CHAPTER 22

FIRM SUPPLY

In this chapter we will see how to derive the supply curve of a competitive firm from its cost function using the model of profit maximization. The first thing we have to do is to describe the market environment in which the firm operates.

22.1 Market Environments

Every firm faces two important decisions: choosing how much it should produce and choosing what price it should set. If there were no constraints on a profit-maximizing firm, it would set an arbitrarily high price and produce an arbitrarily large amount of output. But no firm exists in such an unconstrained environment. In general, the firm faces two sorts of constraints on its actions.

First, it faces the **technological constraints** summarized by the production function. There are only certain feasible combinations of inputs and outputs, and even the most profit-hungry firm has to respect the realities of the physical world. We have already discussed how we can summarize the technological constraints, and we've seen how the technological constraints lead to the **economic constraints** summarized by the cost function.

But now we bring in a new constraint—or at least an old constraint from a different perspective. This is the **market constraint**. A firm can produce whatever is physically feasible, and it can set whatever price it wants ... but it can only sell as much as people are willing to buy.

If it sets a certain price p it will sell a certain amount of output x. We call the relationship between the price a firm sets and the amount that it sells the **demand curve facing the firm**.

If there were only one firm in the market, the demand curve facing the firm would be very simple to describe: it is just the market demand curve described in earlier chapters on consumer behavior. For the market demand curve measures how much of the good people want to buy at each price. Thus the demand curve summarizes the market constraints facing a firm that has a market all to itself.

But if there are other firms in the market, the constraints facing an individual firm will be different. In this case, the firm has to guess how the *other* firms in the market will behave when it chooses its price and output.

This is not an easy problem to solve, either for firms or for economists. There are a lot of different possibilities, and we will try to examine them in a systematic way. We'll use the term **market environment** to describe the ways that firms respond to each other when they make their pricing and output decisions.

In this chapter we'll examine the simplest market environment, that of **pure competition**. This is a good comparison point for many other environments, and it is of considerable interest in its own right. First let's give the economist's definition of pure competition, and then we'll try to justify it.

22.2 Pure Competition

To a lay person, "competition" has the connotation of intense rivalry. That's why students are often surprised that the economist's definition of competition seems so passive: we say that a market is **purely competitive** if each firm assumes that the market price is independent of its own level of output. Thus, in a competitive market, each firm only has to worry about how much output it wants to produce. Whatever it produces can only be sold at one price: the going market price.

In what sort of environment might this be a reasonable assumption for a firm to make? Well, suppose that we have an industry composed of many firms that produce an identical product, and that each firm is a small part of the market. A good example would be the market for wheat. There are thousands of wheat farmers in the United States, and even the largest of them produces only an infinitesimal fraction of the total supply. It is reasonable in this case for any one firm in the industry to take the market price as being predetermined. A wheat farmer doesn't have to worry about what price to set for his wheat—if he wants to sell any at all, he has to sell it at the market price. He is a **price taker**: the price is given as far as he is concerned; all he has to worry about is how much to produce.

This kind of situation—an identical product and many small firms—is a classic example of a situation where price-taking behavior is sensible. But it is not the only case where price-taking behavior is possible. Even if there are only a few firms in the market, they may still treat the market price as being outside their control.

Think of a case where there is a fixed supply of a perishable good: say fresh fish or cut flowers in a marketplace. Even if there are only 3 or 4 firms in the market, each firm may have to take the *other* firms' prices as given. If the customers in the market only buy at the lowest price, then the lowest price being offered is the market price. If one of the other firms wants to sell anything at all, it will have to sell at the market price. So in this sort of situation competitive behavior—taking the market price as outside of your control—seems plausible as well.

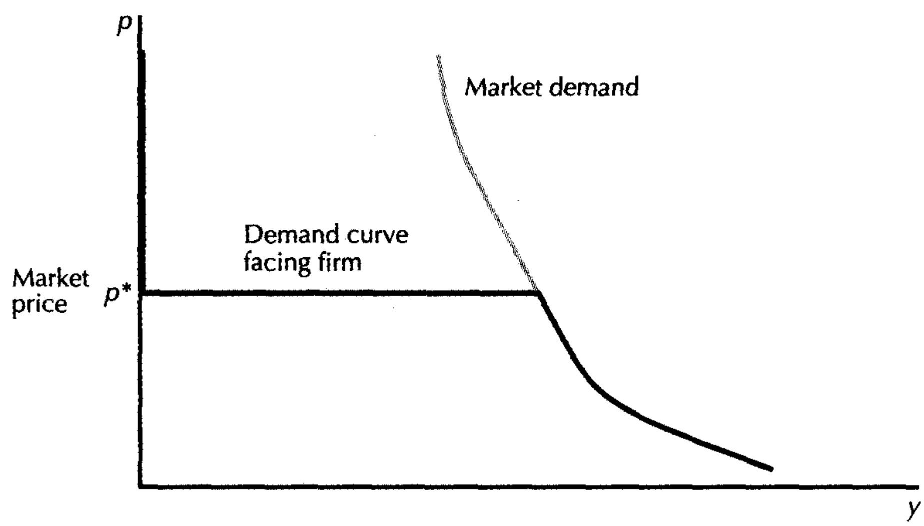
We can describe the relationship between price and quantity perceived by a competitive firm in terms of a diagram as in Figure 22.1. As you can see, this demand curve is very simple. A competitive firm believes that it will sell nothing if it charges a price higher than the market price. If it sells at the market price, it can sell whatever amount it wants, and if it sells below the market price, it will get the entire market demand at that price.

