tend to disappear in the long run. In Figure 8-14 the monopolist is on to a good thing for ever.

Price-setting Whereas the competitive firm is a *price-taker*, taking as given the equilibrium price determined by the interaction of market supply and market demand, the monopolist actually sets prices and is a *price-setter*. Having decided to produce $Q_1$ in Figure 8-14, what the monopolist actually does is to quote a price $P_1$ knowing that customers will then demand exactly $Q_1$ units of output.

Elasticity and Marginal Revenue In Chapter 4 we saw that when the (own-price) elasticity of demand lies between 0 and $-1$ demand is inelastic and an increase in output will reduce total revenue. Marginal revenue is negative. In percentage terms, the fall in price exceeds the rise in quantity. All outputs to the right of $Q_2$ in Figure 8-14 have negative *MR*. The demand curve is inelastic at quantities above $Q_2$. At quantities below $Q_2$ the demand curve is elastic. Higher output leads to higher revenue.

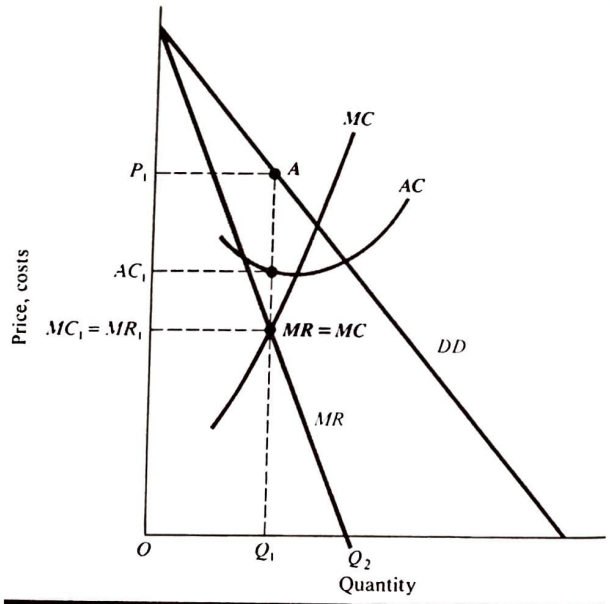



**FIGURE 8-14 THE MONOPOLY EQUILIBRIUM**: *MC = MR*. Applying the usual marginal condition, a profit-maximizing monopolist produces the output level $Q_1$ at which marginal cost *MC* equals marginal revenue *MR*. Then it must check that price covers average cost. In this figure, $Q_1$ can be sold at a price $P_1$ in excess of average costs $AC_1$. Monopoly profits are the shaded area $(P_1 - AC_1) \times Q_1$.



The monopolist sets *MC* equal to *MR*. Since *MC* must be positive, so must *MR*. The chosen output must lie to the left of $Q_2$. Hence, we say that *a monopolist will never produce on the inelastic part of the demand curve.*

Price, Marginal Cost, and Monopoly Power At any output, price exceeds the monopolist's marginal revenue since the demand curve slopes down. Hence, in setting *MR* equal to *MC* the monopolist sets a price that *exceeds* marginal cost. In contrast, a competitive firm always *equates* price and marginal cost, since its price is also its marginal revenue. This suggests that we might view the excess of price over marginal cost as a measure of *monopoly power*. The competitive firm cannot raise price above marginal cost and has no monopoly power.

### Comparative Statics for a Monopolist

Figure 8-14 may also be used to analyse the effect of changes in costs or demand. Suppose there is a change in costs, for example an increase in input prices, which shifts the *MC* and *AC* curves upwards. The higher *MC* curve must cross the *MR* curve at a lower level of output. Provided the monopolist can sell this output at a price that covers average costs, the effect of the cost increase must be to reduce output. Since the demand curve slopes down, this reduction in output will be accompanied by an increase in the equilibrium price.

Now suppose for the original cost curves shown in Figure 8-14 that there is an outward shift in demand and marginal revenue curves. *MR* must now cross *MC* at a higher level of output. Thus an increase in demand leads the monopolist to increase output as we should expect.

## 8-8 OUTPUT AND PRICE UNDER MONOPOLY AND COMPETITION

We now compare a perfectly competitive industry with a monopoly. For this comparison to be of interest the two industries must face the same demand and cost conditions. We are interested in how the *same* industry would behave if it were organized first as a competitive industry then as a monopoly.

