



This is “Producer Theory: Costs”, chapter 9 from the book [Beginning Economic Analysis \(index.html\)](#) (v. 1.0).

This book is licensed under a [Creative Commons by-nc-sa 3.0](#) (<http://creativecommons.org/licenses/by-nc-sa/3.0/>) license. See the license for more details, but that basically means you can share this book as long as you credit the author (but see below), don't make money from it, and do make it available to everyone else under the same terms.

This content was accessible as of December 29, 2012, and it was downloaded then by [Andy Schmitz](#) (<http://lardbucket.org>) in an effort to preserve the availability of this book.

Normally, the author and publisher would be credited here. However, the publisher has asked for the customary Creative Commons attribution to the original publisher, authors, title, and book URI to be removed. Additionally, per the publisher's request, their name has been removed in some passages. More information is available on this project's [attribution page](http://2012books.lardbucket.org/attribution.html?utm_source=header) (http://2012books.lardbucket.org/attribution.html?utm_source=header).

For more information on the source of this book, or why it is available for free, please see [the project's home page](#) (<http://2012books.lardbucket.org/>). You can browse or download additional books there.

Chapter 9

Producer Theory: Costs

The basic theory of the firm regards the firm as a mechanism for transforming productive inputs into final and intermediate goods and services. This process is known as production. For instance, the smelting of copper or gold removes impurities and makes the resulting product into a more valuable substance. Silicon Valley transforms silicon, along with a thousand other chemicals and metals, into computer chips used in everything from computers to toasters. Cooking transforms raw ingredients into food, by adding flavor and killing bacteria. Moving materials to locations where they have higher value is a form of production. Moving stone to a building site to construct an exterior wall, bringing the King Tut museum exhibit to Chicago, or qualifying a basketball team for the league playoffs are all examples of different types of production. According to this simple view, a firm is comprised of a technology or set of technologies for transforming materials into valuable goods and to maximize profits. This “production function” view of the firm is appropriate for some environments, when products and services are standardized and technologies are widely available. However, in other settings, this way of thinking about the firm is misleading, especially when the internal organization of the firm is important. Nevertheless, the “production function” model of the firm is a natural starting point to begin our investigation of competition.

9.1 Types of Firms

LEARNING OBJECTIVE

1. What types of companies exist?

There are four varieties of firms created in law, although these types have several subtypes. At one end is the **proprietorship**¹, which is a firm owned by a single individual (the proprietor) or perhaps by a family. The family farm and many “mom-and-pop” restaurants and convenience stores are operated proprietorships. Debts accrued by the proprietorship are the personal responsibility of the proprietor. Legal and accounting firms are often organized as **partnerships**². Partnerships are firms owned by several individuals who share profits as well as liabilities of the firm according to a specified formula that varies by the relative contribution and potential cost of each partner in the firm. Thus, if a partner in a law firm steals a client’s money and disappears, the other partners may be responsible for absorbing some portion of the loss. In contrast, a **corporation**³ is treated legally as a single entity owned by shareholders. Like a person, a corporation can incur debt and is therefore responsible for repayment. This stands in contrast to a partnership where particular individuals may be liable for debts incurred. Hence, when the energy trader company Enron collapsed, the Enron shareholders lost the value of their stock; however, they were not responsible for repaying the debt that the corporation had incurred. Moreover, company executives are also not financially responsible for debts of the corporation, provided they act prudentially. If a meteor strikes a manufacturer and destroys the corporation, the executives are not responsible for the damage or for the loans that may not be repaid as a result. On the other hand, executives are not permitted to break the law, as was the case with corporate officers at Archer Daniels Midland, the large agricultural firm who colluded to fix the price of lysine and were subsequently fined and jailed for their misdeeds, as was the corporation.

1. A firm is owned by a single individual (the proprietor) or a family.
2. Firms owned by several individuals who share profits as well as liabilities of the firm according to a specified formula that varies by the relative contribution and potential cost of each partner.
3. A single entity owned by shareholders.

Corporations who shield company executives and shareholders from fines and punishments are said to offer “limited liability.” So why would anyone in his or her right mind organize a firm as a proprietorship or a partnership? The explanation is that it is costly to incorporate businesses—about \$1,000 per year at the time of this writing—and corporations are taxed, so that many small businesses find it less costly to be organized as proprietorships. Moreover, it may not be possible for a family-owned corporation to borrow money to open a restaurant; for example, potential lenders may fear not being repaid in the event of bankruptcy, so they insist that owners accept some personal liability. So why are professional groups

organized as partnerships and not as corporations? The short answer is that a large variety of hybrid, organizational forms exist instead. The distinctions are blurred, and organizations like “Chapter S Corporations” and “Limited Liability Partnerships” offer the advantages of partnerships (including avoidance of taxation) and corporations. The disadvantages of these hybrids are the larger legal fees and greater restrictions on ownership and freedom to operate that exist in certain states and regions.

Usually proprietorships are smaller than partnerships, and partnerships are smaller than corporations, though there are some very large partnerships (e.g., the Big Four accounting firms) as well as some tiny corporations. The fourth kind of firm may be of any size. It is distinguished primarily by its source of revenue and not by how it is internally organized. The **nonprofit firm**⁴ is prohibited from distributing a profit to its owners. Religious organizations, academic associations, environmental groups, most zoos, industry associations, lobbying groups, many hospitals, credit unions (a type of bank), labor unions, private universities, and charities are all organized as nonprofit corporations. The major advantage of nonprofit firms is their tax-free status. In exchange for avoiding taxes, nonprofits must be engaged in government-approved activities, suggesting that nonprofits operate for the social benefit of some segment of society. So why can't you establish your own nonprofit that operates for your personal benefit in order to avoid taxes? Generally, you alone isn't enough of a socially worthy purpose to meet the requirements to form a nonprofit. Certainly some of the nonprofit religious organizations created by televangelists suggest that the nonprofit established for the benefit of a single individual isn't too far-fetched. Moreover, you can't establish a nonprofit for a worthy goal and not serve that goal but just pay yourself all the money the corporation raises, because nonprofits are prohibited from overpaying their managers. Overpaying the manager means not serving the worthy corporate goal as well as possible. Finally, commercial activities of nonprofits are taxable. Thus, when the nonprofit zoo sells stuffed animals in the gift shop, generally the zoo collects sales tax and is potentially subject to corporate taxes.

The modern corporation is a surprisingly recent invention. Prior to World War I, companies were typically organized as a pyramid with a president at the top and vice presidents who reported to him at the next lower level, and so on. In a pyramid, there is a well-defined chain of command, and no one is ever below two distinct managers of the same level. The problem with a pyramid is that two retail stores that want to coordinate have to contact their managers, and possibly their manager's managers, and so on up the pyramid until a common manager is reached. There are circumstances where such rigid decision making is unwieldy, and the larger the operation of a corporation, the more unwieldy it gets.