As usual we can think of this kind of demand curve in two ways. If we think of quantity as a function of price, this curve says that you can sell any amount you want at or below the market price. If we think of price as a function of quantity, it says that no matter how much you sell, the market price will be independent of your sales.

(Of course, this doesn't have to be true for literally *any* amount. Price has to be independent of your output for any amount you might consider selling. In the case of the cut-flower seller, the price has to be independent of how much she sells for any amount up to her stock on hand—the maximum that she could consider selling.)

It is important to understand the difference between the "demand curve facing a firm" and the "market demand curve." The market demand curve measures the relationship between the market price and the total amount of output sold. The demand curve facing a firm measures the relationship between the market price and the output of *that particular firm.*

The market demand curve depends on consumers' behavior. The demand curve facing a firm not only depends on consumers' behavior but it also depends on the behavior of the other firms. The usual justification for the competitive model is that when there are many small firms in the market, each one faces a demand curve that is essentially flat. But even if there are only two firms in the market, and one insists on charging a fixed price no matter what, then the other firm in the market will face a competitive demand curve like the one depicted in Figure 22.1. Thus the competitive model may hold in a wider variety of circumstances than is apparent at first glance.



The demand curve facing a competitive firm. The firm's demand is horizontal at the market price. At higher prices, the firm sells nothing, and below the market price it faces the entire market demand curve.

22.3 The Supply Decision of a Competitive Firm

Let us use the facts we have discovered about cost curves to figure out the supply curve of a competitive firm. By definition a competitive firm ignores its influence on the market price. Thus the maximization problem facing a competitive firm is

$$\max_{y} py - c(y).$$

This just says that the competitive firm wants to maximize its profits: the difference between its revenue, py, and its costs, $c(y)$.

What level of output will a competitive firm choose to produce? Answer: it will operate where marginal revenue equals marginal cost—where the extra revenue gained by one more unit of output just equals the extra cost of producing another unit. If this condition did not hold, the firm could always increase its profits by changing its level of output.

In the case of a competitive firm, marginal revenue is simply the price. To see this, ask how much extra revenue a competitive firm gets when it increases its output by Δy. We have

$$\Delta R = p\Delta y$$

since by hypothesis p doesn't change. Thus the extra revenue per unit of output is given by

$$\frac{\Delta R}{\Delta y} = p,$$

which is the expression for marginal revenue.

Thus a competitive firm will choose a level of output y where the marginal cost that it faces at y is just equal to the market price. In symbols:

$$p = MC(y).$$

For a given market price, p, we want to find the level of output where profits are maximal. If price is greater than marginal cost at some level of output y, then the firm can increase its profits by producing a little more output. For price greater than marginal costs means

$$p - \frac{\Delta c}{\Delta y} > 0.$$

So increasing output by Δy means that

$$p\Delta y - \frac{\Delta c}{\Delta y}\Delta y > 0.$$

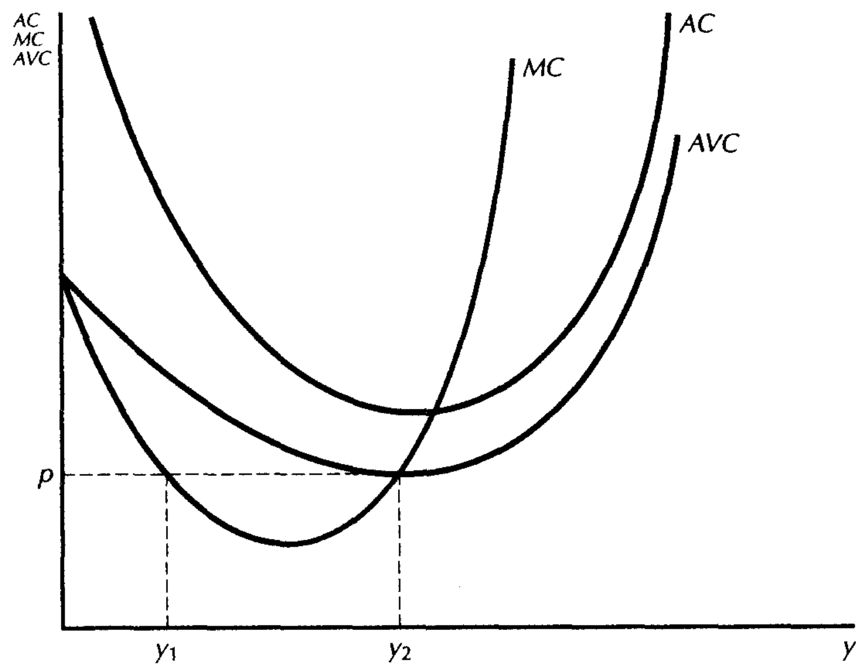
Simplifying we find that

$$p\Delta y - \Delta c > 0,$$

which means that the increase in revenues from the extra output exceeds the increase in costs. Thus profits must increase.

