Short-Run Production and Costs

The purpose of this section is to discuss the underlying work of firms in the short-run – the production of goods and services. Why is understanding production important to understanding firm behavior? Recall that firms are profit maximizers. We learned in the last chapter that profit equals the difference between total revenue and total costs. As you will learn below, production affects costs. Hence, the first part of this chapter discusses production in the short-run while in the second part we discuss how production affects costs and what those costs consist of in the short-run.

I. Production in the Short-Run

Recall that the short-run is a period of time during which the firm may be able to change some of its inputs but cannot change all of them. At least one input, often capital but not always, is fixed in the short-run, which limits the choices that firms may make. In this section, we will discuss in detail how production works in the short-run.

A. The Production Function

Recall that in the previous chapter we discussed the production function, which illustrates the relationship between output (Q) and inputs (L, K, N, E):

\[ Q = f(L, K, N, E) \] (1)

The main point of this section is to discuss the exact nature of the relationship between the inputs and the output. To do that, we must first simplify the production function. The first step of the simplification is to assume that we are talking about a very simple production process. The one we will use is cleaning out irrigation ditches. Cleaning irrigation ditches is a production process that I was involved in as a teenager and requires only two inputs – capital (shovels) and labor (teenagers). Since we are in the short-run, at least one input must be fixed. For our example, we will assume that the number of shovels is fixed while the firm can change the amount of labor. Thus, our production function becomes:

\[ Q = f(L, \bar{K}) \] (2)

The bar over the capital sign (\( \bar{K} \)) indicates that this input while present and used in production cannot be changed in the short-run. The production function in this simple example, where labor is the only input that can be changed, is sometimes referred to as the total product of labor (TP\(_L\)).\(^1\)

---

\(^1\) Some books refer to the production function as the total physical product of Labor or TP\(_P\). These two terms, TP\(_L\) and TP\(_P\), are interchangeable. Both simply refer to the short-run production function in equation 2.
Now that we only have two variables that can change, labor (L) and output (Q), we can discuss the relationship between the two. It seems reasonable that using more labor results in more output so perhaps the production function simply looks like the one presented in Graph 1.

The production function in Graph 1 is simply an upward sloping line. While this may seem reasonable, the question mark in Graph 1 is present because we have not yet really determined whether or not the production function is really this simple – use more labor, get more output. In order to determine whether or not the production function is as shown in Graph 1, we must first discuss a new concept, marginal productivity.

B. The Marginal Product of Labor

Recall from our discussion of consumers and utility, we had already defined one marginal concept – marginal utility. During that discussion we pointed out that such marginal concepts are both numerous in economics and quite important. Marginal Product of Labor (MP_L) is the second, but certainly will not be the last, marginal concept presented in microeconomics principles.

We noted previously that all marginal values have a definition that is similar, although not exactly the same, to marginal utility. An actor in the market, typically a household or a firm, is engaging in an activity. The activity yields or causes some other variable to change. For marginal utility, households are engaging in consumption, which causes their total utility to change. The marginal concept simply measures the extra amount of the variable, such as utility, caused by a one-unit change in the activity, such as consumption of a good.

For marginal product of labor, the firm is engaged in hiring labor in order to change output. Thus, marginal product of labor simply equals the extra amount of output gained by the firm by hiring one more unit of labor. Furthermore, recall that mathematically marginal utility was given by:

\[
\text{Marginal Utility} = \frac{\Delta TU}{\Delta Q}
\]  

That is, the marginal concept mathematically equals the change in the variable affected by the activity divided by the change in the activity itself. Since utility is affected by the change in the activity, consumption, marginal utility equals the ratio of the change in utility to change in consumption.

For marginal production of labor, the activity is hiring labor whereas hiring labor causes output to change. Thus, marginal product of labor is given by:

\[
\text{Marginal Product of Labor}(MP_L) = \frac{\Delta Q}{\Delta L}
\]  

We can see these concepts the most clearly simply by looking at an example:
Notice that even though output rises from 0 to 200, marginal product is only 100. This is because the firm hired two more workers and we want to know how much output would rise if they only hired one additional worker.

