CHAPTER 9

Profit Maximization in Perfectly Competitive Markets

By effectively implementing our strategies and delivering value, we have achieved one of the strongest growth records in the publishing industry.

EBITA (earnings before interest, taxes, and amortization of intangible assets) has increased at a compound annual rate of 26%.

The Class A Common Stock price has grown in value at a compound annual rate of 27%.

How does a firm's optimal output choice in a perfectly competitive market structure respond to changes in price and cost?

Chapter Outline

Learning Objectives

- Outline the conditions that characterize perfect competition.
- Explain why it is appropriate to assume profit maximization on the part of firms.
- Show why the fact that a competitive firm is a price taker implies that the demand curve facing the firm is perfectly horizontal.
- Explore a competitive firm's optimal output choice in the short run and how the firm's short-run supply curve may be derived through this output selection.
- Delineate how the short-run industry supply curve is determined from individual firms' short-run supply curves.
- Define the conditions characterizing long-run competitive equilibrium.
- Understand how the long-run industry supply curve describes the relationship between price and industry output over the long run, taking into account how input prices may be affected by an industry's expansion/contraction.
- Analyze the extent to which the competitive market model applies.

As we have seen, the basic determinants of cost are the prices and productivities of inputs. But a knowledge of cost conditions alone does not explain a firm's output level. Cost curves identify only the minimum cost at which the firm can produce various outputs.

For a firm interested in maximizing profit, cost and demand conditions jointly determine the optimal output level. So, to complete the model of output determination, we need to specify the demand curve confronting the firm. The demand curve determines the sales revenue at different volumes of output. In this chapter we concentrate on perfect competition and the demand curve facing a firm operating in such a market structure. Later chapters focus on the demand curve confronting a firm and the firm's optimal output when competition is imperfect.

We will explore how a firm's optimal output within a perfectly competitive market structure responds to changes in price and cost. This information can be used to derive the firm's supply curve and, in turn, the industry supply curve. We also address the long-run outcome in perfect competition and contrast it with short-run responses.

9.1 THE ASSUMPTIONS OF PERFECT COMPETITION

In common usage, *competition* refers to intense rivalry among businesses. Microsoft and Sun Microsystems, Nike and Reebok, Pepsi and Coke are all competitors in this sense. Each firm makes a business decision—whether to introduce a new product, advertise existing products more forcefully, or enhance product quality—only after considering the effect on competitors and their likely response.

The economist's formal model of perfect competition bears little resemblance to this picture. **Perfect competition** is distinguished largely by its impersonal nature. More specifically, four conditions characterize a perfectly competitive industry.

PERFECT COMPETITION
an economic model characterized by the assumption of (1) a large number of buyers and sellers, (2) free entry and exit, (3) product homogeneity, and (4) perfect information

FREE ENTRY AND EXIT
a situation in which there are no differential impediments across firms in the mobility of resources into and out of an industry

HOMOGENEOUS PRODUCTS
standardized products that, in the eyes of consumers, are perfect substitutes for one another

1. *Large numbers of buyers and sellers*. The presence of a great many independent participants on each side of the market, none of whom is large in relation to total industry sales or purchases, normally guarantees that individual participants' actions will not significantly affect the market price and overall industry output. In a market with many firms, each firm recognizes that its impact on the overall market is negligible, and consequently does not view other firms as personal rivals.

2. *Free entry and exit*. Industry adjustments to changing market conditions are always accompanied by resources entering or leaving the industry. As an industry expands, it uses more labor, capital, and so on; resources enter the industry. Similarly, resources leave a contracting industry. A perfectly competitive market requires that there be no differential impediments across firms in the mobility of resources into, around, and out of an industry. This condition is sometimes called **free entry and exit**. Examples of barriers to entry and exit include an incumbent firm with an exclusive government patent or operating license and economies of scale that impede the entry of new firms.

3. *Product homogeneity*. All the firms in the industry must be producing a standardized or **homogeneous product**. In consumers' eyes, the goods produced by the industry's firms are perfect substitutes for one another. This assumption allows us to add the outputs of the separate firms and talk meaningfully about the industry and its total output. It also contributes to the establishment of a uniform price for the product. One farmer will be unable to sell corn for a higher price than another if the products are viewed as interchangeable, because consumers will always purchase from the lower-priced source.

4. *Perfect information*. Firms, consumers, and resource owners must have all the information necessary to make the correct economic decisions. For firms, for example, the relevant information is knowledge of the production technology, input prices, and the price at which the product can be sold. For consumers, the relevant information is a knowledge of their own preferences and the prices of the various goods of interest to them. Moreover, the consumers, in their role as suppliers of inputs, must know the remuneration they can receive for supplying productive services.

Probably no industry completely satisfies all four conditions. Agricultural markets come close, although government involvement in such markets keeps them from fully satisfying the four conditions. Most industries satisfy some conditions well but not others. Even though the number of market participants in the gasoline retailing business is large and entry into the business is fairly easy, for example, not all gasoline brands are the same. Some brands have higher octane and more detergents, and are better for the environment. Certain stations are closer to particular consumers and thereby more convenient, or offer better complements such as full service, food-marts, and pumps that allow customers to pay for their purchases by inserting a credit card. Moreover, consumers are rarely perfectly informed about the prices all retailers are charging.

The fact that only a few industries may fully satisfy the four conditions does not mean that the study of perfect competition is unwarranted. Many industries come close enough to satisfying the four conditions to make the perfectly competitive model quite useful. Take the case of gasoline retailing. Although product homogeneity and perfect information may not fully apply, the extent to which an individual gas station has some choice over what price to charge per gallon is probably limited to a very narrow band of just a few cents. Such a narrow pricing power band is pretty close to having no significant impact over price, as predicted by the competitive model.

9.2 PROFIT MAXIMIZATION

PROFIT MAXIMIZATION
the assumption that firms select an output level so as to maximize profit

In perfectly (and imperfectly) competitive markets, is it appropriate to assume **profit maximization** on the part of firms? At the outset we should recognize that any profit realized by a business belongs to the business owner(s). For the millions of small businesses with only one owner-manager, decisions concerning what products to carry, whom to employ, what price to charge, and so on, will be heavily influenced by the way the owner's profit is affected. Owners of such businesses may well have goals such as early retirement or expensive educations for their children. These goals, however, are not inconsistent with the assumption of profit maximization. Since money is a means to many ends, early retirement or college educations can more easily be afforded when the owner makes more money.

A possible problem with assuming profit maximization is that the owner-manager cannot have detailed knowledge of the cost and revenue associated with each action that could be taken to maximize profit. Economic theory, however, does not require that firms actually know or think in terms of marginal cost and revenue, only that they behave as if they did.

Firms may come close enough to maximizing profit by trial and error, emulation of successful firms, following rules of thumb, or blind luck for the assumption to be a fruitful one.

When we move from the small, owner-managed firm to the large, modern corporation, another potential criticism of the profit maximization assumption arises. A characteristic of most large corporations is that the stockholder-owners themselves do not make the day-to-day decisions about price, employment, advertising, and so on. Instead, salaried personnel of the corporation—managers—make these decisions. And so there is a separation of ownership and control in the corporation; managers control the firm, but stockholders own it. It is safe to assume that stockholders wish to make as much money on their investment as possible, but it is virtually impossible for them to constantly monitor their managers' actions. Therefore, managers will have some discretion, and some of their decisions may conflict with the stockholder-owners' profit-maximizing goals.

While managers may have some discretion to deviate from the profit-maximizing goals of firms' shareholder-owners, several factors limit the exercise of such discretion. For example, stockholder-owners often link business managers' compensation to profits, sometimes paying them in part with shares of stock or stock options, in order to give an incentive to pursue profits more actively. In addition, the profitability managers achieve in a given enterprise will affect their job prospects with others. And, finally, if managers do not make as large a profit as possible, stock prices, which tend to reflect profitability (especially projected profitability), will be lower than need be. Undervalued stock creates an incentive for outsiders, or "raiders," to buy up a controlling interest in the firm and replace the inefficient management team. A firm that neglects profit opportunities too often leaves itself open to such a takeover bid, a fairly common occurrence in the corporate world.

Operating in a competitive market provides yet another reason it is safe to assume that a firm will pursue profit maximization as a goal. Consider firms in the hospitality industry, such as Hilton and Marriott. Suppose some of them virtually ignore profit, either through ignorance, negligence, bad luck, or intention. Their cost of labor is too high, they fail to minimize waste in their food service, they neglect training for the staff who serve customers, and so forth. Other firms, whether through superior management, close attention to costs, or the good luck of being closest to a newly enlarged convention center, produce the right type of lodging in the appropriate quantity and in the least costly way. What will happen when these firms compete for customers? Clearly, the firms that come closer to maximizing profit will make money and prosper; the others will suffer losses. In general, in competitive markets, firms that do not approximate profit-maximizing behavior fail; the survivors will be the firms that, intentionally or not, make the appropriate, profit-maximizing decisions. This observation is called the **survivor principle** and provides a practical defense for the assumption of profit maximization.

SURVIVOR PRINCIPLE
the observation that in competitive markets, firms that do not approximate profit-maximizing behavior fail, and that survivors are those firms that, intentionally or not, make the appropriate profit-maximizing decisions

While it goes without saying that firms must pay close attention to profit, some deviation from single-minded profit maximization may still occur. Most economists, however, believe that the profit maximization assumption provides a close enough approximation to be useful in analyzing many problems, and it has become the standard assumption regarding the firm's behavior. While the assumption may not adequately explain why Billy, the company president, hires his ne'er-do-well brother-in-law, Roger, or why RJR Nabisco appears to have an excessively large fleet of corporate jets, it does not pretend to try. Instead, it is designed to explain how a firm's output will respond to a higher or a lower price, a tax or a government regulation, a cost change, and so on. Recall that the ultimate test of a theory is whether it explains and predicts well, and theories based on the assumption of profit maximization have passed that test.

Application 9.1 Are American Executives Underpaid?

In recent years, the pay received by senior business executives in the United States has grown much more rapidly (in percentage terms) than the average worker's pay.[1] Moreover, counting bonuses and stock options, the pay received by some executives has been truly spectacular. For example, Oracle's CEO Larry Ellison and Microsoft's Bill Gates earned billions of dollars from their stock options in the 1990s.

Although the high level and rate of increase in senior executive compensation have received much criticism in the media and the political arena, leading academic studies suggest that the monetary incentives offered by firms to their managers may still be inadequate. A study by Michael Jensen and Kevin Murphy of the Harvard Business School, for instance, finds that CEO wealth rises by just $3.25 for each $1,000 increase in shareholder wealth. Jensen and Murphy argue that their estimated sensitivity figures are too small to entirely solve the **principal–agent problem** confronting a firm's stockholders: namely, stockholders' agents (the managers or principals of the firm) do not have sufficient incentive to maximize the firm's profit if they receive only $3.25 per $1,000 of stockholder wealth that is generated. Based on the Jensen-Murphy results, a CEO would find it worthwhile to use $1 million of the shareholders' potential wealth on a private jet or rare artwork if such an item generated $3,250 in purely personal pleasure for the CEO. Jensen and Murphy propose that CEOs' pay should be tied more closely to stockholders' wealth to improve corporate performance, but they acknowledge that opposition in the media and the political arena may limit such a change.

PRINCIPAL–AGENT PROBLEM
a situation in which a principal's agent (such as the manager of a firm representing the firm's stockholders) may have an incentive that is not entirely consonant with the interests of the principal

Support for Jensen and Murphy's proposal is provided by a study by John Abowd of Cornell. Abowd examined 16,000 executives in 250 large U.S. corporations over 1981–1986. He found that increases in the sensitivity of executive pay to corporate performance improved corporate performance in terms of the rate of return earned by a firm's stockholders. In conjunction with the Jensen-Murphy results, Abowd's findings suggest that linking managerial compensation more closely to corporate performance can increase shareholder wealth. The existing link between managerial compensation and corporate performance, that is, is not yet sufficiently strong to completely eliminate the principal-agent problem confronted by the stockholders of U.S. firms.

[1]This application is based on Paul Milgrom and John Roberts, *Economics, Organization and Management* (Englewood Cliffs, NJ: Prentice Hall, 1992); Michael C. Jensen and Kevin J. Murphy, "Performance Pay and Top-Management Incentives," *Journal of Political Economy*, 98 No. 2 (April 1990), pp. 225–264; and John M. Abowd, "Does Performance-Based Managerial Compensation Affect Corporate Performance?" *Industrial and Labor Relations Review*, 43 No. 1 (February 1990), pp. 52S–73S.

9.3 The Demand Curve Facing the Competitive Firm

Assuming that firms are interested in maximizing profit, let's examine the implications of this assumption in a competitive market setting. Since a competitive market is characterized by a large number of firms selling the same product, each firm supplies only a small fraction of the entire industry output. For example, one farmer may account for only one-millionth of the entire corn industry's output or one building contractor may supply 1 percent of the total construction services in a particular city.

The nature of the demand curve confronting a single competitive firm follows directly from the relatively insignificant contribution its output makes toward total supply. The product's price is determined by the interaction of the market supply of and demand for the product. Because each firm produces such a small portion of the total supply, its output

PRICE TAKER
a firm that takes prices as given and does not expect its output decisions to affect price

AVERAGE REVENUE (AR)
total revenue divided by output

decisions have a small effect on the market price. For simplicity the "small" effect is taken to be a "zero" effect, and the demand curve facing a competitive firm is drawn to be perfectly horizontal. A horizontal demand curve means the firm can sell as much output as it wants without affecting the product's price. Stated differently, a competitive firm is a **price taker**: the firm takes the price as given and does not expect its output decisions to affect price.

Figure 9.1 clarifies why the firm's demand curve is drawn horizontally. In Figure 9.1b, the total output of corn is measured horizontally, and the per-unit price is measured vertically. The market demand curve is shown as *D*. The premise is that the total quantity offered for sale by all farms together interacts with this demand curve to determine price. If the combined output of all corn farms is 15 billion bushels, the market price per bushel is $3.

Assume that the Costner farm is one of 1 million identical farms supplying corn, and it is currently selling 15,000 bushels for $3 each, as shown in Figure 9.1a. The farm's demand curve is drawn horizontally as *d* in Figure 9.1a because its output variations will not have an appreciable effect on the market output and price. That is, if the Costner farm produced 15,000 fewer (an output of zero bushels) or more bushels (an output of 30,000 bushels), it would have an insignificant influence on the total industry output of 15 billion bushels. Consequently, the market price will not be altered by the output actions of this single farm. It is stuck with having to charge the market price of $3 no matter what output level it selects.

