# Microeconomics with Ethics

by Steven Suranovic (2022) – George Washington University

# Chapter 13
# Perfectly Competitive Markets: Part 2

In this chapter, we will show how the market demand curve for a product is derived from the decisions of the individual representative firms. We will also explain the conditions necessary to cause the representative firms to shut down and stop producing and supplying a product to the market. Next, we will demonstrate the long-run adjustments likely to occur in a perfectly competitive market. Positive profits earned by the representative firms will signal other firms to enter the market and compete with incumbent firms, thereby expanding total supply of the product and ultimately reducing the market price and lowering profit of the firms. Similarly, losses incurred by representative firms will result in exit of firms and a reduction in market supply causing an increase in the price and an increase in profits for the remaining firms. The long-run equilibrium is characterized by a market price equal to the minimum average cost to produce the product and zero economic profit earned by the firms. Finally, we show how the market welfare attained in a long-run perfectly competitive market will exceed the total market welfare attained in a monopoly market. This is a key reason competition in markets is considered superior to the market concentration that occurs with monopoly or oligopoly.

## 13.1 Deriving the Market Supply Curve

### Learning Objectives

1. Learn how the market supply curve is derived from the marginal cost curves of individual firms
2. Learn the condition that determines when a firm should shut down and cease producing.

In order to determine how much to produce to maximize profit, the representative perfectly competitive (PC) firm must know the price of it product in the market. To proceed it is useful for the reader to accept on faith one important point; namely, that the equilibrium market price will be that price which equalizes market supply and market demand. In the next chapter, we will explain in more detail why the logic of the model requires that outcome.

In Chapter 7 we discussed individual demand determinants and showed how market demand for a product is derived from it. In this section, we will show how the market supply curve is derived from the individual supplies of small representative firms.

Consider the two graphs presented side-by-side in Figure 13.1. The left-side shows the cost curves for a representative firm in a PC market producing some product measured in hundreds of units. In Chapter 12, it was shown that the PC firm would maximize profit by choosing the output level such that the market price equals the marginal cost. Thus, if the market price were

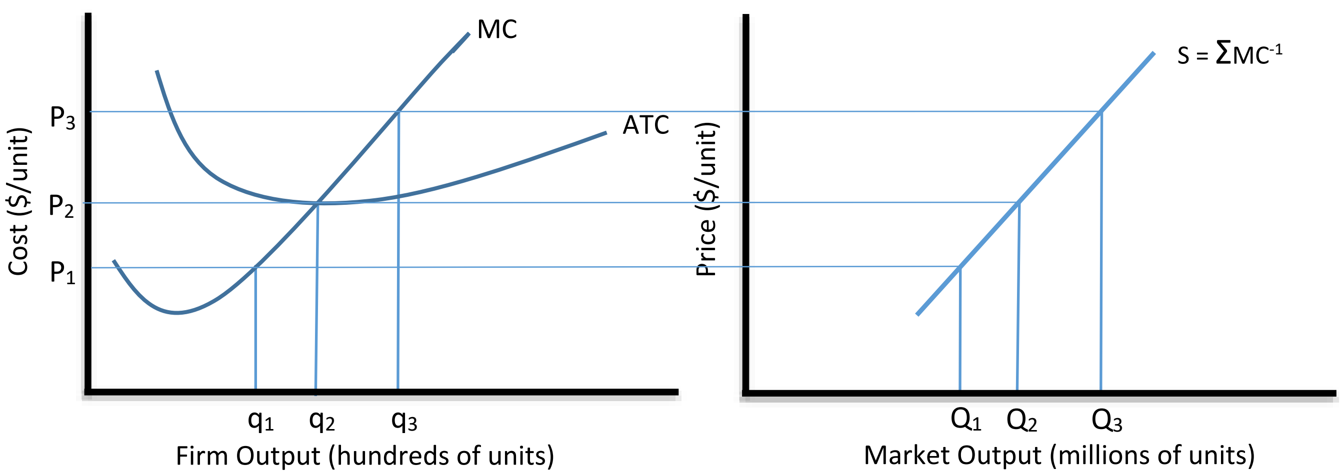
$P_1$, then the firm maximizes profit by producing output quantity $q_1$. Similarly, if the market price were $P_2$ (or $P_3$), then the firm maximizes profit by producing output quantity $q_2$ (or $q_3$).