Clearly this is a tricky comparison. In the next chapter we develop a theory of market structure that aims to explain why some industries are competitive but others are monopolies. If this theory has any content, can it be legitimate to assume that the same industry could be competitive or monopolized? The answer turns out to be yes in some circumstances but no in other circumstances. We now distinguish these two cases.

### Comparing a Competitive Industry and a Multi-plant Monopolist

Consider a competitive industry in which all firms and potential entrants have the same cost curves. From our earlier discussion of the horizontal *LRSS* curve for a competitive industry we know this case can be analysed using Figure 8-15.

Facing the demand curve *DD*, the industry is in long-run equilibrium at *A* where the price is $P_1$ and total output is $Q_1$. The industry *LRSS* curve is horizontal at $P_1$, the lowest point on the *LAC* curve of each firm. Any other price would eventually lead to infinite entry or exit from the industry. We can regard *LRSS* as the industry's long-run marginal cost curve $LMC_1$ of expanding output by enticing new firms into the industry.

Each firm is producing at the lowest point on its *LAC* curve and breaking even. This point is also the lowest point on the firm's *SATC* curve for the level of factors it now has. Since marginal cost curves pass through the point of minimum average costs, each firm is also on its *SMC* and *LMC* curves. Horizontally adding these *SMC* curves, the supply curves of each firm in the short run, we obtain *SRSS*, the short-run industry supply curve. We can regard this as the industry's short-run marginal cost curve $SMC_1$ of expanding output from existing firms with temporarily fixed factors. Since *SRSS* crosses the demand curve at $P_1$, the

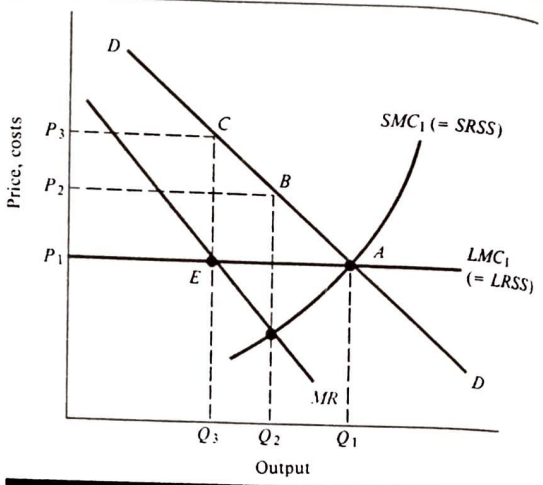



**FIGURE 8-15 RELATIVE TO A COMPETITIVE INDUSTRY A MONOPOLIST PRODUCES A LOWER OUTPUT AT A HIGHER PRICE.** Long-run equilibrium in a competitive industry occurs at *A*. Total output is $Q_1$ and the price $P_1$. On taking over the industry, a monopolist sets *MR* equal to $SMC_1$, restricting output to $Q_2$ and increasing price to $P_2$. In the long run the monopolist sets *MR equal to* $LMC_1$, reducing output to $Q_3$ and increasing the price again to $P_3$. There are no entrants to compete away supernormal profits $P_3CEP_1$ by increasing the industry output.

industry is both in short-run and long-run equilibrium.

Beginning from this position, suppose the competitive industry became a monopoly. The monopolist takes over each plant (firm) but makes centralized pricing and output decisions. For example, in 1967 the UK steel industry was nationalized. British Steel bought out all the private steel producers and had a legal monopoly on UK steel production. If we could ignore international trade in steel and assume that British Steel were instructed to maximize profits, how would the monopolization of the steel industry affect pricing and output decisions?

Overnight the monopolist still has the same number of factories (ex-firms) as in the competitive industry. Since the firm and the industry now coincide, $SMC_1$ remains the short-run marginal cost curve for the monopolist taking all plants together.[4] However, the monopolist makes centralized decisions which recognize that higher output reduces the revenue earned from previous units.

In the short run the monopolist equates $SMC_1$ and *MR*, reaching equilibrium at *B*. $Q_2$ units are produced at a price $P_2$. Relative to short-run competitive equilibrium at *A*, *the monopolist raises price and reduces quantity.*

In the long run the monopolist can enter or set up new factories and can exit or close down existing factories. Even though the monopolist may be making short-run profits at *B* (we need to draw in the *SATC* curve to confirm this) nevertheless, in complete contrast to a competitive industry, the monopolist will decide to exit or retire some factories from the industry in the long run.