A similar argument can be made when price is less than marginal cost. Then reducing output will increase profits, since the lost revenues are more than compensated for by the reduced costs.

So at the optimal level of output, a firm must be producing where price equals marginal costs. Whatever the level of the market price p, the firm will choose a level of output y where $p = MC(y)$. Thus the marginal cost curve of a competitive firm is precisely its supply curve. Or put another way, the market price is precisely marginal cost—as long as each firm is producing at its profit-maximizing level.



Marginal cost and supply. Although there are two levels of output where price equals marginal cost, the profit-maximizing quantity supplied can lie only on the upward-sloping part of the marginal cost curve.

22.4 An Exception

Well ... maybe not *precisely*. There are two troublesome cases. The first case is when there are several levels of output where price equals marginal cost, such as the case depicted in Figure 22.2. Here there are two levels of output where price equals marginal cost. Which one will the firm choose?

It is not hard to see the answer. Consider the first intersection, where the marginal cost curve is sloping down. Now if we increase output a little bit here, the costs of each additional unit of output will decrease. That's what it means to say that the marginal cost curve is decreasing. But the market price will stay the same. Thus profits must definitely go up.

So we can rule out levels of output where the marginal cost curve slopes downward. At those points an increase in output must always increase profits. The supply curve of a competitive firm must lie along the upward-sloping part of the marginal cost curve. This means that the supply curve itself must always be upward sloping. The "Giffen good" phenomenon cannot arise for supply curves.

Price equals marginal cost is a *necessary* condition for profit maximization. It is not in general a *sufficient* condition. Just because we find a point where price equals marginal cost doesn't mean that we've found the maximum profit point. But if we find the maximum profit point, we know that price must equal marginal cost.

22.5 Another Exception

This discussion is assuming that it is profitable to produce something. After all it could be that the best thing for a firm to do is to produce zero output. Since it is always possible to produce a zero level of output, we have to compare our candidate for profit maximization with the choice of doing nothing at all.

If a firm produces zero output it still has to pay its fixed costs, F. Thus the profits from producing zero units of output are just $-F$. The profits from producing a level of output y are $py - c_v(y) - F$. The firm is better off going out of business when

$$-F > py - c_v(y) - F,$$

that is, when the "profits" from producing nothing, and just paying the fixed costs, exceed the profits from producing where price equals marginal cost. Rearranging this equation gives us the **shutdown condition**:

$$AVC(y) = \frac{c_v(y)}{y} > p.$$

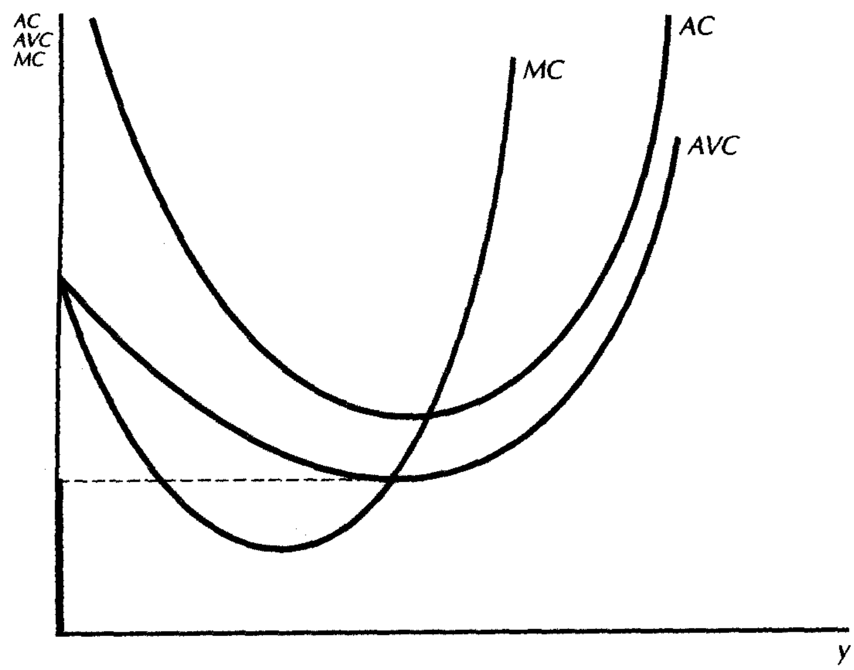
If average variable costs are greater than p, the firm would be better off producing zero units of output. This makes good sense, since it says that the revenues from selling the output y don't even cover the *variable* costs of production, $c_v(y)$. In this case the firm might as well go out of business. If it produces nothing it will lose its fixed costs, but it would lose even more if it continued to produce.

