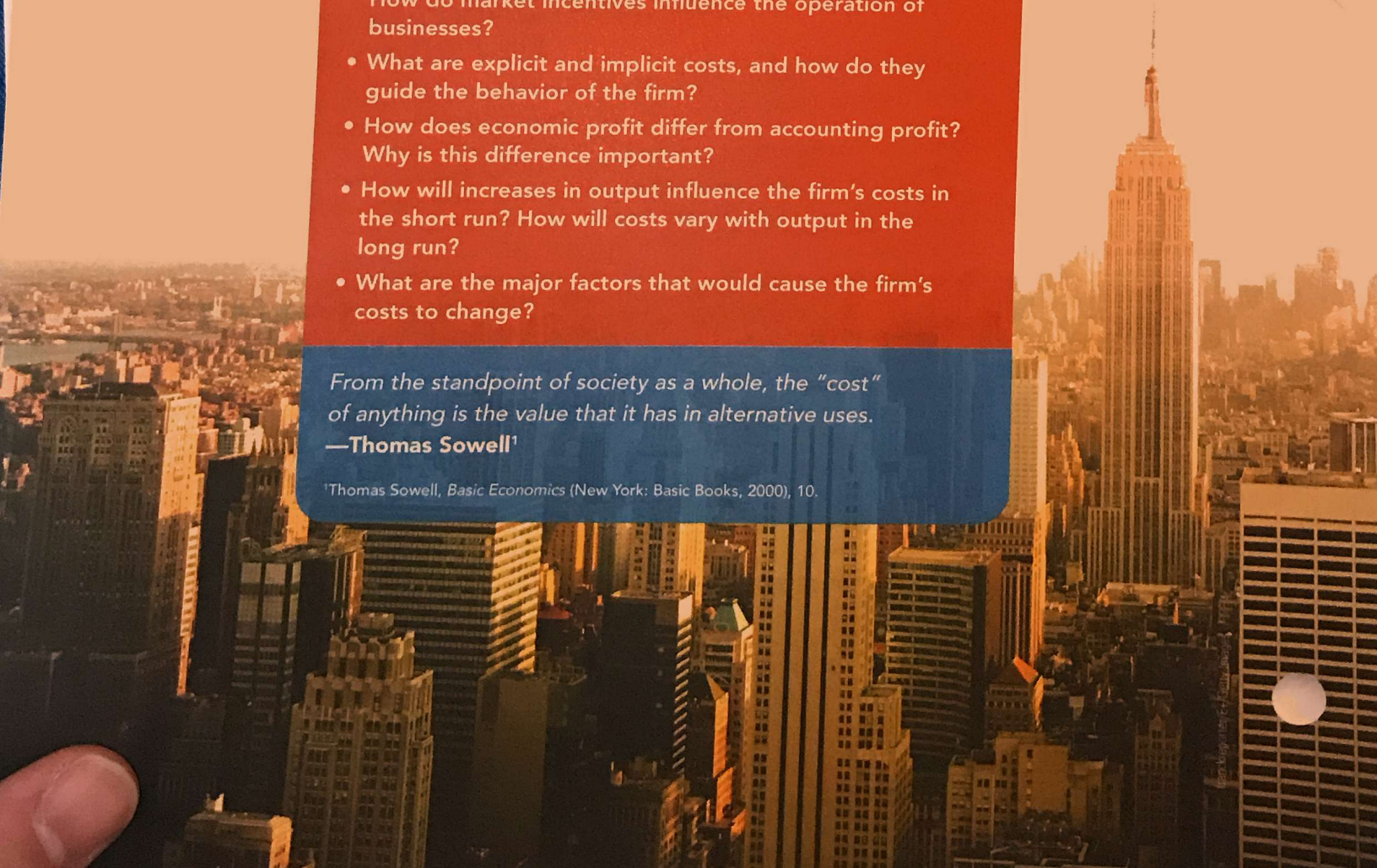
CHAPTER 8
Costs and the Supply of Goods

FOCUS

- Why are business firms used to organize production? How do market incentives influence the operation of businesses?
- What are explicit and implicit costs, and how do they guide the behavior of the firm?
- How does economic profit differ from accounting profit? Why is this difference important?
- How will increases in output influence the firm's costs in the short run? How will costs vary with output in the long run?
- What are the major factors that would cause the firm's costs to change?

From the standpoint of society as a whole, the "cost" of anything is the value that it has in alternative uses.
—Thomas Sowell[1]

[1]Thomas Sowell, *Basic Economics* (New York: Basic Books, 2000), 10.



Demand and supply interact to determine the market price of a product. In the preceding chapter, we showed that the demand for a product reflects the strength of consumer desire for that product. In this chapter, we will focus on the cost of production. The resources needed to produce one good could be used to produce other goods instead. As Thomas Sowell says in the quotation that begins this chapter, the cost to society of anything is the value that it has in alternative uses. The market price for resources makes that cost clear to producers as they must bid resources away from alternative uses. The maker of soccer balls, for example, must compete against producers of other goods when purchasing the machines, raw materials, and labor needed to produce the balls.

Costs carry an important message: They tell producers the value of the resources if left in their alternative uses. If the per-unit cost of producing a good exceeds its price, producers will suffer losses, reduce output, and some may go out of business. Only when a producer can generate enough value for consumers to allow the price to exceed production costs, will the firm be profitable and survive. This chapter lays the foundation for a detailed investigation of the links between costs, business output, and market supply.

8-1 THE ORGANIZATION OF THE BUSINESS FIRM

The business firm is an entity designed to organize raw materials, labor, and machines with the goal of producing goods and/or services. Firms (1) purchase productive resources from households and other firms, (2) transform them into a different commodity, and (3) sell the transformed product or service to consumers. In market economies, business firms choose their own price, output level, and methods of production. They not only reap the benefits of sales revenues, but they also must pay the costs of the resources they use.

8-1a INCENTIVES, COOPERATION, AND THE NATURE OF THE FIRM

In privately owned firms, owners risk their wealth on the success of the business. If the firm is successful and earns profits, these financial gains go to the owners. Conversely, if the firm suffers losses, the owners must bear the consequences.

The property right of owners to the residual income of the firm plays a very important role: It provides owners with a strong incentive to organize and operate their business in a manner that will maximize the value of their output to consumers while keeping the cost of producing output low. The wealth of these **residual claimants** is directly influenced by the success or failure of the firm. Thus, they have both the authority and a strong incentive to see that resources under their direction are used efficiently and directed toward production of goods that are valued more highly than their costs.