The monopolist wants to cut back output to force up the price. Yet in the long run it makes sense to operate each factory at the lowest point on its *LAC* curve. To reduce total output some

[4] In a competitive industry each firm equates the given price to its own marginal cost. Hence firms produce at the same marginal cost. Thus we horizontally add individual *SMC* curves (i.e. at the same price) to get the industry *SMC* curve. A multi-plant monopolist need not equate *MC* across all plants but will always find it profitable to do so. Why? If marginal costs in two plants differed, the monopolist could always produce the same total output more cheaply by producing an extra unit in the low *MC* plant and one less unit in the high *MC* plant. Thus the $SMC_1$ curve for the monopolist across all plants remains the horizontal sum of the *SMC* curves for individual plants, as in a competitive industry.



factories must be retired. In the long run the monopolist sets $LMC_1$ equal to *MR* and reaches the equilibrium position *C*. *Price has risen yet further to* $P_3$ *and output has fallen to* $Q_3$. Long-run supernormal profits are given by the area $P_3CEP_1$ since $P_1$ remains the long-run average cost when all plants are producing at the lowest point on their *LAC* curve.

Although it is the recognition that *MR* is less than price that provides the incentive for a monopolist to produce less than a competitive industry and charge a higher price, in this example it is the legal prohibition on entry by competitors that allows the monopolist to succeed in the long run. In a competitive industry supernormal profits are competed away by the new entrants that they attract to the industry. That is why we have insisted that the absence of entry is intrinsic to the model of monopoly we have developed.

**The Social Cost of Monopoly** Should society mind that a monopolist restricts output and drives price above marginal cost? This is not an issue in positive economics, the description of actual behaviour, but rather an issue in normative economics, which deals in recommendations and policy prescriptions. We deal with such questions at length in Part 3 of this book.

At this point we merely sketch how that argument might go. The marginal cost measures the resources used up to make the last unit of the good. Since consumers voluntarily buy the good, the price of the good must measure the marginal benefit to consumers of buying the last unit of the good. If the marginal benefit were higher than the price, consumers would buy even more at that price. If the marginal benefit were less than the price, consumers would not demand that last unit at that price.

Society should want to equate the marginal cost of the good and its marginal benefit. If marginal cost is less than marginal benefit, society will be better off with more of the good. Whereas a competitive industry automatically sets marginal cost equal to price (equal to presumed marginal consumer benefit) monopoly does not. It sets marginal cost less than price and, by implication, produces less of the good than society might wish. Whether this simple argument is generally correct is one of the issues we explore in Part 3.

## Comparing a Single-plant Monopolist with a Competitive Industry

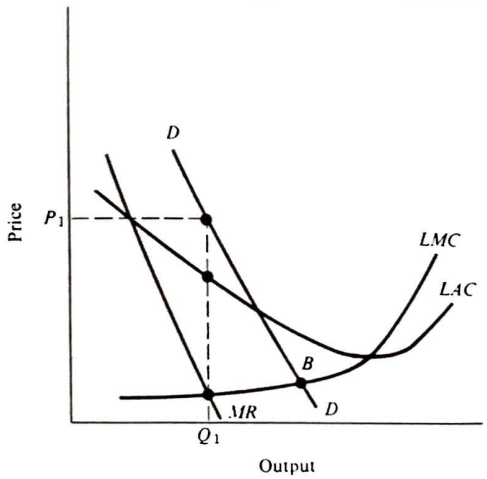
In the previous example we examined a multi-plant monopolist who took over a large number of previously competitive firms. Now we examine a monopolist meeting the entire industry demand from a single plant. This is most plausible when there are large economies of scale. There are huge costs in setting up a national telephone network and nobody would be interested in a partial network. Yet the cost of connecting a marginal subscriber is low once the network has been set up.

Monopolies enjoying huge economies of scale – falling *LAC* curves over the entire range of output – are called *natural monopolies.* As we shall see in the next chapter, large scale-economies may explain why there is a sole supplier who need not worry about entry. Smaller new entrants would be at a prohibitive cost disadvantage.

Figure 8-16 illustrates the long-run equilibrium for a natural monopoly. In the long run the natural monopoly faces average and marginal cost curves *LAC* and *LMC*. Given the position of the demand curve, long-run average cost is declining over the entire range of outputs that might be demanded. The monopoly produces at *LMC* equal to *MR*, selling an output $Q_1$ for a price $P_1$. At this output, price exceeds *LAC*. The monopoly makes supernormal profits and is happy to remain in business.