C. The Law of Diminishing Returns

Let’s now address the issue of what the two concepts we’ve introduced so far, total product and marginal product of labor, look like graphically. To understand these issues we will focus on marginal productivity. To understand this concept simply consider the production process from our example above, cleaning out ditches. We only have two inputs, teenagers (labor) and shovels (capital). However, because we are in the short-run the number of shovels is fixed – let’s suppose that the firm has five, and only five, shovels on hand to clean ditches.

We can change the number of teenagers we hire, however. What happens to marginal product of labor as the amount of labor hired increases? Clearly, when we hire no workers then we get no output – ditches don’t get cleaned with shovels alone. Thus, with the first worker marginal product of labor must be positive (as was shown in Table 1 above). What happens when we hire the second worker? One of the possibilities is that the amount this second worker produces (her marginal product of labor) may actually be larger than the amount produced by the first worker.

That is, marginal product may actually rise as the firm hires more teenagers. This increase in marginal product of labor often happens because of specialization of labor. As we discussed at the beginning of the course, specialization of labor – workers who specialize in only one part of the production process – often allows those workers to increase their productivity. Recall that this was one of the reasons why Henry Ford became wealthy. He recognized the potential gains from specialization by using an assembly line to manufacture cars.

How long do returns to specialization last as the firm hires more labor? Well, this depends entirely upon the particular situation of the firm – its technology, its workers, its production process, etc. Thus, the third worker might produce even more and the fourth worker even more.

<table>
<thead>
<tr>
<th>Labor Hired</th>
<th>Total Product of Labor (Q)</th>
<th>Marginal Product of Labor (ΔQ/ΔL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>500</td>
<td>150</td>
</tr>
<tr>
<td>6</td>
<td>700</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>800</td>
<td>50</td>
</tr>
</tbody>
</table>
But what must happen eventually in the short-run as the firm hires more and more labor? Remember that the short-run means that the amount of capital we have is limited; in our example we only have five shovels. Suppose we hire a sixth worker so that now there are not enough shovels given the number of workers. Perhaps the sixth worker specializes in management and directs the work of the other workers, increasing marginal product even more. But as more and more workers are added, still with only five shovels, the firm eventually has workers standing around, perhaps even congesting the workplace and reducing the productivity of workers with shovels. Thus, even though marginal product may initially rise due to specialization it must eventually fall due to this type of congestion.

This result is known as the law of diminishing returns (or the law of diminishing marginal productivity): 
Eventually as the use of a resource increases, the marginal product of that resource must eventually fall, ceteris paribus. Holding everything else constant (ceteris paribus) in our example means that we are in short-run where only labor can change.

Graph 2 shows us what the marginal product of labor looks like, graphically. Initially, the marginal product rises as the firm hires labor because of the returns to specialization. Eventually, however, diminishing returns begin to take effect and the marginal product falls as labor increases. In fact, the marginal product may even be negative for high enough levels of labor, as shown in the graph.

D. Average Product of Labor

For most students, understanding the concept of average product is much easier than marginal product. We all are used to calculating averages in our lives. For example, we know how to calculate the average height of people in a room. Simply measure each person, add up the height of each person and then divide by the number of people in the room.

Average product of labor is a similar concept. However, rather than measuring the height of each person we are measuring how much they produce. When we add these all up we get the total output – given by the production function. Again, we divide by the total number of people, the total number of workers the firm has hired. Thus, average product of labor is simply given by:

$$\text{Average Product of Labor (AP}_L) = \frac{Q}{L} \quad (5)$$

Table 2 illustrates the calculation of average product of labor using another example.
<table>
<thead>
<tr>
<th>Labor Hired</th>
<th>TP&lt;sub&gt;L&lt;/sub&gt; (Q)</th>
<th>MP&lt;sub&gt;L&lt;/sub&gt; (∆Q/∆L)</th>
<th>AP&lt;sub&gt;L&lt;/sub&gt; (Q/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>--</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>70</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>86</td>
<td>16</td>
<td>14.33</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>14</td>
<td>14.29</td>
</tr>
<tr>
<td>8</td>
<td>112</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>122</td>
<td>10</td>
<td>13.56</td>
</tr>
<tr>
<td>10</td>
<td>130</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>11</td>
<td>136</td>
<td>6</td>
<td>12.36</td>
</tr>
<tr>
<td>12</td>
<td>140</td>
<td>4</td>
<td>11.67</td>
</tr>
<tr>
<td>13</td>
<td>142</td>
<td>2</td>
<td>10.92</td>
</tr>
<tr>
<td>14</td>
<td>142</td>
<td>0</td>
<td>10.14</td>
</tr>
<tr>
<td>15</td>
<td>140</td>
<td>-2</td>
<td>9.33</td>
</tr>
<tr>
<td>16</td>
<td>136</td>
<td>-4</td>
<td>8.5</td>
</tr>
<tr>
<td>17</td>
<td>130</td>
<td>-6</td>
<td>7.65</td>
</tr>
<tr>
<td>18</td>
<td>122</td>
<td>-8</td>
<td>6.78</td>
</tr>
</tbody>
</table>
E. The Relationship between MP\textsubscript{L}, AP\textsubscript{L}, and TP\textsubscript{L}

The relationship between MP\textsubscript{L} and AP\textsubscript{L} is fairly straightforward and can be simply illustrated using our example of height. We find the average height of a group of people by adding up each person’s height and dividing by the number of people. What is the marginal height? If we add one more person to the group, that person’s height is the marginal height.

Suppose that the group’s average height is 5’ 8”. Now, a person is added to the group who is 6’ 4” tall. What happens to the average height? Clearly, if a person who is taller than the average were included in the group, then the average height would rise. Just as clearly, if a person who is shorter than the average joins the group, then the group’s average height would fall. A similar relationship exists between marginal and average productivity.

\[
\begin{align*}
\text{If } MP_L > AP_L & \Rightarrow AP_L \uparrow \text{ as Labor } \uparrow \\
\text{If } MP_L < AP_L & \Rightarrow AP_L \downarrow \text{ as Labor } \uparrow \\
\text{If } MP_L = AP_L & \Rightarrow AP_L \to \text{ as Labor } \uparrow
\end{align*}
\]

We can see this relationship between average and marginal product of labor in Table 2. Graphically, just as is shown in Table 2, average and marginal product are initially the same but marginal product rises faster and is greater than average product. Hence, average product rises but eventually falls. The relationship between average product and marginal product is shown in Graph 3. Notice that average product is at a maximum when it equals marginal product. Prior to that point, average product is rising because marginal product is higher. After that point, just the opposite occurs, marginal product is below average product so average product is falling as the firm hires more labor.

Recall that one of the purposes of investigating marginal product of labor and diminishing returns was to illustrate the relationship between the amount of labor a firm hires and its total output. That is, we want to know what the production function – the total product of labor – looks like graphically. But how does diminishing returns help us understand the total product of labor curve?

Mathematically, the marginal product of labor is given by equation 4 above; the change in total output is divided by the change in the amount of labor hired by the firm. However, as Graph 1 illustrates, the production function is simply the relationship between labor and total output. Thus, the marginal product of labor is the slope of the production function.

Thus, as Graph 3 illustrates, when the marginal product is positive this means that total output rises as the firm hires more labor. However, when the marginal product is negative total output falls as the firm hires more labor. Graph 4 shows the relationship between the total product, the marginal product and the average product curves. Notice, that just as shown in Graph 3, total output rises and falls when marginal product curve is positive and negative.
Also notice that when returns to specialization occur—prior to the first dotted line—the production function is getting steeper. That simply makes sense because, after all, the marginal product is getting larger and it is the slope of the production function. When the slope of a curve gets larger this simply means that the curve is getting steeper.