There are two ways of organizing productive activity: contracting and **team production**. In principle, all production could be accomplished solely through contracting. For example, a builder might have a house built by contracting with one person to pour the concrete, another to construct the wooden part of the house, a third to install the roofing, a fourth to do the electrical wiring, and so on. No employees would have to be involved in such a project. More commonly, though, goods and services are produced with some combination of contracting and the use of team production.

Team production involves the employment of workers operating under the supervision of the owner, or the owner's representative—a manager. While team production can often reduce transaction costs, it leads to another set of problems. Team members—the employees working for the firm—must be monitored and given incentives to avoid **shirking**, or working at less than the expected rate of productivity. Taking long work breaks, paying more attention to their own convenience than to work results, and wasting time when diligence is called for are examples of shirking. Hired managers, even including those at the top, must be monitored and given incentives to avoid shirking.

Residual claimants
Individuals who personally receive the excess, if any, of revenues over costs. Residual claimants gain if the firm's costs are reduced or revenues increase.

Team production
A production process in which employees work together under the supervision of the owner or the owner's representative.

Shirking
Working at less than the expected rate of productivity, which reduces output. Shirking is more likely when workers are not monitored, so that the cost of lower output falls on others.

Principal–agent problem
The incentive problem that occurs when the purchaser of services (the principal) lacks full information about the circumstances faced by the seller (the agent) and cannot know how well the agent performs the purchased services. The agent may to some extent work toward objectives other than those sought by the principal paying for the service.

Imperfect monitoring and imperfect incentives are always a problem with team production. It is part of a larger class of what economists call **principal–agent problems**. If you have ever taken a car to an auto mechanic, you have confronted this problem. The mechanic wants to get the job done quickly and make as much money as possible. The car owner not only wants to get the job done quickly also, but wants the problem fixed in a lasting way, at the lowest possible cost. Because the mechanic typically knows far more about the job than the customer, it is hard for the customer to monitor the mechanic's work. Therefore, the mechanic (the agent) may not act in the best interest of the customer (the principal).

The owner of a firm is in a similar situation. It is often difficult to monitor the performance of individual employees and motivate them to work together productively. If it is going to keep costs low and the value of output high, a firm must discover and adopt an incentive structure that motivates executives, managers, and workers to cooperate productively and discourages shirking. An effective incentive structure will align the costs and benefits faced by employees making decisions with those of the firm as a whole. Ultimately, it is the job of the owners, as residual claimants, to develop an effective incentive structure that will minimize the principal–agent problem.

8-1b THREE TYPES OF BUSINESS FIRMS

Business firms can be organized in one of three primary ways: as a proprietorship, a partnership, or a corporation. The structure chosen determines how the owners share the risks and liabilities of the firm and how they participate in making decisions.

Proprietorship
A business firm owned by an individual who possesses the ownership right to the firm's profits and is personally liable for the firm's debts.

A **proprietorship** is a business firm owned by a single individual who is fully liable for the debts of the firm. In addition to assuming the responsibilities of ownership, the proprietor often works directly for the firm, providing managerial and other labor services. Many small businesses, including restaurants, barbershops, and farms, are business proprietorships. As Exhibit 1 shows, proprietorships account for 72 percent of the business firms in the United States. Because most proprietorships are small, however, they account for only 4 percent of all business revenues.

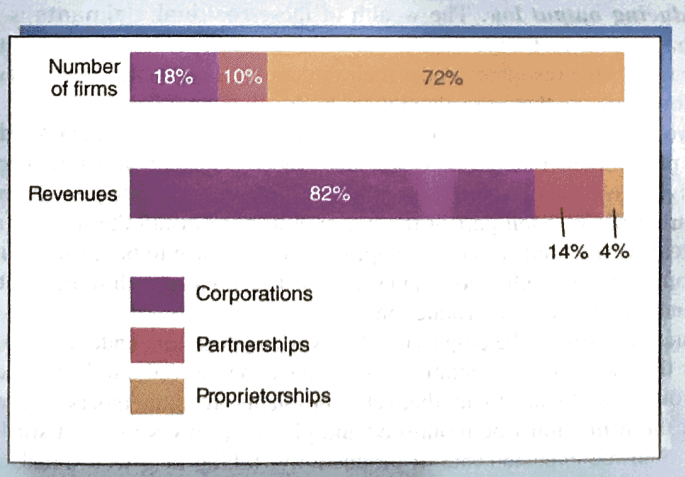
Partnership
A business firm owned by two or more individuals who possess ownership rights to the firm's profits and are personally liable for the debts of the firm.

A **partnership** consists of two or more people who are co-owners of a business firm. The partners share risks and responsibilities in an agreed-upon manner. There is no difference between a proprietorship and a partnership in terms of owner liability. In both cases, the owners are fully liable for all business debts incurred by the firm. Many law, medical,

EXHIBIT 1

How Business Firms Are Organized

Nearly three out of every four firms are proprietorships, but only 4 percent of all business revenue is generated by proprietorships. Corporations account for about one out of every five firms but generate 82 percent of all revenues.



Source: *Statistics of Income Tax Stats Integrated Business Data*, Internal Revenue Service, Table 1. (Data are for 2008.)

and accounting firms are organized along partnership lines. However, this form of business structure accounts for only 10 percent of the total number of firms and 14 percent of all business revenues.

The business firms that are **corporations** account for 82 percent of total business revenue, even though they constitute only 18 percent of all firms. What accounts for the attractiveness of this business structure? From its start, by an act of the British Parliament in 1862, the corporation, or "joint stock company," as it is also called, grew in importance for two main reasons. First, although the stockholders of the corporation are the legal owners, their liability is limited to the value of their shares of the corporation. If a corporation owes you money, you cannot directly sue the stockholders. Of course, you can sue the corporation. However, if a corporation goes bankrupt, you and others to whom the firm owes money may simply be out of luck. This limited liability makes it possible for corporations to attract investment funds from a large number of "owners" who do not participate in the day-to-day management of the firm.

Corporation
A business firm owned by shareholders who possess ownership rights to the firm's profits, but whose liability is limited to the amount of their investment in the firm.

Second, ownership can easily be transferred under the corporate structure. The shares, or ownership rights, of an owner who dies can be sold by the heirs to another owner without disrupting the business firm. Because of this, the corporation is an ongoing concern. Similarly, stockholders who become unhappy with the way a corporation is run can bail out simply by selling their stock.

While there are advantages to the corporate form of business organization, large corporations with many stockholders—millions in some cases—are also more likely to suffer from principal–agent problems. The stockholders elect a board of directors, which in turn appoints the company's high-level managers. Internal corporate policies and competition for control of the firm by outsiders can be used to reduce these principal–agent problems.