The diagram on the right in Figure 13.1 shows the market supply function. It is derived simply by adding up, or aggregating, all of the quantities produced by the representative firms. Thus, if there were N firms producing this product in the industry, (where N is a large number measuring perhaps in the tens of thousands or more), then the market output at the price $P_1$ would be $Q_1 = N^*q_1$. Similarly, at the price P2 (or P3), market output would equal $Q_2 = N^*q_2$ (or $Q_3 = N^*q_3$).

Notice that the scale on the vertical axis is measured the same on both graphs since unit costs and prices share the same units. However, on the horizontal axis the scales are very different. This is reflected by measuring the representative firm outputs in hundreds of units and measuring market output in millions of units. We also maintain the convention that firm output is represented with a small q and market output with a capital Q.

Thus the *market supply function* shows the total quantity of the product that would be supplied to the market at all possible prices that might arise. It is derived by aggregating the supply choices of the numerous identical firms operating in the market. The quantities displayed in the market supply curve are said to be optimal because they are derived from the assumption that the individual firms choose output to maximize their individual profit.

Figure 13.1 Deriving the Market Supply Curve




**Side Note:** Sometimes the market supply curve is described as the horizontal summation of the representative firms' marginal cost curves. This is indicated by showing that the market supply curve S, equals the horizontal summation of the marginal cost functions ($S = \Sigma MC^{-1}$). In this equation, the greek letter sigma, $\Sigma$, is used in mathematics to represent a sum of values over some range. The range is often indicated with an index number whose starting value is written below the $\Sigma$, for example, $i = 1$, and whose ending value is shown above the $\Sigma$, for example, N. A more explicit representation would look like this,

$$S(P) = \sum_{i=1}^{N} MC_i^{-1}(P) = MC_1^{-1}(P) + MC_2^{-1}(P) + \ldots + MC_N^{-1}(P)$$

The -1 superscript above each MC is used to indicated that what we are adding is the inverse of the marginal cost functions, not the marginal cost themselves. Without the -1, we would be adding each firms MC at every output level to obtain the vertical summation of the MC curves, because MC is measured on the horizontal axis. Instead we need to add each firm's output level for every P = MC. That's what $MC^{-1}$ generates. Thus $MC_1^{-1}(P)$ is the output of firm 1 at the market price P, $MC_2^{-1}(P)$ is the output of firm 2 at the price P, $MC_N^{-1}(P)$ is the output of firm N at the price P and the … is used to indicate all the firms between 3 and N not explicitly written in the equation.

The key point to this side note is to provide a detailed explanation of a common economics reference (namely that the supply curve is the horizontal summation of the MC curves) and to illustrate how we can used mathematical notation to describe precisely what is being aggregated and how.

## The Firm's Decision to Shut down Production

One question we can ask at this stage is what market price would induce the firm to close down operations and do something else instead. Simple logic would suggest that anytime a firm loses money, or has negative profit, then it would shut down production. However, it turns out, that intuition is wrong.

The reason is because of the presence of fixed costs which is one of the assumptions generating the cost curves shown.

Consider the following profit relationships. In general, when the firm produces a positive quantity, we can write the firm's profit, $\Pi_{Prod}$, as the difference between total revenue, TR, and total cost, TC, and since total cost is the sum of total fixed cost, TFC, and total variable cost, TVC, we get the following,

$$\Pi_{Prod} = TR - TFC - TVC$$

If the firm does not produce, then it would earn no revenue and incur no variable costs, but it would still incur fixed costs. Therefore profit in this scenario, $\Pi_{NotProd}$, can be written as,

$$\Pi_{NotProd} = -TFC$$

In general, it makes logical sense that a firm should produce whenever the profit from producing exceeds the profit when not producing. Thus, produce whenever,

$$\Pi_{Prod} > \Pi_{NotProd}$$

or,

$$TR - TFC - TVC > -TFC$$

or,

$$TR - TVC > 0 \quad \text{or} \quad TR > TVC$$

If we divide this inequality by the quantity produced, then TR/Q = P and TVC/Q = AVC therefore the condition to produce becomes,