4. Firm that is prohibited from distributing a profit to its owners.

Four companies—Sears, DuPont, General Motors, and Standard Oil of New Jersey (Exxon)—found that the pyramid didn't work well for them. Sears found that its separate businesses of retail stores and mail order required a mix of shared inputs (purchased goods) but distinct marketing and warehousing of these goods. Consequently, retail stores and mail order needed to be separate business units, but the purchasing unit has to service both of them. Similarly, DuPont's military business (e.g., explosives) and consumer chemicals were very different operations serving different customers yet often selling the same products so that again the inputs needed to be centrally produced and to coordinate with two separate corporate divisions. General Motors's many car divisions are "friendly rivals," in which technology and parts are shared across the divisions, but the divisions compete in marketing their cars to consumers. Again, technology can't be housed under just one division, but instead is common to all. Finally, Standard Oil of New Jersey was attempting to create a company that managed oil products from oil exploration all the way through to pumping gasoline into automobile gas tanks. With such varied operations all over the globe, Standard Oil of New Jersey required extensive coordination and found that the old business model needed to be replaced. These four companies independently invented the modern corporation, which is organized into separate business units. These business units run as semiautonomous companies themselves, with one business unit purchasing inputs at a negotiated price from another unit and selling outputs to a third unit. The study of the internal organization of firms and its ramifications for competitiveness is fascinating but beyond the scope of this book. If you want to know more about organization theory, I happily recommend *Competitive Solutions: The Strategist's Toolkit*, by R. Preston McAfee, Princeton: Princeton University Press, 2002.

KEY TAKEAWAYS

- The most basic theory of the firm regards the firm as a means of transforming materials into other, more intermediate and final goods. This is known as production.
- A proprietorship is a firm owned by an individual (the proprietor) or perhaps a family and operated by a small number of people. The family farm and many restaurants and convenience stores are operated in this way. Debts accrued by the proprietorship are the personal responsibility of the proprietor.
- Attorneys and accountants are often organized as partnerships. Partnerships share profits according to a formula and are usually liable for losses incurred by the partnership.
- A corporation is a legal entity that may incur debt, and has the responsibility for repayment of that debt. The owners or officers of the corporation are not liable for the repayment.
- A large variety of hybrid, organizational forms exist.
- The nonprofit firm is prohibited from distributing a profit to its owners. In exchange for avoiding taxes, nonprofits must be engaged in government-approved activities that operate for the benefit of some segment of society. Commercial activities of nonprofits are taxable.
- Large corporations are often organized into business units that are semiautonomous affiliated companies.

9.2 Production Functions

LEARNING OBJECTIVES

1. How is the output of companies modeled?
2. Are there any convenient functional forms?

The firm transforms inputs into outputs. For example, a bakery takes inputs like flour, water, yeast, labor, and heat and makes loaves of bread. An earth moving company combines capital equipment, ranging from shovels to bulldozers with labor in order to dig holes. A computer manufacturer buys parts “off-the-shelf” like disk drives and memory, with cases and keyboards, and combines them with labor to produce computers. Starbucks takes coffee beans, water, some capital equipment, and labor to brew coffee.

Many firms produce several outputs. However, we can view a firm that is producing multiple outputs as employing distinct production processes. Hence, it is useful to begin by considering a firm that produces only one output. We can describe this firm as buying an amount x_1 of the first input, x_2 of the second input, and so on (we’ll use x_n to denote the last input), and producing a quantity of the output. The production function that describes this process is given by $y = f(x_1, x_2, \dots, x_n)$.

The **production function**⁵ is the mapping from inputs to an output or outputs.

For the most part we will focus on two inputs in this section, although the analyses with more than inputs is “straightforward.”

Example: The **Cobb-Douglas production function**⁶ is the product of each input, x , raised to a given power. It takes the form $f(x_1, x_2, \dots, x_n) = a_0 x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$.

The constants a_1 through a_n are typically positive numbers less than one. For example, with two goods, capital K and labor L , the Cobb-Douglas function becomes $a_0 K^{a_1} L^{a_2}$. We will use this example frequently. It is illustrated, for $a_0 = 1$, $a_1 = 1/3$, and $a_2 = 2/3$, in [Figure 9.1 "Cobb-Douglas isoquants"](#).

5. The mapping from inputs to an output or outputs.

6. A production function that is the product of each input, x , raised to a given power.

Figure 9.1 Cobb-Douglas isoquants

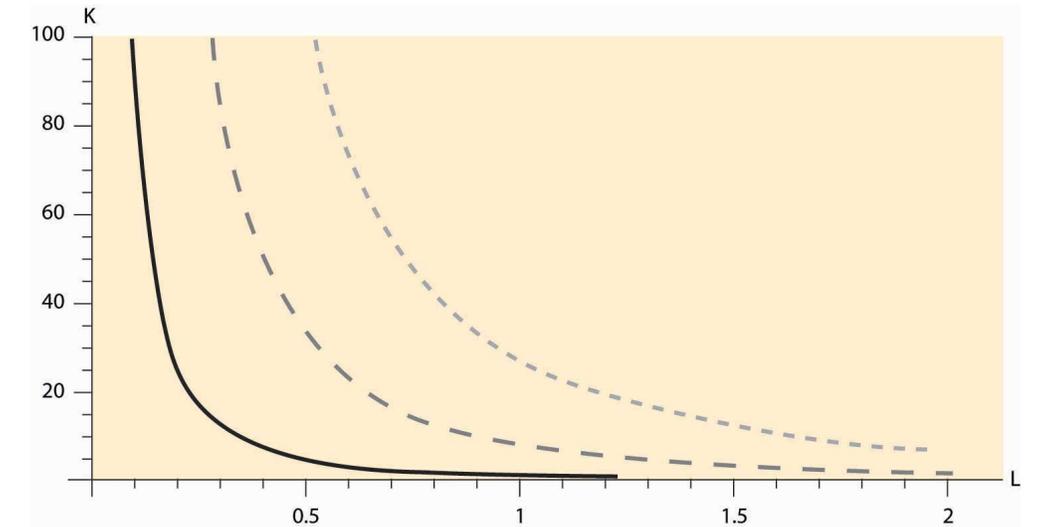
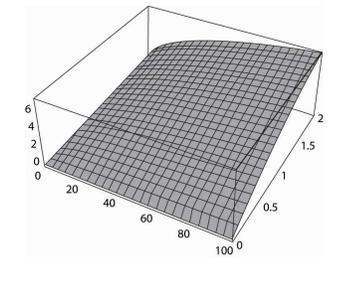


Figure 9.1 "Cobb-Douglas isoquants" illustrates three isoquants for the Cobb-Douglas production function. An **isoquant**⁷, which means "equal quantity," is a curve that describes all the combinations of inputs that produce the same level of output. In this case, given $a = 1/3$ and $b = 2/3$, we can solve $y = K^a L^b$ for K to obtain $K = y^3 L^{-2}$. Thus, $K = L^{-2}$ gives the combinations of inputs yielding an output of 1, which is denoted by the dark, solid line in Figure 9.1 "Cobb-Douglas isoquants". The middle, gray dashed line represents an output of 2, and the dotted light-gray line represents an output of 3. Isoquants are familiar contour plots used, for example, to show the height of terrain or temperature on a map. Temperature isoquants are, not surprisingly, called isotherms.