Another disadvantage of the corporate form of business organization is higher taxation. While the incomes of sole proprietors and partnerships are taxed as ordinary personal income, the income of corporations is taxed twice—once as corporate income and again as personal income when it accrues to the owners in the form of dividends or capital gains. Since 1977, an increasing number of states allow a hybrid form of business organization called the *limited liability company* (LLC), which combines the advantages of limited liability of a corporation but the tax advantages of a sole proprietorship or partnership.

8-2 HOW WELL DOES THE CORPORATE STRUCTURE WORK?

Do corporations serve the interests of consumers? Are corporate managers in a position to serve themselves at the expense of their customers and stockholders? When thinking about these questions, keep three points in mind.

First, in a market economy, firms must compete for investment funds and for customers. Investors are free to buy and sell the shares of corporations. Similarly, consumers are free to choose among firms, including both corporate and noncorporate businesses. This competitive process greatly limits the ability of managers to benefit personally at the expense of either customers or stockholders. If a business firm is not managed in a manner that serves the interests of both its shareholders and customers, they will choose other options. Moreover, investor-driven changes in a company's stock price and consumer-driven changes in sales revenue provide channels through which managerial performance can be judged, and managers can be held accountable, by shareholders and the board of directors.

Second, the compensation of managers can be structured in a manner that will bring their interests into harmony with those of the shareholder-owners. This is generally the case. The salary increases and bonuses of most high-level managers are directly related to the firm's profitability and the price of its shares. In recent years, salaries have constituted only about 10 percent of the compensation of chief executive officers (CEOs). The other 90 percent has been in the form of bonuses, often stock awards and stock options (the right to buy shares at a certain price) related to company performance.

1P3 PRESS/MAXPPP/Newscom

Corporate managers have a strong incentive to serve the interests of customers and stockholders because of the following: competition for investment funds and consumer sales, the linkage of their compensation to company performance, and the threat of takeover if the company is run poorly.

Third, the threat of corporate takeover helps keep current managers from straying too far from a profit-maximization strategy. Managers who do not serve the interests of their shareholders leave the firm vulnerable to a takeover, a move by an outside person or group to gain control of the firm. Poor management will cause market value of the firm's stock to decline, making it an attractive prospect for takeover specialists shopping for a poorly run business, the value of which could be substantially increased by a new and better management team.

The prevalence of the corporate form of business organization provides strong evidence that it is an effective form of business organization in many sectors of the economy. Rival forms of business organization, including proprietorships, partnerships, consumer cooperatives, employee ownership, and mutually owned companies, can and do compete in the marketplace for investment funds and customers. Nonetheless, the corporate structure is the dominant form of business organization (see Exhibit 1). If the corporation was not generally a cost-efficient and consumer-sensitive form of organization, this would not be the case.

8-3 THE ECONOMIC ROLE OF COSTS

Consumers would like to have more economic goods, but resources to produce them are scarce. How much of each desired good should be produced? In a market economy, consumer demand and production costs are central to performing this balancing function. ***The demand for a product represents the voice of consumers instructing firms to produce the good. Conversely, a firm's costs represent the desire of consumers not to sacrifice goods that could be produced if the same resources were employed elsewhere.*** A profit-seeking firm will try to produce only those units of output for which buyers are willing to pay full cost. Proper measurement and interpretation of costs by the firm are critical to both the firm's profitability and its efficient use of resources.

8-3a CALCULATING ECONOMIC COSTS AND PROFITS

Profit directs the actions of business firms. Profit is simply the firm's total revenue minus its total costs. But to calculate profit correctly, costs must be measured properly. Most people think of costs as amounts paid for raw materials, labor, machines, and similar inputs. However, this concept of cost, which stems from accounting procedures, excludes some important components.

The key to understanding the economist's concept of profit is to remember the idea of *opportunity cost*—the highest valued alternative forgone by the resource owner when the resource is used. These costs may be either explicit or implicit. **Explicit costs** result when the firm makes a monetary payment to resource owners. Money wages, interest, and rental payments are a measure of what the firm gives up to employ the services of labor and capital resources. These are relatively easy to track. But firms also incur **implicit costs**—those associated with the use of resources owned by the firm. For example, the owners of small proprietorships often work for their own businesses, for little or no pay. These businesses incur an implicit cost—an opportunity cost—associated with the use of this resource (the owners' labor services). The highest valued alternative forgone in this case is the maximum amount of money the owners could have earned doing something else. The **total cost** of production is the sum of these explicit and implicit costs incurred by the employment of all resources involved in the production process.

Explicit costs
Payments by a firm to purchase the services of productive resources.

Implicit costs
The opportunity costs associated with a firm's use of resources that it owns. These costs do not involve a direct money payment. Examples include wage income and interest forgone by the owner of a firm who also provides labor services and equity capital to the firm.

Total cost
The costs, both explicit and implicit, of all the resources used by the firm. Total cost includes a normal rate of return for the firm's equity capital.

Accounting statements also generally omit the implicit cost of equity capital—the cost of funds supplied by the firm's owners. If a firm borrows financial capital from a bank or other private source, it will have to pay interest. Accountants properly record this interest expense as a cost. In contrast, when the firm acquires financial capital by issuing shares of stock, accountants don't record this as an expense. Essentially, this is because the stockholders *are* the firm's owners. Either way, acquiring capital has an opportunity cost. Banks will demand interest payments, and shareholders will expect a return from their investment in the form of dividend payments or rising share value.

Economists use the normal return on financial capital as a basis for determining the implicit **opportunity cost of equity capital**. If the normal rate of return on financial capital is 10 percent, for example, investors will not continue to supply equity capital unless they can earn this normal return. Thus, it is an opportunity cost of equity capital.

Opportunity cost of equity capital
The rate of return that must be earned by investors to induce them to supply financial capital to the firm.

8-3b HOW DO ECONOMIC AND ACCOUNTING PROFIT DIFFER?

Economic profit is total revenues minus total costs, including both the explicit and implicit cost components. Economic profit will be positive only if the earnings of the business exceed the opportunity cost of all the resources used by the firm, *including the opportunity cost of assets owned by the firm and any unpaid labor services supplied by the owner*. In contrast, economic losses result when the earnings of the firm are insufficient to cover explicit and implicit costs. That is why the **normal profit rate** is zero economic profit, yielding just the competitive rate of return on the capital (and labor) of owners. A higher rate would draw more competitors and their investors into the market; a lower rate would cause competitors and their investors to exit the market.

Economic profit
The difference between the firm's total revenues and its total costs, including both the explicit and implicit cost components.