Likewise when diminishing returns sets in, marginal product is falling, meaning that the slope of the production function should be getting smaller. That is, the production function is getting flatter. Sure enough, we observe exactly this phenomenon in graph 4 in the area between the two dotted lines.

Notice also that when the marginal product is equal to zero that the production function is at its maximum. Thereafter, since the marginal product is negative, the production function decreases with increases in the amount of labor hired by the firm.

Graph 4 is an important graph to remember and understand. We will use it many times through the course of the semester. Students should carefully study both Graph 4 and Table 2. Both show the same relationships between the three production curves—total, marginal, and average product of labor.

II. Production Costs in the Short-Run

Given that firms are profit maximizers, our interest in the costs of production is easy to understand. Below we will carefully define what we mean by the various types of costs and develop the graphical measures of costs that we will later use to help us understand the behavior of profit maximizing firms.

A. What are production costs?

Recall from the previous topic that the total costs of production are simply what the firm must pay for the resources that it uses in the production of its goods. These costs may be explicit, paid for directly, or implicit, not paid for directly but estimated via opportunity costs.

Even in the short-run, total costs remain the sum of explicit and implicit costs. However, in the short-run it is actually more useful to divide total costs into two other types of costs because there are two different types of inputs in the short run: variable inputs that the firm can change and fixed inputs that the firm cannot change.

- Total Variable Costs (TVC) = the cost of the variable inputs that the firm uses in production. These costs may be explicit or implicit.
- Total Fixed Costs (TFC) = the cost of the fixed inputs that the firm uses in production. Again, these costs may be explicit or implicit.

One of the important equations that we will use throughout this discussion relates these
three different types of costs, total costs (TC), total variable costs and total fixed costs.

\[ TC = TVC + TFC \]  

Equation 6 uses an identity sign, the three horizontal lines in the equation, to relate total costs to total variable and total fixed costs. An identity sign has the same properties as an equal sign – to find total cost one adds up total variable costs and total fixed costs. However, the identity also means that this relationship between the different types of total costs is **always** true. We will use the identity in equation 6 below to understand costs in more detail.

B. What do the cost curves look like graphically?

Now, we want to graph the impact of extra production, Q, on the three types of total costs, TFC, TVC, and TC. The easiest total cost to graph is total fixed cost. Recall that fixed costs are paying for the fixed inputs – fixed means exactly that, the fixed inputs do not change in the short-run. Hence, the payment for those fixed inputs does not change either, regardless of how much output is being produced. If TFC equals 100 dollars when quantity equals ten it also equals 100 dollars when quantity equals 100, 1000, or even zero. TFC, then, is simply a horizontal line at the level of fixed costs as shown in Graph 5.

Total Cost is also relatively easy to calculate; it’s just the sum of TFC and TVC. Hence, once one has calculated both of these curves simply adding them up for each level of production yields total costs. But what does total variable cost look like graphically?

Graph 5 illustrates the three types of costs based upon the assumption that TVC simply rises by the same amount whenever output increases by one unit. Is this a reasonable assumption about variable costs? The answer as it turns out is – no, it is not. To see why TVC and, hence, TC, does **not** look as shown in Graph 5 we must return to both our definition of total variable costs and our short-run production function.

The simple short-run production function, where the firm uses only two inputs with labor being the variable input and capital the fixed input at five units, is shown in Graph 6. Graph 6 shows what happens to output as the firm hires more or less of the variable input, labor. What we want to show in a graph similar to Graph 5, is what happens to the cost of that labor, known as total variable cost, as the firm produces more or less output.

How do we get from what we already known in Graph 6 to what we want to measure in
Graph 5? In other words, we need to know how to translate the amount of labor the firm hires into the amount of money the firm pays for the labor they hire. Of course, that is quite simple to do; simply multiply the amount of labor we hire by the price paid to each unit of labor, the wage rate.

\[ \text{Total Variable Cost} = w \times L \]  
(7)

Thus, if the wage rate equals 50 dollars, then the cost of producing 20 units of output as shown in graph 6 is 500 dollars, 50 times the 10 workers hired to produce the output.