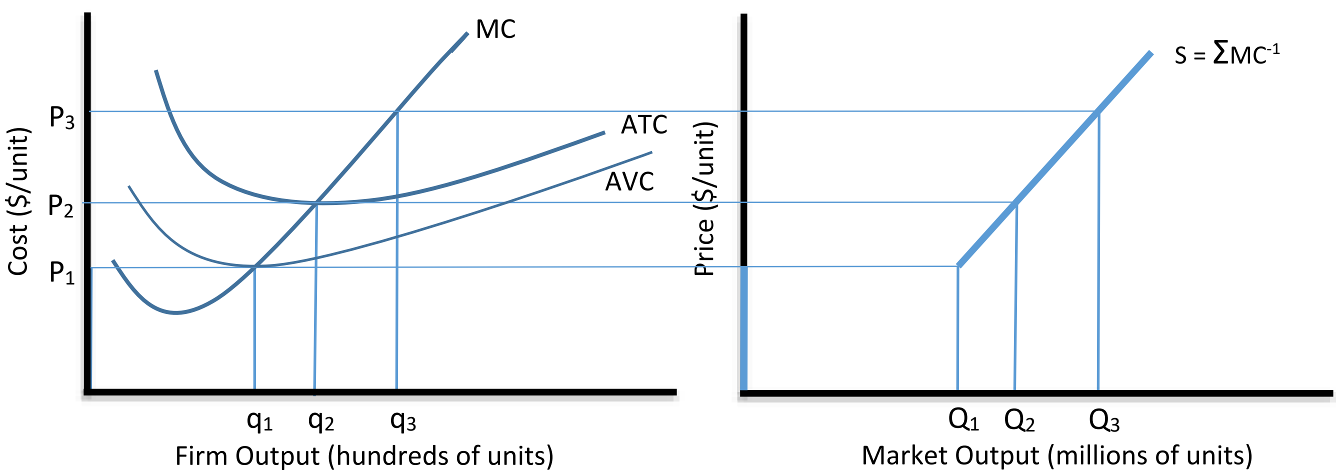
$$P > AVC$$

In other words, it will be more profitable for a firm to produce some quantity, Q, only if the market price of the product exceeds the AVC at the optimal Q chosen by the firm.

Note that in Figure 13.2, the optimal quantity for the firm at price P1 is q1 which also lies on the AVC curve. That means that at $q_1$, P = AVC and the firm should shut down production and produce 0 instead of $q_1$. For any price at $P_1$ or less, P intersects MC at a point below the AVC curve which means that optimal production is zero, rather than being read off the MC curve. Note also that for all values of P above $P_1$, such as at $P_2$ or $P_3$, the price exceeds AVC at quantities $q_2$ and $q_3$ so production will be positive at those values.

With respect to the market supply curve on the right, for all prices at P1 and above, the quantities supplied will be $Q_1$, $Q_2$, and $Q_3$ respectively, but for any price at or below P1 the quantity supplied will be zero. This means that the market supply curve is discontinuous at piece $P_1$.

Figure 13.2 To Produce of Not Produce? That is the Question




The defining characteristic of the break point is that P1 is the price that's equal to the minimum average cost of production. Thus, another way to state this condition is that the firm should produce a positive quantity whenever the market price is above the minimum AVC.

For prices between $P_1$ and $P_2$, profit for the firms will be negative and still they will stay in the market and produce at a loss. The reason is the presence of fixed costs. For example, suppose fixed costs were say $200. In this case, a market price that generates a profit maximizing loss of $100 will be better than shutting down, because zero output would incur a loss of $200, which is worse. It is better to lose $100 than lose $200, and that's why positive production will continue even with small losses.

### Key Takeaways

1. The market supply curve is derived as the horizontal summation of the individual firms marginal cost curves.
2. A firm should cease producing a product and shut down whenever the market price is less than the minimum average variable cost.

## 13.2 Long-Run Adjustments in Perfect Competition

### Learning Objectives

1. Learn the distinction between the sort-run and the long-run in a perfectly competitive market

Below we will consider several different situations and explain what can be expected to happen as a result of entry or exit of new firms in the industry in the long-run. Each situation can arise due to shifts in either supply or demand, but these shifts will be analyzed in the next chapter. Thus, at this juncture we will not ask how a market gets into each situation, only what happens afterwards.

