CHAPTER 15

The Demand for Factors of Production

In the preceding 14 chapters, we have studied markets for consumption goods. In this and the next two chapters, we will study markets for factors of production (also called inputs). Factors of production, such as labor and capital, are supplied by individual households and demanded by firms, which use them to produce output for consumption. In this chapter, we will study the firm's demand for inputs.

Firms demand inputs only because they can be used to produce output. Therefore, the value of those inputs depends on conditions in the output market. For example, a farmer's demand for fertilizer depends on the price at which he can sell his crops. The need to take account of conditions in the output market means that the derivation of the firm's demand for factors will be more subtle than the derivation of the consumer's demand for consumption goods.

The firm's income is paid out to the various factors of production. Workers receive wages, the owners of capital receive rental payments for the use of their facilities, and so forth. In the last section of this chapter, we will use our understanding of the firm's factor demand curves to see what determines how the firm's income is distributed.

15.1 The Firm's Demand for Factors in the Short Run

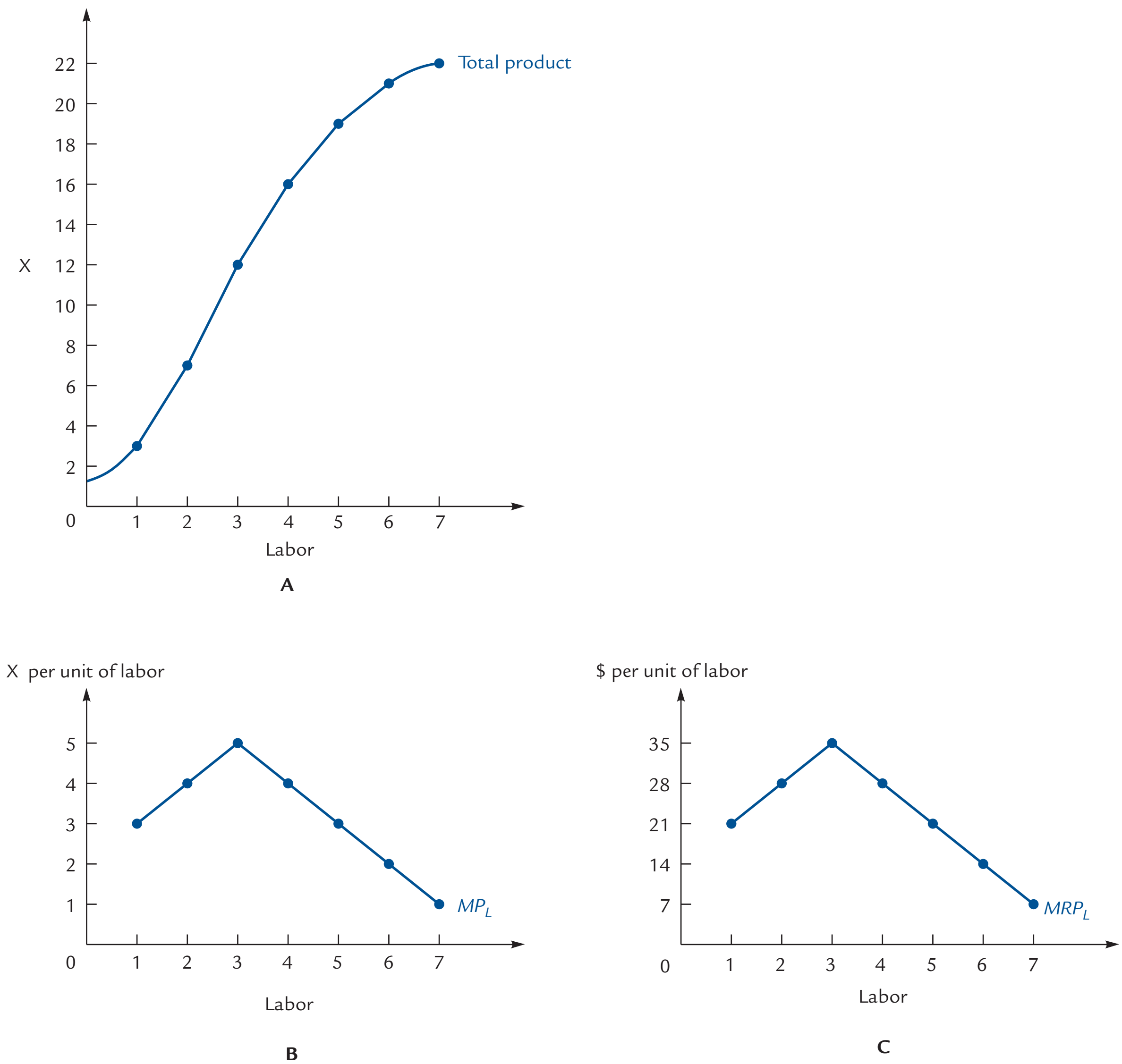
In the short run, only one factor of production is variable, and we will assume that factor to be labor. Thus, we will study the demand for labor on the assumption that the firm uses some fixed quantity of capital.

The Marginal Revenue Product of Labor

Recall from Chapter 6 that the total and marginal product of labor curves are typically shaped like those in the first two panels of Exhibit 15.1. We will also be interested in the **marginal revenue product of labor** ***(MRP_L)***, defined as the additional revenue earned by the firm when one additional unit of labor is employed. The marginal revenue product of labor is measured in dollars per unit of labor, whereas the marginal product of labor is measured in units of output per unit of labor.

Marginal revenue product of labor (MRP_L)
The additional revenue that a firm earns when it employs one more unit of labor.

EXHIBIT 15.1 **The Total, Marginal, and Marginal Revenue Products of Labor**



The total product and marginal product of labor (MP_L) curves are as in Exhibit 6.1. The marginal product of labor increases until diminishing marginal returns set in at $L = 3$, and it decreases thereafter. If the firm is competitive and sells its output at \$7 per unit, then the marginal revenue product of labor (MRP_L) is given by

$$MRP_L = \$7 \times MP_L$$

Thus, the MRP_L curve can be constructed from the MP_L curve by simply changing the units on the vertical axis, as shown in panel C.

For a firm in a competitive industry, selling output at a going price P_X, the marginal revenue product of labor is given by:

$$MRP_L = P_X \cdot MP_L$$

Given the MP_L curve from Exhibit 15.1 and given the price of output (say, \$7 per unit), we can construct the MRP_L curve simply by changing the units on the vertical axis. We have done so in panel C of the exhibit.

Exercise 15.1 If the firm in question were a monopolist in the output market, how would the MRP_L curve differ?

Suppose that the firm can hire labor at a going wage rate of \$25 per unit of labor. How much labor will it hire? As long as additional units of labor yield marginal revenue products in excess of \$25, it will continue hiring. As soon as the MRP_L reaches \$25, it will stop. Therefore, we see from Exhibit 15.1 that the firm will hire 4½ units of labor. In general, at any given wage rate, the firm will want to hire a quantity of labor read from the downward-sloping portion of the MRP_L curve. We can summarize this by saying

The firm's short-run demand curve for labor coincides with the downward-sloping portion of the MRP_L curve.

The Algebra of Profit Maximization

The amount of labor needed to produce one more unit of output is $1/MP_L$. The cost of that labor is the price per unit of labor (P_L) times the quantity of labor ($1/MP_L$), or P_L/MP_L. Therefore, the marginal cost of producing another unit of output is given by:

$$MC = \frac{P_L}{MP_L}$$

When firms maximize profit, they set the price of output P_X equal to marginal cost, or:

$$P_X = MC$$

Combining the two displayed equations, we find that profit maximization requires:

$$P_X = \frac{P_L}{MP_L}$$

or

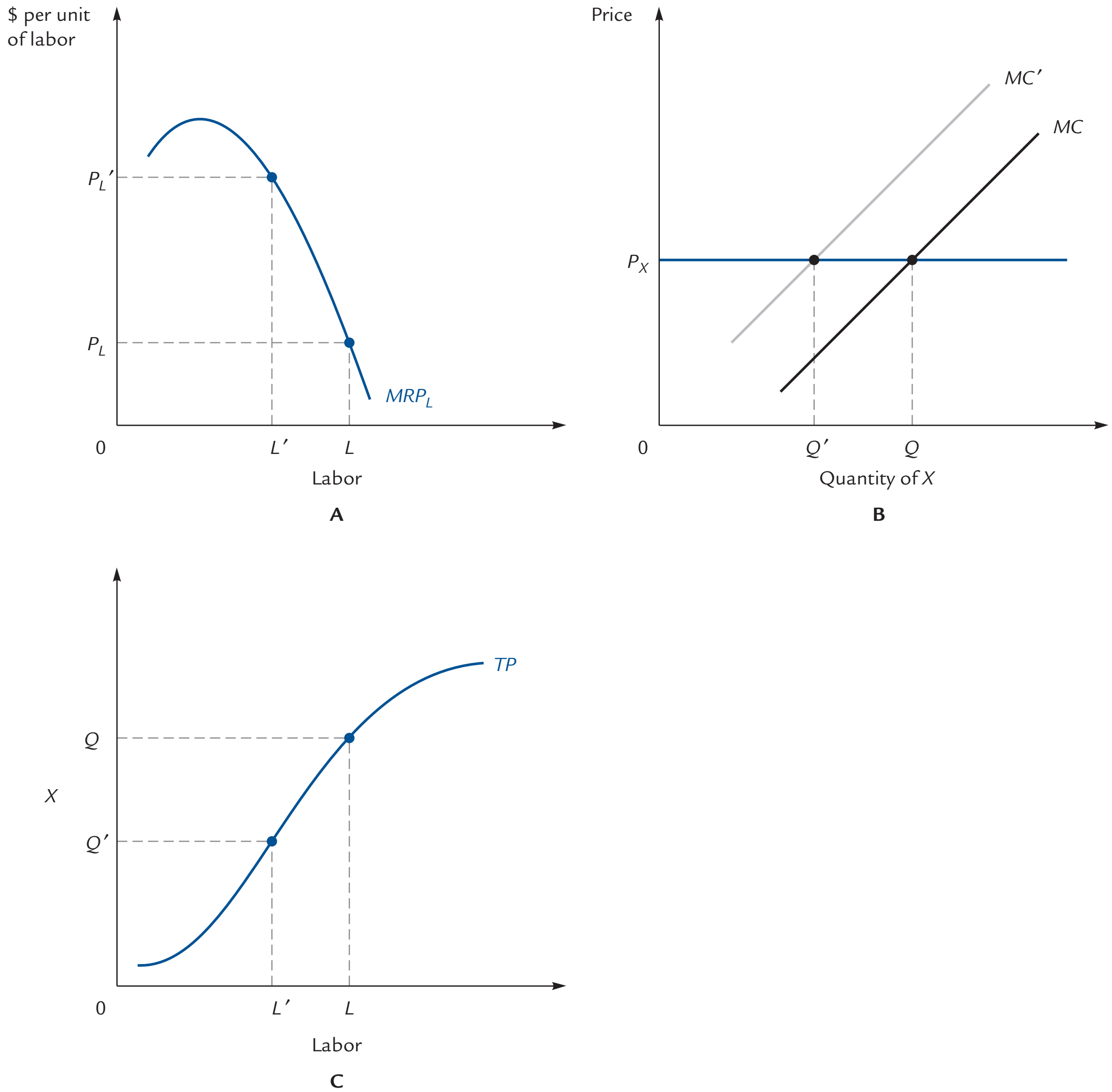
$$P_L = P_X \cdot MP_L = MRP_L$$

This confirms that a profit-maximizing firm wants to operate where the wage rate of labor is equal to its marginal product; in other words, the firm's demand curve for labor coincides with the MRP_L curve, as we have already determined.

These equations enable us to relate the firm's behavior in the labor and output markets. First suppose that the wage rate of labor P_L goes up. The equation $MC = P_L/MP_L$ tells us that the firm's marginal cost curve must go up as well. With a higher marginal cost curve, the firm produces less output and so hires less labor. This confirms yet again that the demand curve for labor is downward sloping.

Exhibit 15.2 shows the picture. When the wage rate increases from P_L to P_L' in panel A, the marginal cost curve increases from MC to MC' in panel B. Output falls from Q to Q', and the amount of labor that the firm needs to hire falls from L to L' in panel C. This fall in the quantity of labor demanded could be read equally well directly off the demand for labor curve in panel A.

EXHIBIT 15.2 **The Market for Labor and the Market for Output**



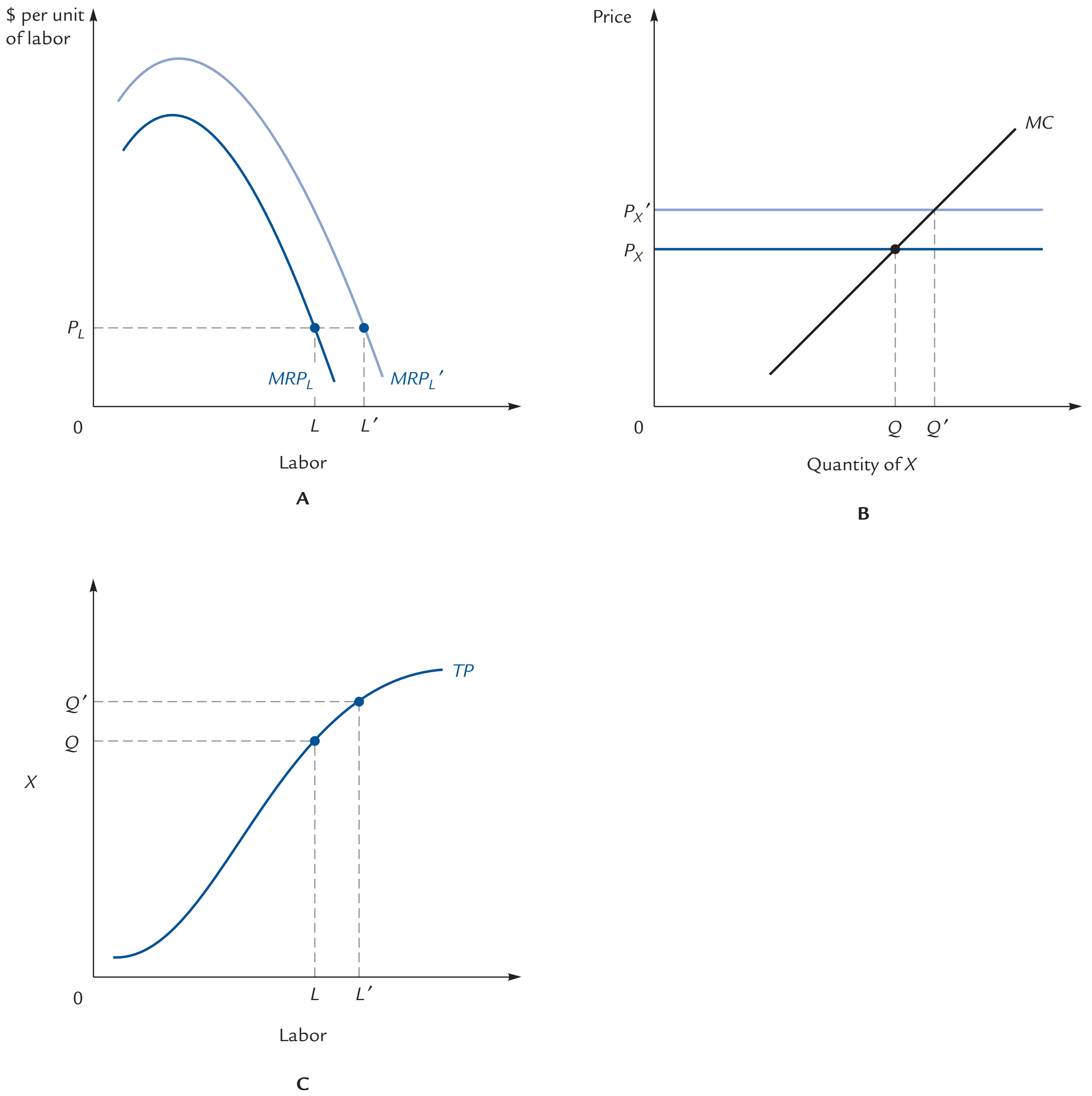
The marginal cost curve, *MC,* in panel B is derived from knowledge of the wage rate of labor, P_L, and the total product of labor curve, *TP,* in panel C. The derivation was given in Chapter 6. Thus, each graph contains some information that is also encoded in the other graphs.

To see the interrelations, notice that when the wage rate is P_L, panel A shows that the firm hires *L* units of labor, panel C shows that *L* units of labor will produce *Q* units of output, and panel B confirms that the firm's output is *Q.* If the wage rate rises to P_L', the marginal cost curve rises to *MC′*. Now panel A shows that the firm hires *L′* units of labor, panel C shows that the firm produces *Q′* units of output, and panel B confirms this.

For an alternative exercise, imagine an increase in the price of output P_X (with the wage rate of labor P_L held fixed). Because $P_L = P_X \cdot \mathrm{MP}_L$ it follows that MP_L must go down, which requires that L go up.

Exhibit 15.3 shows the picture. The increase in price from P_X to P_X' in panel B yields an increase in output from Q to Q'. This requires more labor, as seen in panel C where the

EXHIBIT 15.3 **A Rise in the Price of Output**



Initially, the price of output is P_X and the wage rate of labor is P_L. The firm hires L units of labor and produces Q units of output. When the price of output rises to P_X', the MRP_L curve shifts out to MRP_L', employment rises to L', and output increases to Q'.