This discussion indicates that only the portions of the marginal cost curve that lie above the average variable cost curve are possible points on the supply curve. If a point where price equals marginal cost is beneath the average variable cost curve, the firm would optimally choose to produce zero units of output.

We now have a picture for the supply curve like that in Figure 22.3. The competitive firm produces along the part of the marginal cost curve that is upward sloping and lies above the average variable cost curve.

EXAMPLE: Pricing Operating Systems

A computer requires an operating system in order to run, and most hardware manufacturers sell their computers with the operating systems already



Average variable cost and supply. The supply curve is the upward-sloping part of the marginal cost curve that lies above the average variable cost curve. The firm will not operate on those points on the marginal cost curve below the average cost curve since it could have greater profits (less losses) by shutting down.

installed. In the early 1980s several operating system producers were fighting for supremacy in the IBM-PC-compatible microcomputer market. The common practice at that time was for the producer of the operating system to charge the computer manufacturer for each copy of the operating system that was *installed* on a microcomputer that it sold.

Microsoft Corporation offered an alternative plan in which the charge to the manufacturer was based on the number of microcomputers that were *built* by the manufacturer. Microsoft set their licensing fee low enough that this plan was attractive to the producers.

Note the clever nature of Microsoft's pricing strategy: once the contract with a manufacturer was signed, the marginal cost of installing MS-DOS on an already-built computer was zero. Installing a competing operating system, on the other hand, could cost \$50 to \$100. The hardware manufacturer (and ultimately the user) paid Microsoft for the operating system, but the structure of the pricing contract made MS-DOS very attractive relative to the competition. As a result, Microsoft ended up being the default operating system installed on microcomputers and achieved a market penetration of over 90 percent.

22.6 The Inverse Supply Function

We have seen that the supply curve of a competitive firm is determined by the condition that price equals marginal cost. As before we can express this relation between price and output in two ways: we can either think of output as a function of price, as we usually do, or we can think of the "inverse supply function" that gives price as a function of output. There is a certain insight to be gained by looking at it in the latter way. Since price equals marginal cost at each point on the supply curve, the market price must be a measure of marginal cost for every firm operating in the industry. A firm that produces a lot of output and a firm that produces only a little output must have the *same marginal cost,* if they are both maximizing profits. The total cost of production of each firm can be very different, but the marginal cost of production must be the same.

The equation $p = MC(y)$ gives us the inverse supply function: price as a function of output. This way of expressing the supply curve can be very useful.

22.7 Profits and Producer's Surplus

Given the market price we can now compute the optimal operating position for the firm from the condition that $p = MC(y)$. Given the optimal operating position we can compute the profits of the firm. In Figure 22.4 the area of the box is just p^*y^*, or total revenue. The area $y^*AC(y^*)$ is total costs since

$$yAC(y) = y\frac{c(y)}{y} = c(y).$$

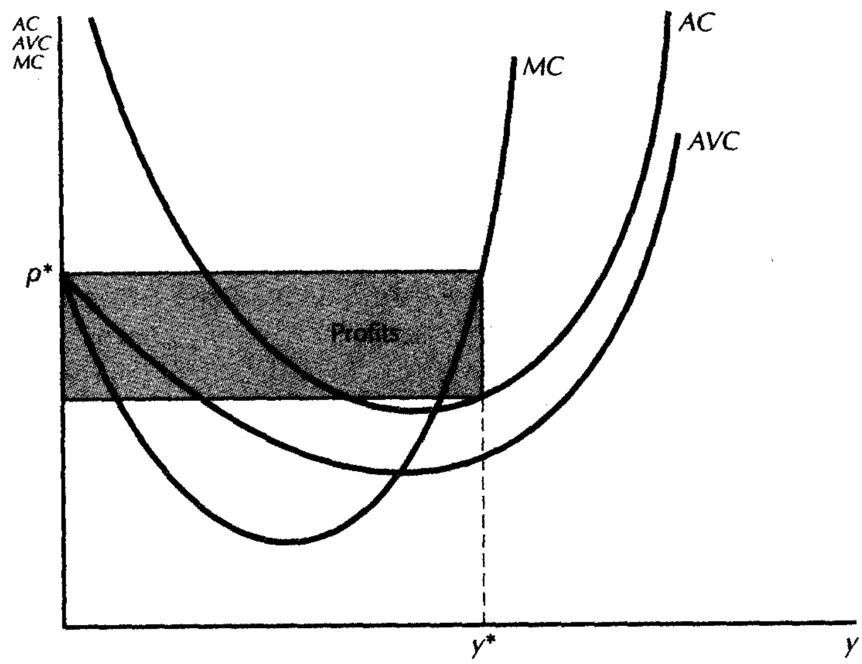
Profits are simply the difference between these two areas.

Recall our discussion of **producer's surplus** in Chapter 14. We defined producer's surplus to be the area to the left of the supply curve, in analogy to consumer's surplus, which was the area to the left of the demand curve. It turns out that producer's surplus is closely related to the profits of a firm. More precisely, producer's surplus is equal to revenues minus variable costs, or equivalently, profits plus the fixed costs:

$$\text{profits} = py - c_v(y) - F$$

$$\text{producer's surplus} = py - c_v(y).$$

The most direct way to measure producer's surplus is to look at the difference between the revenue box and the box $y^*AVC(y^*)$, as in Figure 22.5A. But there are other ways to measure producer's surplus by using the marginal cost curve itself.