Isoquants provide a natural way of looking at production functions and are a bit more useful to examine than three-dimensional plots like the one provided in Figure 9.2 "The production function".

Figure 9.2 The production function



The fixed-proportions production function comes in the form

$$f(x_1, x_2, \dots, x_n) = \text{Min} \{a_1 x_1, a_2 x_2, \dots, a_n x_n\}.$$

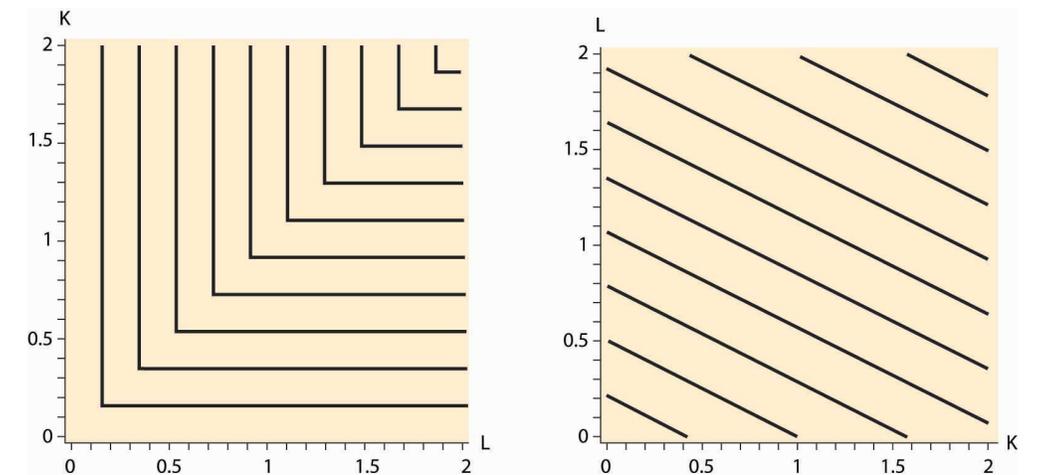
The **fixed-proportions production function**⁸ is a production function that requires inputs be used in fixed proportions to produce output. It has the property that adding more units of one input in isolation does not necessarily increase the quantity produced. For example, the productive value of having more than one

7. Curves that describe all the combinations of inputs that produce the same level of output.

8. A production function that requires inputs be used in fixed proportions to produce output.

shovel per worker is pretty low, so that shovels and diggers are reasonably modeled as producing holes using a fixed-proportions production function. Moreover, without a shovel or other digging implement like a backhoe, a barehanded worker is able to dig so little that he is virtually useless. Ultimately, the size of the holes is determined by $\min \{\text{number of shovels, number of diggers}\}$. Figure 9.3 "Fixed-proportions and perfect substitutes" illustrates the isoquants for fixed proportions. As we will see, fixed proportions make the inputs “perfect complements.”

Figure 9.3 Fixed-proportions and perfect substitutes



Two inputs K and L are perfect substitutes in a production function f if they enter as a sum; so that $f(K, L, x_3, \dots, x_n) = g(K + cL, x_3, \dots, x_n)$ for a constant c . Another way of thinking of **perfect substitutes**⁹ is that they are two goods that can be substituted for each other at a constant rate while maintaining the same output level. With an appropriate scaling of the units of one of the variables, all that matters is the sum of the two variables, not their individual values. In this case, the isoquants are straight lines that are parallel to each other, as illustrated in Figure 9.3 "Fixed-proportions and perfect substitutes".

The **marginal product**¹⁰ of an input is just the derivative of the production function with respect to that input. This is a partial derivative, since it holds the other inputs fixed. Partial derivatives are denoted with the symbol δ . An important property of marginal product is that it may be affected by the level of other inputs employed. For example, in the Cobb-Douglas case with two inputs the symbol α is the Greek letter “alpha.” The symbol β is the Greek letter “beta.” These are the first two letters of the Greek alphabet, and the word alphabet itself originates from these two letters. and for constant A ,

$$f(K, L) = AK^\alpha L^\beta,$$

- 9. Two goods that can be substituted for each other at a constant rate while maintaining the same output level.
- 10. The derivative of the production function with respect to an input.

the marginal product of capital is

$$\frac{\partial f}{\partial K}(K, L) = \alpha AK^{\alpha-1}L^{\beta}.$$

If α and β are between zero and one (the usual case), then the marginal product of capital is increasing in the amount of labor, and it is decreasing in the amount of capital employed. For example, an extra computer is very productive when there are many workers and a few computers, but it is not so productive where there are many computers and a few people to operate them.

The **value of the marginal product**¹¹ of an input is the marginal product times the price of the output. If the value of the marginal product of an input exceeds the cost of that input, it is profitable to use more of the input.

Some inputs are easier to change than others. It can take 5 years or more to obtain new passenger aircraft, and 4 years to build an electricity generation facility or a pulp and paper mill. Very skilled labor such as experienced engineers, animators, and patent attorneys are often hard to find and challenging to hire. It usually requires one to spend 3 to 5 years to hire even a small number of academic economists. On the other hand, it is possible to buy shovels, telephones, and computers or to hire a variety of temporary workers rapidly, in a day or two. Moreover, additional hours of work can be obtained from an existing labor force simply by enlisting them to work “overtime,” at least on a temporary basis. The amount of water or electricity that a production facility uses can be varied each second. A dishwasher at a restaurant may easily use extra water one evening to wash dishes if required. An employer who starts the morning with a few workers can obtain additional labor for the evening by paying existing workers overtime for their hours of work. It will likely take a few days or more to hire additional waiters and waitresses, and perhaps several days to hire a skilled chef. You can typically buy more ingredients, plates, and silverware in one day, whereas arranging for a larger space may take a month or longer.

