Chapter 3

Producer Theory

The goal of producer theory is to derive the supply curve and assess how firms adjust their inputs and outputs with changing market conditions. This chapter covers production functions and profit maximization whereas the subsequent chapter outlines cost minimization which ultimately leads to the supply function.

To a certain degree, there are similarities between consumer and producer theory. Consumer preferences and budget constraints correspond to production technology and cost constraints faced by firms, respectively. Although firms do not have a fixed cost constraint but more flexibility depending on the production level. Indifference curves and budget lines in consumer theory translate into isoquants and isocost lines in producer theory. What differentiates producer from consumer theory are short- and long-run aspects of production or, more specifically, input use.

3.1 Production Functions

Production is the process of combining inputs (e.g., wood, nails) to produce goods (e.g., chair). A production function indicates the maximum amount of output a firm can produce over a given time period for each combination of inputs. Common examples of inputs are labor, capital, or energy. Production functions can be thought of very broadly. For example, the production of an exam score for a university student requires the inputs time and ability (among others). An example of a production function with one input and one output is shown in Panel "Fertilizer Application" (Figure 3.1) It shows crop yield as a function of nitrogen application, i.e., $y = f(N)$. This type of production function/relationship is of interest to farmers who need to determine the nitrogen use. The goal is not necessarily to maximize yields but to maximize profits given fertilizer and crop prices. This production function can also be used to

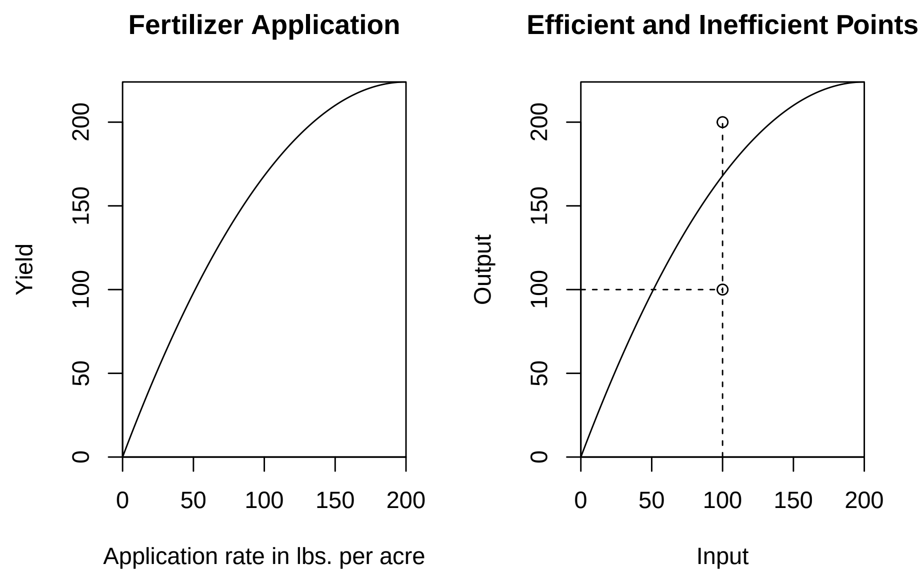

Figure 3.1: Production functions: Panel (a) on the left shows crop yield as a function of fertilizer application. Panel (b) on the right illustrates the concept of inefficient and unattainable production points.

assess the effects to a water treatment plant (e.g., Citizens Energy Group in Indianapolis, Des Moines Water Works) for its operations to remove excess nutrients. Everything below a production function is attainable but inefficient since the same inputs can be used to achieve a higher level of production or fewer inputs can be used to produce the same quantity (Panel "Efficient and Inefficient Points," Figure 3.1). Anything above the production function is not attainable given the current technology of the firm.

