Unit 5.1: Production

Michael Malcolm

June 18, 2011

1 Production Inputs

A *firm* is an organization that combines inputs to produce output. Typically, we divide firm inputs into:

- *Labor* inputs, that are variable in both the short run and in the long run
- *Capital* inputs, that are fixed in the short run but variable over the long run

The amount of a capital input that the firm uses cannot be changed in the short run. For example, a car manufacturer likely cannot change the number of factories that it employs in the short run. However, the amount of labor inputs used can be changed both in the short-run and in the long-run. By convention, "labor" includes other inputs like raw materials that can be added in the short-run.

As such, the *long run* is defined as a period of time that is long enough for a firm to change all its inputs. Conversely, the *short run* is a period of time in which at least one input is fixed. Notice that the period of time that corresponds to the long run and the short run might be different for different firms.

2 Production Functions

A *production function* describes the amount of output q that can be produced with given inputs. Where L designates labor and K designates capital, then the production function $q = f(L, K)$ describes the number of units of output that can be produced when the firm employs L units of labor and K units of capital. The following is an example of a production function.

$$q = 0.1LK + 3L^2K - 0.1L^3K$$

The *short run production function* holds $K = \bar{K}$ fixed and considers how output varies with changes in L. For example, if the firm with the production function above has $\bar{K} = 10$ units of capital in the short-run, then its short run production function is:

$$q = 0.1L(10) + 3L^2(10) - 0.1L^3(10)$$
$$q = L + 30L^2 - L^3$$

The production function function gives the *total product* − the output produced by L workers. The *marginal product of labor* (MPL) is the additional output produced by one more worker. As such, the MPL is the derivative of the production function with respect to L.

$$MPL = \frac{\partial q}{\partial L} = 1 + 60L - 3L^2$$

The *average product of labor* (APL) is the average output produced per worker. For this production function:

$$APL = \frac{q}{L} = 1 + 30L - L^2$$

Notice that the marginal and average product of capital or any other input could be defined similarly.

3 Diminishing Marginal Returns

As L rises, output rises, so in general the marginal product is never negative. As the first few workers are added, marginal product may rise initially because of specialization. Typically, however, the additional output added by each worker will start to decline as more workers are hired.

The *law of diminishing marginal returns* states that, as a firm increases more of any one input **while holding other inputs fixed**, then the marginal product of the input being added will eventually decline. So, as a firm adds workers while holding capital fixed, the marginal product of each worker will eventually fall. We are not saying that output falls as more workers are added, just that each successive worker adds less than the previous worker.

The reason for diminishing marginal returns is straightforward. When more labor is added without adding more capital, there are increasingly more workers "sharing" each unit of capital, so eventually each worker will be less productive.

Mathematically, diminishing marginal returns implies that the production function is concave – output rises as output rises, but at a slowing rate. Using calculus, this means that the second derivative is negative. For our example, diminishing marginal returns holds when

$$\begin{aligned} \frac{\partial^2 q}{\partial L^2} &< 0 \\ 60 - 6L &< 0 \\ L &> 10 \end{aligned}$$

So, for our example, diminishing marginal returns sets in any time more than 10 workers are hired. Past this point, marginal product of each worker falls as more workers are hired.

4 Total, Average and Marginal Product

When marginal product is higher than average product, this means that an additional worker produces more output than the average worker employed by the firm. This means that average product will *rise* when additional workers are hired. Conversely, if marginal product is lower than average product then average product is *falling* since each additional worker adds less output than average, and so additional workers pull down the average.

Geometrically, this means that APL slopes upwards when $MPL > APL$ but slopes downwards if $MPL < APL$. Note that this implies that MPL crosses APL at the maximum point of the APL curve.

The total product curve may rise quickly initially as MPL rises when the first workers are added. However, eventually it starts to rise more slowly as workers are added (MPL falling). Note that the point where the concavity of the total product function switches from convex to concave is precisely the point where MPL begins to fall.

Figure 1 shows the relationships between the total, average and marginal product functions. Note that you should never show total product on the same diagram as average and marginal product. Total product measures total output over all workers, whereas average and marginal product are measuring output in per-worker terms.