The fact that some inputs can be varied more rapidly than others leads to the notions of the long run and the short run. In the short run, only some inputs can be adjusted, while in the long run all inputs can be adjusted. Traditionally, economists viewed labor as quickly adjustable and capital equipment as more difficult to adjust. That is certainly right for airlines—obtaining new aircraft is a very slow process—for large complex factories, and for relatively low-skilled, and hence substitutable, labor. On the other hand, obtaining workers with unusual skills is a slower process than obtaining warehouse or office space. Generally speaking, the long-run inputs are those that are expensive to adjust quickly, while the short-run

11. The marginal product times the price of the output.

factors can be adjusted in a relatively short time frame. What factors belong in which category is dependent on the context or application under consideration.

KEY TAKEAWAYS

- Firms transform inputs into outputs.
- The functional relationship between inputs and outputs is the production function.
- The Cobb-Douglas production function is the product of the *inputs* raised to powers and comes in the form $f(x_1, x_2, \dots, x_n) = a_0 x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ for positive constants a_1, \dots, a_n .
- An isoquant is a curve or surface that traces out the inputs leaving the output constant.
- The fixed-proportions production function comes in the form $f(x_1, x_2, \dots, x_n) = \min \{a_1 x_1, a_2 x_2, \dots, a_n x_n\}$.
- Fixed proportions make the inputs “perfect complements.”
- Two inputs K and L are perfect substitutes in a production function f if they enter as a sum; that is, $f(K, L, x_3, \dots, x_n) = g(K + cL, x_3, \dots, x_n)$, for a constant c .
- The marginal product of an input is just the derivative of the production function with respect to that input. An important aspect of marginal products is that they are affected by the level of other inputs.
- The value of the marginal product of an input is just the marginal product times the price of the output. If the value of the marginal product of an input exceeds the cost of that input, it is profitable to use more of the input.
- Some inputs are more readily changed than others.
- In the short run, only some inputs can be adjusted, while in the long run all inputs can be adjusted.
- Traditionally, economists viewed labor as quickly adjustable and capital equipment as more difficult to adjust.
- Generally speaking, the long-run inputs are those that are expensive to adjust quickly, while the short-run factors can be adjusted in a relatively short time frame.

EXERCISE

1. For the Cobb-Douglas production function, suppose there are two inputs K and L , and the sum of the exponents is one. Show that, if each input is paid the value of the marginal product per unit of the input, the entire output is just exhausted. That is, for this production function, show
$$K \frac{\partial f}{\partial K} + L \frac{\partial f}{\partial L} = f(K, L).$$

9.3 Profit Maximization

LEARNING OBJECTIVE

1. If firms maximize profits, how will they behave?

Consider an entrepreneur who would like to maximize profit, perhaps by running a delivery service. The entrepreneur uses two inputs, capital K (e.g., trucks) and labor L (e.g., drivers), and rents the capital at cost r per dollar of capital. The wage rate for drivers is w . The production function is $F(K, L)$ —that is, given inputs K and L , the output is $F(K, L)$. Suppose p is the price of the output. This gives a profit of $\pi = pF(K, L) - rK - wL$. Economists often use the symbol π , the Greek letter “pi,” to stand for profit. There is little risk of confusion because economics doesn’t use the ratio of the circumference to the diameter of a circle very often. On the other hand, the other two named constants, Euler’s e and i , the square root of -1 , appear fairly frequently in economic analysis.

$$\pi = pF(K, L) - rK - wL.$$

First, consider the case of a fixed level of K . The entrepreneur chooses L to maximize profit. The value L^* of L that maximizes the function π must satisfy

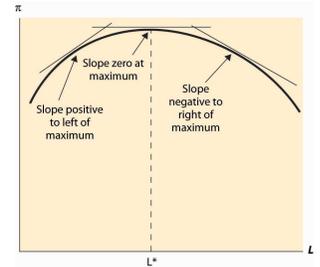
$$0 = \frac{\partial \pi}{\partial L} = p \frac{\partial F}{\partial L}(K, L^*) - w.$$

This expression is known as a **first-order condition**¹², a mathematical condition for optimization stating that the first derivative is zero. It is possible that $L = 0$ is the best that an entrepreneur can do. In this case, the derivative of profit with respect to L is not necessarily zero. The first-order condition instead would be either $0 = \frac{\partial \pi}{\partial L}$, or $L = 0$, and $0 \geq \frac{\partial \pi}{\partial L}$. The latter pair of conditions reflects the logic that either the derivative is zero and we are at a maximum, or $L = 0$, in which case a small increase in L must not cause π to increase. The first-order condition recommends that we add workers to the production process up to the point where the last worker’s marginal product is equal to his wage (or cost).

12. A mathematical condition for optimization stating that the first derivative is zero.

In addition, a second characteristic of a maximum is that the second derivative is negative (or nonpositive). This arises because, at a maximum, the slope goes from positive (since the function is increasing up to the maximum) to zero (at the maximum) to being negative (because the function is falling as the variable rises past the maximum). This means that the derivative is falling or that the second derivative is negative. This logic is illustrated in Figure 9.4 "Profit-maximizing labor input".

Figure 9.4 Profit-maximizing labor input



The second property is known as the **second-order condition**¹³, a mathematical condition for maximization stating that the second derivative is nonpositive. The orders refer to considering small, but positive, terms Δ , which are sent to zero to reach derivatives. The value Δ^2 , the second-order term, goes to zero faster than Δ , the first-order term. It is expressed as

$$0 \geq \frac{\partial^2 \pi}{(\partial L)^2} = p \frac{\partial^2 F}{(\partial L)^2} (K, L^*).$$

This is enough of a mathematical treatment to establish comparative statics on the demand for labor. Here we treat the choice L^* as a function of another parameter—the price p , the wage w , or the level of capital K . For example, to find the effect of the wage on the labor demanded by the entrepreneur, we can write

$$0 = p \frac{\partial F}{\partial L} (K, L^* (w)) - w.$$

This expression recognizes that the choice L^* that the entrepreneur makes satisfies the first-order condition and results in a value that depends on w . But how does it depend on w ? We can differentiate this expression to obtain

$$0 = p \frac{\partial^2 F}{(\partial L)^2} (K, L^* (w)) L^{*'} (w) - 1,$$

or

$$L^{*'} (w) = \frac{1}{p \frac{\partial^2 F}{(\partial L)^2} (K, L^* (w))} \leq 0.$$

13. A mathematical condition for maximization stating that the second derivative is nonpositive.

The second-order condition enables one to sign the derivative. This form of argument assumes that the choice L^* is differentiable, which is not necessarily true.