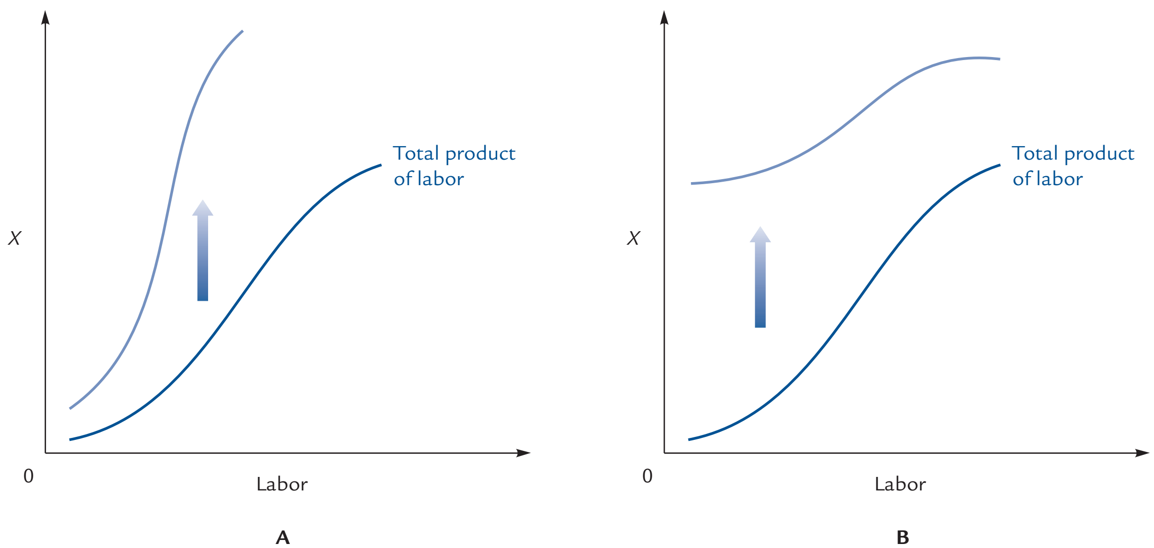
quantity of labor must rise from L to L'. Alternatively, we can argue that the increase in P_X causes an outward shift in MRP_L (because $MRP_L = P_X \cdot MP_L$), as seen in panel A. The quantity of labor demanded rises from L to L', just as we have already seen in panel A.

The Effect of Plant Size

Our entire short-run analysis assumes a fixed plant size (that is, we assume that the firm does not vary its capital usage). It makes a difference what fixed plant size we assume. The marginal product of the 40th doctor in a major hospital equipped with the latest multimillion-dollar technology is different from the marginal product of the 40th doctor in a small practice with two offices and one examining room.

Suppose that the firm increases its capital usage. Then any number of workers will certainly be able to produce at least as much as before (they can always just continue what they were doing before, ignoring the new machinery) and will probably be able to produce more. Therefore, the total product curve can be expected to rise. This does not necessarily imply that the marginal product of labor will rise. In the two panels of Exhibit 15.4, we show two possibilities. In panel A the total product of labor rises while becoming steeper at each level of output. In this case, the marginal product of labor rises, and therefore so does the competitive firm's demand curve for labor. In panel B, the total product of labor rises while becoming shallower at each level of output. This leads to a fall in the marginal product of labor and so to a fall in the competitive firm's labor demand.

EXHIBIT **15.4** **An Increase in Plant Size**



Following an increase in plant size, any quantity of labor can produce more than it did before. Thus, the total product curve shifts upward. Typically, it also becomes steeper, as in panel A, so that the marginal product of labor increases as well. In this case, we say that capital and labor are complements in production. But conceivably the total product could rise but become shallower, as in panel B. In this case, the marginal product of labor falls because of the increase in plant size; we say that capital and labor are substitutes in production.

In the first case, which is the typical one, we say that labor and capital are **complements in production.** When labor and capital are complements in production, increases in capital make workers more productive at the margin and lead to increases in the demand for labor. In the second case, we say that labor and capital are **substitutes in production.** When labor and capital are substitutes in production, an increase in capital leads to a fall in labor's marginal productivity and decreases the demand for labor. People who worry about "automation" reducing the demand for workers believe that capital and labor are substitutes in production. As an empirical matter, this case seems to be much rarer than it is often believed to be.

Complements in production
Two factors with the property that an increase in the employment of one raises the marginal product of the other.

A change in plant size is a long-run phenomenon. Thus, when we talk about the marginal product of labor before and after the capital adjustment, we are comparing one initial short-run situation with the new short-run situation that holds following a long-run adjustment.

Dangerous Curve

Substitutes in production
Two factors with the property that an increase in the employment of one lowers the marginal product of the other.

15.2 The Firm's Demand for Factors in the Long Run

Next we will study the demand for labor in the long run, with both labor and capital treated as variables. (To study the demand for capital, simply interchange the words *capital* and *labor* throughout this section.)

Constructing the Long-Run Labor Demand Curve

Now we will construct the firm's long-run labor demand curve. Throughout the discussion the following are held fixed:

The technology available to the firm (that is, its isoquant diagram).

The rental rate on capital, which we denote by P_K.

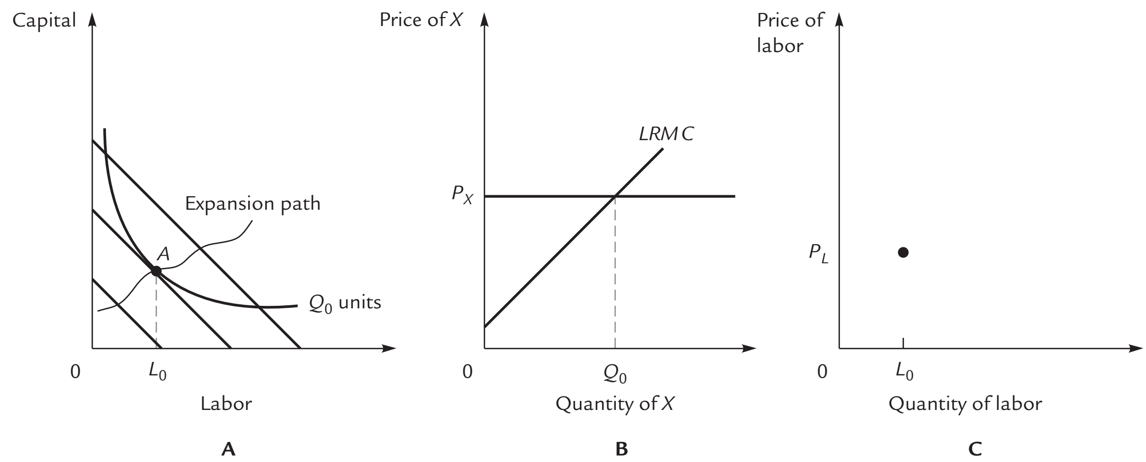
The market price of output, which we denote by P_X.

Constructing a Point on the Curve

To find a point on the labor demand curve, we will take a particular wage rate, P_L, as given and see how much labor the firm chooses to employ.

The wage rate P_L determines the slope of the firm's isocosts, which is $-P_L/P_K$. This allows us to draw in the family of isocosts and so to construct the expansion path as in panel A of Exhibit 15.5. In Section 6.3 we saw how the expansion path determines the firm's (long-run) total and marginal cost curves. The long-run marginal cost curve *(LRMC)* in panel B of Exhibit 15.5 is the one that arises from that process. The firm chooses a level of output, Q_0, so as to maximize its profits. It then looks to the Q_0-unit isoquant and finds the least-cost way of producing Q_0 units. That least-cost way is the basket labeled *A* in panel A. The firm hires the basket of inputs represented by *A*. This basket includes L_0 units of labor. Therefore, a wage rate of P_L leads to the firm's demanding L_0 units of labor. This entire process allows us to construct a single point on the firm's demand curve for labor, shown in panel C of the exhibit.

EXHIBIT 15.5 **Constructing a Point on the Labor Demand Curve**



The graphs illustrate the construction of a single point on the firm's demand curve for labor, shown in panel C. The isoquant in panel A and the output price, P_X, shown in panel B are given and are independent of the wage rate. Now we assume a wage rate P_L. This enables us to draw the isocosts in the first panel, which have slope $-P_L P_K$. These in turn determine the expansion path, also shown in panel A. Using panel A, we can derive the firm's long-run marginal cost (=long-run supply) curve, *LRMC,* using the methods of Section 6.3. Panel B determines the firm's output, which is Q_0. We now return to panel A to see that when the firm produces the quantity Q_0, it chooses the basket of inputs *A,* and this basket contains L_0 units of labor. Finally, we conclude that the wage rate P_L corresponds to the quantity of labor L_0, and we record this fact in panel C.

The Demand for Inputs versus the Demand for Output

The construction of a firm's demand curve for a factor is similar in spirit to that of the consumer's demand curve for an output, but it is also more complicated. The key difference is that a consumer has a budget constraint. Given prices, we can determine that budget constraint and find the basket he consumes. A firm, by contrast, has no budget constraint. Instead, it has an infinite family of isocost lines, and it could choose to operate on any one of them. In order to find out what basket of inputs the firm chooses, we must refer to another market, the market for output (that is, we must use panel B in Exhibit 15.5). The firm's demand curve for a factor of production is called **derived demand,** because it is partly derived from information external to the market for the factor itself.