The response we will show involves the entry of new firms into a market when there are positive profits being made, and exit of firms from the market when losses are being made. We will define the *short-run*, and hence refer to short-run equilibria, as the outcomes expected BEFORE free entry or exit of firms occurs. We define the *long-run*, and hence long-run equilibria, as the outcomes expected AFTER the process of entry or exit is complete.

**The Long-Run Adjustment to Positive Profits**

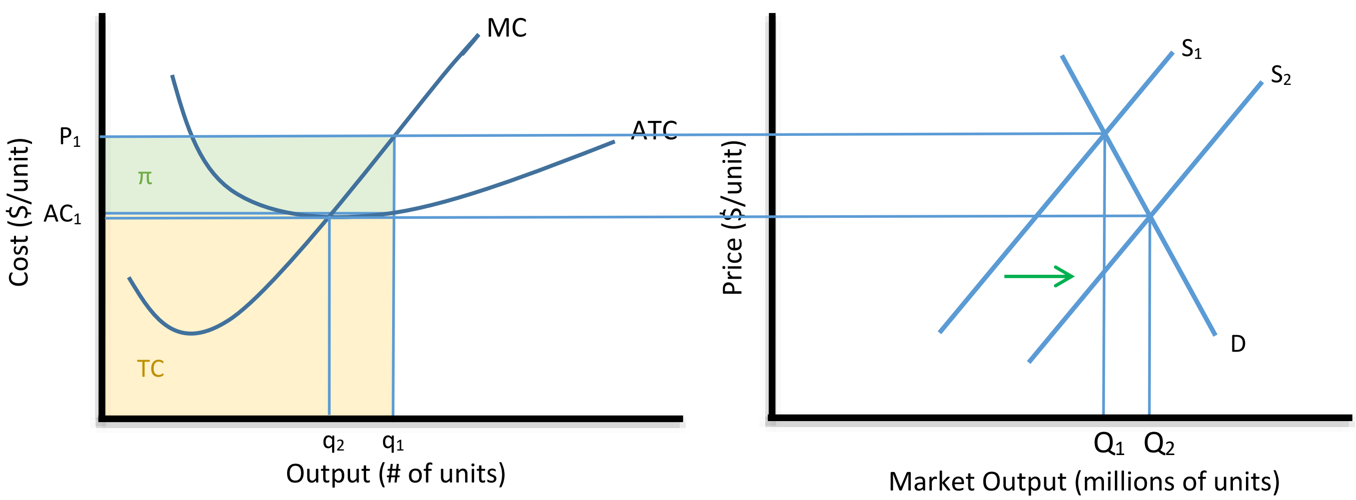
Consider the situation shown in Figure 13.3. Suppose the initial market supply and market demand curves are $S_1$ and D respectively, shown in the graph on the right. These curves generate the market price $P_1$. On the left-side graph is shown the cost curve for a representative firm in the industry, and at the price $P_1$ the firm is clearly making positive profit as indicated by the green area in the graph.

Because we assume there is perfect information, it means that everyone, both inside and outside the market knows about the positive profit for every firm operating in this industry. Next we assume that except for the standard fixed costs necessary to produce the product, there are no other costs or impediments to opening up a new business to produce this product. If someone is producing a product in some other industry and is making negative profit, then knowledge that this industry is generating positive profit is sufficient inducement to switch industries.

As new firms enter the industry, each one maximizing profit like the incumbent firms, the aggregate supply at each price will increase because the number of firms, N, is rising. This causes the aggregate supply function to shift to the right as shown on the right-side diagram. That rightward shift causes the market price to fall from P1 to a new lower level. As a result each representative firm, facing a lower market price, will slide down along its MC curve and reduce its output accordingly. Although each firm reduces output, the larger number of firms compensates and causes aggregate output to rise. Logic dictates that as long as positive profit is being made by firms in the industry, new entry will continue, and the market price will continue to fall. This process will stop as soon as firms are no longer making profit, that is, profit equals zero.

That equilibrium outcome is shown on the diagram. When market supply rises to $S_2$, the price will fall to the level equal to the minimum average cost of production (located just slightly below $AC_1$ in the left-side graph) and profit for all firms will fall to zero. With zero profit, firms will stop entering the market and there will be no reason for the outcome to change. This is the long-run equilibrium.