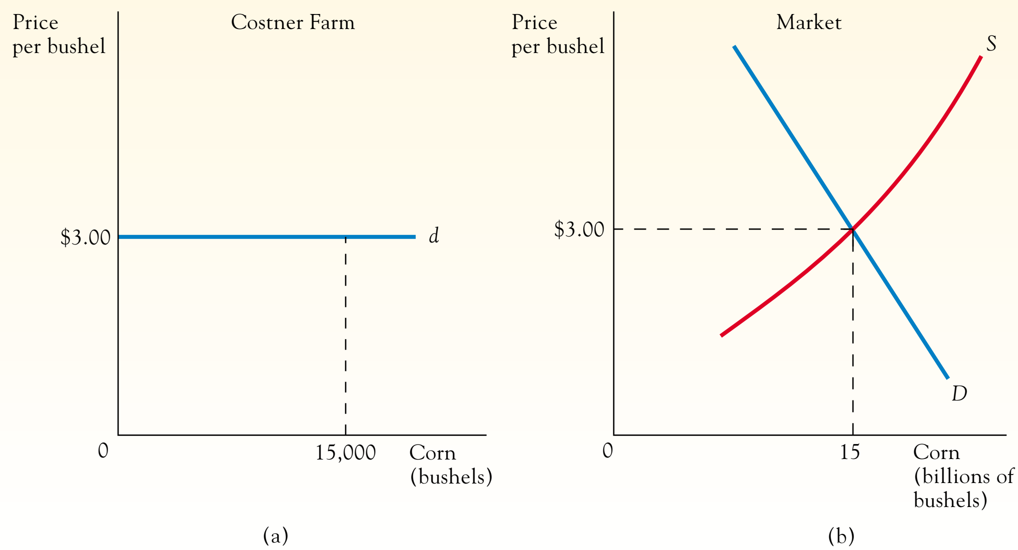
A horizontal demand curve has an elasticity of infinity. Since there are many homogeneous substitutes for any farm's output and customers are perfectly informed in a perfectly competitive market, the quantity of corn demanded from the Costner farm equals zero if the farm attempts to charge even a penny over the prevailing market price.

A firm's **average revenue** (*AR*), or total revenue divided by output, is the same as the prevailing market price. As Figure 9.1 shows, if the market price is $3, then the Costner farm will on average make $3 per bushel.

FIGURE 9.1

The Competitive Farm's Demand Curve

(a) Because an individual farm supplies only a small portion of the market output, its demand curve, *d*, is perfectly elastic. (b) The interaction of market supply, *S*, and demand, *D*, determine the prevailing market price ($3) and output (15 billion bushels).



MARGINAL REVENUE (MR)
the change in total revenue when there is a one-unit change in output

When a firm faces a horizontal demand curve, the market price also equals the firm's marginal revenue. **Marginal revenue** (*MR*) is defined as the change in total revenue when there is a one-unit change in output. A firm in a competitive market can sell one more unit of output without reducing the price it receives for its previous units, so total revenue will rise by an amount equal to the price. For example, if a farm is selling 15,000 bushels of corn at a price of $3 per bushel, total revenue is $45,000. If the farm sells 15,001 bushels at a price of $3 per bushel, as it can with a horizontal demand curve, total revenue rises from $45,000 to $45,003, or by $3. Once again the familiar average-marginal relationship can be seen to apply. Where the average revenue (*AR*) is flat or constant, the marginal revenue equals the average revenue (*MR* = *AR*).

Note that the assumption of a horizontal demand curve confronting a competitive firm does not mean that the price never changes. It just means that the firm, acting by itself, cannot affect the going price. The market price may vary from time to time due to changes in consumers' incomes, technology, consumers' preferences, and so on, but not because of changes in the amount sold by a particular firm.

9.4 SHORT-RUN PROFIT MAXIMIZATION[2]

TOTAL REVENUE (TR)
price times the quantity sold

In the short run, a competitive firm operating with a fixed plant can vary its output by altering its employment of variable inputs. To see how a profit-maximizing firm decides on what level of output to produce, let's begin with a numerical illustration. The information in Table 9.1 allows us to identify the output (*q*) that maximizes profit for a hypothetical competitive firm; it includes the short-run cost of production and revenue from the sale of output. We know the firm is selling in a competitive market because the price (*P*) is constant at $12 regardless of the output level. Note that **total revenue** (*TR*) is equal to price times the quantity sold, and it rises in proportion to output since the price is constant. Total cost (*TC*) rises with output in the familiar fashion, slowly at first and then more rapidly as the plant becomes more fully utilized and marginal cost rises. Total cost, when output is zero, is $15, reflecting the total fixed cost.

TABLE 9.1

SHORT-RUN COST AND REVENUE OF A COMPETITIVE FIRM (IN DOLLARS)

q	*P*	*TR*	*TC*	*TVC*	*ATC*	*AVC*	π	π/q	*MC*	*MR*	
0	12	0	15	0	—	—	−15	—	—	—	
1	12	12	25	10	25	10	−13	−13	10	12	
2	12	24	33	18	16.50	9	−9	−4.50	8	12	
3	12	36	40	25	13.30	8.30	−4	−1.30	7	12	
4	12	48	46	31	11.50	7.80	2	0.50	6	12	MC < MR
5	12	60	54	39	10.80	7.80	6	1.20	8	12	
6	12	72	63	48	10.50	6	9	1.50	9	12	
7	12	84	73	58	10.40	8.30	11	1.57	10	12	
8	12	96	84.90	69.90	10.61	8.70	11.10	1.39	11.90	12	MC ≈ MR
9	12	108	98	83	10.90	9.20	10	1.25	13.10	12	
10	12	120	113	98	11.30	9.80	7	0.70	15	12	MC > MR
11	12	132	132	117	12	10.60	0	0	19	12	

[2]A mathematical treatment of some of the material in this section is given in the appendix at the back of the book (pages xxx–xxx).

TOTAL PROFIT (π)
the difference between total revenue and total cost

AVERAGE PROFIT PER UNIT (π/q)
total profit divided by number of units sold

Total profit (π) is the difference between total revenue and total cost. At low and high rates of output, profit is negative; that is, the firm would suffer losses. In particular, note that the firm loses \$15 if it produces no output at all because it still must pay its fixed cost when it shuts down. At an intermediate rate of output in this example, profit is positive. The firm, however, wishes to make as large a profit as possible, and maximum profit occurs at an output of eight units where profit equals \$11.10.

Note that maximizing total profit is generally not the same thing as maximizing **average profit per unit (π/q)** sold. The firm's goal is to maximize its total profit, and that is achieved at an output of 8 units. Profit per unit at that output is \$1.39, but it could have an even higher average profit, \$1.57, by producing just 7 units. Total profit is profit per unit times the number of units sold, so a lower average profit can correspond to a higher total profit if enough additional units are sold, as is true in the example of Table 9.1.

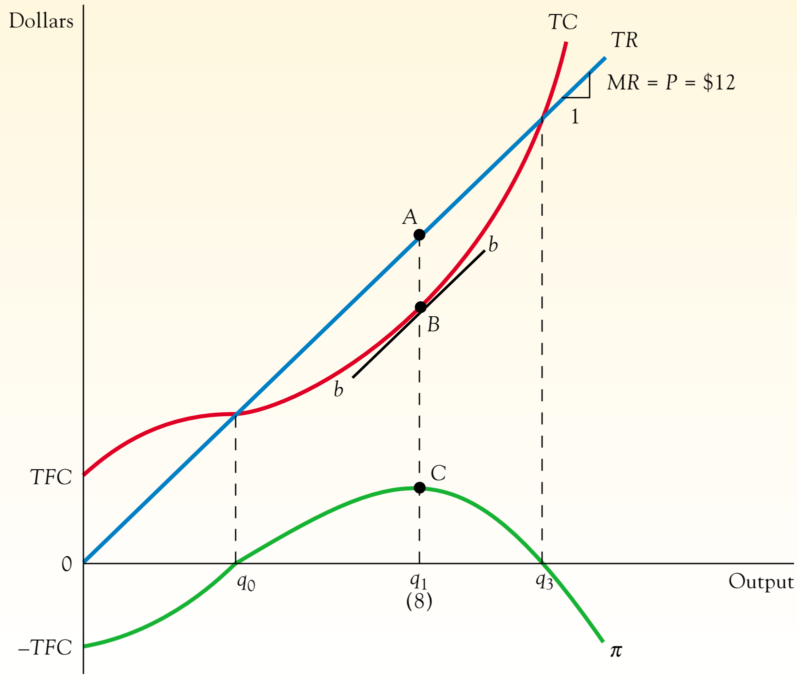
Figure 9.2 shows how we identify the most profitable level of output by using the total revenue and total cost curves. The total revenue curve is a new relationship, but it is a relatively simple one when we are dealing with a competitive firm. With the price per unit constant, total revenue rises in proportion to output and is, therefore, drawn as a straight line emanating from the origin. Its slope, showing how much total revenue rises when output changes by one unit, is marginal revenue.

In terms of Figure 9.2, the firm wishes to select the output level where total revenue exceeds total cost by the largest possible amount—that is, where profit is greatest. This situation occurs at output q_1, where total revenue, Aq_1, exceeds total cost, Bq_1, by AB. The vertical distance AB is total profit at q_1. At lower and higher output levels, total profit is lower than AB. Note that at a lower output level, q_0, for example, the TR and TC curves are diverging (becoming farther apart) as output rises, indicating that profit is greater at a higher output. This reflects the fact that marginal revenue (the slope of TR) is greater than marginal cost (the slope of TC) over this range. At q_1, when the curves are farthest apart (with

FIGURE 9.2

Short-Run Profit Maximization: Total Curves

Profit is maximized at the output where total revenue (TR) exceeds total cost (TC) by the largest possible amount. This occurs at output q_1, where profit is equal to AB. The total profit curve (π), which plots total profit explicitly at each rate of output, also shows the point of profit maximization.



revenue above cost), the slopes of *TR* and *TC* (the slope of *TC* at *B* is equal to the slope of *bb*) are equal, reflecting an equality between marginal revenue and marginal cost.

In Figure 9.2, the level of total profit at each rate of output is also shown explicitly by the total profit curve (π), which reaches a maximum at q_1. The total profit curve is derived graphically by plotting the difference between *TR* and *TC* at each output level. For example, *AB* is equal to Cq_1; alternatively, when output is zero, profit is negative and equal to minus total fixed cost (*TFC*).

Short-Run Profit Maximization Using Per-Unit Curves

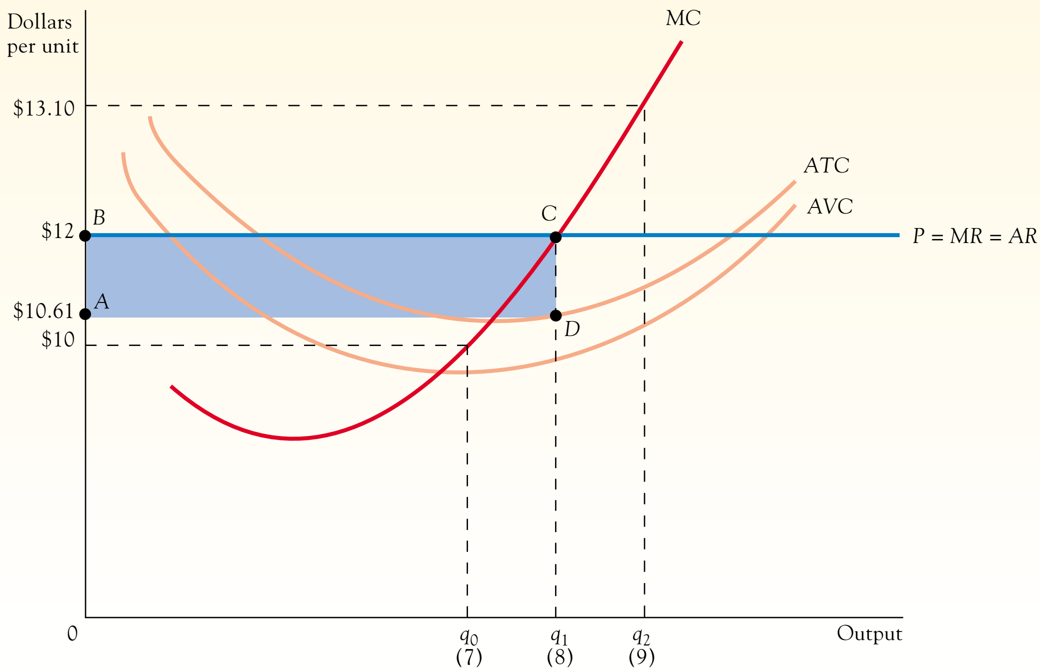
Figure 9.3 presents the same information shown in Figure 9.2, but now we use the familiar per-unit cost and revenue relationships. With the vertical axis measuring dollars per unit of output, the firm's demand curve is shown as a horizontal line, because the firm may sell any number of units at the $12 price. The figure also shows the average total cost (*ATC*), average variable cost (*AVC*), and marginal cost (*MC*) curves. The most profitable output level occurs where marginal cost and marginal revenue are equal, at q_1 in the figure (eight units in Table 9.1). The shaded rectangle *BCDA* shows total profit for that output. The height of the rectangle, *CD*, is average revenue ($12) minus average total cost ($10.61 ≈ $84.9/8), or the average profit per unit of output ($1.39 = $12 − $10.61). Multiplying the average profit per unit by the number of units sold (the length of the rectangle) yields total profit ($11.10).

Understanding why the firm's profit is at a maximum where marginal cost and marginal revenue are equal is very important. Consider what it would mean if the firm were operating

FIGURE 9.3

Short-Run Profit Maximization: Per-Unit Curves

In terms of the per-unit curves, the output that maximizes profit in the short run is where *MC* = *MR*, or q_1. At lower levels of output such as q_0, *MC* < *MR*, and the firm can increase profit by expanding output. At higher outputs like q_2, *MC* > *MR*, and the firm can increase profit by reducing output.



at a lower output where MC ≤ *MR*, such as q_0 (7 units) in Figure 9.3. At q_0, marginal revenue is still $12, but marginal cost is lower ($10). Consequently, the seventh unit of output adds $12 to revenue but increases total cost by only $10; with revenue increasing more than cost, profit (the difference) will rise from $9 (at 6 units) to $11 when the seventh unit is sold. At any rate of output where marginal revenue is greater than marginal cost, the firm can increase its profit by increasing output. As output expands, marginal cost rises, so the addition to profit from each successive unit becomes smaller (but is still positive) until q_1 is reached, where MC = *MR*.