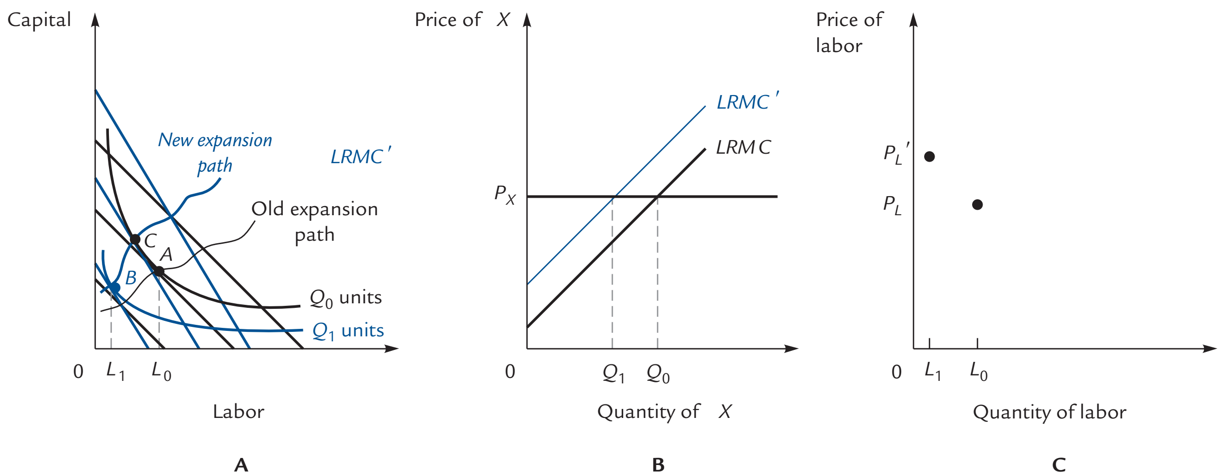
Derived demand

Demand for an input, which depends on conditions in the output market.

A Change in the Wage Rate

Continuing with the example of Exhibit 15.5, suppose that the price of labor rises, to P_L'. This causes all of the isocosts to become steeper, as in panel A of Exhibit 15.6, yielding a new expansion path shown in blue. The new expansion path leads to new (long-run) total and marginal cost curves. Suppose that the new marginal cost curve is the curve *LRMC′* in panel B of Exhibit 15.6. Then the firm reduces output to Q_1 and chooses an input basket where the Q_1-unit isoquant is tangent to an isocost. The new basket is the one labeled *B* in panel A of Exhibit 15.6. The quantity of labor demanded

EXHIBIT 15.6 **A Rise in the Wage Rate**



A rise in the wage rates raises the firm's long-run total cost curve, *TC,* to a new level, *TC′*. Usually, *TC′* is steeper than *TC,* as in panel A. In this case, the firm's long-run marginal cost curve moves upward and output decreases, as in panel B of Exhibit 15.6. However, it is possible that *TC′* could be shallower than *TC,* as in panel B. In this case, long-run marginal cost is reduced and output increases. When the latter case occurs, we say that labor is a regressive factor.

is L_1. This gives a second point on the firm's demand curve for labor, shown in panel C of the exhibit.

Continuing in this way, we can generate as many points as we want and can connect them to get the firm's labor demand curve.

Substitution and Scale Effects

In Exhibit 15.6, when the price of labor rises from P_L to P_L', the firm moves from input basket *A* to input basket *B*. In particular, it reduces its employment of labor. This reduction comes about for two quite different reasons.

One reason is that labor is now more expensive relative to capital, so it pays to use less labor and more capital in producing any given quantity of output. In other words, the expansion path in panel A of the exhibit has shifted upward and to the left. (Instead of passing through *A,* it now passes through *B* and *C.*) This is called the **substitution effect** of the wage change.

Substitution effect

When the price of an input changes, that part of the effect on employment that results from the firm's substitution toward other inputs.

The other reason is that the firm now faces higher costs and consequently produces less output, so that it wants less of every factor of production, including labor. We see this in panel B of the exhibit, where the higher marginal cost curve causes output to fall. This is called the **scale effect** of the wage change.

Scale effect

When the price of an input changes, that part of the effect on employment that results from changes in the firm's output.

The substitution and scale effects of a change in the wage rate are closely analogous to the substitution and income effects that a consumer experiences in response to a change in the price of a consumption good.

An Imaginary Experiment

In order to separate the substitution effect from the scale effect, we can conduct a hypothetical experiment. Suppose that the price of labor were to rise from P_L to P_L' but that the firm kept its output fixed at Q_0. (The experiment is hypothetical because the firm would *not,* in fact, keep its output fixed at Q_0.) In that case, where would the firm operate? It would want to be on its new expansion path but to remain on the Q_0-unit isoquant. That is, it would move to point *C* in panel A of Exhibit 15.6. The movement from point *A* to point *C* is a pure substitution effect. The scale effect, which results from changes in the firm's output level, has been totally eliminated by assuming that the firm holds its output level constant.

Now, in fact, the firm does not hold its output level constant. Instead it moves to point *B*. The "move" from the hypothetical point *C* to the firm's actual new basket *B* is due entirely to the change in output from Q_0 to Q_1. It is the scale effect. To summarize:

The firm's movement from *A* to *B* can be thought of as a movement along the isoquant from *A* to *C* (called the *substitution effect*), followed by a movement along the expansion path from *C* to *B* (called the *scale effect*).

Direction of the Substitution Effect

When the price of labor rises, the substitution effect is a movement along an isoquant to a tangency with a new, steeper isocost. It must be a movement to the left. This is because isoquants become steeper to the left and shallower to the right. In panel A of Exhibit 15.6 this means that point *C* is to the left of point *A* and this represents a basket with less labor.

The substitution effect of a rise in the wage always reduces the firm's employment of labor.

Direction of the Scale Effect

An increase in the wage rate raises the firm's long-run total cost curve. However, this could happen in either of two ways. The long-run total cost curve could both rise and become steeper. In this case, because marginal cost is equal to the slope of total cost, and because that slope has increased, long-run marginal cost will rise. Alternatively, the long-run total cost curve could rise and become shallower, in which case long-run marginal cost will fall. The two possibilities are illustrated in Exhibit 15.7.

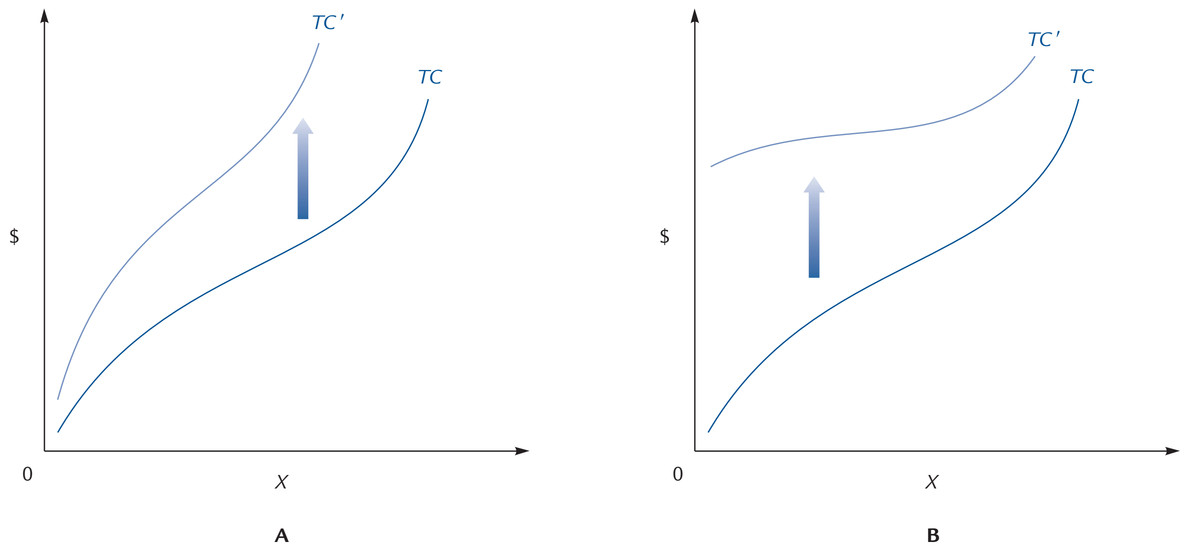
Panel A of Exhibit 15.7 is by far the more usual case. Here a rise in the wage leads to a rise in marginal cost, as was assumed in Exhibit 15.6. Thus, in Exhibit 15.6 output falls, from Q_0 to Q_1. Therefore, the scale effect is a movement along the expansion path to a lower isoquant and so must be a movement to the left. Recall that in Exhibit 15.6 the scale effect is the movement from point *C* to point *B*. Because *B* is to the left of *C,* the scale effect reduces the employment of labor, thereby reinforcing the substitution effect.

However, it is also possible that the rise in the wage rate could lead to an increase in total cost of the sort shown in panel B of Exhibit 15.7 and hence to a fall in the marginal cost curve. If so, we say that labor is a **regressive factor.** For example, the rise in wages might make it profitable for the firm to build a highly automated factory, allowing it to produce at very low marginal cost. This case is shown in Exhibit 15.8, where output rises from Q_0 to Q_2 in panel B. Because of the rise in output, the scale effect is a rightward move, from point *C* to point *B'* in panel A. That is, the scale effect causes the firm to employ more labor than it otherwise would.

Regressive factor

A factor with the property that an increase in its wage rate lowers the firm's long-run marginal cost curve.

EXHIBIT **15.7** **Two Possible Effects of a Rise in the Wage Rate**



A rise in the wage rates raises the firm's long-run total cost curve, *TC,* to a new level, *TC′*. Usually, *TC′* is steeper than *TC,* as in panel A. In this case, the firm's long-run marginal cost curve moves upward and output decreases, as in panel B of Exhibit 15.6. However, it is possible that *TC′* could be shallower than *TC,* as in panel B. In this case, long-run marginal cost is reduced and output increases. When the latter case occurs, we say that labor is a regressive factor.