Normal profit rate
Zero economic profit, providing just the competitive rate of return on the capital (and labor) of owners. An above-normal profit will draw more entry into the market, whereas a below-normal profit will lead to an exit of investors and capital.

Remember, zero economic profits do not imply that the firm is about to go out of business. On the contrary, they indicate that the owners are receiving exactly the normal profit rate, or the competitive market rate of return on their investment.

Whenever accounting procedures omit implicit costs, like those associated with owner-provided labor services or equity capital, the firm's opportunity costs of production will be understated. This understatement of cost leads to an overstatement of profits. Therefore, the **accounting profits** of a firm are generally greater than the firm's economic profits (see the Applications in Economics feature on accounting costs). For most large corporations, though, omitting the implicit costs of services provided by an owner isn't an issue. In this case, the accounting profits approximate the returns to the firm's equity capital. High accounting profits (measured as a rate of return on a firm's assets), relative to those of other firms, suggest that a firm is earning an economic profit. Correspondingly, a low rate of accounting profit implies economic losses. Either positive or negative economic profits, of course, call for a change in output. Such a change, however, will take time.

Accounting profits
The sales revenues minus the expenses of a firm over a designated time period, usually one year. Accounting profits typically make allowances for changes in the firm's inventories and depreciation of its assets. No allowance is made, however, for the opportunity cost of the equity capital of the firm's owners, or other implicit costs.

APPLICATIONS IN ECONOMICS

Economic and Accounting Costs: A Hypothetical Example

The revenue–cost statement for a corner grocery store owned and operated by Emily Blake is presented here.

TOTAL REVENUE	
Sales (groceries)	$170,000
Costs (explicit)	
Groceries, wholesale	$76,000
Utilities	4,000
Taxes	6,000
Advertising	2,000
Labor services (employees)	12,000
Total (explicit) costs	$100,000
Net (accounting) profit	$70,000
Additional (implicit) costs	
Interest (personal investment)	$7,000
Rent (Emily's building)	18,000
Salary (Emily's labor)	50,000
Total (implicit) costs	$75,000
TOTAL EXPLICIT AND IMPLICIT COSTS	$175,000
ECONOMIC PROFIT (TOTAL REVENUE MINUS EXPLICIT AND IMPLICIT COSTS)	–$5,000

Emily works full-time as the manager, chief cashier, and janitor. She has $140,000 worth of refrigeration and other equipment invested in the store. Last year, her total sales were $170,000, and suppliers and employees were paid $100,000. Emily's revenues therefore exceeded explicit costs by $70,000. This is what was recorded on the accounting statements as profit.

But did Emily really make a profit last year? Let's look at her opportunity costs and see: If Emily didn't have $140,000 of her own money invested in equipment, she could be earning 5 percent interest on the money in the bank, which would add up to $7,000 each year. Similarly, if the building she owns weren't being used as a grocery store, it could be rented to someone else for $1,500 per month. Rental income forgone is therefore $18,000 per year. In addition, because Emily is tied up working in the grocery store, a $50,000 managerial position she could hold at another local grocery store is forgone. Considering the interest, rental, and salary income that Emily had to forgo to operate the grocery store last year, her implicit costs were $75,000. This makes her total costs—both explicit and implicit—$175,000. (Recall that explicit costs were $100,000.) That's $5,000 *less* than her actual revenues of $170,000. As a result, Emily incurred an economic loss of $5,000, despite the accounting profit of $70,000 recorded on the store's books.

8-4 SHORT-RUN AND LONG-RUN TIME PERIODS

Short run (in production)
A time period so short that a firm is unable to vary some of its factors of production. The firm's plant size typically cannot be altered in the short run.

Time plays an important role in the production process. All of a firm's resources can be expanded (or contracted) over time, but for specialized equipment, expanding (and contracting) availability quickly is likely to be very expensive or even impossible. Economists often speak of the **short run** as a time period so short that the firm is unable to alter its present plant size. In the short run, the firm is typically stuck with its existing plant and heavy equipment. These assets are "fixed" for a given time period, in other words. The firm can alter output, however, by applying larger or smaller amounts of "variable" resources, like labor and raw materials. In this way, the existing plant capacity can be used more or less intensively in the short run.

How long is the short run? *The short run is that period of time during which at least one factor of production, usually the size of the firm's plant, cannot be changed.* The length varies across industries and firms. A trucking firm might be able to hire more drivers and buy or rent more trucks and double its hauling capacity in a few months. In other industries, particularly those that use assembly lines and mass-production techniques, increasing production capacity may take several years.

Long run (in production)
A time period long enough to allow the firm to vary all of its factors of production.

The **long run** is a time period long enough for existing firms to alter the size of their plants and for new firms to enter (or exit) the market. *All of the firm's resources are variable in the long run.* In the long run, firms can expand their output by increasing the sizes of their plants—perhaps by adding on to them or by constructing entirely new facilities.

An example may help you understand the distinction between the short- and long-run time periods: If a tablet computer manufacturer hired 200 additional workers and ordered

more raw materials in order to squeeze more production out of its existing plant, it would be making a short-run adjustment. In contrast, if the manufacturer built an additional plant (or expanded the size of its current facility) and installed additional heavy equipment, it would be making a long-run adjustment.

8-5 CATEGORIES OF COSTS

In the short run, we can break a firm's costs into two categories—fixed and variable. Each category of costs behaves differently. Seeing that behavior graphically will help us understand how the profit-maximizing level of the firm's output is determined. It will also be important to distinguish between a firm's total costs and its per-unit costs, which we call "average" costs.

Each of the firm's fixed costs, and their sum, called **total fixed cost** (*TFC*), will remain unchanged when output rises or falls in the short run. For example, a firm's insurance premiums; its property taxes; and, most significantly, the opportunity cost of using its fixed assets will be present whether the firm produces a large or small amount of output. These costs will not vary with output. They can be avoided only if the firm goes out of business.

What will happen to **average fixed cost** (*AFC*), which is fixed cost *per unit*, as output expands? Remember that the firm's fixed cost will be the same whether output is 1, or 100, or 1,000. The *AFC* is simply fixed cost divided by output. As output increases, *AFC* declines because the fixed cost will be spread over more and more units (see part a of Exhibit 2).

Some costs vary with output. For example, additional output can usually be produced by hiring more workers and buying more raw materials. The sum of those and other costs that rise as output increases is the firm's **total variable cost** (*TVC*). At any given level of output, the firm's **average variable cost** (*AVC*), which is variable cost *per unit*, is the total variable cost divided by output.