At output levels beyond q_1 the firm would be producing too much output. At q_2 (9 units in Table 9.1), for example, marginal cost is $13.10 and is greater than marginal revenue ($12). Thus, total profit could be increased by reducing output. If the firm produces 1 unit less, it loses $12 in revenue. But cost falls by $13.10, so the net effect is a $1.10 increase in profit. Thus, the firm's profit will increase by decreasing output if *MC* > *MR*.

In summary, the rule for profit maximization is to produce where MC = *MR*. Because marginal revenue equals price for a competitive firm, we can also express this condition as MC = *P*. This rule does *not* mean that the firm intentionally sets price equal to marginal cost since, for a competitive firm, price is given and beyond its control. Instead, the firm will adjust its production until the marginal cost is brought into equality with price.[3]

APPLICATION 9.2 THE IMPORTANCE OF THINKING AT THE RELEVANT MARGIN

Learning to think at the margin is one of the important lessons economists strive to teach.[4] A corollary of that lesson that often does not get the attention it merits, is that the relevant margin over which revenues (benefits) and costs should be compared depends on the nature of the activity involved. The relevant margin also depends on the costs associated with actually calculating marginal revenue and marginal cost as one strives to measure production in ever finer units. For example, for the typical corn farmer, the relevant margin may involve units of 1,000 bushels and comparing the marginal revenue and marginal cost associated with expanding output by 1,000 bushels. Making a similar comparison for each 1-bushel increment may be computationally too expensive an endeavor.

To drive home the importance of focusing on the relevant margin, economist Gary Galles of Pepperdine University likes to tell his students the following story:

> Suppose you are walking through Central Park at two in the morning and you have $200 in your wallet. Suppose further that a mugger pulls a knife on you and tells you to give him all your money. How do you respond? Do you say, "I'll tell you what: I'll give you $200 if you leave me alone; $150 if you rough me up a little bit; $100 if you cut me a little; $50 if you put me in the hospital; but nothing if you kill me?" Of course not, because you would be focusing on the wrong margin for decision-making; you are faced with an all-or-none decision, not one in which marginal adjustments can be made (that is, the relevant marginal decision is the all-or-none decision of how to respond to the mugger—fight or hand over the money. Once that decision is made, marginal ones

[4]This application is based on *Great Ideas for Teaching Economics*, 3rd ed., by Ralph T. Byrns and Gerald W. Stone (Glenview, IL: Scott, Foresman and Company, 1987); and Thomas L. Friedman, *The Lexus and the Olive Tree* (New York: Farrar, Straus and Giroux, 1999).

[3]Because the MC curve is U-shaped, MC may equal *P* at two different output levels. In this case the lower level is not the profit-maximizing output; in fact, it is the minimum-profit (or maximum-loss) output. If it is necessary to distinguish these two outputs, the profit-maximizing output is where MC = *P and* MC cuts *P* from below (that is, MC is rising).

may follow, like how hard to fight or how much money to try to withhold.)

The relevant margins, can, of course, change over time for certain activities. Take the case of farming where it is now possible to attach microchip-based sensors to tractors that are capable of measuring the exact amount of a crop coming off every foot of a farm. The sensors, furthermore, are connected via a transmitter to a global positioning satellite (GPS) system in outer space that allows a farmer to precisely track the output from every location on a farm. The benefits associated with being able to more finely specify the relevant marginal units using technological improvements can be substantial. As noted by Pulitzer-Prize-winning reporter Thomas Friedman, in recounting the story of one so-called precision farmer, Gary Wagner of Minnesota:

> "What I found out was something of a surprise," said Wagner. "It was a common belief that yield did not vary much from one area of your field to another. You looked out at your field and it all looked uniform to the eye. But once we generated a price yield map from this program, we discovered that there was a major difference in yield between some acres and others, as much as $150 an acre, which can be the difference between a profit and a loss on that acre. Once I had that information it was worth its weight in gold to me, because we have a choice every season of different varieties of crops to plant. By using these sorts of computerized yield monitors we can tell exactly which varieties of crops grow best in which soils under our management." [Wagner] could also program . . . all the [relevant] information [soil, moisture, slope of land] into his fertilizer applicator, and then tie that applicator in with the GPS system. Once he had that system together, he could drive down his sugar-beet field and the GPS system would know which acre he was on, the software program would know just how much fertilizer the specific acre liked and the fertilizer applicator would automatically dispense the exact amount of nitrates—more in some places less in others—demanded by that specific area. It saved on fertilizer, which was good for the environment, and maximized his yields, which was good for his pocketbook. "Instead of having to work with information that was based on averages for the region for the average farmer, we were able to tailor everything to ourselves [at the appropriate margin]," said Wagner.

Operating at a Loss in the Short Run

A firm in a perfectly competitive market may find itself in the unenviable situation of suffering a loss no matter what output level it produces. In that event the firm has two alternatives: it can continue to operate at a loss, or it can shut down. (Recall that we are dealing with a short-run setting; halting production may be only a temporary move until market conditions improve and is not necessarily the same as going out of business.) Yet shutting down will not avoid a loss since the firm remains liable for its fixed cost whether or not it operates. The relevant question is whether the firm will lose *less* by continuing to operate or by shutting down.

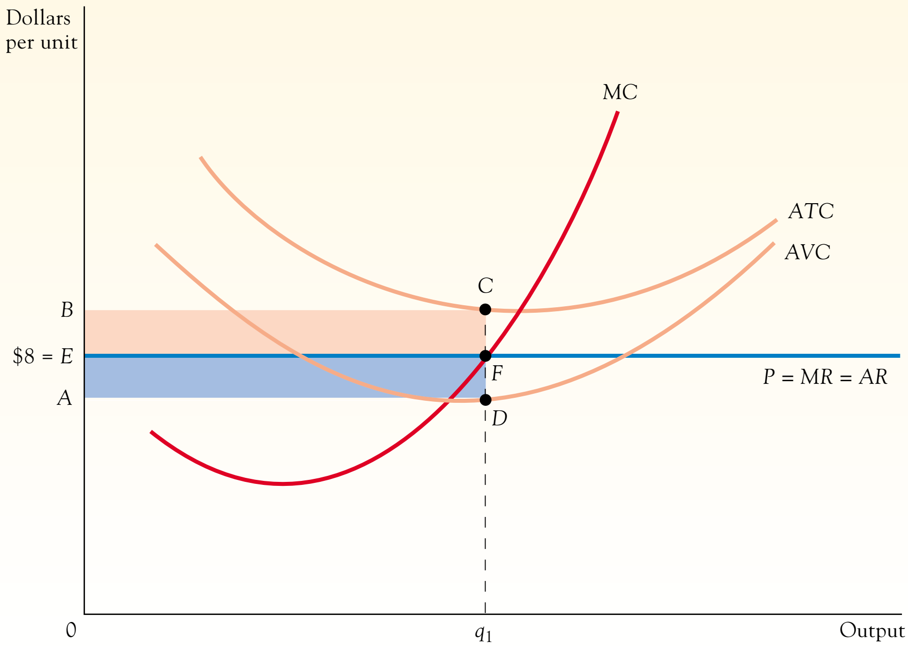
Figure 9.4 illustrates the case where the firm's best option is to operate at a loss rather than shut down. With the price at $8 per unit and the cost curves as shown, the firm's "most profitable" output (which can also mean its "least unprofitable" output, as it does here) is the point where price equals marginal cost at q_1. Because the average total cost curve lies above the price line everywhere, the firm incurs losses at all output levels. But at q_1 the loss is the smallest, as indicated by the rectangle *BCFE*. The height of this rectangle, *CF*, is the difference between average cost and average revenue, or the "negative profit margin" per unit, and the length of the rectangle, *EF*, is the number of units sold at a loss.

How do we know that the firm loses less by producing at q_1 than by shutting down? Consider the larger rectangle *BCDA*. The height of the rectangle, *CD*, is the difference between *ATC* and *AVC* at q_1. Thus, *CD* measures average fixed cost. Recall that $AVC + AFC = ATC$; therefore, $ATC - AVC = AFC$, and average fixed cost multiplied by the number of units produced (the length of the rectangle) equals total fixed cost (or the area of rectangle

FIGURE 9.4

Operating at a Loss in the Short Run

In the short run, a firm may continue to produce even though its best output yields a loss. As long as average revenue covers average variable cost, it is in the firm's interest to continue operating. Output will be q_1 even though the firm's loss is *BCFE*, because the loss would be even greater—*BCDA*—if it shut down.



BCDA). Even if the firm shuts down, it will still incur a loss, namely, total fixed cost. But because that loss (*BCDA*) is larger than the loss (*BCFE*) incurred if the firm continues to operate, the firm loses less by producing q_1 than by shutting down. It is better to operate and lose $100 per week than to shut down and lose $200 per week.

If the price falls sufficiently, however, the firm may lose less by shutting down, as we will see in the following section. Moreover, even when it is in the firm's interest to produce at a loss, as in Figure 9.4, this equilibrium can be only temporary (short run). If the price remains at $8, the firm will ultimately go out of business. The point here is that a firm will not immediately liquidate its assets the moment it begins to suffer losses.

9.5 THE PERFECTLY COMPETITIVE FIRM'S SHORT-RUN SUPPLY CURVE

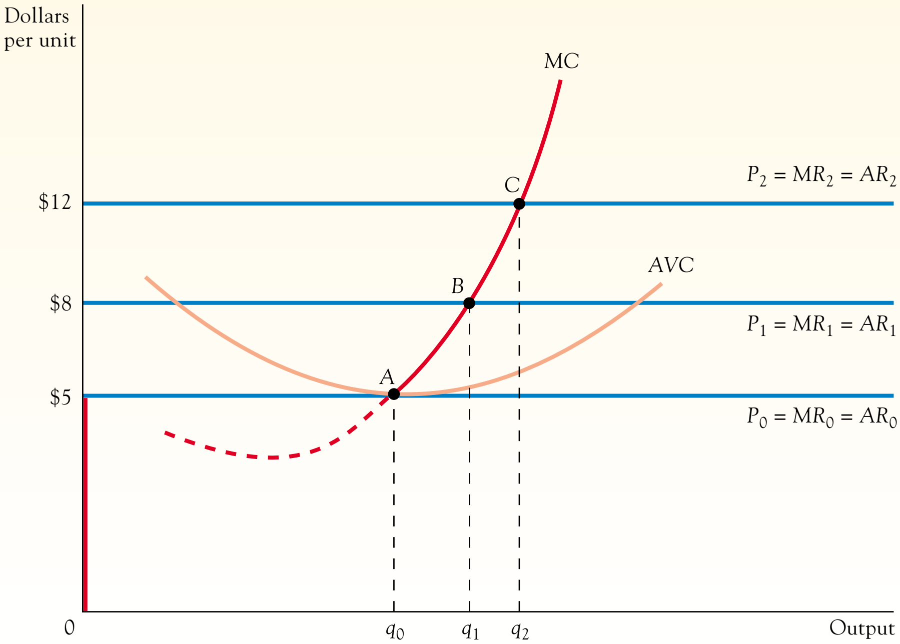
SHORT-RUN FIRM SUPPLY CURVE
a graph of the systematic relationship between a product's price and a firm's most profitable output level

The previous discussion implies a systematic relationship between a product's price and a firm's most profitable output. We can analyze this relationship to derive the **short-run firm supply curve** in a perfectly competitive industry. Figure 9.5 depicts the firm's average variable and marginal cost curves. Note that the average total cost curve is not shown because it is not needed to identify the most profitable output level. The *ATC* curve is primarily useful for showing total profit or loss. If the per-unit price is $8, the firm will produce at point *B*, where marginal cost equals $8, so output will be q_1. If the price rises from $8 to $12, then the firm's profit would change, as would the most profitable output. Suppose that the firm maintains an output of q_1 when the price rises to $12. Its profit will increase because revenue

FIGURE 9.5

The Competitive Firm's Short-Run Supply Curve

The part of the *MC* curve lying above the minimum point on the *AVC* curve identifies the firm's most profitable output at prices above \$5. Below a price of \$5, the firm will be better off shutting down. Above a price of \$5, the firm maximizes profit by producing where price equals marginal cost. Because the *MC* curve slopes upward, the firm increases output at a higher price. The solid, dark red curve represents the competitive firm's short-run supply curve.



rises, but its cost will not change. The firm, however, can increase profit by increasing output. At q_1, marginal revenue (\$12) now exceeds marginal cost (\$8), signaling that an output increase will add more to revenue than to cost. The new profit-maximizing output occurs where output has expanded until marginal cost equals the \$12 price, at point C on the MC curve. Thus, a higher price gives the firm an incentive to expand output from q_1 to q_2.

The MC = P (= MR) rule for maximizing profit in a competitive market structure therefore implies that a firm will produce more at a higher price because increased production becomes profitable at a higher price. In fact, we can think of the MC curve as the firm's supply curve in the short run, since it identifies the most profitable output for each possible price. For example, at a price of \$8 output is q_1 at point *B* on MC, at a price of \$12 output is q_2 at point C on MC, and so on.

SHUTDOWN POINT the minimum level of average variable cost below which the firm will cease operations

One important qualification to this proposition should be mentioned. If price is too low, the firm will shut down. At a price of \$5 the firm can just cover its variable cost by producing at point A, where average variable cost equals the price. At this point the firm would be operating at a short-run loss; in fact, the loss would be exactly equal to its total fixed cost, because revenue just covers variable cost, leaving nothing to set against fixed cost. At a price below \$5 the firm is unable to cover its variable cost and would find it best to shut down. Thus, point A, the minimum level of average variable cost, is effectively the **shutdown point**: if price falls below that level, the firm will cease operations.

As a consequence of this qualification, only the segment of the marginal cost curve that lies above the point of minimum average variable cost is relevant. Stated differently, the marginal cost curve above point A identifies the firm's output at prices above $5; at any lower price output will be zero.

APPLICATION 9.3 THE METHOD TO MOTHBALLING

Average variable cost curves tend to differ in height across an industry's firms as well as across the various operating units of some firms. Thus, we should expect to see a particular sequential pattern in which firms or operating units are shut down as the prevailing price in an industry falls. Namely, firms or operating units having higher minimum levels of average variable cost will be taken out of operations first as the prevailing price in the industry declines. For example, a fall in crude oil prices typically first results in the shutdown of "stripper" wells—wells that are relatively more costly to operate because they produce less than 10 barrels a day. A decline in fares in the domestic airline industry, as occurred in the wake of the terrorist acts of September 11, 2001, led to airline companies mothballing older, less fuel-efficient planes first. The air parks of Arizona, where the desert climate is conducive to the storage of aircraft, began to fill up disproportionately in late 2001 with older models (four-engine 747s and three-engine 727s) as opposed to newer types of aircraft (two-engine 767s, 757s, and MD80s).