Of course, more interesting problems and trade-offs can be analyzed in the case of two inputs (Figure 3.2). As in consumer theory, a Cobb-Douglas production function is used with an added parameter compared to a utility function:

$$Q = f(K, L) = A \cdot L^{\alpha} \cdot K^{\beta}$$

Here, L and K represent labor and capital input, respectively, whereas A represents so-called total factor productivity (TFP). The TFP is the efficiency parameter A that scales output independently of input quantities, capturing technology, knowledge, and other factors not explained by labor and capital. Note that this term is missing from the Cobb-Douglas utility function since in consumer theory only the ordinal ranking matters. In producer theory, however, input use and output (Q) production are cardinal measures and thus, need to

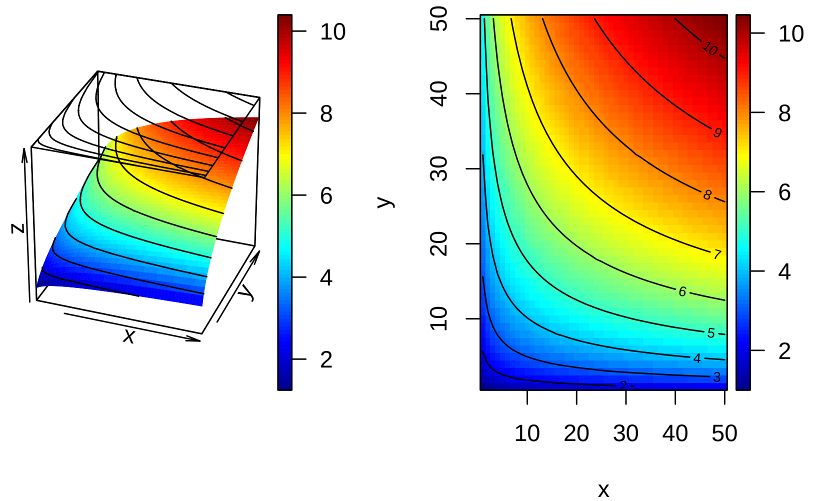

Figure 3.2: Production function with two inputs, i.e., x and y. Output is denoted with z. Note that in the right panel, the curves correspond to identical combinations of x and y to produce the same amount of z.

be scaled accordingly with A. The meaning and use of L and K becomes clear later in this chapter.

One of the reasons that the Cobb-Douglas function is commonly used in economics is its ability to model so-called returns to scale, which can be either constant, increasing, or decreasing. Constant returns to scale represent a production technology that leads to the proportional increase of output due to an increase in inputs. For example, a doubling of inputs leads to a doubling in output. In the case of a single input (x), this translates into the relationship $t \cdot Q = f(t \cdot x)$ where t is some positive number. Increasing returns to scale are formalized as $t \cdot Q > f(t \cdot x)$ whereas decreasing returns to scale are written as $t \cdot Q < f(t \cdot x)$. For example, if inputs double, then output more (less) than doubles under increasing (decreasing) returns to scale. In the case of the Cobb-Douglas function, $\alpha + \beta = 1$ represents a constant returns to scale production function. Increasing and decreasing returns to scale are obtained in the case of $\alpha + \beta > 1$ and $\alpha + \beta < 1$, respectively (Table 3.1).

Similar to consumer theory, the majority of results from producer theory can be obtained from a production function with two inputs. When teaching consumer theory, the case of two goods is often used because it provides a simple and intuitive graphical framework. With two goods, we can draw indifference curves

α	β	L	K	Q	t	$f(t \cdot K, t \cdot L)$	$f(t \cdot K, t \cdot L)/Q$
0.4	0.6	10	20	151.57	2	303.14	2.00
0.8	0.4	10	20	209.13	2	480.45	2.30
0.2	0.3	10	20	38.93	2	55.06	1.41

Table 3.1: Illustration of economies of scale for a Cobb-Douglas production function.

and budget lines in two-dimensional space, which makes it easy to illustrate concepts like preferences, marginal rates of substitution, and optimal consumption choices. The underlying assumption is that consumers allocate their limited budget across goods to maximize utility, and two goods are sufficient to capture the trade-offs without complicating the analysis. There is a difference though for producer theory, particularly when distinguishing between the short- and long-run. That is, while consumer theory with two goods illustrates trade-offs under a fixed budget constraint, producer theory highlights the difference between short-run rigidity and long-run flexibility in input use. The consumer's problem is about allocating expenditure across goods, whereas the producer's problem is about choosing inputs and technologies over different time horizons.