A firm's total cost (*TC*) is the sum of the fixed and variable costs. At zero output a firm has no variable costs, thus total cost will be equal to total fixed cost. As output expands from zero, variable costs begin to increase, causing total cost to rise with output even though fixed costs are remaining unchanged. **Average total cost** (*ATC*), sometimes referred to as unit cost, can be found by dividing total cost by the total number of units produced. *ATC* is also equal to the sum of the average fixed and average variable costs. One way to look at ATC is that it is the amount of revenue needed per unit of output to cover total cost.

The economic way of thinking focuses on what happens "at the margin." How much does it cost to produce an additional unit? **Marginal cost** (*MC*) is the change in

Total fixed cost
The sum of the costs that do not vary with output. They will be incurred as long as a firm continues in business and the assets have alternative uses.

Average fixed cost
Total fixed cost divided by the number of units produced. It always declines as output increases.

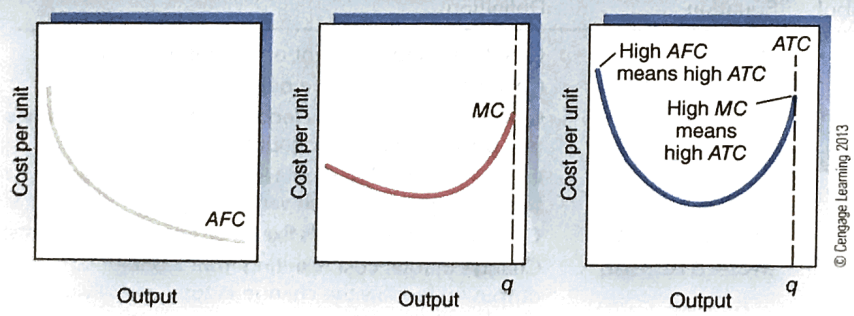
Total variable cost
The sum of those costs that rise as output increases. Examples of variable costs are wages paid to workers and payments for raw materials.

Average variable cost
The total variable cost divided by the number of units produced.

Average total cost
Total cost divided by the number of units produced. It is sometimes called per-unit cost.

Marginal cost
The change in total cost required to produce an additional unit of output.

EXHIBIT 2



The General Characteristics of Short-Run Cost Curves

Average fixed costs (a) will be high for small rates of output, but they will always decline as output expands. Marginal cost (b) will rise sharply as the plant approaches its production capacity, q. As graph (c) shows, ATC will be a U-shaped curve because AFC will be high for small rates of output, and MC will be high as the plant's production capacity is approached.

total cost that results from the production of one additional unit. The profit-conscious decision-maker recognizes *MC* as the addition to cost that must be covered by additional revenue if producing the marginal unit is to be profitable. In the short run, as illustrated by part (b) of Exhibit 2, *MC* will generally decline at first if output is increased, reach a minimum, and then increase. The rising *MC* simply reflects the fact that it becomes increasingly difficult to squeeze additional output from a plant as the facility's maximum capacity (the dotted line of part b of Exhibit 2) is approached. The accompanying **Thumbnail Sketch** summarizes how the firm's various costs are related to one another.

8-6 OUTPUT AND COSTS IN THE SHORT RUN

As a firm changes its rate of output in the short run, how will its unit cost be affected? In the short run, the firm can vary its output by using its fixed plant size more (or less) intensively. Exhibit 2 shows two ways that this can result in high unit costs. First, when the output rate of a plant is small relative to its capacity, the facility is being underutilized, causing *AFC* to be high and *ATC* to be high, too. It will be costly to operate a large plant, with its high fixed costs, substantially below its production capacity. Alternatively, overutilization can also cause high unit costs. An overutilized plant will mean congestion—time spent by workers waiting for machines and similar costly delays. Requiring output beyond the least-cost, or designed, output level of a plant will lead to high *MC* and therefore to high *ATC*.

Thus, the ATC curve will be U-shaped, as pictured in part (c) of Exhibit 2. ATC will be high for both an underutilized plant and an overutilized plant. It will be high for an underutilized plant because average fixed cost will be high. It will be high for an overutilized plant because marginal cost will be high.

ically, in

Thumbnail Sketch

Compact Glossary on Cost

Term	Symbol	Equation	Definition
Fixed cost			Cost that is independent of the output level
Variable cost			Cost that varies with the output level
Total fixed cost	*TFC*		Cost of the fixed inputs (equals sum of quantity times unit price for each fixed input)
Total variable cost	*TVC*		Cost of the variable inputs (equals sum of quantity times unit price for each variable input)
Total cost	*TC*	$TC = TFC + TVC$	Cost of all inputs (equals fixed costs plus variable costs)
Marginal cost	*MC*	$MC = \Delta TC \div \Delta q$	Change in total cost resulting from a one-unit rise in output (q) [equals the change in total cost divided by the change in output]
Average fixed cost	*AFC*	$AFC = TFC \div q$	Total fixed cost per unit of output (equals total fixed cost divided by total output)
Average variable cost	*AVC*	$AVC = TVC \div q$	Total variable cost per unit of output (equals total variable cost divided by total output)
Average total cost	*ATC*	$ATC = AFC + AVC$	Total cost per unit of output (equals average fixed cost plus average variable cost)

8-6a DIMINISHING RETURNS AND PRODUCTION IN THE SHORT RUN

Our analysis of the changes in unit cost as the output rate rises reflects a long-established economic law. This **law of diminishing returns** states that, as more and more units of a variable factor are applied to a fixed amount of other resources, output will eventually increase by smaller and smaller amounts. Therefore, the impact on output of additional units of the variable factor will diminish. The cost per unit of adding the variable factor may be the same, but the added output per dollar spent falls. The impact on cost per unit of output is clear: When the returns to the variable factor are rising, marginal costs (the additions to total variable cost needed to add a unit of output) are falling. Similarly, when the returns to the variable factor are falling, marginal cost is rising.

The law of diminishing returns is as famous in economics as the law of gravity is in physics. It is based on common sense and real-life observation. Have you ever noticed that as you apply a single resource more intensively, the resource eventually tends to accomplish less and less? Consider a wheat farmer who applies fertilizer (a variable resource) more and more intensively to an acre of land (a fixed factor). At some point, the application of additional 100-pound units of fertilizer will expand the wheat yield by successively smaller amounts.

Essentially, the law of diminishing returns is a constraint imposed by nature. If it were not valid, it would be possible to raise all the world's food in a flowerpot. We would be able to increase output simply by applying another unit of labor and fertilizer to the world's most fertile flowerpot! In the real world, of course, this is not the case; the law of diminishing returns is valid, reflecting a constraint we all must face.