Combining the Substitution and Scale Effects

Exhibits 15.6 and 15.8 show two possibilities, corresponding to the two panels of Exhibit 15.7. In each case, the substitution effect, from point *A* to point *C,* is a movement to the left. In Exhibit 15.6, which is the usual case, the scale effect, from *C* to *B,* is a further movement to the left. Thus, we can conclude that *B* must lie to the left of *A,* which is to say that the quantity of labor demanded decreases in response to a rise in the wage rate. That is, in this case the demand curve for labor surely slopes down.

In Exhibit 15.8, where labor is a regressive factor, the substitution and scale effects work in opposite directions. The substitution effect reduces the quantity of labor demanded, whereas the scale effect increases it. That is, *C* is to the left of *A,* but *B′* is to the right of *C*. Where is *B′* with respect to *A*?

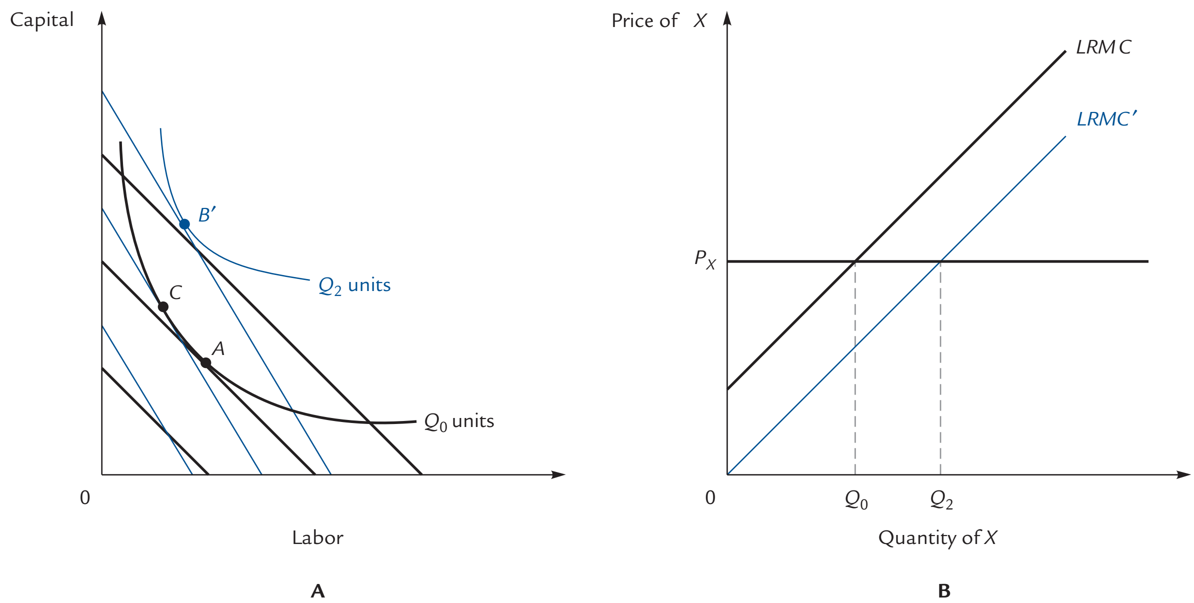
From what we can see in the diagram, there is no way to tell for sure whether *B′* is to the left or to the right of *A*. However, as a matter of mathematical fact, *B′* must lie to the left of *A*. That is, for a regressive factor the substitution effect must be greater than the scale effect. The proof of this is a bit subtle. If you are very talented mathematically, you will learn a lot from trying to discover it.

We can summarize by saying that in any case a rise in the wage rate leads to a fall in the quantity of labor demanded. Put another way:

The competitive firm's demand curve for labor (or any other factor of production) always slopes down.

In fact, the same statement is also true for a monopoly firm's demand curve for labor.

EXHIBIT 15.8 **A Rise in the Wage of a Regressive Factor**



If labor is a regressive factor, then a rise in the wage rate leads to a fall in marginal cost and an increase in output, from Q_0 to Q_2. Therefore, the firm moves from point *A* on the Q_0 isoquant to point *B′* on the higher Q_2 isoquant. The move can be decomposed into a substitution effect (the move from *A* to *C*) and scale effect (the move from *C* to *B′*).

In the case of consumer goods, which we studied in Section 4.3, we had to admit the theoretical possibility of a Giffen good, for which the consumer's demand curve would slope up. However, there is not even a theoretical possibility of a Giffen *factor.* A firm's derived demand curves for factors of production must slope down.

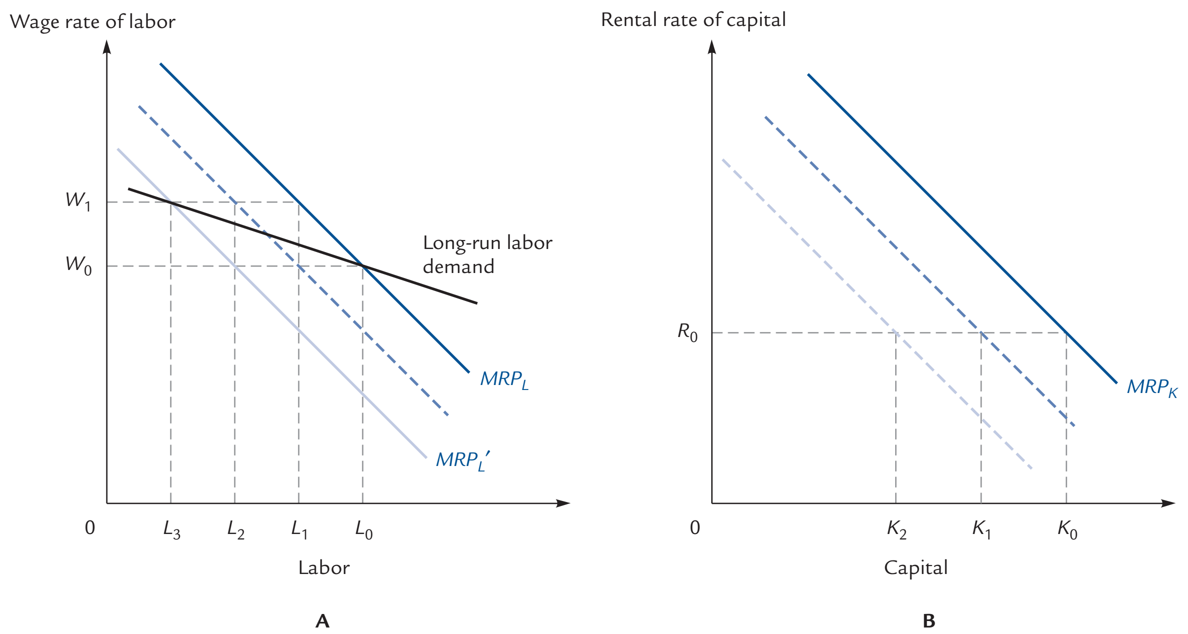
Relationships between the Short Run and the Long Run

We began this section by studying the case in which labor is the only variable input, and we argued that the firm's demand curve for labor is just the downward-sloping part of the MRP_L curve. We then moved on to the more complicated case in which two factors are variable, and we derived the firm's demand curve for labor via the more complicated process depicted in Exhibit 15.6. What is the relationship between these two approaches to labor demand?

The answer is that in the long run the MRP_L curve shifts because of adjustments in the employment of capital. For example, consider the effect of a rise in the wage when labor and capital are complements in production. Exhibit 15.9 shows the adjustment process. Initially, the wage rate of labor is W_0 and the rental rate on capital is R_0. At these prices, the firm hires L_0 units of labor (chosen from the MRP_L curve) and K_0 units of capital (chosen from the MRP_K curve).[1]

[1] Instead of using P_K and P_L for the prices of capital and labor, we are writing *R* and *W.* The only reason for this is that we want to be able to use numerical subscripts, which look ugly when appended to P_K and P_L.

EXHIBIT **15.9** **Labor Demand in the Short Run and the Long Run**



Initially, the wage rate of labor is W_0 and the rental rate on capital is R_0. The firm hires L_0 units of labor and K_0 units of capital.

Now the wage rate rises to W_1. In the short run, the firm reduces its employment of labor to L_1, read off the MRP_L curve. Assuming that capital and labor are complements in production, this causes the MRP_K curve to fall to the level of the middle curve in panel B. The firm reduces its capital employment to K_1.

The reduced capital employment lowers the MRP_L curve to the level of the dashed curve in panel A, causing labor employment to fall to L_2. This lowers the MRP_K still further, causing capital employment to fall to K_2, and the process repeats. Eventually, the MRP_L curve settles at the new level MRP_L'. Here the firm hires L_3 units of labor. Thus, the long-run labor demand curve (in black) shows that a wage of W_1 corresponds to the quantity L_3 of labor employed.