The pattern also works in reverse when the price in an industry increases. That is, firms or operating units having higher minimum levels of average variable cost are brought out of storage last. Consider the oil tanker market following Iraq's invasion of Kuwait in 1990. The invasion resulted in an upward spike in crude oil prices that in turn led to a sharp increase in oil tanker charter rates. As the charter rate or price that could be earned from operating a tanker increased, owners first brought newer, more fuel-efficient tankers—very large or ultra-large diesel-powered tankers—out of dry dock (storage) as opposed to older, usually smaller, steam-powered tankers.

APPLICATION 9.4 THE COMPETITIVE FIRM'S SUPPLY CURVE IN THE VERY SHORT RUN

We have been examining cases in which a competitive firm can vary its output in the short run by altering its employment of variable inputs. What if, in the very short run, it is not possible for the firm to alter its use of any inputs? What does the competitive firm's supply curve look like under such a scenario? For example, consider a producer of shovels immediately following the discovery of gold in 1848 at Sutter's Mill near then sparsely populated Sacramento, California. The discovery sparked a tremendous rush to mine the precious metal and dramatically increased the demand for and price of such mining implements as shovels. A shovel, whose nominal price had been only $1 in the Sacramento area prior to the gold discovery, sold for as much as $50 immediately following the discovery.

In the days just after the gold discovery, a supplier of shovels to the Sacramento area could not appreciably alter its output even though the demand for and price of shovels sharply increased. In 1848, it took more than a few days for the manufacturer to hear news of the gold discovery, manufacture additional shovels, and then ship the additional shovels to the Sacramento area (the relevant communication, manufacturing, and transportation times were much greater 150 years ago).

In a very-short-run scenario when no inputs can be altered, the firm's marginal cost curve is effectively vertical above the shutdown point (not upward-sloping, as it is in Figure 9.5 above point A). With a vertical marginal cost curve, no matter how high the price rises for the product, the firm cannot alter its output. Thus, the firm's supply curve is also vertical or perfectly price inelastic in a very-short-run setting where the firm cannot alter the use of its inputs.

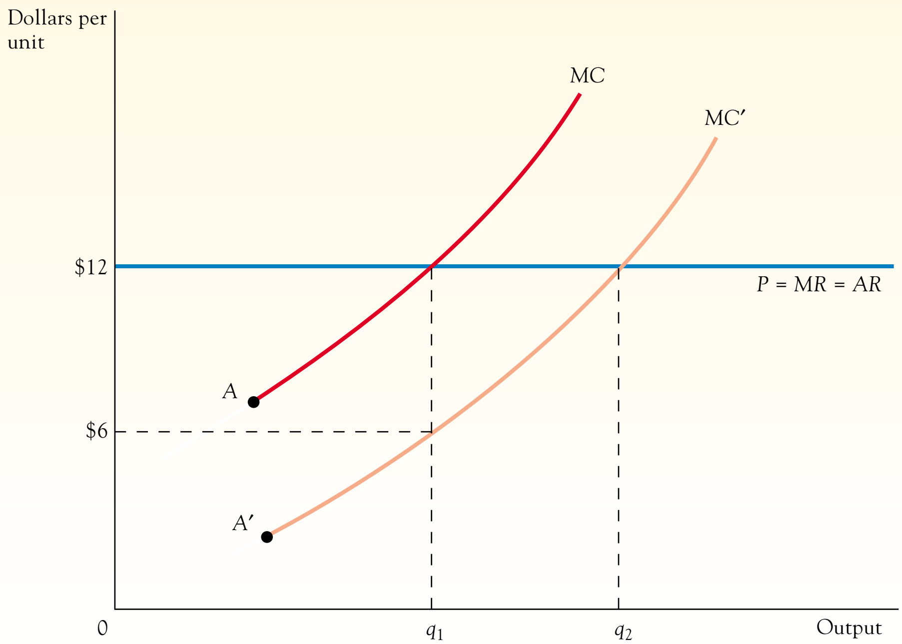
Output Response to a Change in Input Prices

Many factors can affect the competitive firm's output decision, but perhaps the two most common are variations in the product price and variations in the prices of inputs used in production. In reality, product and input prices frequently change at the same time, but at the outset, examining each factor in isolation is best. In the preceding paragraphs we examined the way a product price change affects output when input prices are constant (reflected in the fact that the cost curves did not shift). Now let's study the impact of an input price change while holding constant the product price.

Figure 9.6 illustrates a competitive firm initially producing the output q_1 where marginal cost, shown by MC, equals the \$12 product price. Note that only the marginal cost curve above the shutdown point, point A, is drawn in. Now suppose that the price of one (or more) of the variable inputs falls. As we saw in Chapter 8, lower input prices cause the cost curves to shift downward, indicating a lower cost associated with each rate of output than before. This change is depicted in the diagram by a shift in the marginal cost curve from MC

FIGURE 9.6

How the Firm Responds to Input Price Changes
An input price change, with an unchanged product price, alters the profit-maximizing output. If an input price falls, *MC* shifts to *MC′*, and output increases to q_2, where *MC′* equals the unchanged price.



to MC′. The rule for profit maximization in competitive markets, MC = P, determines the new output level q_2. To proceed to this conclusion somewhat more slowly, note that at the initial output level, a lower input price decreases the marginal cost from \$12 to \$6. After the reduction in cost, the price (\$12) is higher than marginal cost (\$6), indicating that an expansion in output will now (at the lower input price) add more to revenue than to cost and thereby increase profit. Consequently, the firm will expand output along the new MC′ curve until marginal cost once again equals marginal revenue at q_2.

In this section we have reviewed the competitive firm's reactions to changes in its economic environment—changes communicated to the firm through variations in input or product prices. A competitive firm by itself does not influence these prices; it is a price taker. Nonetheless, the combined actions of all the competing firms affect product and input prices, as we will see in the following sections. Each firm's response to price changes is an integral part of the adjustment process by which the overall market strives to reach equilibrium.

9.6 THE SHORT-RUN INDUSTRY SUPPLY CURVE

Moving from the determination of the most profitable output of the competitive firm to the short-run industry supply curve is a short step. In the short run a competitive firm will produce at a point where its marginal cost equals the price, as long as the price is above the minimum point of the average variable cost curve. In other words, each firm's marginal cost curve indicates how much the firm will produce at alternative prices. *As a first approximation, the* **short-run industry supply curve** *is derived by simply adding the quantities produced by each firm—that is, by summing the individual firms' marginal cost curves horizontally.* We assume for now that variable input prices remain constant at all levels of industry output.

SHORT-RUN INDUSTRY SUPPLY CURVE
the horizontal sum of the individual firms' marginal cost curves

Figure 9.7 illustrates how a short-run industry supply curve is derived for the three firms (A, B, and C) composing the cement industry. Although three firms may not be enough to constitute a competitive industry, the derivation is the same regardless of the number of firms involved. The figure shows the portion of each firm's marginal cost (MC) curve lying above its minimum average variable cost (AVC). Because the supply curve identifies the total quantity offered for sale at each price, we will add the outputs each firm chooses to produce individually.

At any price below P_0, the minimum point on firm C's AVC curve, industry output is zero because all three firms shut down. At P_0, firm C begins to produce. Note that its output is the only output on the market until price reaches P_1, because the other companies will not operate at such low prices. Thus, the lower portion of firm C's MC curve reflects the total industry supply curve at low prices. When price reaches P_1, however, firm A begins to produce, and its output must be added to C's—hence, the kink in the industry supply curve at P_1. When price reaches P_2, firm B begins production, and its output is included in the supply curve at prices above P_2. We refer to this derivation of the short-run industry supply curve as a horizontal summation of the individual firms' short-run marginal cost curves because we are horizontally summing quantities across firms. For example, at a price of P_3, total industry supply is Q, the sum of the amounts each company produces at that price ($q_A + q_B + q_C$).

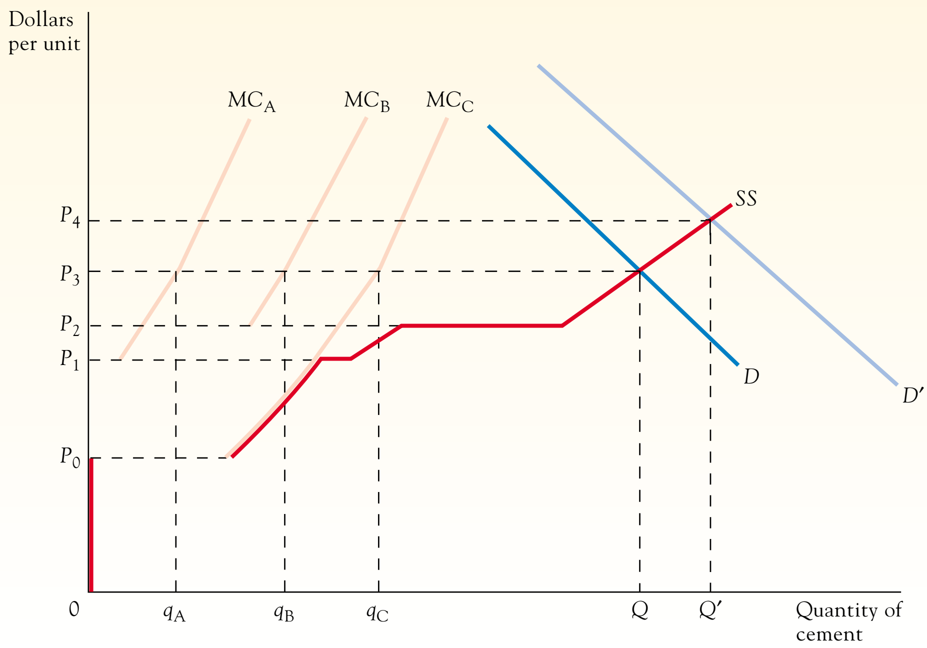
Note that the short-run industry supply curve, SS, slopes upward. Remember that each firm's MC curve slopes upward because it reflects the law of diminishing marginal returns to variable inputs. Thus, the law of diminishing marginal returns is the basic determinant of the shape of the industry's short-run supply curve.

Price and Output Determination in the Short Run

Incorporating the market demand for the product completes the short-run competitive model. Because the market demand relates total purchases (from all firms) to the price of cement, it interacts with the supply curve to determine price and quantity. In Figure 9.7, the

FIGURE 9.7

The Short-Run Competitive Industry Supply Curve
Each firm will produce an output where marginal cost equals price to maximize profit. To derive the short-run industry supply curve, *SS*, we must horizontally sum the amounts produced by the industry's various firms, as shown by their respective marginal cost curves.



intersection of the demand curve *D* with the supply curve identifies the price where total quantity demanded equals total quantity supplied. Thus, price P_3 is the short-run equilibrium price, and total industry output is *Q*. Each firm, taking price as given, produces an output where marginal cost equals P_3. Firm A thus produces q_A; firm B, q_B; firm C, q_C; the combined output, $q_A + q_B + q_C$, equals *Q*.

Given the Figure 9.7 supply and demand relationships, if price were at a level other than P_3 for some reason, familiar market pressures would work to push price toward its equilibrium level. For example, at a price lower than P_3, total quantity demanded by consumers would exceed the total amount supplied, a temporary shortage would exist, and price would be bid up. As price rises toward P_3, quantity demanded becomes smaller while quantity supplied becomes larger, until the two eventually come into balance at price P_3. Conversely, if price were higher than P_3, quantity supplied would exceed quantity demanded, a temporary surplus would exist, and competition among firms would drive the price down.

In the short run an increase in market demand leads to a higher price and a higher output. Suppose that we begin with the equilibrium just described, and then demand increases to *D′*. A shortage will exist at the initial price of P_3, and price will rise. The higher price will elicit a greater output response from the individual firms as they move up along their MC curves. The new short-run equilibrium will be price P_4 and quantity *Q′*.

Note that Figure 9.7 does not explicitly identify the profits, if any, realized by the firms. If we drew in each firm's average total cost curve, we could show them, but we do not need to use the average cost curves to explain the determination of price and quantity in the short run. All the necessary information is contained in the industry supply and demand curves.

APPLICATION 9.5 Why Electric Bills in California Increased After Deregulation

Traditionally, electricity generation was assumed to be subject to economies of scale.[5] To allow consumers to realize the cost-side benefits of increased output by an individual supplier, electric utilities typically were granted exclusive rights to generate electricity for particular service territories. The utilities were placed under price regulation or state ownership to ensure that these lower average costs associated with a single company's generating electricity were passed on to consumers in the form of lower electric bills.

The extent to which there are economies of scale in the generation of electricity, however, has come under scrutiny over the last two decades. Indeed, most economists now believe that such economies are insufficient to justify granting utilities exclusive rights to generate power for particular service territories. Indeed, although there may be important cost-side advantages to having a single, existing utility *distribute* power in a given locality, there also appear to be significant advantages associated with promoting competition between different companies to *generate* the power that will be distributed within the locality. A fitting analogy is the supply of gasoline for your car. It is desirable to have only one fuel system in your car, but competition between gas stations serves to keep your gasoline bills down.

To promote the benefits of competition, most states either already had deregulated electricity or were considering taking such a step by 2000. California initiated its deregulation in 1999, starting with San Diego. Instead of going down, however, electric bills went up. During the summer of 2000, for example, the typical San Diegan's electric bill was two to three times greater than it had been the previous summer.

Consider Figure 9.8a to understand why bills for electricity, measured in kilowatt hours (kWh), increased. Everything else being equal, suppose that deregulation allows for sources of power other than the public utility that has traditionally served San Diego. The short-run supply curve to San Diego subsequent to deregulation, SS, thus is derived by horizontally summing the traditional public utility's marginal cost curve (MC_{PU}) with the marginal cost curve associated with other power companies from outside the traditional service territory (MC_O). As you can see in Figure 9.8a, deregulation puts downward pressure on price. Assuming that prior to deregulation regulators set the price at P_2 (where the demand curve, D, intersects the traditional utility's marginal cost curve, MC_{PU}), competition for electricity generation lowers the price to P_1 (where demand intersects the supply curve, SS).