Exhibit 3 illustrates the law of diminishing returns numerically. Column 1 indicates the quantity of the variable resource, labor in this example, which is combined with a specified amount of the fixed resource. Column 2 shows the **total product** that will result as the utilization rate of labor increases. Column 3 provides data on the **marginal product**, the change in total output associated with each additional unit of labor. Without the application of labor, output will be zero. As additional units of labor are applied, total product (output) rises. As the first three units of labor are applied, total product increases by successively larger amounts (8, then 12, then 14). Beginning with the fourth unit, however, diminishing returns

Law of diminishing returns
The postulate that as more and more units of a variable resource are combined with a fixed amount of other resources, using additional units of the variable resource will eventually increase output only at a decreasing rate. Once diminishing returns are reached, it will take successively larger amounts of the variable factor to expand output by one unit.

Total product
The total output of a good that is associated with each alternative utilization rate of a variable input.

Marginal product
The increase in the total product resulting from a unit increase in the employment of a variable input. Mathematically, it is the ratio of the change in total product to the change in the quantity of the variable input.

EXHIBIT 3

The Law of Diminishing Returns (Hypothetical Data)

(1) Units of the Variable Resource, Labor (per day)	(2) Total Product (output)	(3) Marginal Product	(4) Average Product
0	0		—
		8	
1	8		8.0
		12	
2	20		10.0
		14	
3	34		11.3
		12	
4	46		11.5
		10	
5	56		11.2
		8	
6	64		10.7
		6	
7	70		10.0
		4	
8	74		9.3
		1	
9	75		8.3
		−2	
10	73		7.3

are confronted. When the fourth unit of labor is added, marginal product—the change in the total product—declines to 12 (down from 14, when the third unit was applied). As additional units of labor are applied, marginal product continues to decline. It is increasingly difficult to squeeze a larger total product from the fixed resources (for example, plant size and equipment). Eventually, marginal product becomes negative (beginning with the tenth unit).

Average product
The total product (output) divided by the number of units of the variable input required to produce that output level.

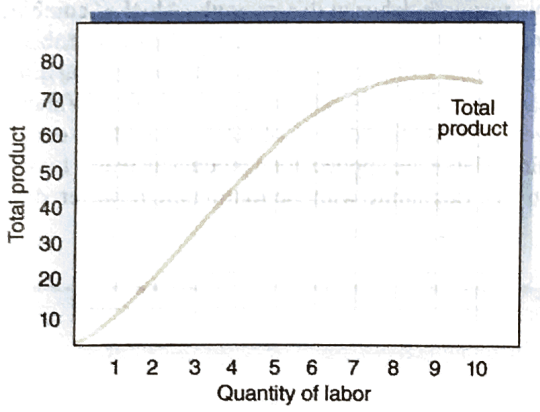
Column 4 of Exhibit 3 provides data for the **average product** of labor, which is simply the total product divided by the units of labor applied. Note that the average product increases as long as the marginal product is greater than the average product. Whenever the marginal unit's contribution is greater than the average, it must cause the average to rise. (A good analogy would be your grade point average. If the grade you get in this course is higher than your overall grade point average, your grade in this class will increase your overall average.) Here, marginal product rises through the first four units. The marginal product of the fifth unit of labor, though, is 10, less than the average product for the first four units of labor (11.5). Therefore, beginning with the fifth unit, the average product declines as additional labor is applied. When marginal productivity is below the average, it brings down the average product.

Using the data from Exhibit 3, Exhibit 4 illustrates the law of diminishing returns graphically. Initially, the total product curve (part a) increases quite rapidly. As diminishing marginal returns are confronted (beginning with the fourth unit of labor), total product increases more slowly. Eventually, a maximum output (75) is reached with the application of the ninth unit of labor. The marginal product curve (part b) reflects the total product curve. Geometrically, marginal product is the slope—the rate of increase—of the total product

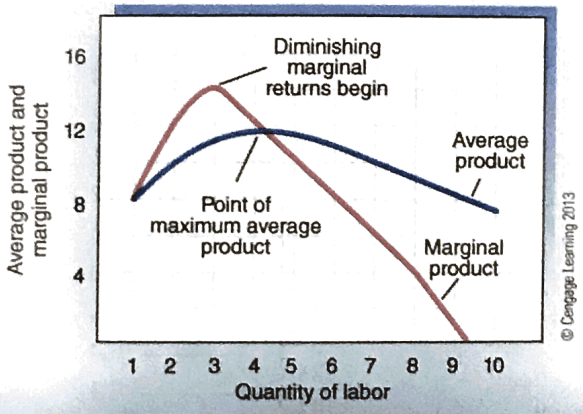
EXHIBIT 4

The Law of Diminishing Returns

As units of variable input (labor) are added to a fixed input, total product will increase, first at an increasing rate and then at a declining rate (a). This will cause both marginal and average product curves (b) to rise at first and then decline. Note that the marginal product curve intersects the average product curve at its maximum (when four units of labor are used). The smooth curves indicate that labor can be increased by amounts of less than a single unit.



(a) Total product curve



(b) Average and marginal product curve

curve. That slope, the marginal product, reaches its maximum here with the application of three units of labor. Beyond three units, diminishing returns are present. Eventually, at ten units of labor, the marginal product becomes negative. When marginal product becomes negative, total product is necessarily declining. The average product curve rises as long as the marginal product curve is above it, because each added unit of labor is raising the average. The average product reaches its maximum at four units of labor. Beyond that, each additional unit of labor brings down the average product, and the curve slopes downward.

8-6b DIMINISHING RETURNS AND THE SHAPE OF THE COST CURVES

What impact will diminishing returns have on a firm's costs? Once a firm confronts diminishing returns, larger and larger additions of the variable factor are required to expand output by one unit. This will cause marginal cost (MC) to rise. As MC continues to increase, eventually it will exceed average total cost. Until that point, MC is below ATC, bringing ATC down. When MC is greater than ATC, the additional units cost more than the average, and ATC must increase. Think about what happens when you get a grade on an exam above your current class average. Your class average goes up. What happens if a unit of above-average cost is added to output? Average total cost rises. The firm's *MC* curve therefore crosses the *ATC* curve at the *ATC*'s lowest point. For output rates beyond the minimum *ATC*, the rising *MC* causes *ATC* to increase.

Exhibit 5 numerically illustrates the effect of the law of diminishing returns on a firm's short-run cost curve. Here, we assume that Royal Roller Blades, Inc., combines units of a variable input with a fixed factor to produce units of output (pairs of inline skates). Columns 2, 3, and 4 indicate how the total cost schedules vary as output is expanded. Total fixed costs (*TFC*), representing the opportunity cost of the fixed factors of production, are $50 per day at all levels of output. For the first four units of output, total variable costs (*TVC*) increase at a decreasing rate—by $15 with the production of the first unit, $10 with the production of the second unit, $9 with the third, and so on. Why? In this range, there are increasing returns to the variable input. Beginning with the fifth unit of output, however, diminishing marginal returns are present. From this point on, *TVC* and *TC* increase by successively larger amounts as output is expanded.