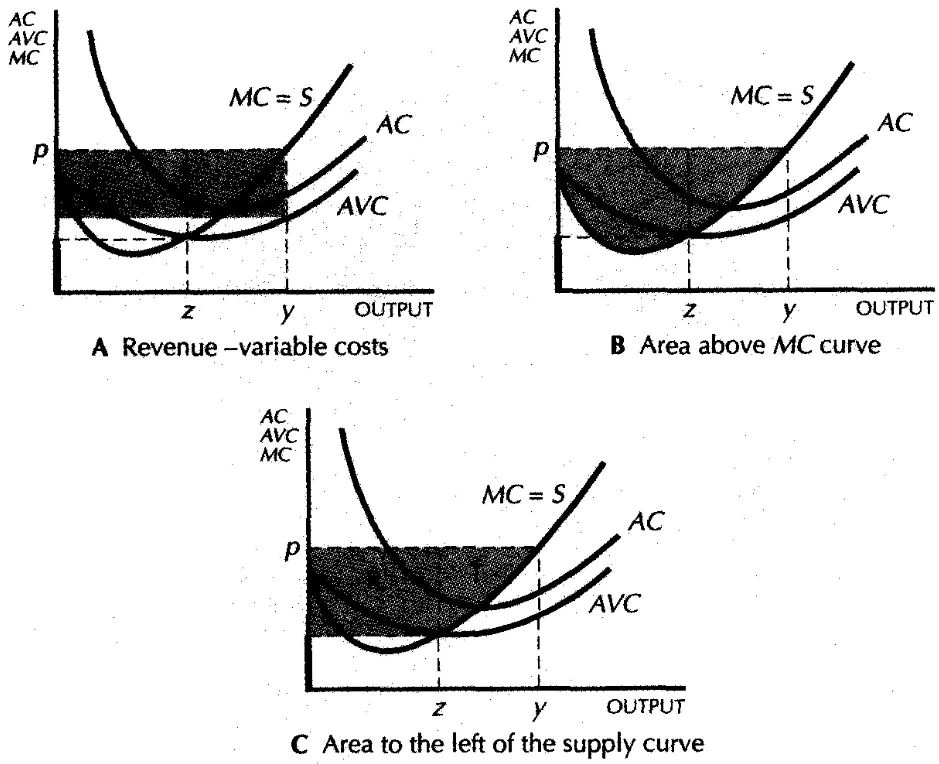
Profits. Profits are the difference between total revenue and total costs, as shown by the colored rectangle.

We know from Chapter 21 that the area under the marginal cost curve measures the total variable costs. This is true because the area under the marginal cost curve is the cost of producing the first unit plus the cost of producing the second unit, and so on. So to get producer's surplus, we can subtract the area under the marginal cost curve from the revenue box and get the area depicted in Figure 22.5B.

Finally, we can combine the two ways of measuring producer's surplus. Use the "box" definition up to the point where marginal cost equals average variable cost, and then use the area above the marginal cost curve, as shown in Figure 22.5C. This latter way is the most convenient for most applications since it is just the area to the left of the supply curve. Note that this is consistent with definition of producer's surplus given in Chapter 14.

We are seldom interested in the *total* amount of producer's surplus; more often it is the *change* in producer's surplus that is of interest. The change in producer's surplus when the firm moves from output y^* to output y' will generally be a trapezoidal shaped region like that depicted in Figure 22.6.

Note that the change in producer's surplus in moving from y^* to y' is just the change in profits in moving from y^* to y', since by definition the fixed costs don't change. Thus we can measure the impact on profits of a change in output from the information contained in the marginal cost curve, without having to refer to the average cost curve at all.



Producer's surplus. Three equivalent ways to measure producer's surplus. Panel A depicts a box measuring revenue minus variable cost. Panel B depicts the area above the marginal cost curve. Panel C uses the box up until output z (area R) and then uses the area above the marginal cost curve (area T).