In reality, however, deregulation failed to introduce sizable downward pressure on price because California didn't encourage entry by new suppliers as actively as did some other states. For example, Pennsylvania, with an electricity market less than 12 percent the size of California's, aggressively promoted competition and attracted more than 100 new suppliers. By contrast, California drew only a handful. The added potential output from the new suppliers (represented by MC_O) thus is small relative to the existing supply curve from traditional sources (represented by MC_{PU}). Consequently, the potential decrease in price brought about by deregulation, from P_2 to P_1 in Figure 9.8a, is smaller than it would have been had competition been more aggressively promoted.

More importantly, two changes in the market put appreciable upward pressure on electric bills. First, the price of fossil fuels such as oil, natural gas, and coal increased significantly due to cutbacks in the supply of crude oil by the members of the Organization of Petroleum Exporting Countries (OPEC). Since utilities rely extensively on fossil fuels to generate electricity, the increase in the price of these inputs shifted the relevant marginal cost curves upward from MC_{PU} to MC'_{PU} and MC_O to MC'_O in Figure 9.8b. As a result of the increase in the cost of generating electricity by both traditional and outside suppliers, the short-run supply curve shifted leftward from SS to SS'.

Second, demand for electricity in California had been growing nearly 5 percent a year, but no new power plants were built in the state for more than 10 years prior to deregulation due to environmental concerns and a cumbersome regulatory approval process. The increase in demand fueled by the economic boom of the

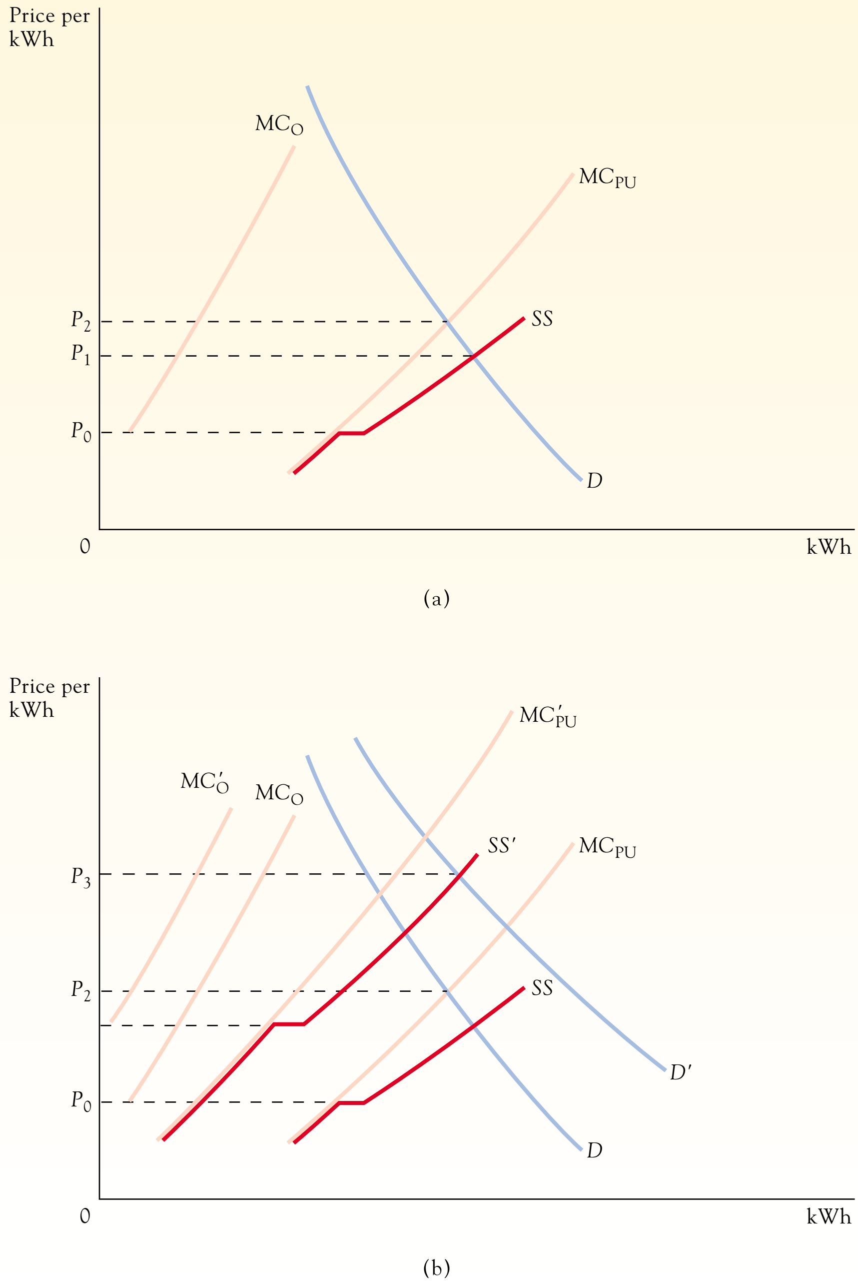
[5]This application is based on: Paul L. Joskow and Richard Schmalensee, *Markets for Power* (Cambridge, MA: MIT Press, 1983); James Flanigan, "Deregulation Is the Answer, Not the Problem," *latimes.com*, August 27, 2000; and Betsy Streisand, "Power to the People," *U.S. News Online*, August 21, 2000.

FIGURE 9.8

The Short-Run Electricity Supply Curve in California and Why the Price of Electricity Increased in 2000

(a) The short-run electricity supply curve is derived by horizontally summing the supply curves from traditional public utility sources (MC_{PU}) and new outside providers (MC_O). In and of itself, allowing in outside suppliers puts downward pressure on the price of electricity from P_2 to P_1.

(b) Instead of decreasing subsequent to deregulation, the price of electricity ends up increasing due to a simultaneous shift in demand (from D to D') and a leftward shift in the short-run supply curve (from SS to SS').



late 1990s shifted the demand for electricity rightward—from D to D' in Figure 9.8b.

Between the rightward shift in demand and the leftward shift in short-run supply, the price of electricity ends up rising to P_3 in Figure 9.8b (where SS' and D' intersect). And, as we can now see, the increase in the price of electricity reflects the workings of the familiar forces of supply and demand rather than a consequence of deregulation.

Of course, knowing that the forces of supply and demand are at work is little solace to a consumer whose electricity bill has suddenly doubled. As has happened in other cases where the price of a good has increased appreciably, pressure quickly mounted in California political circles to force utilities to roll back rates and/or to lower electricity prices through taxpayer subsidies. Moreover, in keeping with the law of demand, consumers quickly began to look for ways to lower the amount of electricity they demanded in the wake of the price increase. Even Sea World's star attraction, Shamu, began to feel the heat. The chillers that keep the killer whale's 7-million-gallon tank at a comfortable 55 degrees began to be shut down periodically to save on the park's $2.5 million annual electric bill.

9.7 LONG-RUN COMPETITIVE EQUILIBRIUM[6]

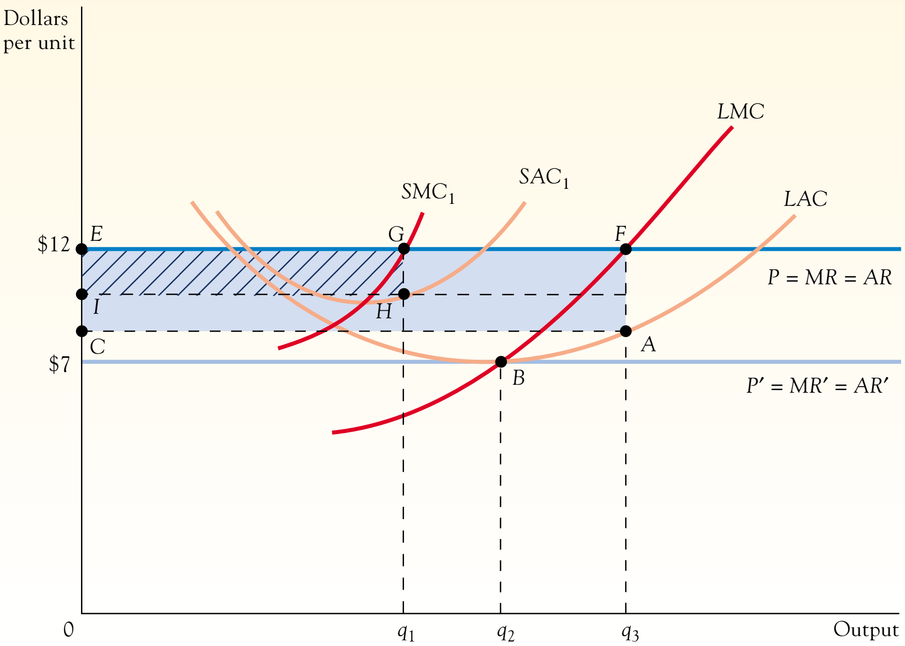
At any time, the short-run scale of plant a firm has built reflects a previous long-run decision. As a result, we must consider how the goal of profit maximization guides firms' long-run decisions and what that implies, in turn, about long-run competitive equilibrium. The same principles we used for the short-run setting apply to long-run profit maximization, but now we employ long-run cost curves, which allow a sufficient period of time to vary all inputs.

Figure 9.9 shows the long-run cost curves *LAC* and *LMC* for a representative firm in a perfectly competitive industry. If the firm believes the price will remain at $12, it will want to select an output level of q_3, where long-run marginal cost equals price. Producing q_3 in-

FIGURE 9.9

Long-Run Profit Maximization

The competitive firm maximizes its profit by producing where $LMC = P$. With a price of $12, output will be q_3. By building a larger-scale plant and adjusting its output from q_1 to q_3, the firm increases its profit from *EGHI* to *EFAC*.



[6]A mathematical treatment of some of the material in this section is given in the appendix at the back of the book (pages xxx–xxx).

volves building a scale of plant having a related short-run average cost curve (not depicted) tangent to *LAC* at point *A*. After building the plant, the firm realizes a total profit shown by area *EFAC*. No other size plant yields as much profit, since the long-run cost curves reflect all possible scales of plants the firm can select.

Figure 9.9 also illustrates why a short-run profit-maximizing outcome may be only a temporary equilibrium. Suppose that the price is $12 but the representative firm currently has a scale of plant with the short-run cost curves SAC_1 and SMC_1. With that size plant the firm's short-run profit-maximizing equilibrium is q_1, with the firm earning profit equal to the striped rectangular area *EGHI*. If the firm expects the price of $12 to persist, however, it would immediately begin making plans to enlarge its scale of plant, because it could earn significantly higher profit, area *EFAC*, by building the appropriate long-run scale of plant identified by point *A*. By its very nature, a short-run equilibrium is only temporary unless the firm's scale of plant is consistent with the price that is expected to persist in the market.

Suppose now that the expected price over the long run is $7 instead of $12. The horizontal demand curve at this price is just tangent to *LAC* at point *B*. For this demand curve the most profitable long-run output is q_2, where price equals long-run marginal cost. Total profit, however, is zero since price just equals the average cost of production.

The firm's long-run equilibrium at q_2, with price just covering average cost, has great significance. For one thing, the q_2 equilibrium indicates **zero economic profit**. Since the long-run cost curves take into account the opportunity costs of inputs, this implies that the return from employing inputs in the firm is as high as from the best alternative use of those inputs. In the case of capital, for instance, a firm's owners must be earning as great a return as the one they could make if their capital was employed elsewhere. Although the firm's books may show a positive accounting profit, the q_2 equilibrium involves a "normal" economic return being earned by inputs such as capital and thus zero economic profit.

ZERO ECONOMIC PROFIT
the point at which total profit is zero since price equals the average cost of production

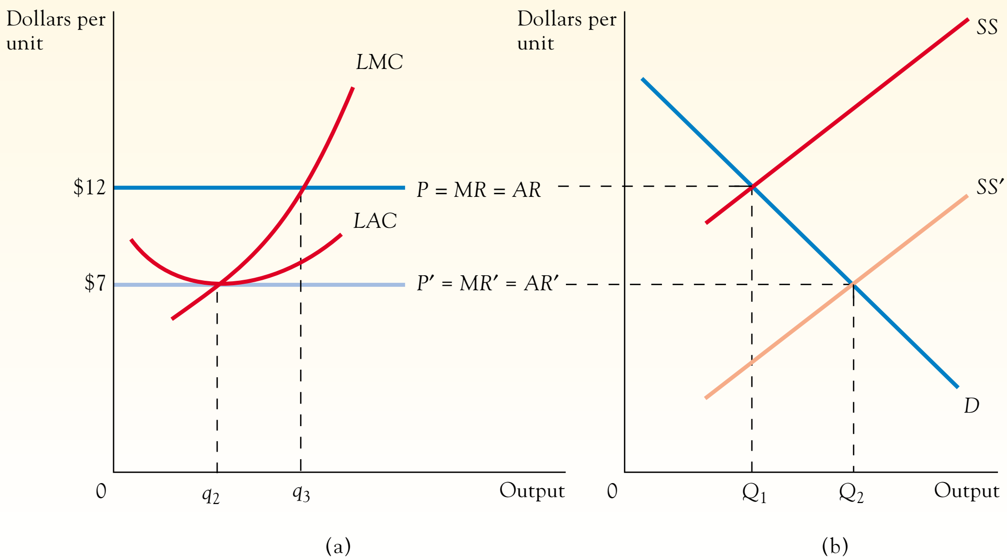
Should a firm stay in business if it is earning zero economic profit at an output such as q_2 in Figure 9.9? The answer to this question is an emphatic *yes*. Zero economic profit means that the various inputs, including capital provided by the owners, can earn just as much somewhere else. Although this is true, the relevant consideration is that they can't earn any *more* elsewhere, which is why the firm has an incentive to continue production.

The q_2 equilibrium is also significant because it is precisely the type of equilibrium competitive markets tend toward in the long run. To see why, remember that free entry and exit of resources is one of the characteristics of a perfectly competitive market. Figure 9.9 indicates how a price of $12 creates an incentive for the firm to expand output from q_2 to q_3 and provides the firm with a positive economic profit. Positive economic profit implies that resources invested in the industry generate a return higher than what could be earned elsewhere. Without barriers to entry, the abnormally high return results in new firms entering the industry; investors can make more money by shifting their resources into an industry affording positive economic profits. As shown in Figure 9.10, however, new entry also results in the industry short-run supply curve shifting to the right and a decrease in price. This process continues until the market demand curve and the new industry supply curve (SS') intersect at the same price (P') where long-run marginal cost equals the minimum point on the representative firm's long-run average total cost curve. Entry continues, in other words, until any positive economic profit signal and hence incentive for entry are eliminated. There is no incentive to enter the industry (that is, economic profit equals zero) when long-run average cost equals average revenue where long-run marginal cost equals price.