Figure 13.3 Long-Run Adjustments to Positive Profits




## Reality Check

The assumption of perfect information, tied with no entry costs beyond the standard fixed costs, is a very strong and unrealistic assumption. Perfect information here means that knowledge of production methods and profits being earned is known to everyone in the economy. No entry or exit costs means it takes no time or effort to close one business and open an entirely different one tomorrow.

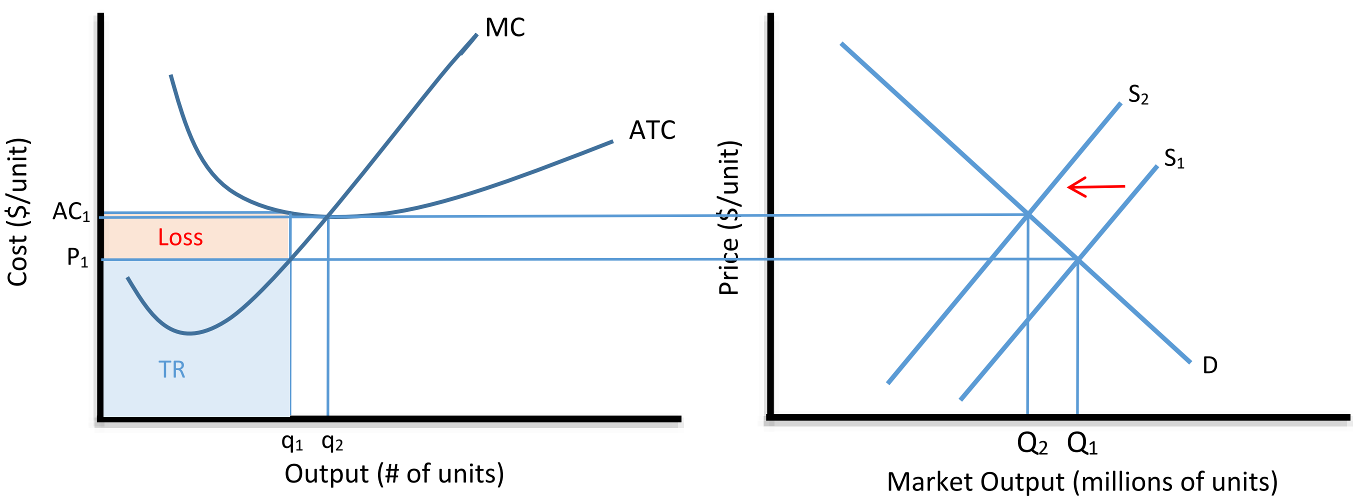
However, in the real world we expect owners, managers, and workers in an industry will have specific knowledge tied to their industry because they have learned the techniques and practiced producing the product. For example, the bakers in the pastry industry knows how to produce pastries, but if the bakers suddenly learn that more money could be made making furniture, it is unlikely they could close their bakery today and open up a furniture factory tomorrow. And yet, that is what's being imagined in the long-run adjustments in the PC model.

Incorporating realistic assumptions would make the modeling vastly more complicated and so one reason these assumptions are made is to keep the model simple and tractable. However, the lack of realism also means we must be careful how to interpret the results shown. One way to make some sense of it is to recognize that information is often eventually discoverable if enough time passes and that costs of adjustment can be incurred and expensed over time as well. Thus, while the model envisions somewhat instantaneous adjustments to the long-run equilibria, it is more realistic to imagine a real world adjustment process that might take months or even years. Positive profits will not be instantaneously known, but eventually observers will be able to determine where profits are high, especially when the profits are persistent. For example, owners in that industry may become noticeably wealthier. Production techniques cannot be instantaneously learned, but with study and time, smart people can figure out how to learn new skills and apply them in new industries. That means that the long-run equilibria described here can offer a guide to the direction of changes likely to occur, but real world market frictions will mean that the long-run really would take a long time in coming.

### The Long-Run Adjustment to Negative Profits

Consider the situation shown in Figure 13.4. Suppose the initial market supply and market demand curves are $S_1$ and D respectively, shown in the graph on the right. These curves generate the market price $P_1$. On the left-side graph is shown the cost curve for a representative firm in the industry, and at the price $P_1$ the firm is clearly making losses as indicated by the red area in the graph.