The easiest way to see the relationship between total variable costs and output is to simply assume that the wage rate is equal to one dollar. In that case, the amount of labor hired to produce a given output is equal to the total variable costs. In this case, the relationship between output and total variable costs is the same as the relationship between labor and output, the production function.

Suppose that the wage rate doubles to $2 per unit of labor. In this case, as shown by graph 7, the cost of producing any given level output doubles but the shape of the total variable cost curve remains the same. Graph 7 and our discussion clearly demonstrate that the total variable cost and total cost curves originally shown in Graph 5 are not correct.

Given the above discussion, what do all the cost curves look like? There are two steps we must take to show what the cost curves look like. First, we must switch the axes. That is in Graph 7, where we are showing the total variable cost curve, the curve has costs on the horizontal axis and output on the vertical axis. However, what we want to show is just the reverse – costs on the vertical axis and output on the horizontal axis as in Graph 5. Second, we must also include both total costs and total fixed costs with total variable costs.

Switching the two axes results in a total variable cost curve as shown in Graph 8. Graph 8 shows exactly the same relationship between output and total variable costs as is shown in Graph 7, except that the two axes have been switched. Generally speaking, we will ignore the top part of total variable costs, where TVC rise and quantity falls because firms will never operate in this area of the cost curves. Recall that this portion of the TVC curve occurs because output falls as firms hire more labor.

We now know what the TFC and TVC curves look like. Given that TFC and TVC must sum up to total costs, we can also calculate the total cost curve as well. These three curves are
shown in Graph 9. The TC curve has the same shape as the TVC curve just moved up by the amount of Total Fixed Costs.

C. Average and Marginal Costs

We now define four more costs, three types of average costs associated with the three different types of total costs, and marginal costs.

The average costs are fairly straightforward. We just want to know how much each of the different types of total costs – total costs, total variable costs, and total fixed costs – cost the firm on average per unit of output. That is, if a firm is currently producing 100 units of output and its total costs equal $2,000, how much is the firm spending on average for each unit of output? Clearly, the answer equals $20 and is found by dividing total costs of $2000 by output of 100 units. In fact, each average cost is found similarly by dividing the associated total cost by the output the firm is currently producing. Thus:

\[
\text{Average Total Cost (ATC)} = \frac{\text{TC}}{Q} \quad (8)
\]

\[
\text{Average Variable Cost (AVC)} = \frac{\text{TVC}}{Q} \quad (9)
\]

\[
\text{Average Fixed Cost (AFC)} = \frac{\text{TFC}}{Q} \quad (10)
\]

Marginal cost (MC) is similar to the other marginal concepts that have previously been introduced and equals the extra cost the firm incurs by increasing their output by one additional unit. Mathematically, marginal cost is given by:

\[
\text{Marginal Cost (MC)} = \frac{\Delta \text{TC}}{\Delta Q} = \frac{\Delta \text{TVC}}{\Delta Q} \quad (11)
\]

Why is marginal cost equal to both a change in total cost divided by a change in quantity and a change in total variable cost divided by a change in quantity? Actually, the reason is quite straightforward. We know from our discussion above, see equation 6 above that total costs consist of the sum of total variable costs and total fixed costs. Thus, total costs can only change if either of these two, TVC or TFC, changes. But we also know from our discussion above, see Graph 5, that TFC never changes. Thus, the only possible source for a change in total costs is a change in total variable costs. Table 3 presents all of the production and cost numbers for the example originally presented in Table 2.
Table 3
The relationship between Productivity and Costs - An Example