3.2 Production in the Short- and Long-Run

In producer theory, we differentiate between the short-run and the long-run. In the short-run, some inputs are fixed and cannot be adjusted to changes in their price. For example, a restaurant that is renting a space is unable to adjust its area in the short-run since the rental contract is for a specific time period and rent is due every month even if the number of customers served is zero. The same is true for an airline that cannot adjust the number of planes from one day to the next because it takes time to acquire new planes or sell them. Even if the airline reduces its flight connections, there are still costs associated with the planes. In producer theory, those fixed inputs (e.g., machinery, factories) are labeled as capital (K). In the short-run, the firm can only adjust its variable inputs which are labeled as labor (L). For example, the number of cooks and servers in the case of the restaurant and the (ground) crew in the airline case can be adjusted in the short-run. Inputs are flexible only in the long-run.

3.2.1 Short-Run Production

In a first case, production in the short-run is considered with capital being fixed at $\bar{K}$. That is, the short-run production function is written as $Q = f(\bar{K}, L)$. For example, consider the following Cobb-Douglas production function:

$$Q = 10 \cdot K^{0.5} L^{0.5} = 10 \cdot \sqrt{K \cdot L}$$

If capital is fixed, for example, at $\bar{K} = 9$, then short-run production is only a function of L:

$$Q = 30 \cdot \sqrt{L}$$

This function represents the short-run production when only labor can be adjusted in order to change output. This is similar to the airline that can only adjust the number of pilots, flight attendants, and ground crew in the short-run but not the number of planes which are considered capital.

A production function of the following form is used to introduce the concepts of average and marginal product of labor:

$$Q = L + \alpha \cdot L^2 + \beta \cdot L^3$$

This production function results in increasing returns to scale for low levels of L and increasing returns to scale for higher levels of L. The marginal product of labor (MPL) is the additional output produced if one more worker is hired, i.e.,

$$MPL = \frac{\Delta Q}{\Delta L}$$

where Δ represents the change in quantity or labor. Put differently, the marginal product of labor is simply the slope of the production function at a particular input level of L (Figure 3.3). The average product (AP) is defined as the average quantity produced per worker:

$$AP = \frac{Q}{L}$$

Graphically, the average product of labor at a specific level of labor represents the slope of a line connecting the origin of the production function and the point on the production function at that labor level. This is related to the concept of rise-over-run where the rise represents the output Q and run represents the level of labor L.

The law of diminishing marginal product states that as more of any input is added to a fixed amount of other inputs, its marginal product eventually declines. Note that for low inputs, we can have an increasing marginal product to labor, i.e., MPL increases as more labor is hired. But in most cases, we are faced with diminishing marginal product of labor.

3.2.2 Long-Run Production

The more interesting case is production in the long-run when all inputs are variable. The optimal input combination of capital and labor needs to consider the cost associated with those factors of production. Let w be the wage per worker (per unit of L) and let r be the rental rate of capital per unit of K. The total cost of production given a certain level of K and L can be written as