Figure 13.4 Long-Run Adjustments to Negative Profits




In this case, losses will cause firms to exit the industry but it is likely that this will not happen all at once. In the graph on the left-side, average variable cost is not shown, so it is unknown whether, at price $P_1$, the representative firm is losing more than their fixed costs and would shut down. Let's assume that the losses are not that large, in which case the firms would continue operating despite the losses. In chapter 9.2, we noted that fixed costs often arise due to contracts a firm has committed to, such as a property lease on their factory space. When the lease expires, so do their fixed costs. If we now assume that all firms in the industry are identical in their cost functions, but are different in terms of when their leases expire (because the businesses opened at different times in the past), then the fixed costs for some loss-making firms will disappear before other firms. Those firms whose fixed costs disappear first, will exit the industry first. Thus, by imagining different contract terms across firms in an industry, we can introduce a simple mechanism in which firms leave the industry gradually over time, rather than all at once.

As firms gradually exit the industry, the aggregate supply at each price will decrease because the number of firms, N, is falling. This causes the aggregate supply function to shift to the left as shown on the right-side diagram. That leftward shift causes the market price to rise from P1 to a new higher level. As a result each representative firm, facing a higher market price, will slide up along its MC curve and increase its output accordingly. Although each firm raises output, the smaller number of firms producing offsets this and causes aggregate output to fall. Logic dictates that as long as negative profit continues for firms in the industry, exit will continue, and the market price will rise. This process will stop as soon as firms are no longer making losses and profit equals zero.

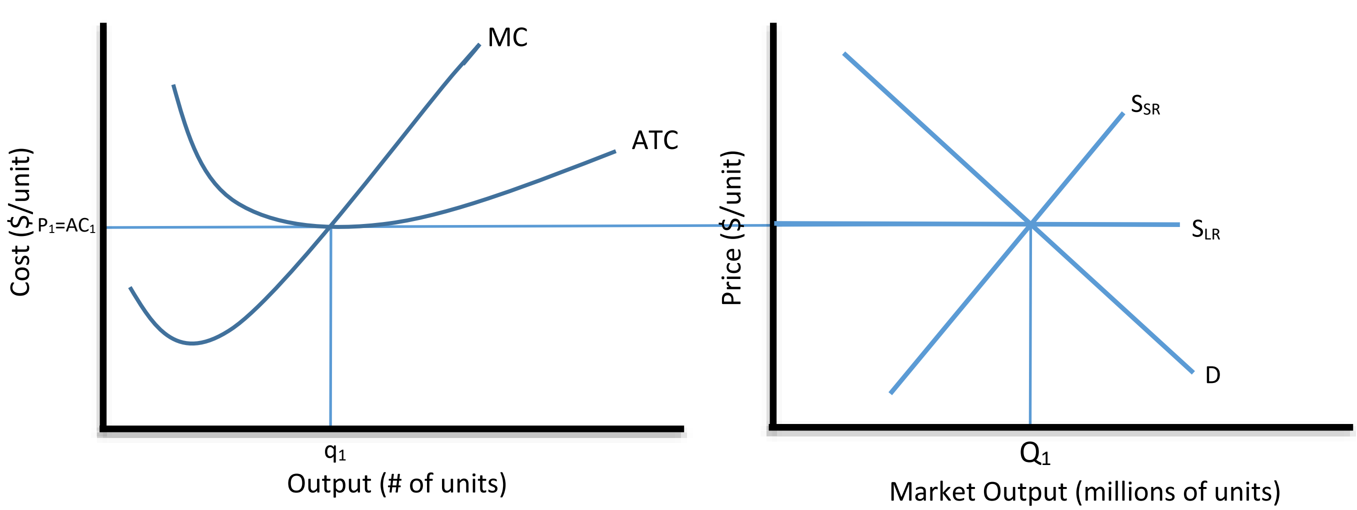
That equilibrium outcome is shown on the diagram. When market supply falls to $S_2$, the price will rise to the level equal to the minimum average cost of production (located just slightly below $AC_1$ in the left-side graph) and profit for all firms will rise to zero. With zero profit, firms will stop exiting the market and there will be no reason for the outcome to change. This is the long-run equilibrium.