EXAMPLE: The Supply Curve for a Specific Cost Function

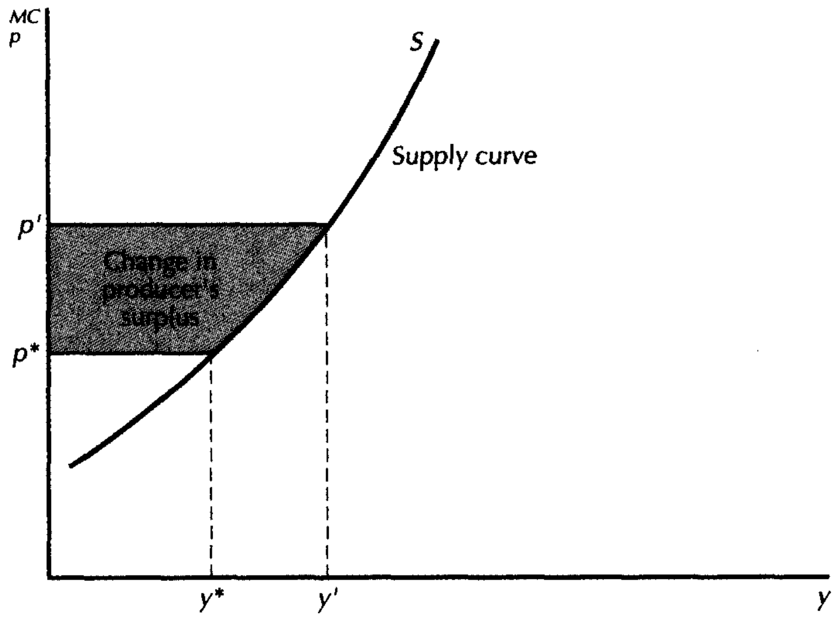
What does the supply curve look like for the example given in the last chapter where $c(y) = y^2 + 1$? In that example the marginal cost curve was always above the average variable cost curve, and it always sloped upward. So "price equals marginal costs" gives us the supply curve directly. Substituting $2y$ for marginal cost we get the formula

$$p = 2y.$$

This gives us the inverse supply curve, or price as a function of output. Solving for output as a function of price we have

$$S(p) = y = \frac{p}{2}$$

as our formula for the supply curve. This is depicted in Figure 22.7.



The change in producer's surplus. Since the supply curve coincides with the upward-sloping part of the marginal cost curve, the change in producer's surplus will typically have a roughly trapezoidal shape.

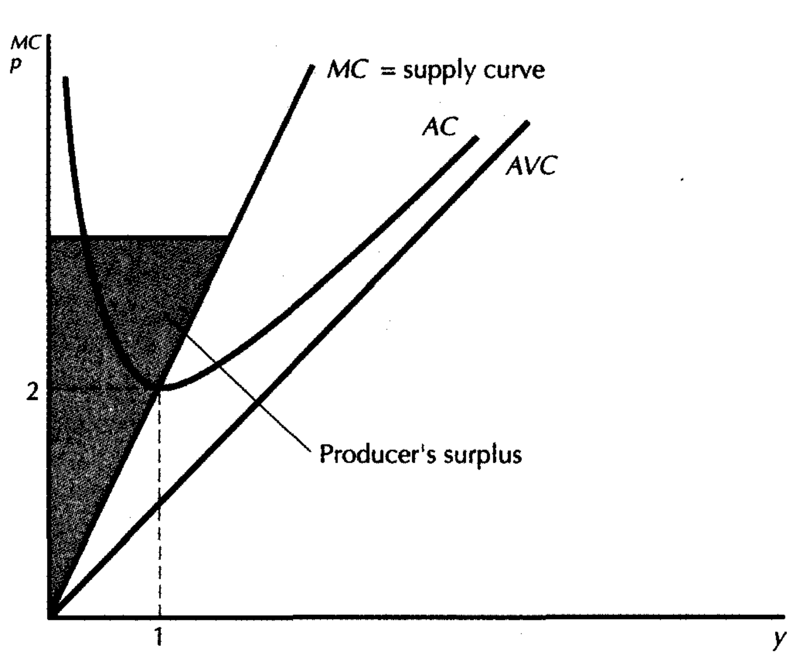
If we substitute this supply function into the definition of profits, we can calculate the maximum profits for each price p. Performing the calculation we have:

$$\begin{aligned}\pi(p) &= py - c(y)\\ &= p\frac{p}{2} - \left(\frac{p}{2}\right)^2 - 1\\ &= \frac{p^2}{4} - 1.\end{aligned}$$

How do the maximum profits relate to producer's surplus? In Figure 22.7 we see that producer's surplus—the area to the left of the supply curve between a price of zero and a price of p—will be a triangle with a base of $y = p/2$ and a height of p. The area of this triangle is

$$A = \left(\frac{1}{2}\right)\left(\frac{p}{2}\right)p = \frac{p^2}{4}.$$

Comparing this with the profits expression, we see that producer's surplus equals profits plus fixed costs, as claimed.



A specific example of a supply curve. The supply curve and producer's surplus for the cost function $c(y) = y^2 + 1$.

22.8 The Long-Run Supply Curve of a Firm

The long-run supply function for the firm measures how much the firm would optimally produce when it is allowed to adjust plant size (or whatever factors are fixed in the short run). That is, the long-run supply curve will be given by