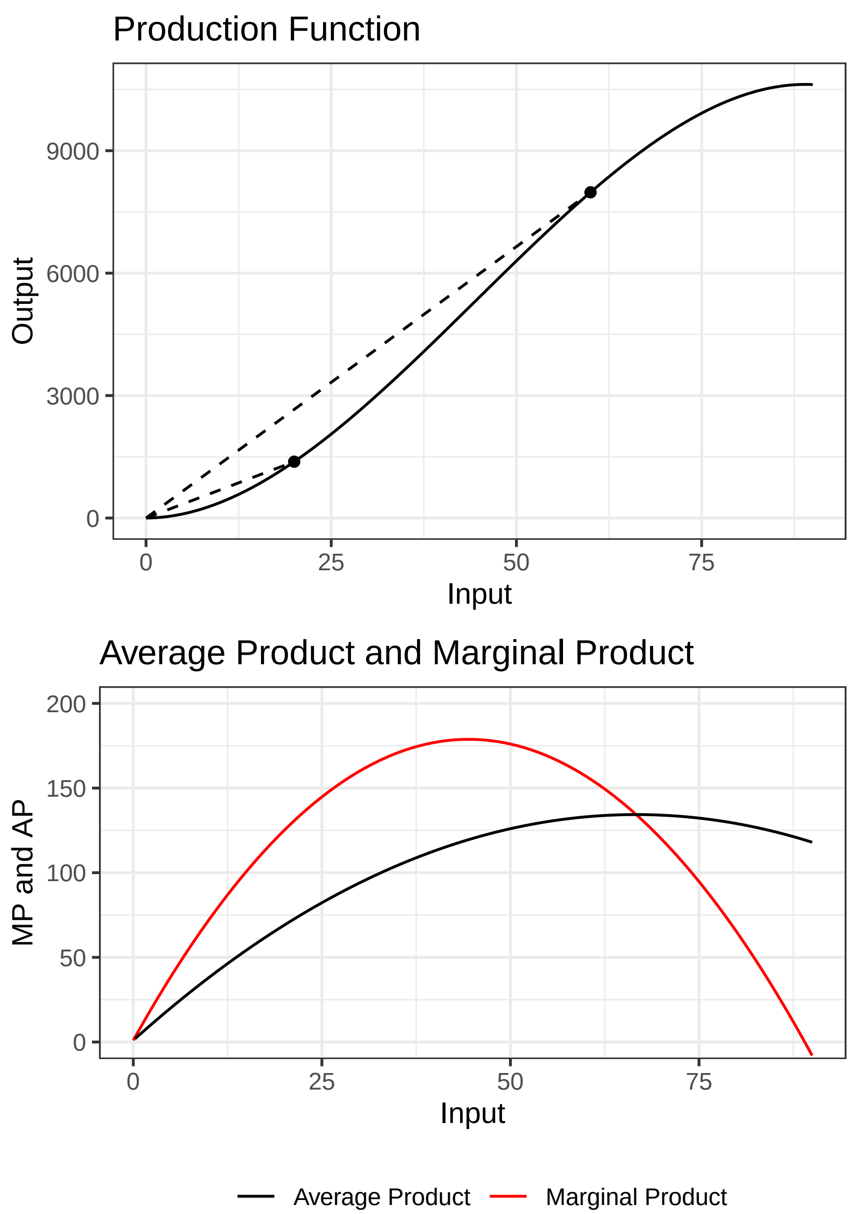

Figure 3.3: Relationship between production function, average product, and marginal product

follows:

$$TC = w \cdot L + r \cdot K$$
$$K = \frac{TC}{r} - \frac{w}{r} \cdot L$$

This is very similar to a budget constraint from consumer theory with one major difference. For consumers, the income is fixed whereas for a firm, the total cost TC is not fixed but can be adjusted based on output quantity. Note that in this section, the optimal output quantity is not covered and there is also no discussion about the output price, i.e., the price at which the produced good is sold. The expression $K = f(L)$ above represents all the combinations of K and L that result in the same cost for a given set TC, w, and r. This line is called the isocost line.

Figure 3.2 shows the so-called isoquant lines. Those lines (similar to indifference curves) represent all the combinations of K and L resulting in the same output quantity. The Marginal Rate of Technical Substitution (MRTS) is the rate at which a firm can substitute one input for another while keeping output constant. The MRTS decreases as we move rightward along an isoquant. It is also the slope of the isoquant, which represents all the combinations of capital and labor that result in the same production quantity Q.

$$MRTS = -\frac{\Delta K}{\Delta L} = \frac{MP_L}{MP_K}$$

Recall that

$$MP_K = \frac{\Delta Q}{\Delta K} \quad \text{and} \quad MP_L = \frac{\Delta Q}{\Delta L}$$

The optimal choice is the point where an isocost line is tangent to the isoquant for that output level (Figure 3.4)

$$MRTS = -\frac{\Delta K}{\Delta L} = \frac{p_L}{p_K} \frac{MP_L}{p_L} = \frac{MP_K}{P_K}$$

That is, the marginal product per dollar of any input will be equal to the marginal product per dollar of any other input.