It does not make sense to compare this equilibrium with how the industry would behave if it were competitive. With such economies of scale, there should be only one firm in the industry. *LAC* is the cost curve for each possible firm. If there were only one firm, it would be crazy not to recognize that its output decisions affected price. If it were stupid enough to try to set price equal to *LMC* it would reach the point *B*, conclude that it was not covering average costs, and leave the

**FIGURE 8-16 A NATURAL MONOPOLY WITH ECONOMIES OF SCALE.** The *LAC* curve is falling throughout the relevant range of output levels. Economies of scale are large relative to the market size. The monopoly produces $Q_1$ at a price $P_1$ and makes profits. If it tried to behave like a price-taking competitive firm it would produce at *B* where price equals *LMC* and make losses. By recognizing the effect of output on price the single firm monopoly can do much better. This industry cannot support a lot of small firms. Each would have very high average costs at low output. This cannot be a competitive industry.




industry.[5] If a lot of small firms produced a small fraction each of total demand, their average costs would be enormous. A single large firm could undercut them and wipe them out. This industry must have a sole supplier, and that natural monopoly will maximize profits only by recognizing that its marginal revenue is not its price.

It is this insight that we develop in the next chapter to provide a general theory of market structure. Turning to the normative issue, society may still be interested in forcing the monopolist to produce at a price closer to marginal cost. In the extreme case, society might order the monopoly to price at marginal cost, produce at the point *B*, make losses, and receive a government subsidy. We return to this issue in Part 3 when discussing nationalized industries and more general forms of government regulations of monopolies.

[5] From Chapter 7 we know that marginal cost lies below average cost at all points left of the point of minimum average cost. Since the *LAC* curve is still falling in Figure 8-16 it must lie above *LMC*. Pricing at marginal cost *must* yield losses at point *B*.

## 8-9 THE ABSENCE OF A SUPPLY CURVE UNDER MONOPOLY

A competitive firm sets price equal to marginal cost if it supplies at all. If we know its marginal cost curve we know how much it supplies at each price. Aggregating across firms, we also know how much the industry supplies at each price. We can draw the supply curve without knowing anything about the market demand curve. Confronting the supply curve with the market demand curve, we then analyse how supply and demand interact to determine equilibrium price and quantity.

The monopolist recognizes that output affects marginal cost and marginal revenue simultaneously. Figure 8-17 shows a given *LMC* curve. How much will the monopolist produce at the price $P_1$? It all depends on demand and marginal revenue. When demand is *DD* and the corresponding marginal revenue *MR*, the monopolist produces $Q_1$ and charges a price $P_1$. However, when demand is *D'D'* and marginal revenue *MR'*, the monopolist produces $Q_2$ but still charges $P_1$.

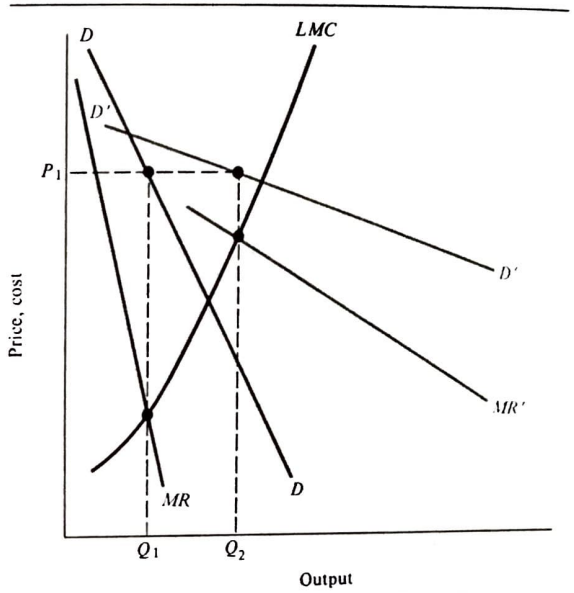
The monopolist does not have a supply curve independent of demand conditions. What we can say is that the monopolist simultaneously examines demand (hence marginal revenue) and cost (hence marginal cost) when deciding how much to produce and what to charge.