The long-run equilibrium shown in Figure 9.10 has three characteristics. First, the representative firm is maximizing profit and producing where *LMC* equals price. This condition must hold for a simple reason. If firms are not producing the appropriate amounts, they have an incentive to alter their output levels to increase profit. A change in firms' outputs, however, affects the total quantity offered for sale by the industry, which in turn changes the

FIGURE 9.10

Long-Run Competitive Equilibrium
(b) An industry in long-run competitive equilibrium where *SS′* and *D* intersect.
(a) The equilibrium from the perspective of a representative firm. The firm is producing the most profitable output (*LMC* = *P′*) and is making zero economic profit (*LAC* = *AR′*).



price at which the output is sold. Thus, the initial price cannot be an equilibrium price if firms are not producing where *LMC* equals price.

Second, there must be no incentive for firms to enter or leave the industry. This condition occurs when firms are making zero economic profits. If profits were higher, other firms seeking higher returns would enter the industry; if profits were lower, firms would leave the industry because they could do better elsewhere. This entry or exit would affect the level of industry output and change the price. If, however, profits in this industry are comparable to profits in other industries, there is no reason for entry or exit to occur, and price and output will remain stable.

Third, the combined quantity of output of all the firms at the prevailing price equals the total quantity consumers wish to purchase at that price. If this condition were violated, there would be either excess demand or excess supply at the prevailing price, so the price would not be an equilibrium one.

Zero Profit When Firms' Cost Curves Differ?

When all firms in a competitive industry have identical cost curves, every firm must be making zero economic profit in long-run equilibrium. Cost curves, however, may differ among firms. Some businesses may have access to superior technology, patents, or more skilled workers vis-à-vis other firms. These differences will be reflected as differences in production costs. While firms' *LAC* curves may be shaped and positioned differently, their minimum points will still end up at the same cost per unit through the process described below.

If firms' cost curves differ, we must reconsider the proposition that *every* firm makes zero economic profit in a long-run equilibrium. As it turns out, this proposition is still correct, but the reason it is deserves attention. To see why, consider the case of two firms that are both in the business of raising corporate executives' IQs, Densa and Mensa. Let's suppose that Mensa has some especially productive input that Densa does not have. For example, suppose that Mensa and Densa both purchase office space to begin their operations. Through foresight or chance, Mensa's selected office site proves to be more favorable. Suppose that government policymakers unexpectedly encourage businesses to set up shop in Mensa's vicinity while locating a nuclear waste dump adjacent to Densa, thereby scaring off potential corporate neighbors. Because of the government actions, Mensa's location is more productive, even though both firms originally may have paid the same price for the office space. The better location accounts for Mensa's lower production cost.

Due to the unexpectedly productive location, Mensa's income statement will indicate a positive accounting profit. The original price Mensa paid for the space understates its real economic value; that is, the cost of using the space is greater than its purchase price. Remember, however, that cost reflects forgone opportunities, and once it is known that Mensa's location is highly productive, its market value—what other firms would pay for the space—will rise. Provided that input markets are perfectly competitive, the opportunity cost to Mensa of its office space will be bid up to reflect its full economic value. Once Mensa's office space is valued at its true opportunity cost, Mensa's economic profit falls to zero.

In this analysis it makes no difference whether Mensa rents or owns the office space. If Mensa rents the space, and it becomes apparent that the space is unusually productive, the rental cost will go up because other firms will be willing to pay more for highly productive office space. If Mensa owns the office space, the same principle applies; the original purchase price is irrelevant. What counts is how much the office space is worth to other potential users—that is, its opportunity cost. In this case the cost of using the office space is an implicit cost, and, as explained in an earlier chapter, we include implicit costs in the cost curves.

The process by which unusually productive inputs receive higher compensation is not restricted to office space. Suppose that you go to work for a business as a manager and are promised a salary of $40,000. When the firm hires you, it doesn't know whether you'll be any good at your job. Fortunately, you turn out to be brilliant, and even after paying your salary, the firm's net revenue rises by $25,000 due to your extraordinary managerial skills. The firm's owner is delighted, and the firm's books show a $25,000 profit increase. Once your managerial skills are recognized, however, you are in a position to command and receive a $25,000 raise, which will effectively eliminate the firm's accounting profit. You will be in such a position so long as your managerial skills are transferable across firms and there is a competitive market for your managerial talents. Provided these conditions hold, if you don't get a raise, you can resign and accept a position with another firm that will pay you a salary closer to what you're worth—your best alternative—and your former employer's profit will decline to its original level.

Needless to say, this process doesn't work instantaneously or with exact precision. There is a tendency, however, for inputs to receive compensation equal to their opportunity costs, that is, what they are worth to alternative users. This process leads to the zero-profit equilibrium. One implication of this analysis is that the accounting measure of profit may vary widely among firms in an industry even though economic profit is zero for each firm. This variation arises because a firm's assets are frequently valued on the books at their original purchase prices instead of their opportunity costs. In our earlier example, Mensa's accounting profit would be higher than Densa's if they both counted only the purchase price of the office space as a cost.

9.8 The Long-Run Industry Supply Curve

To derive the competitive industry's short-run supply curve we horizontally sum individual firms' supply curves. We cannot similarly derive a competitive industry's long-run supply curve, however, because in the long run firms enter or exit the industry in response to the economic profits being earned. So we do not know which firms' supply curves to sum horizontally. Moreover, as an industry expands or contracts due to firm entry or exit, the prices that firms pay for their inputs may change. For example, as the movie business expands due to firm entry, the prices of actors and directors may be bid up. As the oil producing industry shrinks due to firm exit, the prices of petroleum engineers and drilling rigs may decline.

As we will see, the derivation of a **long-run industry supply curve** in a competitive market depends centrally on what happens to input prices as the industry expands or contracts. Based on the three possible effects of industry size on input prices, competitive industries are classified as **constant-cost**, **increasing-cost**, or **decreasing-cost**. We address each of these cases in turn, starting with the simplest of the three: when input prices remain constant regardless of an industry's overall size.

LONG-RUN INDUSTRY SUPPLY CURVE
the long-run relationship between price and industry output, which depends on whether input prices are constant, increasing, or decreasing as the industry expands or contracts

CONSTANT-COST INDUSTRY
an industry in which expansion of output does not bid up input prices, long-run average production cost per unit remains unchanged, and the long-run industry supply curve is horizontal

INCREASING-COST INDUSTRY
an industry in which expansion of output leads to higher long-run average production costs and the long-run supply curve slopes upward

DECREASING-COST INDUSTRY
a highly unusual situation in which the long-run supply curve is downward-sloping

Constant-Cost Industry

To derive the long-run supply curve for a constant-cost industry, we start from a position of equilibrium and trace the effects of a demand change until the industry once again returns to a long-run equilibrium. As our example, let's assume that the market for MBA education is perfectly competitive, with business schools being the representative firms. We begin by examining the industry when it is in long-run equilibrium (say in 1970).

Figure 9.11b shows the industry's initial long-run equilibrium: when demand is D, output is Q_1, and price is P. Assuming that we start in long-run equilibrium, point A is then one point on the long-run supply curve, LS. Figure 9.11a shows the representative firm's position—say that of New York University (NYU). It is producing its most profitable output, q_1, and making zero economic profit at price P.

To identify other points on the long-run supply curve, let's imagine that there is an unexpected increase in demand to D' and work through the consequences for the industry. Demand for MBA education has grown over the past three decades for a number of reasons, including more women joining the work force and seeking a business education, worldwide economic growth and entrepreneurship (while domestic applications to U.S. MBA programs have increased slightly in recent years, international applications have boomed), and the ever-advancing state of business management knowledge.

Reflecting the increased demand, existing business schools have an immediate short-run response—they increase output by expanding operations in their existing plants and increasing employment of variable inputs such as faculty and staff. The response appears as a movement along the short-run industry supply curve, SS, from point A to point B. (Note that the initial long-run equilibrium is also a short-run equilibrium. Every school is operating its existing plant at the appropriate level—where SMC equals P—so point A is also a point on the short-run supply curve.) In the short run, price rises to P', and output increases to Q_2 as the industry expands along SS.

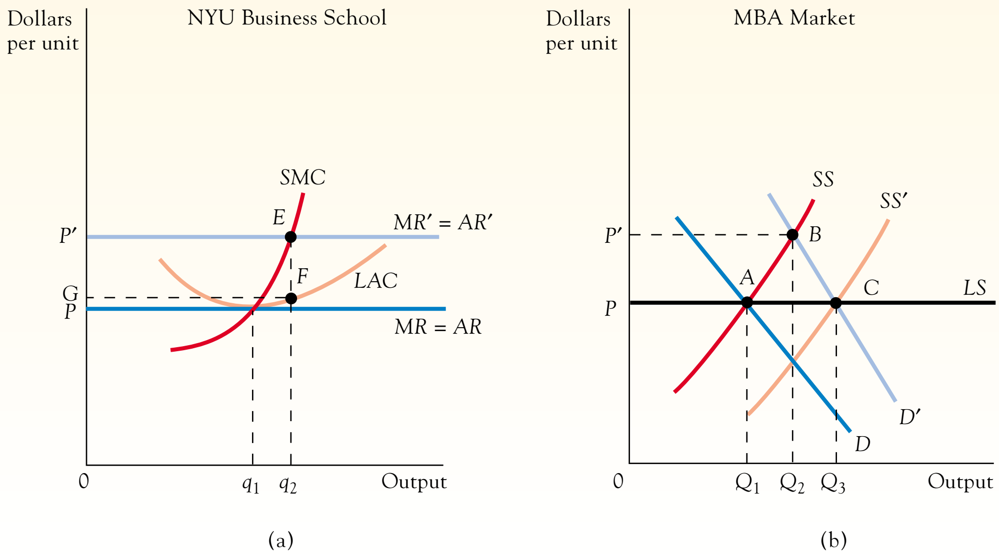
In Figure 9.11a, we can see what this adjustment means for NYU, a representative firm in the industry. The higher price induces NYU to increase output along its short-run marginal cost curve SMC and produce q_2 where $SMC = P'$. (Recall that the summation of all the schools' SMC curves yields the SS curve.) At this point every business school is making an economic profit (equal to area $P'EFG$).

The industry, however, has not fully adjusted to the increase in demand; this position is only a short-run equilibrium. *The key to the long-run adjustment is profit-seeking by firms.* In the short-run equilibrium, existing schools are making economic profits. The return on investment in the MBA education market is now greater than in other industries. Whenever economic profits exist in an industry, investors will realize that they can make more money by moving resources into the industry. If entry is costless, as it is assumed to be in the case of perfect competition, resources will be shifted into the MBA education market, thereby increasing its productive capacity.

FIGURE 9.11

Long-Run Supply in a Constant-Cost Industry

The long-run industry supply curve *LS* of a constant-cost competitive industry is horizontal. Expansion of industry output does not bid up input prices. Thus, when industry output expands from Q_1 to Q_3, firms' cost curves do not shift and the long-run average production cost per unit remains unchanged.



Suppose that new schools enter the industry in an attempt to share in the market's profitability. The entry increases the industry's total output, which, in Figure 9.11b, is shown as a rightward shift in the short-run supply curve. Recall that a short-run supply curve represents a fixed number of schools with given plants. When the number of firms increases, the total output at each price is greater (there are more "member" firms with their corresponding MC curves in the industry "club")—implying a shift in the short-run supply curve. As output expands in response to entry, price falls from its short-run level along D' since the higher output can be sold only at a lower price. This process of firms entering, total output increasing, and price falling continues so long as the industry is generating a positive economic profit signal. In other words, a new long-run equilibrium emerges when schools are once again making only a "normal profit"—that is, zero economic profit.

For a constant-cost industry the increased employment of inputs associated with expanding output occurs without an increase in the price of individual inputs. This means business school cost curves do not shift. Thus, price must fall back to its original level before profits return to a normal level: if price is higher than *P* economic profits will persist, and entry of new schools will continue. Figure 9.11b shows the process of entry as a rightward shift in *SS*; it continues until SS' intersects the demand curve D' at point C and price has returned to its original level *P*.

Point C is a second point on the long-run supply curve *LS*. With demand curve D', industry output expands until it reaches Q_3 and price is *P*. Each school once again makes zero economic profit. In Figure 9.11a, NYU is again confronted with price *P* and can do no better than cover its average cost by producing q_1. The increase in industry output from Q_1 to Q_3 is the result of entry by new schools.

A constant-cost competitive industry is characterized by a horizontal long-run supply curve. Given time to adjust, the industry can expand with no increase in price along *LS*. The crucial assumption producing this result is that input prices are not affected by an industry's size, so new schools can enter the market and produce at the same average cost as existing schools. The term *constant cost* refers to the fact that schools' cost curves do not shift as the industry expands or contracts; it does *not* mean that each school in the industry has a horizontal *LAC* curve.

Increasing-Cost Industry

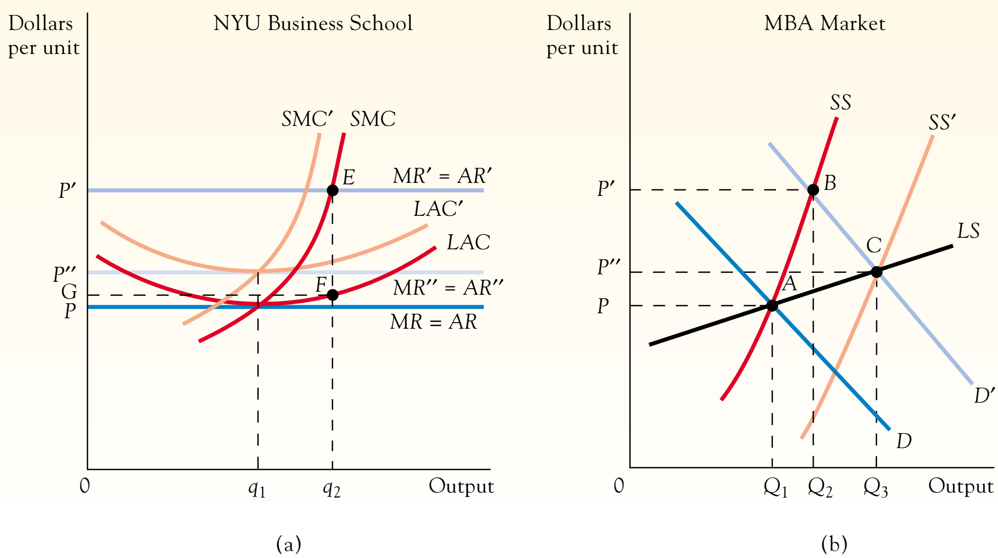
We can derive the long-run supply curve for an increasing-cost industry in the same way we derived it for a constant-cost industry. We assume an initial long-run equilibrium; then the demand curve shifts, and we follow the adjustment process through to its conclusion. Figure 9.12 illustrates this case. The initial industry long-run equilibrium price and quantity are P and Q_1, respectively. The typical firm, NYU, is producing q_1 and just covering its costs, shown by *LAC*, at the market price.