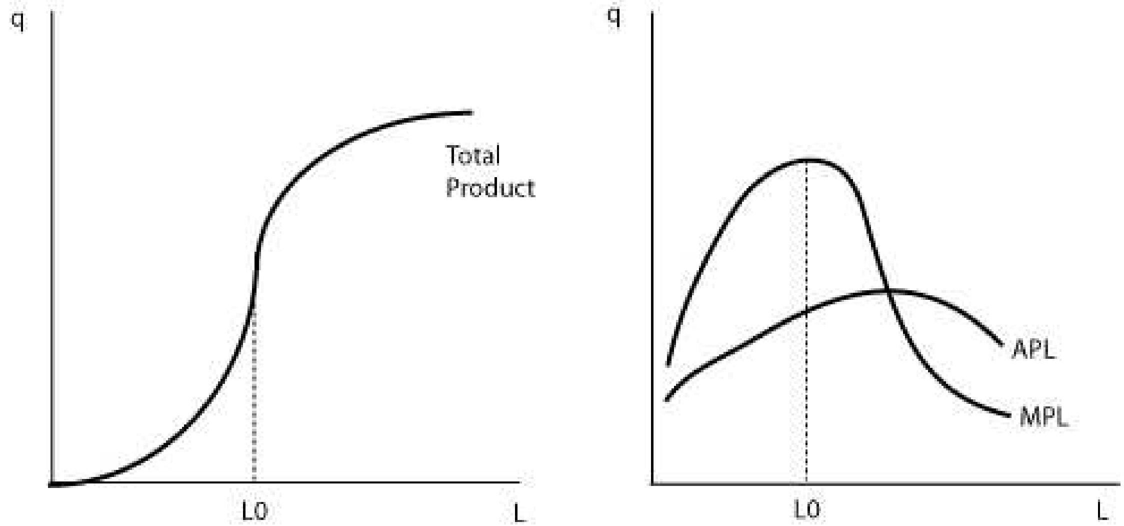

Figure 1: Total, Average and Marginal Product

5 Long-Run Production

In the long-run, both capital and labor can change. Consider the following long-run production function.

$$q = \sqrt{L}\sqrt{K}$$

An *isoquant* describes various combinations of capital and labor that generate the same level of output. For our example, the following combinations all produce $q = 6$ units of output.

$$(L, K) = \{(36, 1), (9, 4), (6, 6), (4, 9), (1, 36)\}$$

The isoquant for $q = 6$, containing all these different combinations, is shown in figure 2.

Isoquants are level sets of a production function in exactly the same way as indifference curves are level sets of a utility function. As such, they obey the same sorts of properties. Higher isoquants, with more of both inputs, correspond to higher levels of output. Isoquants do not cross. Also, isoquants slope downwards – as labor is added, capital must be reduced in order to keep output constant.

The *marginal rate of technical substitution* (MRTS) is the amount of capital needed to replace one unit of labor while keeping output the same. Isoquants typically obey diminishing MRTS as more labor is added. When a firm has only a small amount of labor, each worker is very productive and so a lot of capital is needed to replace a worker and keep output constant. However, as a firm hires more workers, each worker is less productive and so less capital is needed to replace a worker.

Mathematically, the MRTS is the ratio of the marginal product of labor to the marginal product of capital. If $MPL = 10$ and $MPK = 5$, then on the margin 2 units of capital are needed to replace one worker. For our example, where $q = \sqrt{L}\sqrt{K} = L^{\frac{1}{2}}K^{\frac{1}{2}}$, we can calculate the MPL and the MPK:

$$MPL = \frac{\partial q}{\partial L} = \frac{1}{2}L^{-\frac{1}{2}}K^{\frac{1}{2}}$$
$$MPK = \frac{\partial q}{\partial K} = \frac{1}{2}L^{\frac{1}{2}}K^{-\frac{1}{2}}$$

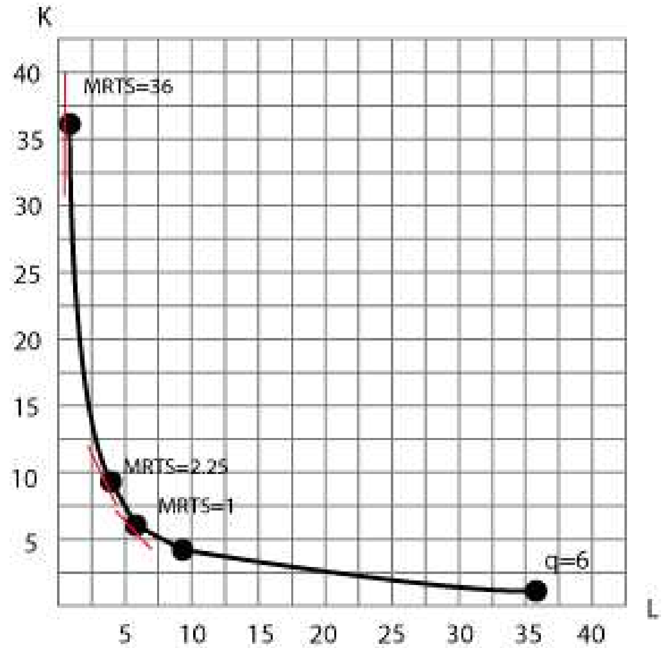

Figure 2: A typical isoquant

Then the MRTS is the ratio:

$$\begin{aligned} MRTS &= \frac{MPL}{MPK} \\ &= \frac{\frac{1}{2}K^{\frac{1}{2}}L^{-\frac{1}{2}}}{\frac{1}{2}K^{-\frac{1}{2}}L^{\frac{1}{2}}} = \frac{K}{L} \end{aligned}$$

Graphically, the MRTS is the slope of the isoquant. At $(L, K) = (1, 36)$, $MRTS = 36$. At $(L, K) = (4, 9)$, $MRTS = 2.25$. At $(L, K) = (6, 6)$, $MRTS = 1$. As expected, as labor rises, the MRTS falls and so the isoquant flattens. The MRTS at various points is shown in figure 2.