$$p = MC_l(y) = MC(y, k(y)).$$

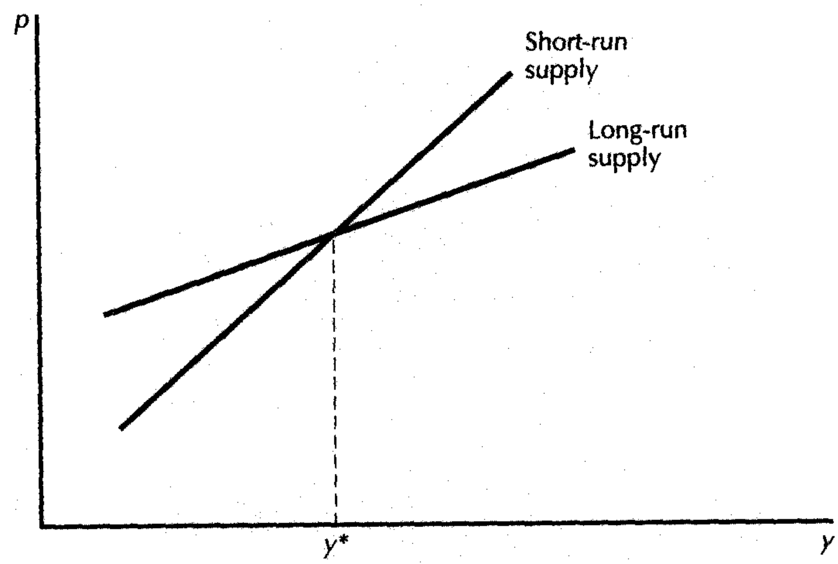
The short-run supply curve is given by price equals marginal cost at some fixed level of k:

$$p = MC(y, k).$$

Note the difference between the two expressions. The short-run supply curve involves the marginal cost of output holding k fixed at a given level of output, while the long-run supply curve involves the marginal cost of output when you adjust k optimally.

Now, we know something about the relationship between short-run and long-run marginal costs: the short-run and the long-run marginal costs coincide at the level of output y^* where the fixed factor choice associated with the short-run marginal cost is the optimal choice, k^*. Thus the short-run and the long-run supply curves of the firm coincide at y^*, as in Figure 22.8.

In the short run the firm has some factors in fixed supply; in the long run these factors are variable. Thus, when the price of output changes, the



The short-run and long-run supply curves. Typically the long-run supply curve will be more elastic than the short-run supply curve.

firm has more choices to adjust in the long run than in the short run. This suggests that the long-run supply curve will be more responsive to price—more elastic—than the short-run supply curve, as illustrated in Figure 22.8.

What else can we say about the long-run supply curve? The long run is defined to be that time period in which the firm is free to adjust all of its inputs. One choice that the firm has is the choice of whether to remain in business. Since in the long run the firm can always get zero profits by going out of business, the profits that the firm makes in long-run equilibrium have to be at least zero:

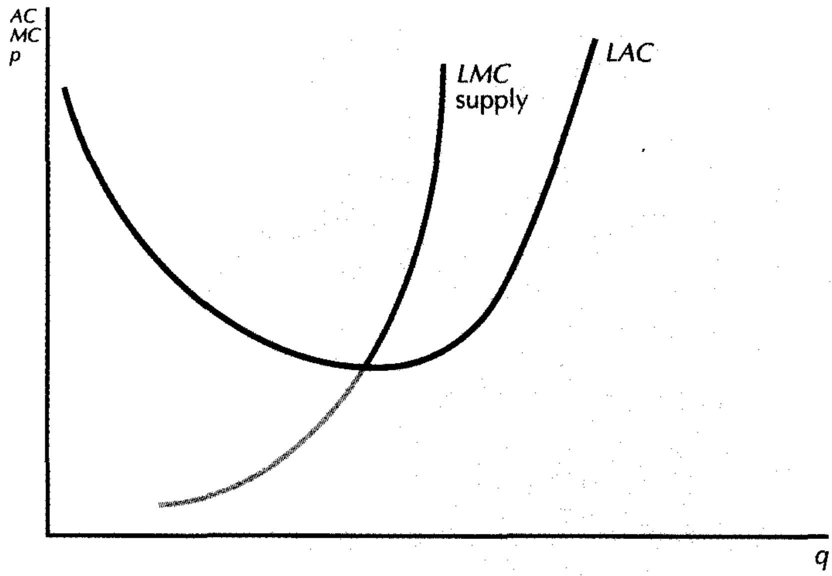
$$py - c(y) \geq 0,$$

which means

$$p \geq \frac{c(y)}{y}.$$

This says that in the long run price has to be at least as large as average cost. Thus the relevant part of the long-run supply curve is the upward-sloping part of the marginal cost curve that lies above the long-run average cost curve, as depicted in Figure 22.9.

This is completely consistent with the short-run story. In the long run all costs are variable costs, so the short-run condition of having price above average variable cost is equivalent to the long-run condition of having price above average cost.