When the wage rises to W_1, the firm's short-run response is to move along the MRP_L curve and reduce the employment of labor to L_1. The reduction in labor reduces the marginal product of capital, so that the MRP_K curve moves down to the middle curve in panel B. In the long run, the firm reduces its capital employment to K_1, causing the marginal product of labor to fall to the dashed curve in panel A. This causes employment to fall further, to L_2. This in turn leads to a further reduction in the marginal product of capital, which leads to even less capital employed, which reduces the marginal product of labor still further, and so on. After many iterations, the marginal product of labor settles down, as indicated in panel A, and the final level of employment is L_3.

In the long run, therefore, the firm hires L_3 units of labor when the wage is W_1. Thus, on the long-run labor demand curve, shown in black, the wage W_0 corresponds to L_0 and the wage W_1 corresponds to L_3.

Dangerous Curve

The adjustment process described here requires, in principle, an infinite number of steps. But because the firm can foresee the outcome of these infinitely many steps, it can simply move directly to the new level of employment without actually stopping at each step along the way.

15.3 The Industry's Demand Curve for Factors of Production

The industry's demand curve for factors of production can be approximated by adding the demand curves of the individual firms. However, this overlooks an important complication. When the wage rate goes up, in the usual case all firms' marginal cost curves move up. As a result, the industry supply curve shifts and the price of output rises. This in turn means that firms will not reduce output by as much as they would if price remained constant. The substitution effect is unchanged, but the scale effect is lessened. Firms reduce their employment of labor by less than Exhibit 15.6 predicts. On similar grounds, a fall in the wage leads to a smaller increase in employment than one would expect from our study of individual firms. The bottom line is that the industry's demand curve for a factor tends to be less elastic than the sum of the demand curves from the individual firms in the industry.

Finally, in any discussion of the demand for labor (or any input), it should be remembered that labor is demanded by many different industries. All the corresponding industry demand curves must be added together to get "the" demand curve for labor.

Monopsony

Throughout this chapter we have assumed that firms take factor prices as given. This is equivalent to saying that for each factor the firm faces a supply curve that is horizontal at the market wage rate. However, there remains the possibility that a single firm could account for a substantial portion of the market for some factor. In this case, the quantity demanded by the firm affects that factor's wage rate. The firm faces an upward-sloping supply curve for that factor.

The most extreme example occurs if there is some factor of production that is demanded by only one firm. In that case, the firm in question is a "single buyer," just as a monopolist might be a "single seller." A single buyer is called a **monopsonist.** However, just as we use the word *monopolist* to describe any seller who faces a downward-sloping demand curve, so we shall use the word *monopsonist* to describe any buyer who faces an upward-sloping supply curve.

Monopsonist

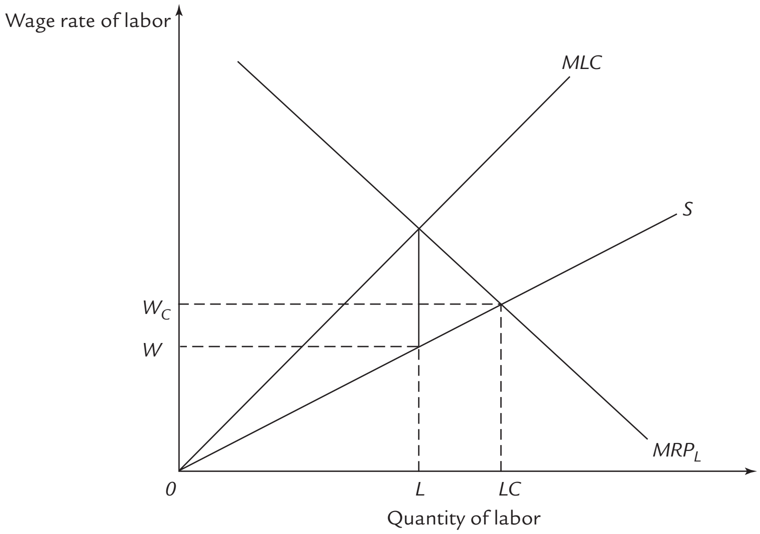
A buyer who faces an upward-sloping supply curve.

To a monopsony demander of labor, the cost of hiring an additional unit of labor exceeds the wage rate. The reason for this is that when the monopsonist hires an additional worker, there are two ways in which his costs increase: (1) he must pay the new worker's wage and (2) he bids up the wages of all workers. As a result, the monopsonist faces a **marginal labor cost (*MLC*)** curve that lies everywhere above the labor supply curve that he faces. He maximizes profits by choosing that quantity where the marginal revenue product of labor and the marginal cost of labor are equal; then he pays a wage read off the supply curve at that quantity. The process is illustrated in Exhibit 15.10.

Marginal labor cost (*MLC*)

The cost of hiring an additional unit of labor.

EXHIBIT **15.10** **Monopsony**



A monopsony demander of labor faces an upward-sloping labor supply curve (S) and a marginal labor cost (MLC) curve that lies everywhere above S. He hires L units of labor (where $MRP_L = MLC$) and pays the wage W that he reads off the supply curve at that quantity.

In an industry with many firms, the going price for labor would be W_C and each firm would face a flat supply curve at this price. L_C units of labor would be hired.

Do not confuse *MLC,* which is the cost of hiring one additional unit of labor, with *MC,* which is the cost of producing one additional unit of output.

Dangerous Curve

The monopsonist hires fewer workers and pays a lower wage than would be the case if many firms competed to hire labor. Under competition there would be a going wage rate of W_C in Exhibit 15.10, and employment would be L_C.

How Widespread Is Monopsony?

In order for a firm to have monopsony power, it must constitute a substantial portion of the demand for some factor. Therefore, even a firm that is unique in its industry has no monopsony power, provided that there are firms in *other* industries competing with it for the use of factors.

For example, suppose that all of the major auto manufacturers were to merge into one giant firm. At first, this firm could well have monopsony power in the market for autoworkers, who would have no other employer competing to hire their valuable skills. However, if the giant auto firm were to exercise this monopsony power to keep wages low, some autoworkers would eventually decide to acquire other skills and to sell their services elsewhere—say, as shipbuilders. In the long run, the single automaker competes in the labor market with all of the firms in the shipbuilding industry and in countless other industries besides.

The same is true when a single employer dominates a certain geographic area. Although the employer may have some monopsony power in the short run, he may be unable to exercise that power without causing some of the area's residents to move elsewhere. Ultimately, he competes for the local workers with employers all over the world.

15.4 The Distribution of Income

Firms hire factors of production and combine them to create output. This output generates revenue, or income, for the firm. Each factor of production receives a portion of this revenue as its payment for participating in the firm's activity. (Economists persist in speaking of payments to factors of production, even though it would often be more accurate to speak of payments to the *owners* of the factors.) After all of these payments are made, any remaining revenue (positive or negative) accrues to the owners of the firm in the form of profit.

Factor Shares and Rents

From the first part of this chapter, we know that when labor markets are competitive, the price of any factor is equal to its marginal revenue product, if a firm or an industry hires L units of labor at a wage rate of $P_L \cdot L$. Therefore, we can say that labor's income is equal to $MRP_L \cdot L$.

If the supply curve of labor to this firm or industry is upward sloping, the suppliers of labor earn a producers' surplus, or rent, equal to area B in Exhibit 15.11. Labor's income is the sum of areas B and C, so that only a portion of this income can be considered rent.

Dangerous Curve

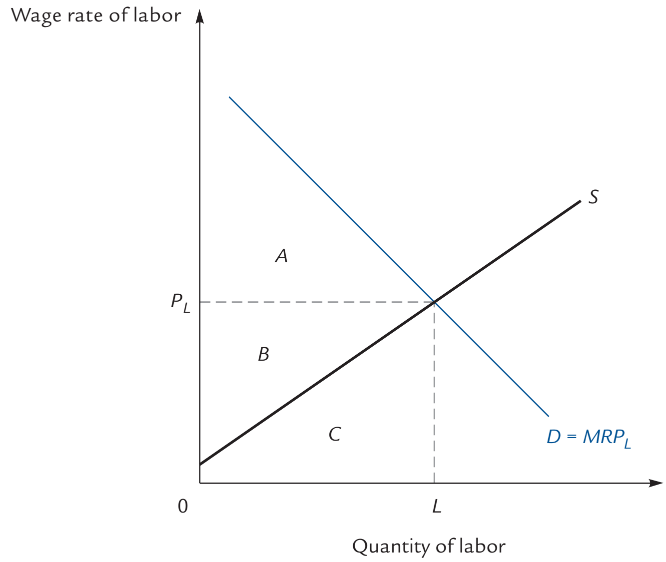
Do not confuse the word *rent,* meaning producers' surplus, with the rental (i.e., wage) paid by the firm to hire a factor of production. The factor earns a producers' surplus equal to the payment it receives from the firm *minus* its opportunity costs. Only when the factor supply curve is perfectly vertical does the rental payment consist entirely of rent.

The area $A + B + C$ under the MRP_L curve in Exhibit 15.11 is equal to the total revenue of the firm or of the industry. (The area can be broken into rectangles representing the revenue from the first unit of labor employed, the revenue from the second unit, and so on.) Because labor receives $B + C$, the remaining area A must represent the sum of the payments to all other factors, plus any profits that are earned.