6 Perfect Complements and Perfect Substitutes

The production function $q = L + K$ represents the case where the two inputs are perfectly substitutable for each other (e.g. a firm where a robot is the same as a worker). One additional unit of capital always adds the same amount of output as one additional unit of labor. Like perfect substitutes in consumer theory, the isoquants are linear, as shown in figure 3.

The production function $q = \min\{L, K\}$ represents the case where the two inputs are perfect complements (e.g. typists and typewriters). Capital and labor must be added together in the same proportion in order to raise output. Again, like perfect complements in consumer theory, the isoquants are right angles, as shown in figure 4.

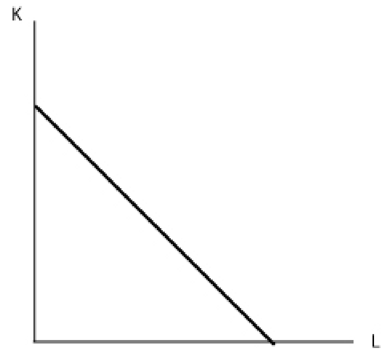

Figure 3: An isoquant for inputs that are perfect substitutes

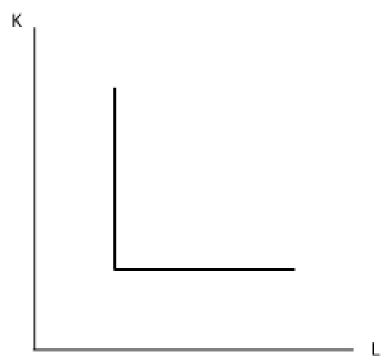

Figure 4: An isoquant for inputs that are perfect complements

7 Returns to Scale

Let $f(L, K)$ be the production function. Suppose that the firm then doubles *all* its inputs, so that its output is $f(2L, 2K)$. In other words, the firm makes an exact copy of its current production process.

- If $f(2L, 2K) = 2f(L, K)$, meaning that doubling all inputs exactly doubles the original output, then we say that the production process obeys *constant returns to scale*.
- If $f(2L, 2K) > 2f(L, K)$, meaning that doubling all inputs more than doubles the original output, then we say that the production process obeys *increasing returns to scale*.
- If $f(2L, 2K) < 2f(L, K)$, meaning that doubling all inputs less than doubles the original output, then we say that the production process obeys *decreasing returns to scale*.

For example, consider the production function $f(L, K) = \sqrt{L}\sqrt{K}$. If we double both L and K:

$$\begin{aligned} f(2L, 2K) &= \sqrt{2L}\sqrt{2K} \\ &= 2^{\frac{1}{2}} L^{\frac{1}{2}} 2^{\frac{1}{2}} K^{\frac{1}{2}} \\ &= 2L^{\frac{1}{2}} K^{\frac{1}{2}} \\ &= 2f(L, K) \end{aligned}$$

This production function obeys constant returns to scale since doubling both inputs resulted in exactly twice the original output.

Now consider the production function $f(L, K) = L^a K^b$. Any production function of this form is known as a *Cobb-Douglas* production function. Doubling both inputs:

$$\begin{aligned} f(2L, 2K) &= (2L)^a (2K)^b \\ &= 2^a L^a 2^b K^b \\ &= 2^{a+b} L^a K^b \\ &= 2^{a+b} f(L, K) \end{aligned}$$

Now, if $a + b < 1$, then $f(2L, 2K) < 2f(L, K)$ and so the production function exhibits decreasing returns to scale. If $a + b > 1$, then $f(2L, 2K) > 2f(L, K)$ and so the production function exhibits increasing returns to scale. Finally, if $a + b = 1$, then $f(2L, 2K) = 2f(L, K)$ and so the production function exhibits constant returns to scale. A Cobb-Douglas production function can have any scale property, depending on the values of the parameters a and b.

It is important to understand the difference between decreasing returns to scale and diminishing marginal returns. Decreasing returns to scale is a statement about the effect of varying *all* inputs, while diminishing marginal returns is a statement about the effect of varying *one* input and holding other inputs fixed. As such, decreasing returns to scale is a long run phenomenon whereas diminishing marginal returns is a short run phenomenon.

It is perfectly possible that a firm with increasing returns to scale when varying *all* its inputs faces diminishing marginal returns when varying a *single* input.

Furthermore, scale properties can be different over different ranges of input usage. Often, firms face increasing returns to scale early on both because of specialization and because of more efficient use of lumpy inputs (e.g. a medical office with 20 doctors will be making more efficient use of its X-ray machine than a medical office with only 1 doctor). However, decreasing returns to scale can set in as the firm gets larger. Inefficiency, mismanagement and bureaucracy can mean that an already large firm that doubles all its inputs will have a difficult time doubling its output.