<table>
<thead>
<tr>
<th>L</th>
<th>Q</th>
<th>AP_L</th>
<th>MP_L</th>
<th>TVC</th>
<th>TFC</th>
<th>TC</th>
<th>AVC</th>
<th>AFC</th>
<th>ATC</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(TP_L)</td>
<td>(Q/L)</td>
<td>(∆Q/∆L)</td>
<td>(w*L)</td>
<td>(TVC+TFC)</td>
<td>(TVC/Q)</td>
<td>(TFC/Q)</td>
<td>(TC/Q)</td>
<td>(∆TC/∆Q)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>10.00</td>
<td>10</td>
<td>50</td>
<td>100</td>
<td>150</td>
<td>5.00</td>
<td>10.00</td>
<td>15.00</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>11.00</td>
<td>12</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td>4.55</td>
<td>4.55</td>
<td>9.09</td>
<td>4.17</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>12.00</td>
<td>14</td>
<td>150</td>
<td>100</td>
<td>250</td>
<td>4.17</td>
<td>2.78</td>
<td>6.94</td>
<td>3.57</td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>13.00</td>
<td>16</td>
<td>200</td>
<td>100</td>
<td>300</td>
<td>3.85</td>
<td>1.92</td>
<td>5.77</td>
<td>3.13</td>
</tr>
<tr>
<td>5</td>
<td>70</td>
<td>14.00</td>
<td>18</td>
<td>250</td>
<td>100</td>
<td>350</td>
<td>3.57</td>
<td>1.43</td>
<td>5.00</td>
<td>2.78</td>
</tr>
<tr>
<td>6</td>
<td>86</td>
<td>14.33</td>
<td>16</td>
<td>300</td>
<td>100</td>
<td>400</td>
<td>3.49</td>
<td>1.16</td>
<td>4.65</td>
<td>3.13</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>14.29</td>
<td>14</td>
<td>350</td>
<td>100</td>
<td>450</td>
<td>3.50</td>
<td>1.00</td>
<td>4.50</td>
<td>3.57</td>
</tr>
<tr>
<td>8</td>
<td>112</td>
<td>14.00</td>
<td>12</td>
<td>400</td>
<td>100</td>
<td>500</td>
<td>3.57</td>
<td>0.89</td>
<td>4.46</td>
<td>4.17</td>
</tr>
<tr>
<td>9</td>
<td>122</td>
<td>13.56</td>
<td>10</td>
<td>450</td>
<td>100</td>
<td>550</td>
<td>3.69</td>
<td>0.82</td>
<td>4.51</td>
<td>5.00</td>
</tr>
<tr>
<td>10</td>
<td>130</td>
<td>13.00</td>
<td>8</td>
<td>500</td>
<td>100</td>
<td>600</td>
<td>3.85</td>
<td>0.77</td>
<td>4.62</td>
<td>6.25</td>
</tr>
<tr>
<td>11</td>
<td>136</td>
<td>12.36</td>
<td>6</td>
<td>550</td>
<td>100</td>
<td>650</td>
<td>4.04</td>
<td>0.74</td>
<td>4.78</td>
<td>8.33</td>
</tr>
<tr>
<td>12</td>
<td>140</td>
<td>11.67</td>
<td>4</td>
<td>600</td>
<td>100</td>
<td>700</td>
<td>4.29</td>
<td>0.71</td>
<td>5.00</td>
<td>12.50</td>
</tr>
<tr>
<td>13</td>
<td>142</td>
<td>10.92</td>
<td>2</td>
<td>650</td>
<td>100</td>
<td>750</td>
<td>4.58</td>
<td>0.70</td>
<td>5.28</td>
<td>25.00</td>
</tr>
<tr>
<td>14</td>
<td>142</td>
<td>10.14</td>
<td>0</td>
<td>700</td>
<td>100</td>
<td>800</td>
<td>4.93</td>
<td>0.70</td>
<td>5.63</td>
<td>-</td>
</tr>
</tbody>
</table>

D. The Average and Marginal Cost Curves Graphically

Graph 10 presents the average and marginal cost curves that are first presented in Table 3. An examination of Table 3 and Graph 10 illustrates that marginal cost and two of the three average cost curves, AVC and ATC, are u-shaped. That is, initially as output increases, these costs fall but eventually they begin to rise. Average fixed costs, on the other hand, always falls as output rises. Students should carefully study Graph 10; it will be used repeatedly in the following chapters discussing the behavior of profit maximizing firms.
The main purpose of this section is to develop an understanding of the average and marginal cost curves depicted in Graph 10. Why do they look as they do; what are their interrelationships? In order to understand these issues, the following discussion will focus on answering four important questions about Graph 10.