Digression: In fact, there is a form of argument that makes the point without calculus and makes it substantially more general. Suppose $w_1 < w_2$ are two wage levels and that the entrepreneur chooses L_1 when the wage is w_1 and L_2 when the wage is w_2 . Then profit maximization requires that these choices are optimal. In particular, when the wage is w_1 , the entrepreneur earns higher profit with L_1 than with L_2 :

$$pf(K, L_1) - rK - w_1L_1 \geq pf(K, L_2) - rK - w_1L_2$$

When the wage is w_2 , the entrepreneur earns higher profit with L_2 than with L_1 :

$$pf(K, L_2) - rK - w_2L_2 \geq pf(K, L_1) - rK - w_2L_1.$$

The sum of the left-hand sides of these two expressions is at least as large as the sum of the right-hand side of the two expressions:

$$\begin{aligned} pf(K, L_1) - rK - w_1L_1 + pf(K, L_2) - rK - w_2L_2 &\geq \\ pf(K, L_1) - rK - w_2L_1 + pf(K, L_2) - rK - w_1L_2 & \end{aligned}$$

A large number of terms cancel to yield the following:

$$-w_1L_1 - w_2L_2 \geq -w_2L_1 - w_1L_2$$

This expression can be rearranged to yield the following:

$$(w_1 - w_2)(L_2 - L_1) \geq 0.$$

This shows that the higher labor choice must be associated with the lower wage. This kind of argument, sometimes known as a **revealed preference**¹⁴ kind of argument, states that choice implies preference. It is called “revealed preference” because choices by consumers were the first place the type of argument was applied. It can be very powerful and general, because issues of differentiability are avoided. However, we will use the more standard differentiability-type argument, because such arguments are usually more readily constructed.

14. Kind that states that choice implies preference.

The effect of an increase in the capital level K on the choice by the entrepreneur can be calculated by considering L^* as a function of the capital level K :

$$0 = p \frac{\partial F}{\partial L} (K, L^*(K)) - w$$

Differentiating this expression with respect to K , we obtain

$$0 = p \frac{\partial^2 F}{\partial K \partial L} (K, L^*(K)) + p \frac{\partial^2 F}{(\partial L)^2} (K, L^*(K)) L^{*'}(K),$$

or

$$L^{*'}(K) = \frac{-\frac{\partial^2 F}{\partial K \partial L} (K, L^*(K))}{\frac{\partial^2 F}{(\partial L)^2} (K, L^*(K))}.$$

We know the denominator of this expression is not positive, thanks to the second-order condition, so the unknown part is the numerator. We then obtain the conclusion that

an increase in capital increases the labor demanded by the entrepreneur if

$$\frac{\partial^2 F}{\partial K \partial L} (K, L^*(K)) > 0, \text{ and decreases the labor demanded if}$$

$$\frac{\partial^2 F}{\partial K \partial L} (K, L^*(K)) < 0.$$

This conclusion looks like gobbledygook but is actually quite intuitive. Note that $\frac{\partial^2 F}{\partial K \partial L} (K, L^*(K)) > 0$ means that an increase in capital increases the derivative of output with respect to labor; that is, an increase in capital increases the marginal product of labor. But this is, in fact, the definition of a complement! That is, $\frac{\partial^2 F}{\partial K \partial L} (K, L^*(K)) > 0$ means that labor and capital are complements in production—an increase in capital increases the marginal productivity of labor. Thus, an increase in capital will increase the demand for labor when labor and capital are complements, and it will decrease the demand for labor when labor and capital are substitutes.

This is an important conclusion because different kinds of capital may be complements or substitutes for labor. Are computers complements or substitutes for labor? Some economists consider that computers are complements to highly skilled workers, increasing the marginal value of the most skilled, but substitutes for lower-skilled workers. In academia, the ratio of secretaries to professors has fallen dramatically since the 1970s as more and more professors are using machines to perform secretarial functions. Computers have increased the marginal product of

professors and reduced the marginal product of secretaries, so the number of professors rose and the number of secretaries fell.

The revealed preference version of the effect of an increase in capital is to posit two capital levels, K_1 and K_2 , with associated profit-maximizing choices L_1 and L_2 . The choices require, for profit maximization, that

$$pF(K_1, L_1) - rK_1 - wL_1 \geq pF(K_1, L_2) - rK_1 - wL_2$$

and

$$pF(K_2, L_2) - rK_2 - wL_2 \geq pF(K_2, L_1) - rK_2 - wL_1.$$

Again, adding the left-hand sides together produces a result at least as large as the sum of the right-hand sides:

$$\begin{aligned} pF(K_1, L_1) - rK_1 - wL_1 + pF(K_2, L_2) - rK_2 - wL_2 &\geq \\ pF(K_2, L_1) - rK_2 - wL_1 + pF(K_1, L_2) - rK_1 - wL_2. \end{aligned}$$

Eliminating redundant terms yields

$$pF(K_1, L_1) + pF(K_2, L_2) \geq pF(K_2, L_1) + pF(K_1, L_2),$$

or

$$F(K_2, L_2) - F(K_1, L_2) \geq F(K_2, L_1) - F(K_1, L_1),$$

or

$$\int_{K_1}^{K_2} \frac{\partial F}{\partial K}(x, L_2) dx \geq \int_{K_1}^{K_2} \frac{\partial F}{\partial K}(x, L_1) dx,$$

Here we use the standard convention that

$$\int_a^b \dots dx = - \int_b^a \dots dx.$$

or

$$\int_{K_1}^{K_2} \frac{\partial F}{\partial K}(x, L_2) - \frac{\partial F}{\partial K}(x, L_1) dx \geq 0,$$

and finally

$$\int_{K_1}^{K_2} \int_{L_1}^{L_2} \frac{\partial^2 F}{\partial K \partial L}(x, y) dy dx \geq 0.$$

Thus, if $K_2 > K_1$ and $\frac{\partial^2 F}{\partial K \partial L}(K, L) > 0$ for all K and L , then $L_2 \geq L_1$; that is, with complementary inputs, an increase in one input increases the optimal choice of the second input. In contrast, with substitutes, an increase in one input decreases the other input. While we still used differentiability of the production function to carry out the revealed preference argument, we did not need to establish that the choice L^* was differentiable to perform the analysis.