EXHIBIT 5

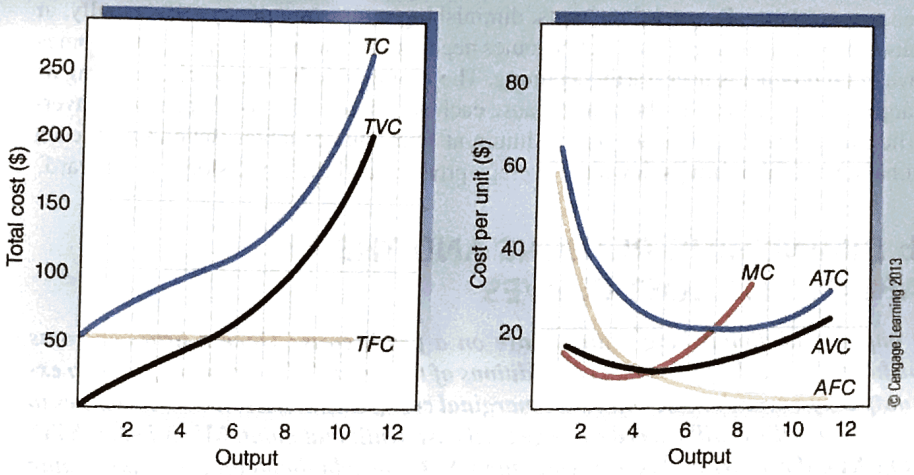
The Numerical Short-Run Cost Schedules of Royal Roller Blades, Inc.

	Total Cost Data (Per Day)			Average/Marginal Cost Data (Per Day)			
(1) Output Per Day	(2) TFC	(3) TVC	(4) TC (2) + (3)	(5) AFC (2) ÷ (1)	(6) AVC (3) ÷ (1)	(7) ATC (4) ÷ (1)	(8) MC Δ(4) ÷ Δ(1)
0	$50	$0	$50	—	—	—	—
1	50	15	65	$50.00	$15.00	$65.00	$15
2	50	25	75	25.00	12.50	37.50	10
3	50	34	84	16.67	11.33	28.00	9
4	50	42	92	12.50	10.50	23.00	8
5	50	52	102	10.00	10.40	20.40	10
6	50	64	114	8.33	10.67	19.00	12
7	50	79	129	7.14	11.29	18.43	15
8	50	98	148	6.25	12.25	18.50	19
9	50	122	172	5.56	13.56	19.11	24
10	50	152	202	5.00	15.20	20.20	30
11	50	202	252	4.55	18.36	22.91	50

EXHIBIT 6

Costs in the Short Run

Using data from Exhibit 5, this exhibit shows the general shape of the firm's short-run total cost curves (a), and average and marginal cost curves (b). Note that when output is small (for example, two units), *ATC* will be high because the *AFC* is so high. Similarly, when output is large (for example, eleven units), per-unit cost (*ATC*) will be high because additional units will be extremely costly to produce at this point. Thus, the short-run *ATC* curve will be U-shaped.



(a) Total cost data

(b) Average and marginal cost data

Columns 5 through 8 of Exhibit 5 are the average and marginal cost schedules. For small output rates, the *ATC* of producing roller blades is high, primarily because of the high *AFC*. Initially, *MC* is less than *ATC*, so *ATC* is falling. When diminishing returns set in for output rates beginning with five units, however, *MC* rises. Beginning with the sixth unit of output, *MC* exceeds *AVC*, causing *AVC* to rise. Beginning with the eighth unit of output, *MC* exceeds *ATC*, causing it also to rise. *ATC* thus reaches its minimum at seven units of output. Look carefully at the data of Exhibit 5 to be sure that you fully understand the relationships among the various cost curves. Do you understand how columns 4 to 8 are derived from columns 1 to 3?

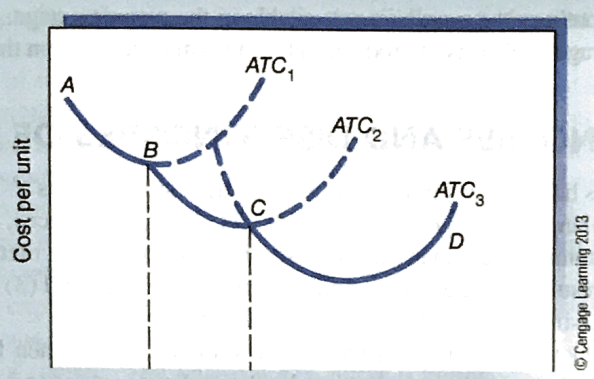
Using the numeric data of Exhibit 5, Exhibit 6 graphically illustrates the total, the average, and the marginal cost curves. Note that the *MC* curve intersects both the *AVC* and *ATC* curves at their minimum points (part b). As *MC*, driven up by diminishing returns, continues to rise above *ATC*, unit costs rise higher and higher as output increases beyond seven units.

In sum, the firm's short-run cost curves reflect the law of diminishing marginal returns. Assuming that the price of the variable resource is constant, *MC* declines so long as the marginal product of the variable input is rising. This results because, in this range, smaller and smaller additions of the variable input are required to produce each extra unit of output. The situation is reversed, however, when diminishing returns are confronted. Once diminishing returns set in, more and more units of the variable factor are required to generate each additional unit of output. *MC* will rise because the marginal product of the variable resource is declining. Eventually, *MC* exceeds *AVC* and *ATC*, causing these costs also to rise. A U-shaped, short-run average total cost curve results.

8-7 OUTPUT AND COSTS IN THE LONG RUN

The short-run analysis relates costs to output *for a specific size of plant*. Firms, though, are not committed forever to their existing plants. In the long run, all resources used by the firm are variable, thus there are no long-run fixed costs.

EXHIBIT 7



Long-Run Average Total Cost

The short-run average total cost curves are shown for three alternative plant sizes. If these three were the only possible plant sizes, the long-run average total cost curve would be *ABCD*.