- Why is average fixed cost always declining?
  To understand why average fixed costs always decline one must simply recall two important facts. First, as illustrated in Table 3 and Graph 9 total fixed costs are always constant regardless of the level of output. Second, equation 10 illustrates that to calculate average fixed costs we simply divide total fixed costs by output. Hence, as the firm’s output rises we are dividing the same number – total fixed costs – by an ever-larger output. The result, average fixed cost, must fall as output rises.

- Why are the remaining cost curves u-shaped?
  While it is relatively easy to understand why average fixed costs fall as output rises, why are the remaining cost curves u-shaped? To understand why these cost curves are u-shaped, we must refer back to productivity. The reason why we discussed productivity previously was to be able to discuss the firm’s costs. Recall that costs are important because, eventually, our purpose will be to discuss the firm’s profits, which equals the difference between the firm’s total revenue and total costs.

  What is the relationship between productivity and costs? Essentially, as productivity rises this enables the firm to produce more with the same amount of resources or, to produce the same level of output; the firm will need fewer resources. Given that using fewer resources translates into lower costs, we can see that the relationship between productivity and costs must be an inverse relationship. As productivity rises, costs will fall and the reverse.

  Thus, the cost curves are u-shaped because the productivity curves, AP_L and MP_L, have an inverted u-shape as illustrated in Graphs 3 and 4. Recall why the productivity curves have this inverted u-shape – the law of diminishing returns. As with productivity the shape of the cost curves in Graph 10 is determined by the law of diminishing returns. Initially, productivity rises, which means that initially the cost curves fall. However, eventually diminishing returns begin and productivity declines, meaning that the cost curves begin to increase.

- How are the average cost curves related?
  Recall from equation 6 that total cost equals the sum of total variable cost and total fixed cost. Dividing both sides of equation 6 by quantity yields the following relationship:

  \[
  \frac{TC}{Q} = \frac{TVC + TFC}{Q} = \frac{TVC}{Q} + \frac{TFC}{Q}
  \]

  Notice that according to our original definitions of the three types of average costs (see equations 8 through 10) that equation 12 actually becomes:

  \[
  ATC = AVC + AFC
  \]

  That is, the three average cost curves are related in a similar manner to the three total cost curves; variable and fixed costs sum up to total costs. Because this was true for the total costs it is also true for the average costs. An examination of both Table 3 and Graph 10 clearly demonstrates that this relationship is correct. Thus, if we know any
two of the average costs, then we can also calculate the third average cost. Moreover, notice that ATC and AVC get closer and closer together as output rises. This is because the difference between these two curves equals AFC and as output rises AFC is getting closer and closer to zero.

Given equation 13 it is easy to calculate any of the average cost curves if one knows the other two average cost curves. As a result, in our discussion of profit maximizing firms in the following chapters below we follow the common convention of only showing two of the three average cost curves, ATC and AVC.

- How are the average cost curves related to marginal cost?

Both Graph 10 and Table 3 illustrate, as discussed above, that AFC always declines as output rises. It doesn’t matter whether MC is rising or falling, AFC is always falling. Thus, MC and AFC are simply not related. The one has no impact on the other.

However, MC is related to ATC and AVC. These relationships exist because MC is calculated from either TC or TVC as discussed above in equation 11. We calculate MC by taking the change in either TC or TVC and dividing that change by the change in output.