3.2.3 Expansion Paths

The expansion path describes how a firm chooses input combinations as output expands, but the distinction between short-run and long-run is critical (Figure 3.4). In the short run, at least one input is fixed, so the firm cannot freely adjust input proportions. As a result, the chosen input bundle may not equate the marginal rate of technical substitution (MRTS) with the ratio of input prices. Instead, the firm is forced to use more of the variable input while holding the fixed input constant, which generally leads to higher total costs than would be possible with full flexibility. In the long run, by contrast, all inputs are variable, and the firm can adjust their proportions so that the isoquant is tangent to the

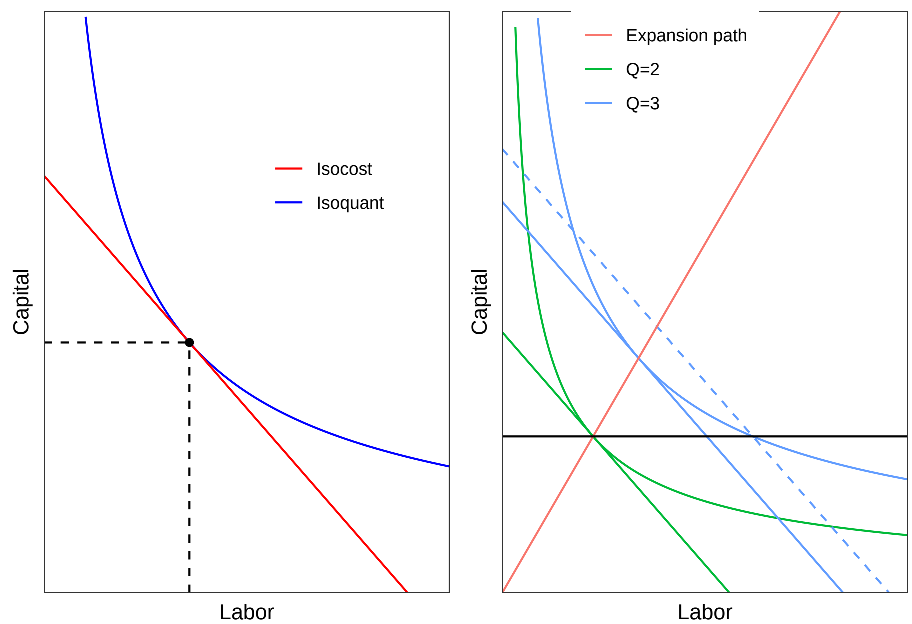

Figure 3.4: Panel (a): Cost minimizing combination of capital and labor for a given output quantity. The isocost line is tangent to the isoquant. Panel (b): Short-run and long-run expansion path

isocost line. This tangency ensures that the MRTS equals the ratio of input prices, which minimizes cost for any given level of output. Consequently, the long-run expansion path reflects truly cost-minimizing input combinations, and total costs are always less than or equal to those incurred in the short run for the same level of output.

3.3 Profit Maximization

This section introduces the concept of profit maximization. Assume we have a single output produced via the following production function:

$$Q = 30 \cdot \sqrt{L}$$

Assume that the labor cost is $w = 10$ and that the output price is $p = 5$. Then the profit maximization problem can be written as:

$$\max \quad p \cdot f(L) - w \cdot L$$

And the solution to this problem can be written as:

$$p \cdot f'(L) = w \quad \Leftrightarrow \quad f'(L) = \frac{w}{p}$$

Put differently $\pi = p \cdot y - w \cdot x$. Solving for y:

$$y = \frac{\pi - w \cdot x}{p} = \frac{\pi}{p} + \frac{w}{p} \cdot x$$

The profit maximization problem as a function of output is $\pi(q) = R(q) - C(q)$ where $R = p \cdot q$. To find the profit maximizing output, we need the marginal revenue

$$\frac{\Delta R}{\Delta q} = p$$

and the marginal cost

$$\frac{\Delta C}{\Delta Q} = MC$$

Hence, the profit maximization condition is: $MC(q) = MR(q)$. This is true for any market structure. What differs across market structures is marginal revenue.