Example (Labor demand with the Cobb-Douglas production function): The Cobb-Douglas production function has the form $F(K, L) = AK^\alpha L^\beta$, for constants A , α , and β , all positive. It is necessary for $\beta < 1$ for the solution to be finite and well defined. The demand for labor satisfies

$$0 = p \frac{\partial F}{\partial L}(K, L^*(K)) - w = p\beta AK^\alpha L^{*\beta-1} - w,$$

or

$$L^* = \left(\frac{p\beta AK^\alpha}{w} \right)^{1/1-\beta}.$$

When $\alpha + \beta = 1$, L is linear in capital. Cobb-Douglas production is necessarily complementary; that is, an increase in capital increases labor demanded by the entrepreneur.

KEY TAKEAWAYS

- Profit maximization arises when the derivative of the profit function with respect to an input is zero. This property is known as a first-order condition.
- Profit maximization arises with regards to an input when the value of the marginal product is equal to the input cost.
- A second characteristic of a maximum is that the second derivative is negative (or nonpositive). This property is known as the second-order condition.
- Differentiating the first-order condition permits one to calculate the effect of a change in the wage on the amount of labor hired.
- Revealed preference arguments permit one to calculate comparative statics without using calculus, under more general assumptions.
- An increase in capital will increase the demand for labor when labor and capital are complements, and it will decrease the demand for labor when labor and capital are substitutes.
- Cobb-Douglas production functions are necessarily complements; hence, any one input increases as the other inputs increase.

EXERCISES

1. For the fixed-proportions production function $\min\{K, L\}$, find labor demand when capital is fixed at K .
2. The demand for hamburgers has a constant elasticity of 1 of the form $x(p) = 8,000 p^{-1}$. Each entrant in this competitive industry has a fixed cost of \$2,000 and produces \sqrt{x} hamburgers per year, where x is the amount of meat in pounds.
 - a. If the price of meat is \$2 per pound, what is the long-run supply of hamburgers?
 - b. Compute the equilibrium number of firms, the quantity supplied by each firm, and the market price of hamburgers.
 - c. Find the short-run industry supply. Does it have constant elasticity?
3. A company that produces software needs two inputs, programmers (x) at a price of p and computers (y) at a price of r . The output is given by $T = 4x^{1/3}y^{1/3}$, measured in pages of code.
 - a. What is the marginal cost?
 - b. Now suppose each programmer needs two computers to do his job. What ratio of p and r would make this input mix optimal?
4. A toy factory costs \$2 million to construct, and the marginal cost of the q^{th} toy is $\max[10, q^2/1,000]$.
 - a. What is average total cost?
 - b. What is short-run supply?
 - c. What is the long-run competitive supply of toys?

9.4 The Shadow Value

LEARNING OBJECTIVE

1. If a firm faces constraints on its behavior, how can we measure the costs of these constraints?

When capital K can't be adjusted in the short run, it creates a constraint, on the profit available, to the entrepreneur—the desire to change K reduces the profit available to the entrepreneur. There is no direct value of capital because capital is fixed. However, that doesn't mean we can't examine its value. The value of capital is called a **shadow value**¹⁵, which refers to the value associated with a constraint. Shadow value is well-established jargon.

What is the shadow value of capital? Let's return to the constrained, short-run optimization problem. The profit of the entrepreneur is

$$\pi = pF(K, L) - rK - wL.$$

The entrepreneur chooses the value L^* to maximize profit; however, he is constrained in the short run with the level of capital inherited from a past decision. The shadow value of capital is the value of capital to profit, given the optimal decision L^* . Because $0 = \frac{\partial \pi}{\partial L} = p \frac{\partial F}{\partial L}(K, L^*) - w$, the shadow value of capital is $\frac{d\pi(K, L^*)}{dK} = \frac{\partial \pi(K, L^*)}{\partial K} = p \frac{\partial F}{\partial K}(K, L^*) - r$.

Note that this value could be negative if the entrepreneur might like to sell some capital but can't, perhaps because it is installed in the factory.

Every constraint has a shadow value. The term refers to the value of relaxing the constraint. The shadow value is zero when the constraint doesn't bind; for example, the shadow value of capital is zero when it is set at the profit-maximizing level. Technology binds the firm; the shadow value of a superior technology is the increase in profit associated with it. For example, parameterize the production technology by a parameter a , so that $aF(K, L)$ is produced. The shadow value of a given level of a is, in the short run,

$$\frac{d\pi(K, L^*)}{da} = \frac{\partial \pi(K, L^*)}{\partial a} = pF(K, L^*).$$

15. The value associated with a constraint.

A term is vanishing in the process of establishing the shadow value. The desired value L^* varies with the other parameters like K and a , but the effect of these parameters on L^* doesn't appear in the expression for the shadow value of the parameter because $0 = \frac{\partial \pi}{\partial L}$ at L^* .

KEY TAKEAWAYS

- When an input is fixed, its marginal value is called a shadow value.
- A shadow value can be negative when an input is fixed at too high a level.
- Every constraint has a shadow value. The term refers to the value of relaxing a constraint. The shadow value is zero when the constraint doesn't bind.
- The effect of a constraint on terms that are optimized may be safely ignored in calculating the shadow value.

9.5 Input Demand

LEARNING OBJECTIVES

1. How much will firms buy?
2. How do they respond to input price changes?

Over a long period of time, an entrepreneur can adjust both the capital and the labor used at the plant. This lets the entrepreneur maximize profit with respect to both variables K and L . We'll use a double star, **, to denote variables in their long-run solution. The approach to maximizing profit over two c separately with respect to each variable, thereby obtaining the conditions

$$0 = p \frac{\partial F}{\partial L} (K^{**}, L^{**}) - w$$

and

$$0 = p \frac{\partial F}{\partial K} (K^{**}, L^{**}) - r.$$

We see that, for both capital and labor, the value of the marginal product is equal to the purchase price of the input.

It is more challenging to carry out comparative statics exercises with two variables, and the general method won't be developed here. If you want to know more, the approach is to arrange the two equations as a vector with $\mathbf{x} = (K, L)$, $\mathbf{z} = (r/p, w/p)$, so that $\mathbf{0} = \mathbf{F}'(\mathbf{x}^{**}) - \mathbf{z}$, and then differentiate to obtain $d\mathbf{x}^{**} = (\mathbf{F}''(\mathbf{x}^{**}))^{-1} d\mathbf{z}$, which can then be solved for each comparative static. However, we can illustrate one example as follows.