What is true of labor is also true of every other factor. Capital earns an income of $MRP_K \cdot K$, of which some portion is rent. There are also intangible factors like "entrepreneurial ability" that are typically supplied by the owners of the firm. Although such factors are not explicitly on the payroll, they should be viewed as implicitly receiving a wage equal to their marginal revenue product. If the owner supplies E units of entrepreneurial ability, with a marginal revenue product of MRP_E, then we think of the firm as paying the owner an income of $MRP_E \cdot E$ in his capacity as a factor of production.

Inputs like entrepreneurial ability are often supplied quite inelastically. The owner of a shoe store has a great deal of knowledge about the specific workings of his own

EXHIBIT 15.11 **Labor's Share of Income**



The firm or industry hires L units of labor at the wage P_L and earns a total revenue of $A + B + C$. Of this revenue, labor receives a share equal to $P_L \cdot L = B + C$. Of this, area C covers workers' opportunity costs and area B is earned as rent.

enterprise. Such knowledge is a factor of production that would be much less valuable in any alternative use. As a result, he might supply almost all of this knowledge to his own business, regardless of whether he earns a high or a low wage by doing so. Thus, the supply curve for the owner's entrepreneurial services is very inelastic, so that a large portion of the income earned by these services tends to be rent.

Profit

The sum of the factor payments may be less than, equal to, or greater than the revenue of the firm. If the factor payments are less than the firm's revenue, then the difference is profit and accrues to the owner of the firm. If the factor payments exceed the firm's revenues, the firm takes a loss, sometimes called a *negative profit,* equal to the difference. This loss comes from the pocket of the firm's owner.

Dangerous Curve

Notice that in our analysis the owner of the firm receives two very different kinds of payments. (They are different to the economist, although an accountant or a businessman would see no reason to distinguish them.) First, there is the income that he earns as the supplier of certain factors of production. Much of this income is usually a rent, or a producer's surplus. Second, there is the profit remaining after the firm has made all of its factor payments (including the ones to the owner).

As was discussed briefly in Chapter 7, many economists would prefer not to think of specialized skills, such as knowledge of the workings of a particular shoe store, as factors of production that are hired by the firm. They would prefer to think of the firm as earning positive profits due to the existence of these factors. The two analyses use different words but describe the same outcomes.

Returns to Scale

In long-run equilibrium, it can be shown mathematically that when production is subject to decreasing returns to scale (that is, when average cost is increasing), factor shares add up to less than the firm's total revenue (so that the firm has a positive profit); when production is subject to constant returns to scale (that is, when average cost is flat), factor shares add up to the firm's revenue exactly (so that profit is zero); and when production is subject to increasing returns to scale (that is, when average cost is decreasing), factor shares add up to more than the firm's revenue (so that profit is negative).

In long-run competitive equilibrium, the firm operates at the minimum point of its average cost curve, where returns to scale are constant. Therefore, profits are zero, as we already know from Chapter 7.

However, Professor Paul Romer of the University of California at Berkeley argues that in many industries firms experience increasing returns to scale over the entire relevant range.[2] The reason is that many important inputs (unlike the labor and capital we have considered in this chapter) are *nonrivalrous:* Once produced, there is no limit to how much they can be used. A firm that produces one specialized software program to install on a manager's computer can allow other managers to install the same program at essentially no additional cost.

A firm that has one specialized software program but doubles all of its other inputs (number of computers, number of managers, number of factories, etc.) might be expected to double its output. If the firm really doubles *all* of its inputs by constructing a second specialized software program, then it should *more* than double its output. This is precisely the definition of increasing returns to scale.

If increasing returns are truly a common phenomenon, they present a major challenge to the standard competitive model of the firm. A competitive firm that experiences increasing returns must earn negative profits after all factors' shares are paid out. In such circumstances, we should not expect to see any competitive firms.

Producers' Surplus

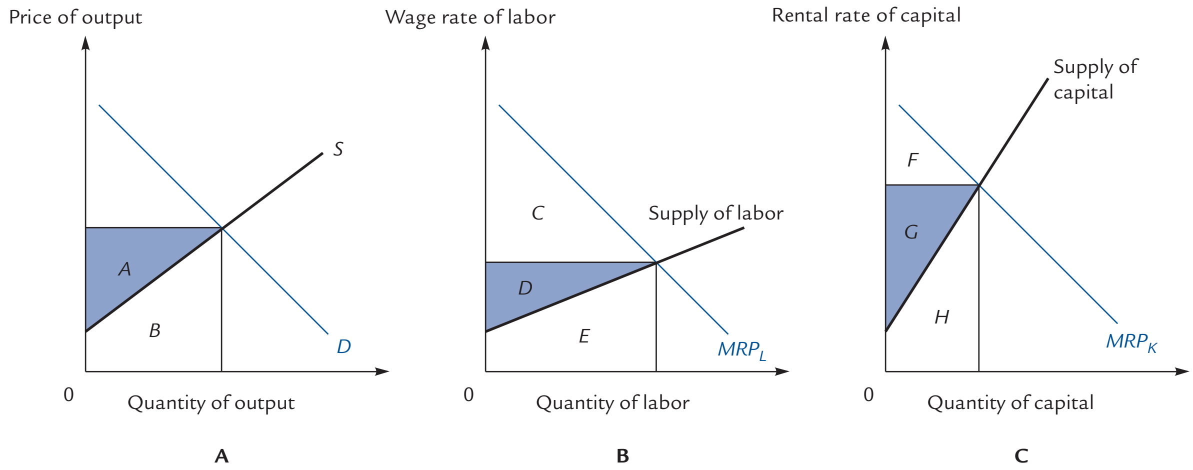
In earlier chapters, we talked about the producer's surplus earned by firms. It is often useful to think of producers' surplus in that way. However, in a more careful analysis, we recognize that at least part of the producers' surplus is actually earned by the factors that the firms employ.

In fact, in long-run competitive equilibrium, firms earn zero profits. This means that all of the producers' surplus that we have previously attributed to the firms is actually paid out to factors.

Exhibit 15.12 shows the relationship between the industry-wide markets for output, labor, and capital when each firm earns zero profits. The firms earn total revenue equal to $A + B$ in the output market, of which A is producers' surplus. (The firms'

[2] P. Romer, "Are Nonconvexities Important for Understanding Growth?" *American Economic Review* 80 (1990): 97–107.

EXHIBIT **15.12** **The Distribution of Rent**



In long-run zero-profits equilibrium, the industry's total revenue (given by $A + B = C + D + E = F + G + H$) is paid out to factors. Because labor's total wages are $D + E$ and the total rental payments to capital are $G + H$, we have $A + B = (D + E) + (G + H)$. Producers' surplus in the industry is equal to A, of which workers get D and owners of capital get G. Therefore, $A = D + G$.

total revenue is also equal to $C + D + E$ in panel B and to $F + G + H$ in panel C.) This revenue is distributed to workers, who earn $D + E$ in panel B, and to the owners of capital, who earn $G + H$ in panel C. Because we assume that firms earn zero profits, these factor payments must exactly account for the firms' total revenue. That is, $(D + E) + (G + H) = A + B$.

The portion of total revenue that is producers' surplus is exactly A, of which D is earned by workers and G is earned by the owners of capital. Therefore, $A = D + G$. If profits were nonzero, then area A would include those profits in addition to $D + G$.

Of course, when there are more than two factors of production, rents are divided among all of them, not just capital and labor.

Dangerous Curve

Who Benefits?

Factors that are supplied relatively inelastically (the most extreme case being a fixed factor) earn more rents than those supplied more elastically. As a result, the more nearly fixed factors have more to gain (or to lose) from changes in the demand for the output of the industry. If the demand for output rises, the derived demand for all inputs rises. This increases producers' surplus by more for those factors with inelastic supply curves than for other factors. By the same reasoning, these factors bear most of the loss when the demand for output falls.

For example, professional football games are produced with many inputs, including professional quarterbacks and footballs. The supply of quarterbacks is quite inelastic, because the particular skills of a quarterback have relatively few alternatives uses that are anywhere near as valuable. Therefore, quarterbacks earn substantial rents. (That is, their wage bills far exceed their opportunity costs.) Footballs are supplied much more elastically, because the skills needed to produce footballs are also useful in a variety of other industries. Therefore, suppliers of footballs earn comparatively little rent. Any change in the public's demand for football games will have a much greater effect on the fortunes of quarterbacks than it will on the fortunes of football manufacturers.

Dangerous Curve

What matters in this example is not the fact that quarterbacks' wages are high, but that their supply curve is inelastic. Suppose, for example, that all quarterbacks could equally well earn $500,000 a year as movie stars. Then, over a substantial range, the supply curve for quarterbacks would be flat (perfectly elastic) at $500,000 per year. In this case, the wage would be high, but there would be no producers' surplus. And, in fact, in this case quarterbacks would not be hurt if the public completely lost interest in football. Changes in the industry's fortunes are felt most by those factors that are inelastically supplied, not by those factors whose wage bills are high.

Some factors are fixed in the short run and variable in the long run. An increase in the price of output benefits these factors more in the short run than in the long run. For example, in the short run there are a fixed number of recording studios capable of producing compact discs. A rise in the price of compact discs will raise revenue in the recording industry, and in the short run this increased revenue will largely be paid as rent to the owners of the recording studios. In the long run, however, more recording studios can be built, and the owners of existing recording studios will not continue to reap this windfall benefit. Short-term rents due to inelastic short-run supply are sometimes called **quasi-rents.**

Quasi-rents

Producers' surplus earned in the short run by factors that are supplied inelastically in the short run.