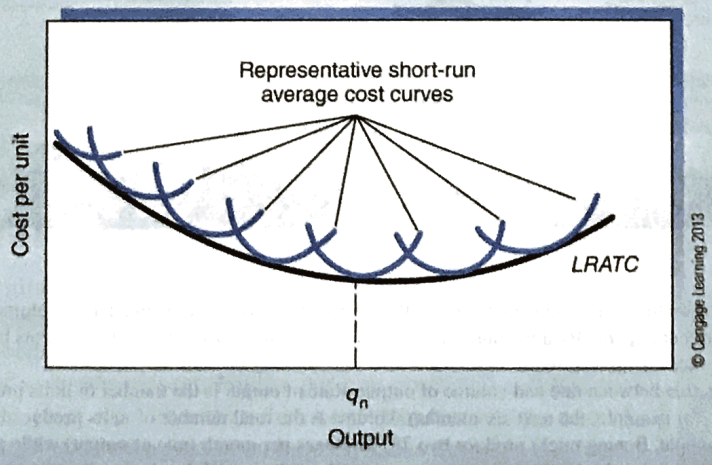
How will the firm's choice of plant size affect per-unit production costs? Exhibit 7 illustrates the short-run *ATC* curves for three different plant sizes, ranging from small to large. If these three plant sizes were the only possible choices, which one should the firm choose as it plans for the future? The answer depends on the rate of output the firm expects to produce from the plant. The smallest plant would have the lowest cost if an output rate of less than q_1 were produced. The medium-size plant would provide the least cost method of producing output rates between q_1 and q_2. For any output level greater than q_2, the largest plant would be the most cost efficient.

The long-run* ATC *curve shows the minimum average cost of producing each output level when the firm is free to choose among all possible plant sizes. It can best be thought of as a planning curve because it reflects the expected per-unit cost of producing alternative rates of output while plants are still in the blueprint stage.

Exhibit 7 illustrates the long-run *ATC* curve when only three plant sizes are possible, and the planning curve *ABCD* is thus mapped out. Of course, given sufficient time, firms can usually choose among many plants of various sizes. Exhibit 8 presents the long-run planning curve under these circumstances. It is a smooth curve, with each short-run *ATC* curve tangent to it.

It is important to keep in mind that no single plant size could produce the alternative output rates at the costs indicated by the planning curve *LRATC* in Exhibit 8. Any of the planning curve options are, of course, available to the firm before a plant size is chosen and the plant is built.

EXHIBIT 8



The Planning Curve (LRATC)

When many different plant sizes are possible, the long-run average total cost curve (*LRATC*) can be mapped out. When firms are able to plan large volumes of output, using mass-production methods will generally lead to lower per-unit costs. This helps explain why the *LRATC* has a downward-sloping portion.

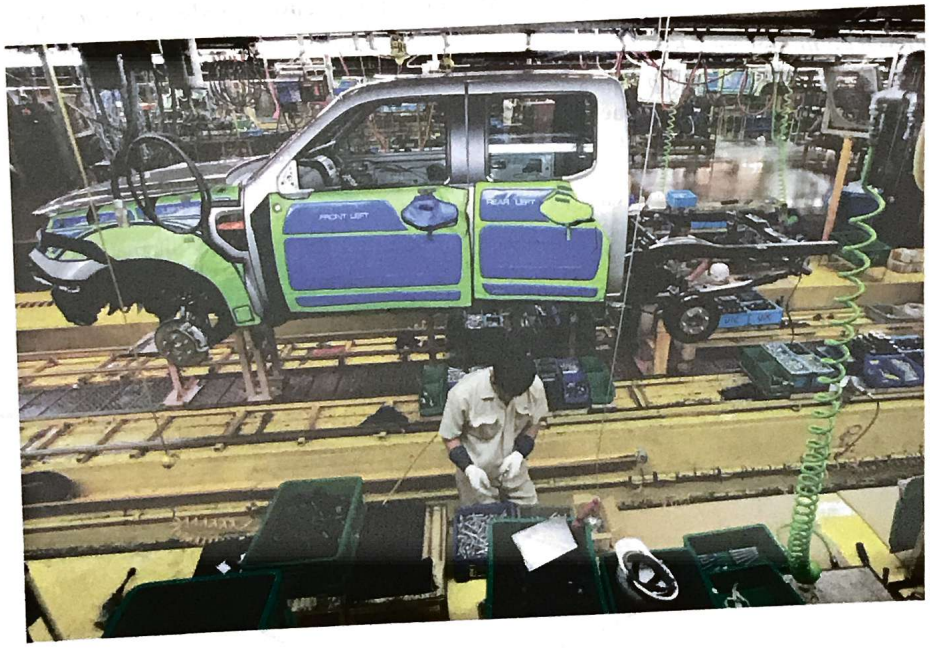
But it can *operate* in the short run only *after* a plant size has been chosen and put in place. The *LRATC* curve outlines the *possibilities* available in the planning stage. It shows the expected output and average total costs of production for the firm depending on the plant size it chooses.

8-7a ECONOMIES AND DISECONOMIES OF SCALE

Do larger firms have lower minimum unit costs than smaller ones?[2] There are three major reasons why planning a larger volume of output generally reduces, at least initially, unit costs: (1) economies accompanying the use of mass-production methods, (2) higher productivity as a result of specialization and "learning by doing," and (3) economies in promotion and purchasing.[3] Let's consider each of these factors.

Mass-production techniques usually are economical only when large volumes of output are planned, because they tend to involve large development and setup costs. Once the production methods are established, though, marginal costs are low. For example, the use of molds, dies, and assembly line production methods reduce the per-unit cost of automobiles only when the planned volume of output is in the millions. High-volume methods, although cheaper to use for high rates of output and high volumes, will typically require high fixed costs and therefore cause unit costs to be far higher for low volumes of production.

Large-scale operation also allows the specialized use of labor and machines. In a giant auto plant, hundreds of different jobs must be done, and many of them require a training period for each worker. In a small plant, a single worker might do ten or twenty of these jobs, so each worker would have a much longer, more costly training period. Even then, the worker doing so many tasks might never fully develop the same level of



REUTERS/SUKREE SUKPLANG/LANDOV

Mass-production methods can often reduce average costs for firms planning higher output volumes.

[2]Throughout this section, we assume that firms with larger plants necessarily plan a larger volume of output than do their smaller counterparts. Reality approximates these conditions. Firms choose large plants because they are planning to produce a large volume.

[3]Note the distinction between rate and volume of output. Rate of output is the number of units produced during a specific period (for example, the next six months). Volume is the total number of units produced during all time periods. For example, Boeing might produce two 787 airplanes per month (rate of output) while planning to produce a volume of two hundred 787s during the expected life of the model. Increasing the rate (reducing the time period during which a given output is produced) tends to raise costs, whereas increasing the volume (total amount produced) tends to lower costs per unit.