## The Long-Run Supply Function

The long-run represents the period of time needed to enable firm entry or exit to cause a change in the market price until zero profit obtains for all firms in the industry. A long-run equilibrium is depicted in Figure 13.5. Market supply and demand intersect at the price $P_1$ which is also the minimum average cost in production, $AC_1$. This, zero profit obtains for all firms in the industry and there is no incentive for firms to enter or exit any further.

In the next chapter, we will analyze changes in a market caused by movements in the underlying exogenous variables affecting supply or demand. These changes will cause profits to rise or fall in the short-run, but in the long-run entry or exit will force profit back to zero and the price back to the minimum average cost. For that reason, we can display a *long-run supply function* as a horizontal line drawn at the minimum average cost for the representative firm. This is labeled $S_{LR}$ in the market diagram on the right and represents all possible supplies at the single price that will prevail in the long-run after entry or exit has completed its process. The original supply curve in the market is now labeled $S_{SR}$ to indicate that this should be interpreted as a short-run supply function. Typically, we will drop the SR notation later, but it is worth remembering that any upward sloping market supply function indicates outcomes that arise before free entry or exit has occurred.

Figure 13.5 A Long-Run Equilibrium Supply Curve



There are some situations that can cause a the long-run supply function to slope either upwards or downwards, if some of the underlying assumptions are changed. For example, $S_{LR}$ can have a positive slope if different firms in the industry have different production costs. In this case, lower cost firms will enter earlier while higher cost firms will enter later. Similarly in reverse, high cost firms will exit the market before the low cost firms, if market prices fall. $S_{LR}$ can also be positively sloped if, as more firms enter the market, there is a substantial increase in demand for inputs in the production process. The increase in input demand could in turn cause input prices for all the firms to rise, thereby requiring higher product prices to induce greater supply. One other situation could cause $S_{LR}$ to slope negatively. If inputs in the production process were subject to economies of scale, then increases in demand in the product market, could increase demand for inputs which via greater production could reduce unit costs and therefore lower input prices in the final product industry. This would mean higher supplies could be maintained with lower prices because input prices had fallen.

This is just another set of examples that when you adjust the assumptions in the model, economic relationships change, graphs are drawn differently, and outcomes are sometimes significantly affected. Going forward we will ignore these complicating factors and assume the conditions needed to assure that $S_{LR}$ is horizontal at minimum average cost.

### Key Takeaways

1. The long-run in perfect competition is the period of time after entry or exit in response to profits or losses has completed.
2. Positive profits in an industry induces entry of new firms which in turn raises market supply, reduces the market price and reduces firm profits.
3. Negative profits, or losses, in an industry induces exit of firms which in turn lowers market supply, and raises the market price and firm profits.
4. In a long-run equilibrium, firm profits are zero and the market price is equal to the minimum average cost to produce the product.

## 13.3 Economic Efficiency Effects of Perfect Competition

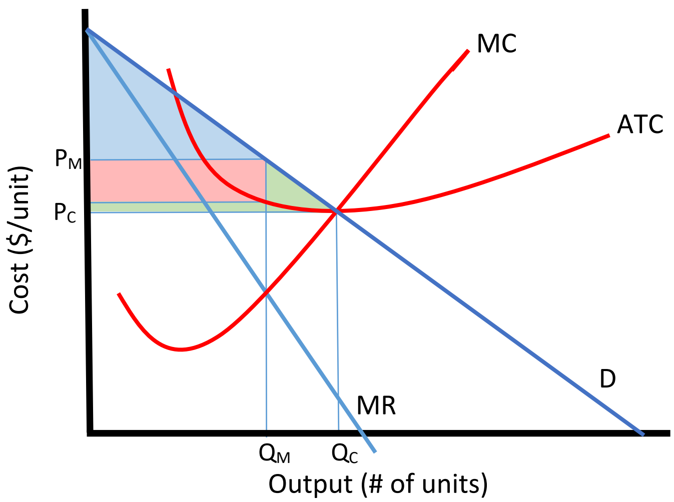
### Learning Objectives

1. Compare market welfare between a perfectly competitive market outcome and a monopoly market outcome.

In this section we will consider the welfare effects in a perfectly competitive market outcome compared with a monopoly outcome. In Figure 13.6 is shown a market demand curve and its associated marginal revenue curve as seen from the vantage point of a monopoly firm that can set the price it charges in the market. An average cost and marginal cost curve is also drawn but we must make some simplifying assumptions to enable the comparison between a monopoly and a perfectly competitive market.