Example: The Cobb-Douglas production function implies choices of capital and labor satisfying the following two first-order conditions: It is necessary for $\alpha + \beta < 1$ for the solution to be finite and well defined.

$$0 = p \frac{\partial F}{\partial L} (K^{**}, L^{**}) - w = p\beta AK^{**\alpha} L^{**\beta-1} - w,$$

$$0 = p \frac{\partial F}{\partial K} (K^{**}, L^{**}) - r = p\alpha AK^{**\alpha-1} L^{**\beta} - r.$$

To solve these expressions, first rewrite them to obtain

$$w = p\beta AK^{**\alpha} L^{**\beta-1}$$

and

$$r = p\alpha AK^{**\alpha-1} L^{**\beta}.$$

Then divide the first expression by the second expression to yield

$$\frac{w}{r} = \frac{\beta K^{**}}{\alpha L^{**}},$$

or

$$K^{**} = \frac{\alpha w}{\beta r} L^{**}.$$

This can be substituted into either equation to obtain

$$L^{**} = \left(\frac{Ap\alpha\beta^{1-\alpha}}{r^\alpha w^{1-\alpha}} \right)^{\frac{1}{1-\alpha-\beta}}$$

and

$$K^{**} = \left(\frac{Ap\alpha^{1-\beta}\beta^\beta}{r^{1-\beta}w^\beta} \right)^{\frac{1}{1-\alpha-\beta}}.$$

While these expressions appear complicated, the dependence on the output price p , and the input prices r and w , is quite straightforward.

How do equilibrium values of capital and labor respond to a change in input prices or output price for the Cobb-Douglas production function? It is useful to cast these changes in percentage terms. It is straightforward to demonstrate that both capital

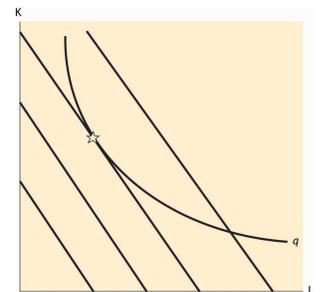
and labor respond to a small percentage change in any of these variables with a constant percentage change.

An important insight of profit maximization is that it implies minimization of costs of yielding the chosen output; that is, *profit maximization entails efficient production*.

The logic is straightforward. The profit of an entrepreneur is revenue minus costs, and the revenue is price times output. For the chosen output, then, the entrepreneur earns the revenue associated with the output, which is fixed since we are considering only the chosen output, minus the costs of producing that output. Thus, for the given output, maximizing profits is equivalent to maximizing a constant (revenue) minus costs. Since maximizing $-C$ is equivalent to minimizing C , the profit-maximizing entrepreneur minimizes costs. This is important because profit-maximization implies not being wasteful in this regard: A profit-maximizing entrepreneur produces at least cost.

There are circumstances where the cost-minimization feature of profit maximization can be used, and this is especially true when a graphical approach is taken. The graphical approach to profit maximization is illustrated in **Figure 9.5 "Tangency and Isoquants"**. The curve represents an isoquant, which holds constant the output. The straight lines represent **isocost**¹⁶ lines, which hold constant the expenditure on inputs. Isocost lines solve the problem $rK + wL = \text{constant}$ and thus have slope $\frac{dK}{dL} = -\frac{w}{r}$. Isocost lines are necessarily parallel—they have the same slope. Moreover, the cost associated with an isocost line rises the farther northeast we go in the graph, or the farther away from the origin.

Figure 9.5 Tangency and Isoquants



What point on an isoquant minimizes total cost? The answer is the point associated with the lowest (most southwest) isocost that intersects the isoquant. This point will be tangent to the isoquant and is denoted by a star. At any lower cost, it isn't possible to produce the desired quantity. At any higher cost, it is possible to lower cost and still produce the quantity.

The fact that cost minimization requires a tangency between the isoquant and the isocost has a useful interpretation. The slope of the isocost is minus the ratio of input prices. The slope of the isoquant measures the substitutability of the inputs in producing the output. Economists call this slope the **marginal rate of technical**

16. Line that holds constant the expenditure on inputs.

substitution¹⁷, which is the amount of one input needed to make up for a decrease in another input while holding output constant. Thus, one feature of cost minimization is that the input price ratio equals the marginal rate of technical substitution.

KEY TAKEAWAYS

- In the long run, all inputs can be optimized, which leads to multiple first-order conditions.
- The solution can be illustrated graphically and computed explicitly for Cobb-Douglas production functions.
- An important implication of profit maximization is cost minimization—output is produced by the most efficient means possible.
- Cost minimization occurs where the ratio of the input prices equals the slope of the isocost curve, known as the marginal rate of technical substitution, which is the amount of one input needed to make up for a decrease in another input and hold output constant.

EXERCISE

1. For the Cobb-Douglas production function $F(K, L) = AK^\alpha L^\beta$, show that

$$\frac{r}{L^{**}} \frac{\partial L^{**}}{\partial r} = -\frac{\alpha}{1-\alpha-\beta}, \quad \frac{w}{L^{**}} \frac{\partial L^{**}}{\partial w} = -\frac{1-\alpha}{1-\alpha-\beta},$$

$$\frac{p}{L^{**}} \frac{\partial L^{**}}{\partial p} = \frac{1}{1-\alpha-\beta}, \quad \frac{r}{K^{**}} \frac{\partial K^{**}}{\partial r} = -\frac{1-\beta}{1-\alpha-\beta},$$

$$\frac{w}{K^{**}} \frac{\partial K^{**}}{\partial w} = -\frac{\beta}{1-\alpha-\beta} \text{ and } \frac{p}{K^{**}} \frac{\partial K^{**}}{\partial p} = \frac{1}{1-\alpha-\beta}.$$

17. The amount of one input needed to make up for a decrease in another input while holding output constant.

9.6 Myriad Costs

LEARNING OBJECTIVE

1. What are the different ways of measuring costs, and how are they related to the amount of time the firm has to change its behavior?

How much does it cost to produce a given quantity q ? We already have a detailed answer to this question, but now we need to focus less on the details and more on the “big picture.” First, let’s focus on the short run and suppose L is adjustable in the short run but K is not. Then the **short-run total cost**¹⁸ of producing q ,—that is, the total cost of output with only short-term factors varying—given the capital level, is

$$\text{SRTC}(q|K) = \min_L rK + wL, \text{ over all } L \text{ satisfying } F(K, L) \geq q.$$

In words, this equation says that the short-run total cost of the quantity q , given the existing level K , is the minimum cost, where L gets to vary (which is denoted by “min over L ”) and where the L considered is large enough to produce q . The vertical line $|$ is used to indicate a condition or conditional requirement; here $|K$ indicates that K is fixed. The minimum lets L vary but not K . Finally, there is a constraint $F(K, L) \geq q$, which indicates that one must be able to produce q with the mix of inputs because we are considering the short-run cost of q .

The short-run total cost of q , given K , has a simple form. First, since we are minimizing cost, the constraint $F(K, L) \geq q$ will be satisfied with equality, $F(K, L) = q$. This equation determines L , since K is fixed; that is, $F(K, L_S(q, K)) = q$ gives the short-run value of L , $L_S(q, K)$. Finally, the cost is then $rK + wL = rK + wL_S(q, K)$.