Once again, we assume an unexpected demand increase to D'. The short-run equilibrium is determined by the intersection of the short-run supply curve *SS* and D'. Price rises to P', and output increases to Q_2. The representative school expands output along its *SMC* curve to q_2 and realizes economic profit. In the long run, the profit attracts resources to the industry.

Up to this point the analysis is identical to the constant-cost case. Moreover, the attainment of a new long-run equilibrium involves further expansion of industry output until economic profits return to zero, just as in a constant-cost industry. However, in an increasing-cost industry the expansion of output leads to an increase in some input prices. To produce more output, schools must increase their demand for inputs, and some inputs such as finance

FIGURE 9.12

Long-Run Supply in an Increasing-Cost Industry
For an increasing-cost competitive industry, the long-run supply curve *LS* slopes upward. Expansion of industry output bids up some input prices. Thus, when industry output expands from Q_1 to Q_3, firms' cost curves shift upward from *LAC* to *LAC*′ and *SMC* to *SMC*′, and the output expansion leads to a higher long-run average production cost.



professors are assumed to be available in larger quantities only at higher prices. This situation contrasts sharply with that of the constant-cost case, where schools can hire larger quantities of all inputs without affecting their prices.

Let's see how this difference affects the long-run adjustment process. At the short-run equilibrium, positive economic profits lead to entry by new schools and an expansion in industry output—a rightward shift in the short-run supply curve. Increased industry output now tends to reduce profits in two ways. First, as we saw earlier, the higher output causes price to fall, which reduces profits. Second, the increased demand for inputs that accompanies the expansion in industry output leads to higher input prices. Higher input prices mean higher production costs, which also reduce profits. Profits are thus caught in a two-way squeeze as the industry expands. The two-way squeeze continues until economic profits equal zero and a new long-run equilibrium is attained.

Figure 9.12 shows the two-way squeeze on economic profits. Starting with the short-run equilibrium at point *B* (panel b), as new schools enter the market, *SS* shifts to the right and price falls. Each school's horizontal demand curve shifts downward as price declines from P' to P'', and this decline reduces profits. At the same time, higher input prices shift the firms' cost curves upward, from *LAC* to *LAC′* and *LMC* to *LMC′*. In Figure 9.12a, NYU's profit (area $P'EFG$) is squeezed from above by the decline in the product price and from below by the rising unit cost of production. A rising cost and falling price eventually eliminate NYU's economic profit. This occurs at price P'', when NYU can just cover its average cost by producing at the minimum point on *LAC′*. Once profit is eliminated, there is no longer an incentive for industry output to increase further, and a new long-run equilibrium is reached. In Figure 9.12b, *SS* has shifted rightward to *SS′*, producing a price of P'' and an output of Q_3. No further shift in the short-run supply curve will occur, because economic profits have fallen to zero. Thus, point C is another point on the long-run industry supply curve.

An increasing-cost competitive industry is characterized by an upward-sloping long-run supply curve. This industry can produce an increased output only if it receives a higher price, because the cost of production rises (cost curves shift upward) as the industry expands. The term *increasing cost* indicates that the cost curves of all schools shift upward as the industry expands and input prices are bid up.

Decreasing-Cost Industry

A *decreasing-cost* competitive industry is one that has a downward-sloping long-run supply curve. You might think that such a situation is impossible, and, in fact, some economists believe that it is. Others claim that a decreasing-cost industry is theoretically conceivable but admit that it is a remote possibility. Because all agree that it is, at best, highly unusual, we will deal with it only briefly here.

A decreasing-cost industry adjusts to an increase in demand by expanding output, just as industries in the other two cases. In this instance, though, the expansion of output by the industry in some way lowers the cost curves of the individual schools, leading to a new long-run equilibrium with a higher output but a lower price. The tricky part is to try to explain why the cost curves shift downward. A downward shift in cost curves usually reflects a decrease in input prices, but for that to happen we have to explain how an increase in demand for some input leads to an increased quantity supplied at a lower price—not an easy task.

Although economists are in agreement about how rare decreasing-cost industries must be, many noneconomists find the concept appealing. The reason for its appeal stems from the observation that prices have declined sharply as output has expanded in some industries. For example, color televisions, VCRs, and pocket calculators have all fallen in price in real terms by 80 to 90 percent since they were first introduced. More recently, there have been dramatic reductions in personal computer prices.

Before concluding that such evidence reflects decreasing-cost industries, we should explore some other possibilities. One common feature in these examples is that the price was high when the product was first marketed but later fell dramatically. This suggests two possible explanations. First, the firm initially marketing the product is a monopoly and thus has some pricing power. The price that the monopoly firm sets may be fairly high. As other firms enter the market and begin to compete, price falls and output increases. This process suggests that what we may be seeing is the rise of competition from an initial monopoly position and not a movement along an industry's long-run supply curve.

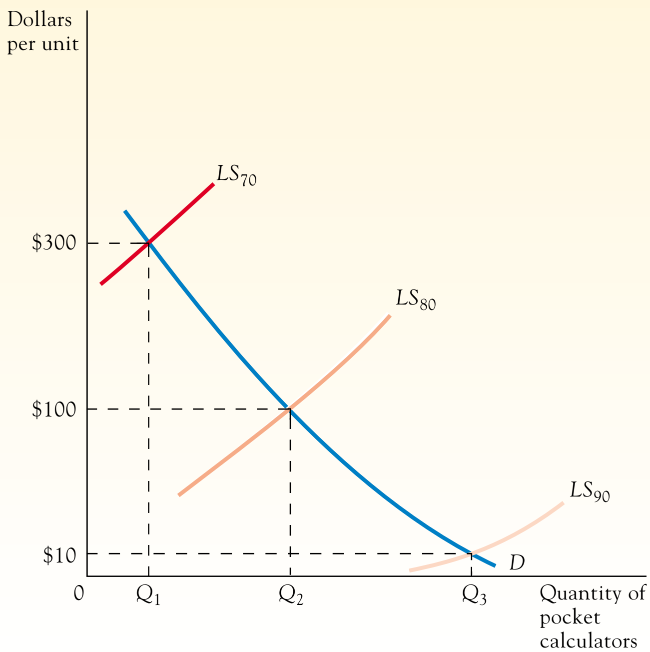
Second, the passage of time after a product's introduction makes technological improvements in production possible. Particularly in the case of new and complex products, technological know-how is frequently rudimentary when firms first market a product. With experience in production over time, technological improvements may occur quickly. We emphasize that technological know-how is assumed to be unchanged along a supply curve. An improvement in technology shifts the entire supply curve to the right.

Consider Figure 9.13. In conjunction with demand, the long-run supply curve for pocket calculators in 1970, LS_{70}, determined a price of \$300. This price was, in fact, the approximate price of a calculator that performed only the basic functions—when expressed in constant 1990 dollars. After 10 years firms developed methods that lowered production cost, and the 1980 long-run supply curve was LS_{80}. We assume that demand is unchanged, although it was probably increasing over the period. This shift in supply led to a price of \$100 in 1980. Further technological improvements shifted the supply curve once again, and price fell to \$10 in 1990. The combination of lower prices and higher outputs over time resulted here from shifts in an upward-sloping long-run supply curve, not from a slide down the negatively-sloped supply curve of a decreasing-cost industry. This explanation of the phenomenon is especially appealing because new high-technology items are known to show rapid improvements in technical knowledge in the first years of their production.

FIGURE 9.13

Technological Advances Shift the Long-Run Supply Curve

A price reduction and output increase over time need not imply that the long-run supply curve has a negative slope. Technological advances can occur over time, which shift the entire long-run supply curve downward, producing lower prices and higher outputs.



These remarks do not rule out the possibility of decreasing-cost industries, but if they exist, like the Giffen good case in demand theory, they are extremely rare. For all practical purposes the increasing- and constant-cost cases are the relevant possibilities.

APPLICATION 9.6 THE BIDDING WAR FOR MIS AND FINANCE PROFESSORS

Each year, the country's 700-plus business schools produce a total of 80,000 MBAs.[7] The number of business schools has more than doubled and the number of MBA graduates per year has increased 15-fold since 1970. As the industry has expanded, certain inputs have been bid up in price. Among the inputs that have been bid up most in price are faculty in disciplines such as management information systems (MIS) and finance, in which student enrollments have witnessed the largest increases. For example, top new finance assistant professors garnered starting salaries of more than $145,000 in 2002—52 percent higher than in 1997. Leading rookie MIS faculty members were earning $130,000 in 2002—53 percent more than in 1997. Annual salaries for the most respected senior faculty members in both fields rose from roughly $120,000 in the mid-1980s to well over $300,000 by 2002.

Of course, as the business school market has expanded, not all inputs have risen in price. Some have remained constant or perhaps even decreased. For example, schools' applicant screening costs have been reduced due to widespread reliance on the Graduate Management Admissions Test (GMAT) as a candidate-screening mechanism.

On average, however, real input prices have increased since the 1960s. This suggests that the long-run supply curve in the MBA education market is upward-sloping and provides at least a partial explanation for the rise in annual tuition at private business schools.[8] For example, annual MBA tuition at USC, a representative private business school, rose from $1,950 in 1970 to $35,000 in 2002. The 2002 annual USC MBA tuition equals roughly $7,550 in constant 1970 dollars. Although other factors (such as increased reliance by universities on their business schools to generate an operating surplus) have certainly affected MBA tuition, the growth of the MBA industry and the attendant increases in demand for industry inputs and the (real) prices of those inputs surely also played a role.

[7]This application is based on *AACSB Newsline*, Vol. 30 No. 3 (2000), pp. 1–5.

[8]Tuition at public business schools is generally controlled and subsidized by the state.

Comments on the Long-Run Supply Curve

An industry's long-run supply curve summarizes the results of a complex and subtle process of adjustment. Once the underlying determinants of the supply curve are understood, it becomes a powerful tool of analysis. To use the supply curve correctly, however, we must have a firm understanding of its underpinnings. To that end, we discuss several frequently misunderstood points.

1. We do not derive the long-run supply curve by summing the long-run marginal cost curves of an industry's firms. Admittedly, every firm is producing where $LMC = P$ at each point on LS, but as the industry adjusts along LS, firms are entering or leaving the market. Therefore, we cannot sum firms' LMC curves as we did in the short run. In addition, for an increasing-cost industry the LMC curves themselves shift because of changes in input prices.

2. Just as we did with a demand curve, we assume certain things remain unchanged at all points on the long-run supply curve. For one we assume technology is constant. An industry expanding along its LS curve is using the same technical know-how but employing more inputs to increase output. A change in technology causes the entire supply curve to shift, as we saw in the example about pocket calculators discussed earlier. We also assume that at all points on the long-run supply curve the supply curves of inputs to the industry remain unchanged. Note that we are not assuming that the prices of inputs are unchanged but rather that the conditions of supply remain constant. In a constant-cost industry, input prices remain constant not because we assume them to be fixed but because the input supply curves facing the industry are horizontal. In contrast, in an increasing-cost industry input prices change because the input supply curves facing the industry are upward-sloping, a condition that gives rise to an upward-sloping, long-run industry supply curve. When relevant, other factors like government regulations or the weather must also be assumed constant along the supply curve.

3. In reality, an industry is not likely to fully attain a position of long-run equilibrium. Real-world industries are continually buffeted by changes. For instance, input supplies, demand, technology, and government regulations frequently change. A long-run adjustment takes time, and if underlying conditions change often, an industry will find itself moving toward a long-run equilibrium that is continually shifting. Recognizing the reality of frequent change, however, does not undermine the usefulness of the long-run equilibrium concept. Although the industry may never attain a long-run equilibrium, the tendency for the industry to move in the indicated direction is what is important, and the outcome is correctly predicted by the theory.

4. Economic profit is zero all along a competitive industry's long-run supply curve. This point is often misunderstood. For example, someone may say that when industry output expands due to an increase in demand, the industry's firms benefit. We have seen, however, that after an increase in demand, all firms will make zero economic profits in the long run. There may be temporary economic profits in the short run, but the benefit is not permanent. Who does benefit as we move along the long-run supply curve? Owners of inputs whose prices are bid up by the industry-wide expansion. With a constant-cost industry no input prices rise, and no input owners receive a permanent gain. If firms own some of the inputs and thus gain as input prices increase, the gain accrues to them on account of their input ownership and not because they are firms in an increasing-cost industry. It is also possible that firms own none of the inputs whose prices rise. It may be, for example, that finance professors whose wages go up as the MBA industry expands are the sole beneficiaries on this market's supply side.

The tendency toward zero economic profit means that the rate of return on invested resources will tend to equalize across industries. If invested resources yield an annual return of 10 percent in the restaurant industry, which is comparable to earnings elsewhere, then the restaurant industry is earning zero economic profit. Accountants, of course, generally call the 10 percent return a "profit," but economists regard the 10 percent return as a cost necessary to attract resources to the restaurant industry.

5. In deriving the long-run supply curve, we assumed that a short-run equilibrium was first established and then long-run forces came into play. Price first went up to a short-run high and then came down to a long-run equilibrium level. The actual process of adjustment to demand changes may not follow this pattern exactly. Identifying a short-run equilibrium and then tracing out long-run effects is merely an expedient way of explaining the determination of the final equilibrium. In fact, following a demand increase, price may never reach its short-run level and may never go above the ultimate long-run level. This can happen if, for example, firms anticipate the demand increase and make adjustments before demand actually rises.

These remarks also suggest that care must be taken in using the short-run supply curve. The short-run supply curve can be used to identify the initial effects of a change in demand under only two conditions: when the demand change is unexpected, and when it is expected to be temporary. If firms anticipate the demand change, they can adjust their scales of plants or move into or out of a market before the demand change actually occurs. An unexpected demand change catches firms unaware, and they must operate temporarily with whatever scales of plants they have at that time. If firms expect a demand change to be temporary, they will not expand capacity or enter a market on the basis of conditions they know will not persist. Thus, any output change will result from the firms utilizing existing plants more or less intensively, and a short-run analysis is appropriate.