Recall that above, we discussed the relationship between MPL and APL and pointed out that it follows the same relationship as is true between any marginal variable and its related average. If the marginal is above the average, the average must rise. However, if the marginal is below the average, the average must fall. This same relationship holds between MC and its related averages, ATC and AVC. Thus:

If $MC > ATC (AVC)$ \Rightarrow $ATC (AVC)$ ↑ as $Q$ ↑

If $MC < ATC (AVC)$ \Rightarrow $ATC (AVC)$ ↓ as $Q$ ↑

If $MC = ATC (AVC)$ \Rightarrow $ATC (AVC)$ → as $Q$ ↑

We see these relationships in both Table 3 and Graph 10. When the MC is below either ATC or AVC, then that curve is falling. The MC curve crosses both ATC and AVC at their minimum points. Finally, when the MC is above either ATC or AVC, then that curve is rising.

E. Relating Costs to Production Formally

Recall our discussion above of why the cost curves in Graph 10 were u-shaped. We noted that they were u-shaped because the cost curves were related to productivity; when productivity rises initially due to the returns to specialization, costs fall and when productivity eventually falls due to diminishing returns, costs must rise. In our discussion we relied upon an intuitive understanding of the relationship between productivity and costs. Now, however, a more formal approach will be used to expand our understanding of this relationship.

Recall that in our simple model we are using only labor and capital to produce the output with labor the variable input and capital the fixed input. Thus, TVC are simply the firm’s labor costs while TFC equal the firm’s capital costs. As shown by equation 7, labor costs simply equal the wage rate (the price of labor) times the number of units of labor hired. We also know that AVC, per equation 9, equals TVC divided by output. Thus we know that:

$$AVC = \frac{TVC}{Q} = \frac{wL}{Q} \text{ or } = w\frac{L}{Q}$$
However, what does $L/Q$ equal? From equation 5, we know that the average product of labor equals its inverse, $Q/L$. Thus, $L/Q$ equals the inverse of the $A_P_L$. As a result, we can conclude that:

$$AVC = \frac{TVC}{Q} = \frac{wL}{Q} \text{ or } \frac{L}{Q} = \frac{w}{A_P_L} \quad (14)$$

The algebraic result developed in equation 14 clearly illustrates the relationship between $AVC$ and $A_P_L$, they are inversely related. The inverse relationship between $AVC$ and $A_P_L$ is also illustrated in Table 3. As $A_P_L$ rises initially, $AVC$ is falling. When $A_P_L$ reaches its maximum, then $AVC$ reaches its minimum. Finally, as $A_P_L$ eventually falls, then $AVC$ is rising.

A similar algebraic manipulation illustrates the relationship between $MC$ and $M_P_L$. Recall that $TVC$ equals the wage times the amount of labor the firm hires and that $MC$ equals the change in $TVC$ divided by the change in output. As a result:

$$MC = \frac{\Delta TVC}{\Delta Q} = \frac{\Delta(wL)}{\Delta Q} \text{ or } w \frac{\Delta L}{\Delta Q}$$

However, what does $\frac{\DeltaL}{\DeltaQ}$ equal? From equation 4, we know that the marginal product of labor equals its inverse, $\frac{\DeltaQ}{\DeltaL}$. Thus, $\frac{\DeltaL}{\DeltaQ}$ equals the inverse of the $M_P_L$. As a result, we can conclude that:

$$MC = \frac{\Delta TVC}{\Delta Q} = \frac{\Delta(wL)}{\Delta Q} \text{ or } w \frac{\Delta L}{\Delta Q} = \frac{w}{M_P_L} \quad (15)$$

Again, equation 15 clearly illustrates the inverse relationship between productivity and costs. As the $M_P_L$ initially rises, as it does in Table 3, then $MC$ must fall. Likewise, when $M_P_L$ reaches its maximum then $MC$ must reach its minimum. Finally, when $M_P_L$ eventually falls due to diminishing returns, then $MC$ must be increasing.

Both equations 14 and 15 clearly illustrate, as does Table 3, that our intuition was correct. The relationship between costs and productivity is an inverse one; increases in productivity cause costs to fall while decreases in productivity cause costs to rise.