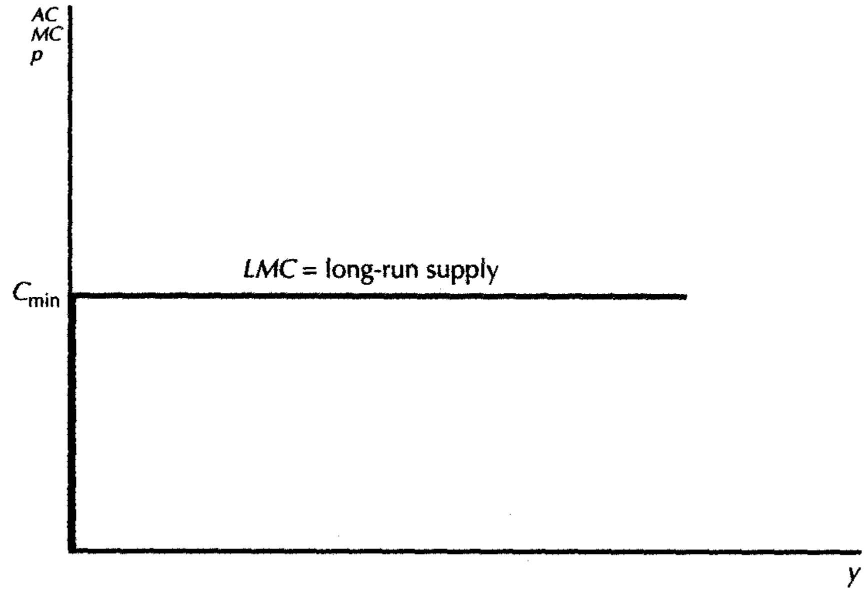
The long-run supply curve. The long-run supply curve will be the upward-sloping part of the long-run marginal cost curve that lies above the average cost curve.

22.9 Long-Run Constant Average Costs

One particular case of interest occurs when the long-run technology of the firm exhibits constant returns to scale. Here the long-run supply curve will be the long-run marginal cost curve, which, in the case of constant average cost, coincides with the long-run average cost curve. Thus we have the situation depicted in Figure 22.10, where the long-run supply curve is a horizontal line at $c_{\min}$, the level of constant average cost.

This supply curve means that the firm is willing to supply any amount of output at $p = c_{\min}$, an arbitrarily large amount of output at $p > c_{\min}$, and zero output at $p < c_{\min}$. When we think about the replication argument for constant returns to scale this makes perfect sense. Constant returns to scale implies that if you can produce 1 unit for $c_{\min}$ dollars, you can produce n units for $nc_{\min}$ dollars. Therefore you will be willing to supply any amount of output at a price equal to $c_{\min}$, and an arbitrarily large amount of output at any price greater than $c_{\min}$.

On the other hand, if $p < c_{\min}$, so that you cannot break even supplying even one unit of output, you will certainly not be able to break even supplying n units of output. Hence, for any price less than $c_{\min}$, you will want to supply zero units of output.



Constant average costs. In the case of constant average costs, the long-run supply curve will be a horizontal line.

Summary

1. The relationship between the price a firm charges and the output that it sells is known as the demand curve facing the firm. By definition, a competitive firm faces a horizontal demand curve whose height is determined by the market price—the price charged by the other firms in the market.

2. The (short-run) supply curve of a competitive firm is that portion of its (short-run) marginal cost curve that is upward sloping and lies above the average variable cost curve.

3. The change in producer's surplus when the market price changes from p_1 to p_2 is the area to the left of the marginal cost curve between p_1 and p_2. It also measures the firm's change in profits.

4. The long-run supply curve of a firm is that portion of its long-run marginal cost curve that is upward sloping and that lies above its long-run average cost curve.

REVIEW QUESTIONS

1. A firm has a cost function given by $c(y) = 10y^2 + 1000$. What is its supply curve?

2. A firm has a cost function given by $c(y) = 10y^2 + 1000$. At what output is average cost minimized?

3. If the supply curve is given by $S(p) = 100 + 20p$, what is the formula for the inverse supply curve?

4. A firm has a supply function given by $S(p) = 4p$. Its fixed costs are 100. If the price changes from 10 to 20, what is the change in its profits?

5. If the long-run cost function is $c(y) = y^2 + 1$, what is the long-run supply curve of the firm?

6. Classify each of the following as either technological or market constraints: the price of inputs, the number of other firms in the market, the quantity of output produced, and the ability to produce more given the current input levels.

7. What is the major assumption that characterizes a purely competitive market?

8. In a purely competitive market a firm's marginal revenue is always equal to what? A profit-maximizing firm in such a market will operate at what level of output?

9. If average variable costs exceed the market price, what level of output should the firm produce? What if there are no fixed costs?

10. Is it ever better for a perfectly competitive firm to produce output even though it is losing money? If so, when?

11. In a perfectly competitive market what is the relationship between the market price and the cost of production for all firms in the industry?

APPENDIX

The discussion in this chapter is very simple if you speak calculus. The profit-maximization problem is

$$\max_{y} \; py - c(y)$$

$$\text{such that } y \geq 0.$$

The necessary conditions for the optimal supply, y^*, are the first-order condition

$$p - c'(y^*) = 0$$

and the second-order condition

$$-c''(y^*) \leq 0.$$

The first-order condition says price equals marginal cost, and the second-order condition says that the marginal cost must be increasing. Of course this is presuming that $y^* > 0$. If price is less than average variable cost at y^*, it will pay the firm to produce a zero level of output. To determine the supply curve of a competitive firm, we must find all the points where the first- and second-order conditions are satisfied and compare them to each other—and to $y = 0$—and pick the one with the largest profits. That's the profit-maximizing supply.