Even when appropriate, a short-run analysis identifies only an industry's temporary resting place, and subsequent long-run adjustments will continue moving the industry toward long-run equilibrium. So in most supply-demand applications we generally use the long-run supply curve.

9.9 When Does the Competitive Model Apply?[9]

The assumptions of perfect competition are stringent and are likely to be satisfied fully in very few real-world markets. Still, many markets come close enough to satisfying the assumptions of perfect competition to make the model quite useful. As we mentioned earlier, product homogeneity and perfect information may not apply in the case of gasoline retailing, yet the industry may be very close to being perfectly competitive since representative firms have only limited pricing power.

How close is "close enough" in terms of the assumptions of perfect competition being satisfied to make the model applicable? This question has no easy answer. For example, consider the assumption that there must be "many" small firms, an assumption that can be readily verified: firms can be counted. The relationship between the number of firms and an industry's competitiveness is not obvious. As we will see in Chapter 13, a relatively large number of firms producing a homogeneous product may collude to fix prices above the competitive level.

Conversely, even if there are only a few firms in an industry, each firm may face a highly elastic demand curve if the market demand curve is very elastic or if the elasticity of supply by rival firms is high. For example, suppose that United is one of four suppliers of air travel between Los Angeles and San Francisco, and that each of the four suppliers accounts for one-fourth of the total air passenger travel between the two cities. Although it is by no means a small player in this market, United may still face a highly elastic demand curve if when it raises its price, the rivals are quick to increase their output in response to United's price increase. A high elasticity of supply by rival airlines can thus severely limit United's pricing power. In Chapter 13 we will explore in greater depth how the output responses of rivals to a firm's price changes affect the elasticity of the firm's demand curve.

With regard to other assumptions of perfect competition, such as free entry and exit of resources, the competitive model frequently can be adapted to take the violations of these assumptions into account. This adaptability is another reason why the model is important: it can be extended to analyze the implications of specified departures from perfect competition. For example, suppose that an industry is in long-run equilibrium and demand increases, but new firms are blocked from entering the market. It is easy to show that output will increase by less in this case, price will be higher, and economic profit will persist after the industry's firms have made the appropriate long-run adjustment.

Consider as well the assumption that firms produce a homogeneous product. This assumption serves two purposes in the competitive model: it implies that a uniform price will

[9]A mathematical treatment of some of the material in this section is given in the appendix at the back of the book (page xxx).

prevail, and it affects the elasticity of demand facing a particular firm. What if the product homogeneity assumption is not fully satisfied? A good example of this is provided by comparison-shopping surveys that find food prices higher in the inner city than in the suburbs. For those who expect the competitive model to be an accurate description of reality, this result violates the uniform-price implication and suggests monopoly pricing in the inner cities. Further investigation, however, reveals profits to be no higher for inner-city food stores than for suburban markets, a result consistent with competition but not with monopoly.

What is going on? Basically, food sold in the inner city and food sold in the suburbs are not homogeneous products. To city residents, food sold in the inner city is worth more than food sold farther away. People are willing to pay something for the convenience of shopping nearby, and when convenience costs the retailer something, the price of the "same" product will differ. In comparison with costs in the suburbs, the cost of operating a food store in the inner city is higher because rent, fire and theft insurance, and the salaries of personnel all tend to be higher there. As a result, inner-city residents pay a higher price for shopping near their homes.

Dropping the homogeneous product assumption means that differentials in price can exist, but that is not often a serious problem and the competitive model can still be used. Suppose, for example, that a city government passed a law that food prices in the inner city cannot exceed prices in the suburbs. The competitive model applied to this setting would predict that output (food sales) would fall in the inner city, shortages would exist, some food stores would go out of business, and many inner-city residents would be worse off because no food was locally available at the mandated lower price—the familiar effects of a price ceiling. These predictions would, we suspect, be borne out in practice.

Finally, the assumption that firms and consumers have all the relevant information is also necessary if markets are to behave exactly as the competitive model predicts. For example, if consumers are ignorant of price, even homogeneous products can sell for different prices.

Dropping the assumption of full information does not mean, however, that we have to replace it with one of complete ignorance. One of the results of real-world markets is that firms and consumers have incentives to acquire the information that is important for their economic decisions. Although they may not become fully informed, because there are costs associated with acquiring information, we can easily suppose that they will become well enough informed that the assumption of complete information is not too great a distortion.

Economists have only recently begun to systematically analyze the acquisition of information and the way "information costs" influence the workings of markets. It is too early to make any sweeping generalizations in this area, but we do not want to suggest that the degree of information is irrelevant to the functioning of a market. Lack of information on the part of consumers may result in market outcomes that deviate significantly from the competitive norm. We will discuss issues related to information in more detail in Chapter 14.

SUMMARY

- A perfectly competitive industry is characterized by a large number of buyers and sellers, free entry and exit of resources, a homogeneous product, and perfect information. Although few industries fully satisfy all these conditions, the model of perfect competition remains quite useful in analyzing many markets.
- Since a competitive firm sells a product similar to that sold by many other firms, the firm's output decision has a

negligible effect on the product's price. Thus, the competitive firm's demand curve can be approximated by a horizontal line at the level of the prevailing price.

- For any firm, the most profitable output occurs where the marginal cost of producing another unit of output just equals the marginal revenue from selling it.
- For a competitive firm with its horizontal demand curve, price equals marginal revenue, so it maximizes profit by producing where marginal cost equals price. Price, however, must be at least as high as average variable cost or the firm suffers a loss in excess of total fixed cost.
- If price does not cover *AVC* in the short run, the firm will shut down; if price does not cover *LAC* in the long run, the firm will go out of business.
- The competitive firm's short-run supply curve is its marginal cost curve, so long as marginal cost exceeds average variable cost.
- The assumption of profit maximization allows us to predict how a competitive firm will respond to changes in the product price or in input prices. An increase in the product price (with unchanged cost curves) will lead the firm to expand output as it moves up its marginal cost curve.
- A reduction in one or more input prices will shift the cost curves downward and so will lead the firm (with an unchanged product price) to expand output.
- The competitive industry's short-run supply curve is the horizontal summation of the short-run marginal cost curves (above *AVC*) of an industry's firms.
- Because the law of diminishing marginal returns implies that each firm's marginal cost curve slopes upward, the short-run industry supply curve also slopes upward.
- In the long run, firms have sufficient time to alter plant capacity and to enter or leave the industry. The long-run supply curve takes these adjustments into account, with firms always guided by the search for profit.
- Two shapes of the long-run industry supply curve are most likely. First, an increasing-cost industry has an upward-sloping long-run supply curve, reflecting the increase in the prices of one or more inputs as the industry expands.
- Second, a constant-cost industry has a horizontal long-run supply curve, reflecting a situation where the industry can expand its use of inputs without affecting their prices.
- At all points along a long-run supply curve, economic profit is zero, because only when profit is zero is there no incentive for firms to enter or leave the industry.
- Input supply curves and technology are assumed as givens in deriving a long-run supply curve; changes in these underlying factors shift the long-run supply curve.

Review Questions and Problems

Questions and problems marked with an asterisk have solutions given in Answers to Selected Problems at the back of the book (page xxx).

9.1. What are the four assumptions of the perfectly competitive model? Critics are fond of pointing out that few, if any, real-world markets satisfy all four conditions, implying that the competitive model has little relevance for real-world markets. How would you respond to these critics?

9.2. How can consumers of apartments in Atlanta have a downward-sloping demand curve for apartments, and yet each Atlanta rental property owner face a horizontal demand curve?

***9.3.** Assume that a competitive firm has the short-run costs given in Table 8.1. What is the firm's most profitable output, and how large is profit if the price per unit of output is (a) \$15? and (b) \$26?

9.4. "IBM should never sell its product for less than it costs to produce." If "costs to produce" is interpreted to mean IBM's average total cost, is this correct? If it is interpreted to mean average variable cost, is the statement correct? If it is interpreted to mean marginal cost, is the statement correct?

9.5. In Table 9.1, suppose that variable input prices increase by 50 percent. Will the firm's profit-maximizing output level change? Illustrate your answer with a graph.

9.6. "If Conagra is a competitive firm, it will never operate at an output where its average total cost curve is downward-sloping." True or false? Explain.

***9.7.** "The difference between price and marginal cost is the amount of profit per unit of output, which the firm wishes to be as large as possible." True, false, or uncertain? Explain.

9.8. Suppose that through laziness and/or sheer stupidity, Densa Inc. always falls 10 percent short of producing the profit-maximizing output. Would a higher product price lead to greater output? Would an increase in input prices lead to a reduction in output? Does this result suggest why economists are not overly concerned about whether the profit-maximizing assumption is exactly correct? Explain.

9.9. Suppose that Keystone is a firm in the perfectly competitive ski resort business. If all of Keystone's input prices unexpectedly double, and at the same time the product price doubles, what will happen to Keystone's profit-maximizing level of output and its profit in the short run? In the long run? (Assume that Keystone begins from a position of long-run equilibrium.)

9.10. Starting from a long-run equilibrium, trace the effects of an unanticipated reduction in demand for (a) a constant-cost industry and (b) an increasing-cost industry. *Note:* This process is just the reverse of our derivation of the supply curves in Section 9.8, but it is very good practice to think through the process.

***9.11.** In a constant-cost industry each firm's MC curve is upward-sloping, yet all the firms together—the industry—have a horizontal supply curve. Explain why there is no contradiction.

9.12. If each business school in the MBA education industry is operating where $LMC = P$, does it follow that the industry is in long-run equilibrium? Explain.

***9.13.** "Suppose that the defense contracting business is perfectly competitive. In long-run equilibrium the price just covers average production cost, and every contractor makes zero profit. Thus, if the price goes down, even a little bit, due to a decrease in government defense spending, all contractors will go out of business." Discuss this statement.

9.14. Assume that the MBA education industry is constant-cost and is in long-run equilibrium. Demand increases, but due to strict accreditation standards, new firms are not permitted to enter the market. Analyze the determination of a new long-run equilibrium, showing the effects for a representative school as well as for the market as a whole.

9.15. For a given increase in demand, will output increase by more if the MBA education industry is constant-cost or increasing-cost? For a given decrease in demand, will output fall by more if the industry is constant-cost or increasing-cost?

9.16. How would each of the following phenomena affect the long-run supply curve of apples?

a. Workers in the apple industry form a union.
b. Consumers find out that apples cause cancer.
c. Hard-to-control bugs that eat apples invade from Mexico.
d. The government passes a law requiring apple trees to be planted at least 60 feet apart.
e. The government sets a maximum legal price (price ceiling) at which apples can be sold.
f. Immigration laws change to permit more itinerant apple pickers to enter the country.
g. The government passes a minimum wage law for apple pickers.

9.17. "Because agricultural demand is inelastic, a technological advance that lowers production costs will reduce total revenue. Thus, farmers have no incentive to introduce such a technique." True, false, or uncertain? Explain.

9.18. Can all schools in the MBA education market face constant returns to firm scale and yet the industry have an upward-sloping long-run supply curve? If the MBA education industry is constant-cost, what, if anything, does this imply about the returns to scale faced by the industry's individual schools?

9.19. As the industry moves from point *B* to *C* in Figure 9.11b, how can the industry's output increase yet NYU's output fall?

9.20. The American Red Cross is a supplier to the perfectly competitive domestic blood market. Unlike the other suppliers, however, the Red Cross is strictly nonprofit—its goal is to sell as much blood as possible without making a profit. Given this goal, is the Red Cross supply curve equal to its average variable cost curve? Explain why or why not.

9.21. Suppose that the gasoline retailing industry is perfectly competitive, constant-cost, and in long-run equilibrium. If the government unexpectedly levies a five-cent tax on every gallon sold by gasoline retailers, depict what will happen to the representative firm's cost curves. What will the effects of the tax be in the short run on industry output and price? Will the price rise by the full five cents in the short run? In the long run? How would your answers change if the industry was increasing-cost?

9.22. Suppose that Mensa Inc. is a representative firm operating in a perfectly competitive industry. Mensa's total cost of production, *TC*, is given by the equation $TC = 5{,}000 + 5q^2$, where q is Mensa's output. Based on this equation, Mensa's marginal cost is $10q$. If the output price is \$100, what is Mensa's short-run profit-maximizing output? How much profit does Mensa make at that output?

9.23. Suppose that oil tankers fall into three classes: medium (tankers capable of carrying between 10,000 and 70,000 deadweight tons DWT of crude oil—approximately 170 million gallons); large (tonnages between 70,000 and 175,000 DWT); and super (tonnages over 175,000 DWT). The total tonnage in each of the three classes is 53,000,000 in medium; 76,000,000 DWT in large; and 171,000,000 in super. Finally, suppose that the constant per-unit annual operating cost per deadweight ton is \$149 for medium; \$107 for large; and \$84 for super. Based on the preceding information, graph the short-run supply curves for each of the three tanker classes as well as for the overall oil shipping industry.

9.24. On October 13, 1993, as news of the potential merger between Bell Atlantic and Tele-Communications Inc. spread, hordes of investors beseiged brokers with "buy" orders for TCI. The trouble was, TCI is the stock symbol for Transcontinental Realty Investors. TCI shares soared nearly 30 percent before the New York Stock Exchange caught on. It halted trading in the stock and sent out an advisory warning to members of the Exchange. Which assumption of perfect competition was violated by this episode?

9.25. Over the last two decades, the price of personal computers in real as well as nominal terms has declined markedly. Does this mean that the personal computer industry is decreasing-cost and that the long-run supply curve for personal computers is downward-sloping? Explain why or why not.

9.26. Suppose that Continental Airlines fills only 60 percent of the seats, on average, on its Boeing 737 flights—a number 20 percent below the industry average. If Continental incurs incremental costs of $2,000 per 737 flight, wouldn't it do better by scheduling fewer 737 flights and thereby increasing the percentage of seats filled, on average, per flight? Why or why not? Explain.

9.27. In the market for cell phones, the amount generated in revenue per unit sold by Motorola roughly equals the associated incremental cost. Given that Motorola is losing money on its cell phone product line, it would be wise for the company to cease manufacturing such a product. True, false, or uncertain? Explain.

9.28. "In a competitive industry, high prices in response to a positive demand shock prevent higher prices." Explain whether this statement is true, false, or uncertain in the context of the competitive industry model developed in this chapter. To help your analysis, think what would happen if government policymakers precluded suppliers from raising their prices in the wake of a positive demand shock (e.g., local suppliers of lumber were prohibited from raising the price of their products in the wake of a tornado ravaging a community and the need to begin rebuilding homes in the community).