The **short-run marginal cost**¹⁹, given K , is just the derivative of short-run total cost with respect to output, q . To establish the short-run marginal cost, note that the equation $F(K, L) = q$ implies

$$\frac{\partial F}{\partial L}(K, L_S(q, K)) dL = dq,$$

18. The total cost of output with only short-run factors varying.

19. The derivative of short-run total cost with respect to output.

or

$$\left. \frac{dL}{dq} \right|_{F=q} = \frac{1}{\frac{\partial F}{\partial L}(K, L_S(q, K))}.$$

The tall vertical line, subscripted with $F = q$, is used to denote the constraint $F(K, L) = q$ that is being differentiated. Thus, the short-run marginal cost is

$$SRMC(q|K) = SRTC'(q) = \frac{d}{dq}(rK + wL) = w \left. \frac{dL}{dq} \right|_{F=q} = \frac{w}{\frac{\partial F}{\partial L}(K, L_S(q, K))}.$$

There are three other short-run costs we require to complete the analysis. First, there is the short-run average cost of production that we obtain by dividing the total cost by the quantity produced:

$$SRAC(q|K) = \frac{SRTC(q|K)}{q}.$$

Second, there is the **short-run variable cost**²⁰ that is the total cost minus the cost of producing zero units—that is, minus the fixed cost—which in this case is rK . Finally, we need one more short-run cost: the short-run average variable cost. The short-run average variable cost is the short-run variable cost divided by quantity, which is given,

$$SRAVC(q|K) = \frac{SRTC(q|K) - SRTC(0|K)}{q} = \frac{wL_S(q|K)}{q}.$$

The **short-run average variable cost**²¹ is the average cost ignoring the investment in capital equipment.

The **short-run average cost**²² could also be called the short-run average total cost, since it is the average of the total cost per unit of output, but “average total” is a bit of an oxymoron. An oxymoron is a word or phrase that is self-contradictory, like “jumbo shrimp,” “stationary orbit,” “virtual reality,” “modern tradition,” or “pretty ugly.” Oxymoron comes from the Greek “oxy,” meaning sharp, and “moros,” meaning dull. Thus, oxymoron is itself an oxymoron, so an oxymoron is self-descriptive. Another word that is self-descriptive is “pentasyllabic.”

Consequently, when total, fixed, or variable is not specified, the convention is to mean total. Note that the marginal variable cost is the same as the marginal total costs, because the difference between variable cost and total cost is a constant—the cost of zero production, also known as the **short-run fixed cost**²³ of production.

20. The total cost minus the cost of producing zero units.

21. The average cost ignoring the investment in capital equipment.

22. The average of the total cost per unit of output.

23. The difference between variable cost and total cost.

At this point, we have identified four distinct costs, all of which are relevant to the short run. These are the total cost, the marginal cost, the average cost, and the average variable cost. In addition, all of these costs may be considered in the long run as well. There are three differences in the long run. First, the long run lets all inputs vary, so the **long-run total cost**²⁴ is the total cost of output with all factors varying. In this case, $LRTC(q) = \min_{L,K} rK + wL$ over all L and K combinations satisfying $F(K, L) \geq q$.

Second, since all inputs can vary, the long-run cost isn't conditioned on K . Finally, the **long-run average variable cost**²⁵ is the long-run total cost divided by output; it is also known as the long-run average total cost. Since a firm could use no inputs in the long run and thus incur no costs, the cost of producing zero is zero. Therefore, in the long run, all costs are variable, and the long-run average variable cost is the long-run average total cost divided by quantity.

Note that the easiest way to find the long-run total cost is to minimize the short-run total cost over K . Since this is a function of one variable, it is straightforward to identify the K that minimizes cost, and then plug that K into the expression for total cost.

One might want to distinguish the very short run from the short run, the medium run, the long run, and the very long run. But a better approach is to view adjustment as a continuous process, with a gradual easing of the constraints. Faster adjustment costs more. Continuous adjustment is a more advanced topic, requiring an Euler equation approach.

24. The total cost of output with all factors varying.

25. The long-run total cost divided by output.

KEY TAKEAWAYS

- The short-run total cost is the minimum cost of producing a given quantity minimized over the inputs variable in the short run. Sometimes the word total is omitted.
- The short-run fixed cost is the short-run total cost at a zero quantity.
- The short-run marginal cost, given K , is just the derivative of the short-run total cost with respect to quantity.
- The short-run average cost is the short-run total cost divided by quantity.
- The short-run average variable cost is the short-run total cost minus the short-run fixed cost, all divided by quantity.
- Marginal variable cost is the same as the marginal total costs, because the difference between total cost and variable cost is the fixed cost of production, a constant.
- The long-run total cost is the minimum cost of producing a given quantity minimized over all inputs. Sometimes the word total is omitted.
- The long-run fixed cost is zero.
- The long-run marginal cost is the derivative of the long-run total cost with respect to quantity.
- The long-run average cost is the long-run total cost divided by quantity.
- The long-run average variable cost equals the long-run average cost.

EXERCISES

1. For the Cobb-Douglas production function $F(K, L) = AK^\alpha L^\beta$, with $\alpha + \beta < 1$, K fixed in the short run but not in the long run, and cost r of capital and w for labor, show

$$\text{SRTC}(q|K) = rK + w \left(\frac{q}{AK^\alpha} \right)^{\frac{1}{\beta}},$$

$$\text{SRAVC}(q|K) = w \frac{q^{\frac{1-\beta}{\beta}}}{(AK^\alpha)^{\frac{1}{\beta}}},$$

$$\text{SRMC}(q|K) = w \frac{q^{\frac{1-\beta}{\beta}}}{\beta(AK^\alpha)^{\frac{1}{\beta}}},$$

$$\text{LRTC}(q|K) = \left(\left(\frac{\alpha}{\beta} \right)^{\frac{\beta}{\alpha+\beta}} + \left(\frac{\beta}{\alpha} \right)^{\frac{\alpha}{\alpha+\beta}} \right) w^{\frac{\beta}{\alpha+\beta}} r^{\frac{\alpha}{\alpha+\beta}} \left(\frac{q}{A} \right)^{\frac{1}{\alpha+\beta}}.$$

2. Consider a cost function of producing an output q of the form $c(q) = q^2 + 2q + 16$. Determine the following:
- Marginal cost
 - Average cost
 - Average variable cost

Graph the long-run supply curve, assuming the cost function is for a single plant and can be replicated without change.