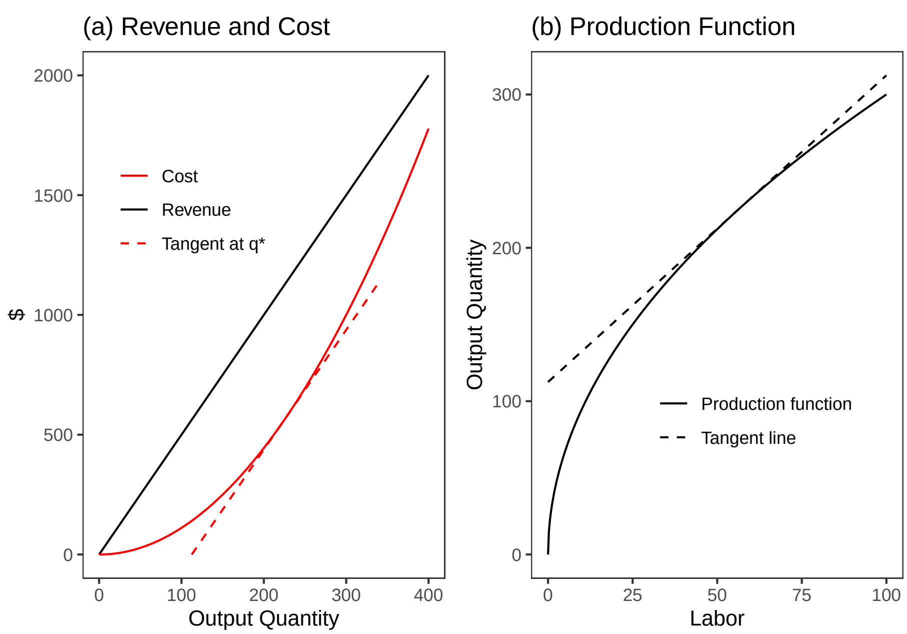

Figure 3.5: Two different approaches to illustrate profit maximization. Panel (a): Revenue and cost approach with setting marginal revenue (slope of the revenue function) equal to marginal cost (slope of the cost function). Panel (b): Profit maximizing input use based on production function.

3.4 Exercises

1. ***Nitrogen Use*** (*): A farmer uses nitrogen (N) as an input to produce corn yield as the output. The production function which relates nitrogen inputs to corn yield is written as $y = f(N)$. With the current management practices, the farmer obtains $y^\star$ bushels of corn with $N^\star = 120$ pounds of nitrogen fertilizer. Given that setup, we have $y^\star < f(120)$. Are the inputs used efficiently? Use a graph to justify your answer.

2. ***Malthus*** (*): The Malthusian Catastrophe was the prediction by the English cleric and scholar Thomas Malthus (1766-1834) that population will grow faster than agricultural production. Use the concept of diminishing returns to labor in the short-run to explain his reasoning. Explain why it did not happen. Use a graph to justify your answer.

3. ***Profit Function*** (**): The production function of a firm is written as $f(x) = 8\sqrt{x}$ where x are the units of input. The per-unit output price is \$100 and the per-unit input cost is \$75. Write down the profit only as a function of x.

4. ***Electricity Producer*** (***): Assume an electricity producer that currently uses a mix of natural gas and coal as inputs to produce electricity as depicted in Figure 3.6. In the short-run, the electricity producer can adjust neither coal nor natural gas as the input for electricity production, i.e., both inputs are fixed. The initial input mix of coal and natural gas as well as the current isoquant are depicted in the figure . The output does not change throughout the exercise.

   a. Reproduce the figure and sketch an isocost for the quantity of electricity produced.
   b. Assume that in the short-run, the price of coal decreases relative to natural gas. What happens to the isocost line? Is the choice of inputs optimal after the price change? Justify your answer.
   c. In the long-run, what will happen to the combination of coal and natural gas used in the production of electricity if the price of coal remains low. Illustrate in your graph.