### Discriminating Monopoly

Thus far we have assumed that all consumers must be charged the same price, although this price will depend on the level of output and the position of the demand curve. Unlike a competitive industry, where competition between firms prevents any individual firm charging more than its competitors,



**FIGURE 8-17 ABSENCE OF A SUPPLY CURVE UNDER MONOPOLY.** Given the demand curve *DD* and the corresponding marginal revenue curve *MR*, the monopolist produces $Q_1$ at a price $P_1$. However, facing *D'D'* and the corresponding schedule *MR'*, the monopolist produces $Q_2$ at a price $P_1$. Knowing the price, we cannot uniquely infer the quantity supplied unless we also know demand and marginal revenue. Because the monopolist knows that output affects both marginal cost and marginal revenue, the two must be considered simultaneously.




a monopolist may be able to charge different prices to different customers. This will be especially attractive when it is possible to identify different types of customer whose demand curves are quite distinct.

Consider an airline monopolizing flights between London and Rome. It has business customers whose demand curve is very inelastic. They have to fly, and the plane fare is a trivial expense for their companies. For this group, demand and marginal revenue curves are very steep.

The airline also carries tourists whose demand curve is much more elastic. If flights to Rome get too expensive tourists can holiday in Athens instead. Tourists have much flatter demand and marginal revenue curves.

Recall why the marginal revenue curve lies below the demand curve. Adding an extra unit of output and sales bids down the price for which existing output can be sold and reduces revenue from existing units of output. The more inelastic is the demand curve the more the marginal revenue curve must lie below the demand curve because the higher will be the reduction in revenue from existing output units.

Suppose the airline charges tourists and business travellers the same price. From the separate demand curves we can read off at each price the number of each type of traveller and add these to obtain the total number of travellers at each price. However since the demand curve of business travellers is less elastic, the marginal revenue obtained from the last business traveller must be lower than the marginal revenue obtained from the last tourist.

Whatever the total number of passengers (and hence total cost of carrying them), the airline is carrying the wong *mix* between tourists and business travellers. Since the marginal revenue from the last tourist exceeds the marginal revenue from the last business traveller the airline would gain revenue without adding to cost by carrying the same number of passengers but carrying more of the group with the higher marginal revenue and

less of the group with the lower marginal revenue. And it will pay to keep changing the mix until the marginal revenue of the two groups is equated.

To do this, the airline must charge the two groups *different* prices. Since tourist demand is elastic the airline wants to charge tourists a low fare to increase tourist revenue. Since business demand is inelastic the airline wants to charge business travellers a high fare to increase business revenue.

Profit-maximizing output will satisfy two separate conditions. First, business travellers with inelastic demand will pay a fare sufficiently higher than tourists with elastic demand that the marginal revenue from the two separate groups is equated. Then there is no incentive to rearrange the mix by altering the price differential between the two groups. Second, the general level of prices and the total number of passengers will be determined to equate the marginal cost of carrying passengers to both these marginal revenues. This ensures that the airline operates on the most profitable scale as well as with the most profitable mix.

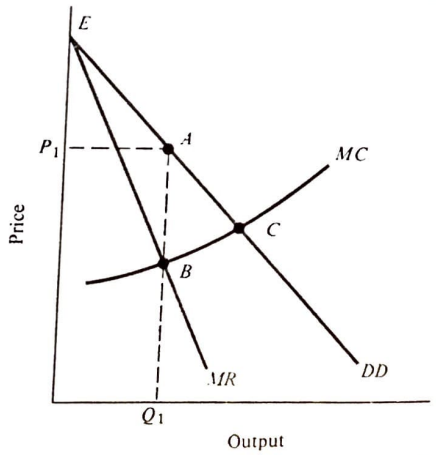
When a producer charges different customers different prices we say the producer *price discriminates*. There are many examples of this in the real world. Air fares per mile between London and Brussels, almost exclusively an expense account business trip, are among the highest in Europe, but package holidays are much cheaper. British Rail charges rush-hour commuters a higher fare than midday shoppers whose demand for trips to the city is much more elastic. Expensive doctors in private practice frequently charge lower prices to less well-off patients but charge very high prices to the very rich whose demand for the best medical care is very inelastic.

It is no accident that many of the best examples of price discrimination refer to services which must be consumed on the spot rather than to goods which can be resold. Price discrimination in a standardized commodity is unlikely to work. The group buying at the lower price have an incentive to resell to the group paying the higher price thus undercutting the monopolist's attempt to charge some customers a higher price. Effective price discrimination is feasible only when the submarkets can be isolated from one another to prevent resale of the product from one group to another.