Suppose there are many small firms each operating a single manufacturing plant. As an example, imagine lots of small bakeries producing bread in a city. Suppose the minimum average cost for each bakery occurs at the price $P_C$. However, because each small bakery produces a small quantity relative to the size of the total market, its output might only number in the hundreds of loaves of bread whereas the market supply shown in Figure 13.6 might number in the tens of millions. This set of assumptions implies that the average cost curve shown in the diagram is the horizontal summation of the many small bakery AC curves, rather than the AC curve for a single large-sized bread mega-factory.

Figure 13.6 Comparing Perfect Competition with Monopoly




Next, in a long-run equilibrium in a perfectly competitive setting, the number of bakeries operating in the industry will adjust via entry or exit, such that the aggregate average cost, as shown in the diagram, will intersect the aggregate demand function at minimum average cost for each firm, which corresponds to the price $P_C$. This means that the perfectly competitive outcome will be a price of $P_C$ and an aggregate output of $Q_C$.

Because in a long-run equilibrium, these competitive bakeries make zero profit, all surplus accrues to the bread consumers. That consumer surplus is measured as the triangular area between the demand curve and the horizontal line drawn at the perfectly competitive market price, $P_C$. In the Figure 13.6 this is all of the shaded areas. This is also market welfare since consumers are the only ones to enjoy surplus in this equilibrium.

Next, let's suppose that a monopoly forms in the following way. Suppose a very wealthy individual purchases every bakery operating in this PC market, with the intention of adjusting output and price in order to raise its profit. If this monopoly only adjusts output by adjusting the output of every bakery proportionally, then its average cost and marginal cost curves will be the same as depicted in Figure 13.6. [Note that the monopoly could also close some of the bakeries and adjust the overall scale of production which would shift its aggregate AC and MC curve and may result in a different outcome. We keep this exercise simpler by assuming the number of bakeries under monopoly control remains fixed.]

The profit maximizing output choice of this monopoly would be made by setting the market marginal revenue (MR) equal to marginal cost (MC) and produce $Q_M$ units of the product, charging a price $P_M$, read off of the demand curve. The cost of producing $Q_M$ would be $AC_M$, not labeled Figure 13.6. Finally, the monopoly's total profit is given by $(P_M – AC_M)Q_M$ and is represented by the reddish area on the diagram.

Consumer surplus at the monopoly price and quantity is shown as the area between the demand curve and the price line at $P_M$ and is represented by the blue area in Figure 13.6. That means that total market welfare is given by the sum of the blue and reddish areas in the diagram.

Note that total market welfare, equal to all of the shaded areas in perfect competition has now fallen to the smaller blue and reddish area only. This means that total market welfare is reduced with the formation of the monopoly, given by the greenish area in the diagram.

The advantage of monopolization is that firm profit rises to a positive level from zero. This is good for the monopoly owner. However, it is bad for the consumers who must pay a higher price for a lower overall quantity of bread. Some consumers may no longer purchase bread and are driven from the market, and/or, bread consumers all purchase a smaller quantity due to the higher price. Regardless, the analysis shows that the total losses to consumers will exceed the benefits that accrue to the monopoly firm resulting in a net loss of benefits to the market participants in total. Economists say that income (another way to measure well-being) is redistributed moving from perfect competition to a monopoly. As it's redistributed, economic efficiency is reduced with the formation of a monopoly.

The bad outcome that arises with monopoly also demonstrates the good outcome that arises due to competition in a market. Competition forces prices down, expands total supply, reduces monopoly profit, and enhances both consumer and overall market well-being. This is the same and saying there is an improvement in economic efficiency. Furthermore, to the extent that the average consumers are less wealthy than a monopoly firm owner, which is quite likely, competition redistributes income away from the wealthy individual towards the average person and can make income more equal overall. These are the key advantages of competitive markets.

## Key Takeaways

1. Perfect competition, relative to a monopoly in the same market, results in a lower market price, higher total output, an increase in consumer surplus, a reduction in firm profit (to zero), and an improvement in overall economic efficiency, or market welfare.
2. Perfect competition, relative to monopoly, because it reduces firm profit and extends greater surplus to consumers, may also result in a more equal distribution of income.