6. ***Copper Mine*** (***): Suppose you are running a copper mine. You can either use machinery or workers to extract the copper. Currently, you are using a 600 horse power train to get the copper out of the mine. You have signed a leasing agreement for that train which you cannot change in the short-run. In Figure 3.7, you will see two isoquant curves and a point indicating the initial (optimal) combination of workers and the train. The input prices remain unchanged throughout the question. Sketch the corresponding isocost line through the initial, optimal point. Assume that the machinery is fixed at 600 horse power and cannot be changed in the short-run. Due to changed demand, you need to expand production to the higher isoquant line. Show the new isocost line. Is it optimal? If yes, why?

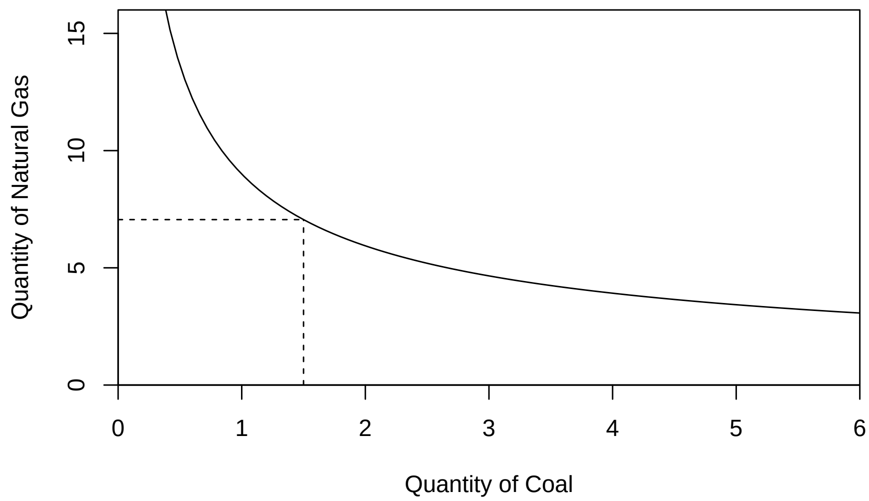

Figure 3.6: Electricity production with natural gas and coal.

If no, why not? You are now able to adjust workers and machinery in the long-run while staying on the higher isoquant curve. What is the effect on cost? What is the effect on the optimal input combination. Support your answer in the graph.

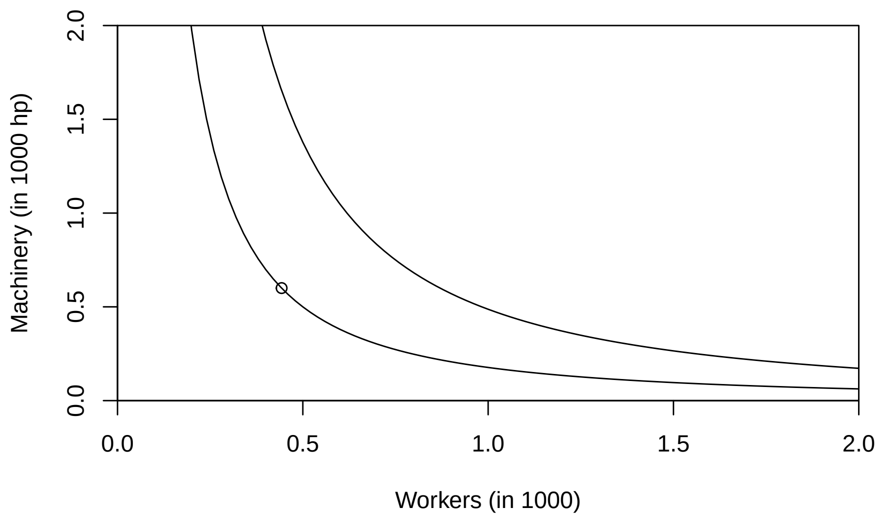

Figure 3.7: Copper Mine: This figure illustrates the initial production point on the lower isoquant and the new isoquant after production has increased.