Finally, we should note that the owners of the factors of production are the same individuals and households that are the consumers in the economy. In earlier chapters, we maintained a careful distinction between the consumers' surplus earned by individuals and the producers' surplus earned by firms. Now we see that the producers' surplus is actually earned by the same individuals who are earning the consumers' surplus. All gains from trade ultimately accrue to individuals. Who else is there to benefit?

Summary

A factor's marginal revenue product is defined as the amount of additional revenue the firm can earn by employing one more unit of that factor. The equimarginal principle implies that the firm's demand curve for the factor will be identical with the downward-sloping portion of the factor's marginal revenue product curve.

An increase in employment of one factor will usually raise the marginal productivity of other factors; hence, it will raise the firm's demand curve for other factors. In this case, we say that the factors are complements in production. It is also possible that an increase in the employment of one factor will reduce the marginal productivity of other factors, in which case we say that the factors are substitutes in production.

In the long run, a change in the wage rate of labor will cause the firm to change its employment of both labor and capital. The firm's marginal cost curve will change, leading to a change in output as well.

In the hypothetical case in which the firm does *not* adjust output, the change in the wage rate leads to a movement along an isoquant, known as the *substitution effect.* The substitution effect is always in the expected direction: A rise in the wage rate reduces the quantity of labor demanded, and a fall in the wage rate increases the quantity demanded.

The scale effect of a wage change is that part of the change in employment that is due to the change in output. It is a movement along the new expansion path. The scale effect is usually in the same direction as the substitution effect, but it can go in the opposite direction, in which case we say that labor is a regressive factor. For a regressive factor, however, the substitution effect is always larger than the scale effect. Thus, even for a regressive factor the firm's labor demand curve must slope downward.

The firm's revenues are paid out to the factors of production, with each factor earning a wage equal to its marginal revenue product. Among these payments may be payments to the firm's owners for the use of specialized factors such as particular skills. After all these payments are made, whatever remains is the firm's profit. In long-run competitive equilibrium, profits are zero, so the factor payments exactly exhaust the firm's income.

Payments to factors minus the factors' opportunity costs are the factors' producers' surplus, or rent. The firm's producer's surplus (the area above the firm's supply curve up to the price and out to the quantity supplied) is the sum of all these factor rents plus the firm's profit, if any. Thus, the producers' surplus that we have attributed to firms in previous chapters is actually distributed as factor rents.

The more inelastically supplied the factor, the greater the percentage of its income that is rent. Thus, inelastically supplied factors benefit the most from the existence of the industry, and they stand to gain or lose the most when the industry's fortunes wax or wane.

Review Questions

R1. What is the relationship between marginal product and marginal revenue product?

R2. Draw total and marginal product diagrams to show how a rise in the price of output affects the employment of labor.

R3. Draw total and marginal product diagrams to show how an increase in plant size affects the employment of labor.

R4. Explain how to construct a point on the firm's long-run demand curve for labor.

R5. Define the substitution and the scale effects of an increase in the wage rate. What can be said about their directions?

R6. Define monopsony. Does a monopsonist employ more or less labor than a firm that hires workers competitively? Why?

R7. In long-run competitive equilibrium, the firm's total revenue is equal to the sum of its factor payments. Why?

R8. What is the relationship between the producers' surplus measured above the firm's supply curve for output and the producers' surpluses measured above the factors' supply curves for their services?

R9. A factor that is supplied perfectly elastically to an industry has nothing to gain or lose from changes in the price of output. Explain why, first using graphs and then giving the verbal interpretation.

R10. A factor that is supplied perfectly inelastically to an industry earns rents equal to its entire wage bill. Thus, such a factor participates heavily in the industry's fortunes, be they good or bad. Explain why, first using graphs and then giving the verbal interpretation.

Numerical Exercises

N1. Consider a firm that produces according to the production function:

$$Q = \sqrt{KL}$$

where Q is the firm's output and K and L are the quantities of capital and labor that it employs. With this production function, the slope of an isoquant at the point (L, K) is given by $-K/L$.

- **a.** Suppose that the going wage rate of labor is W and the going rental rate on capital is R. What is the slope of an isocost? If the firm uses K units of capital and L of labor in long-run equilibrium, derive a formula for K in terms of L, W, and R. Derive a formula for L in terms of K, W, and R. (*Hint:* In long-run equilibrium, the firm operates at a point where the slope of an isocost and the slope of an isoquant are equal.)
- **b.** Using the production function and the result of part (a), write a formula for L in terms of Q, W, and R, and a formula for K in terms of Q, W, and R.
- **c.** Write a formula for the total cost of producing Q units of output.
- **d.** Describe the firm's long-run marginal cost curve.
- **e.** In long-run equilibrium, what must the price of output be? Would you have had enough information to answer this question if your answer to part (d) had been different than it was?
- **f.** In terms of Q, how much does the firm pay out to labor and to capital? What is its total revenue? What is its profit?

N2. Consider a perfectly competitive industry with many identical firms, each producing according to the production function:

$$Q = \sqrt{KL}$$

Labor and capital are supplied to the industry according to the supply curves $L = W$ and $K = 4R$.

- **a.** Suppose that the industry produces Q units of output, using K units of capital and L of labor. Write a formula for L in terms of Q, W, and R and for K in terms of Q, W, and R.
- **b.** Write two equations expressing the conditions of equilibrium in the two factor markets. Use these equations to get a numerical value for W/R. (*Hint:* Divide one equation by the other.)

c. Show that the industry's long-run total cost curve is given by:

$$Q = P$$

(*Hint:* Make use of your answers from N1.)

d. Suppose that the demand curve for the industry's output is given by:

$$Q = 1/5{,}000 - P$$

What are the price and quantity of output? How much labor is hired, and at what wage? How much capital is rented, and at what rental rate?

e. Under the conditions of part (d), calculate the producers' surplus in the output market. How much producers' surplus is earned by labor and how much by capital? How much profit is earned by firms? Is your answer consistent with your answer to Numerical Exercise N1(f)?

Problem Set

1. **True or False:** A rise in the demand for apples has no effect on the productivity of apple-pickers and hence no effect on the demand for apple pickers.
2. **True or False:** If the demand curve for a product is vertical, then any rise in the wage rate could be passed on entirely from firms to customers, without any fall in production. Thus, a rise in the wage rate would not reduce employment, either in the short run or in the long run.
3. **True or False:** If labor and capital are complements in production, then the long-run labor demand curve is more elastic than the short-run labor demand curve.
4. **a.** Prepare graphs like those in Exhibit 15.9 to illustrate the relationships between short-run and long-run labor demand when capital and labor are substitutes in production.

   b. In this case, is the short-run labor demand curve more or less elastic than the long-run labor demand curve?
5. **a.** Use Exhibit 15.9 to show that when labor and capital are the only inputs and when they are complements in production, the long-run labor demand curve must slope downward.

   b. Use the graphs you prepared for Problem 4a to show that when labor and capital are the only inputs and when they are substitutes in production, the long-run labor demand curve must slope downward.
6. **True or False:** The industry demand curve, for a regressive factor is likely to be more elastic than the sum of the firms' demand curves.
7. **True or False:** The isocosts of a monopsonist in the labor market are not straight lines.
8. Use a graph to demonstrate the social welfare consequences of monopsony.
9. **True or False:** If there is monopsony in the labor market, a minimum wage law can lead to increased employment.

10. Suppose that labor and capital are both supplied perfectly inelastically to the U.S. economy.

a. Show the producers' surplus earned by *capital* on a graph of the marginal product of *labor.* Explain where you make use of the fact that the supply of capital is perfectly inelastic.

b. Suppose that General Motors moves one of its plants to South Korea, increasing the number of workers who can be combined with U.S. capital. Show the gains and losses to (1) U.S. workers, (2) U.S. owners of capital, and (3) South Korean workers.

c. Does the plant's relocation help or hurt Americans as a whole?

11. **True or False:** If firms earn zero profits and if labor and capital are the only inputs, then a rise in wages must be bad for the owners of capital.

12. **True or False:** If firms earn zero profits and if labor and capital are the only inputs, then labor and capital must be complements in production. (*Hint:* Make use of your answer to the preceding problem.)

13. Suppose that there are exactly three factors of production: skilled labor, which is represented by unions; unskilled labor, which is not represented by unions; and capital. Currently, skilled labor earns $15 per hour and unskilled labor earns $5 per hour. Legislation has been proposed to establish a minimum wage of $10 per hour for all workers, and this legislation has been strongly endorsed by the unions.

Assuming that the unions act in the best interest of their members, can you determine whether skilled and unskilled labor are complements or substitutes in production? What about capital and unskilled labor? Can you predict how the owners of capital will feel about the legislation?

14. In order to promote economic expansion, the town of Hyde Park has declared certain areas of the city to be "no-tax zones." Businesses located in these areas are exempt from all city taxes. As a result, many new firms have started up, each of which rents offices and machinery and hires many workers.

In the long run, which of the following groups are likely to benefit from the existence of the no-tax zones: the owners of firms, the customers of the firms, landowners in the no-tax zones, the producers of machinery, the workers?