What does price discrimination have to do with the absence of a supply curve under monopoly? Figure 8-18 illustrates this most clearly for the case of *perfect price discrimination* where we assume that it is actually possible to charge each and every customer a different price for the same good.

Suppose first that the monopolist charges every customer the same price. Given the demand curve $DD$ we obtain the usual marginal revenue curve $MR$ which lies below $DD$ precisely because to sell more output the monopolist must reduce the price

**FIGURE 8-18 PERFECT PRICE DISCRIMINATION.** Charging all customers the same price the monopolist will produce at $B$ where $MC = MR$. If each output unit can be sold for a different price, the revenue from existing units is not reduced by cutting the price to sell another unit. The demand curve $DD$ is the marginal revenue curve and the perfectly discriminating monopolist will produce at $C$. Output is higher and profits are higher. By price discrimination the monopolist gains an extra revenue $EP_1A$ from selling $Q_1$ but also increases output beyond this level making a marginal profit of $ABC$ in expanding from $A$ to $C$.






not only on the extra unit of output but also on all existing units. The profit-maximizing output is $Q_1$ where $MR$ equals $MC$ and the corresponding price is $P_1$.

Now suppose the monopolist can perfectly price discriminate, charging a different price for each unit of output sold. The very first can be sold for a price $E$. Having sold this output to the highest bidder, the customer most desperate for the good, the next unit can be sold to the next highest bidder and so on. As we move down the demand curve $DD$ we can read off the price for which each extra unit can be sold. However, in reducing the price to sell that extra unit, the monopolist no longer reduces revenue from previously sold units. *Hence the demand curve is the marginal revenue curve under perfect price discrimination.* The marginal revenue of the last unit is simply the price for which it can be sold.

Treating $DD$ as the marginal revenue curve we conclude that a perfectly price discriminating monopolist will produce at point $C$ where marginal revenue and marginal cost are equal. Two points follow immediately. First, if price discrimination is possible it is profitable to employ it. In moving from the uniform pricing point $A$ to the price discriminating point $C$ the monopolist adds the area $ABC$ to profits. This represents the excess of additional revenue over additional cost when output is increased. But the monopolist makes a second gain from price discrimination. Even the output $Q_1$ now brings in more revenue than under uniform pricing. The monopolist also gains the area $EP_1A$ by being able to charge different prices on the first $Q_1$ units of output rather than the single price $P_1$. In practice, when firms call in economic consultants one of the main ways these consultants manage to increase the profits of the firm is by devising new ways in which the firm can price discriminate.

Second, whether or not the firm is able to price discriminate affects the output it will choose to produce even if demand and cost conditions remain unaltered. Earlier in this section we said there was no unique supply curve relating output to price for a monopolist. We also had to know the elasticity of the demand curve and hence how far marginal revenue would lie below the price. Figure 8-18 shows it is not even sufficient to know the total demand curve facing a firm. In addition we need to know whether the market can be segmented enough to allow price discrimination. Uniform and discriminatory pricing will lead to very different outputs because they affect the marginal revenue obtained from any given total demand curve facing a monopolist.

## 8-10 MONOPOLY AND TECHNICAL CHANGE

In Section 8-8 we compared the behaviour of a monopoly and a perfectly competitive industry. When such a comparison was meaningful we discovered two things: (1) a monopoly will tend to restrict output and drive up prices; and (2) in consequence a monopoly will tend to make economic profits in the short run, and need not fear the erosion of these profits by entrants in the long run.

Joseph Schumpeter (1883–1950) argued that this comparison might be misleading because it ignores the possibility of technical advances, which reduce costs and may allow price reductions and output expansion. If banks are unwilling to lend money for risky research projects, a large monopolist with steady profits may find it much easier to fund internally the research and development (R & D) necessary to make cost-saving breakthroughs. Second, and completely distinct, a monopolist may have a greater *incentive* to undertake R & D.

In a competitive industry a firm with a technical advantage has only a temporary opportunity to earn high profits to recoup its research expenses. Imitation by existing firms and new entrants gradually compete away any super-normal profits. In contrast, by shifting down all its cost curves, a monopoly may be able to enjoy higher supernormal profits for ever. Schumpeter argued that these two forces – greater resources available for R & D and a higher potential return on any successful venture – tend to make monopolies more inno-