CHAPTER

12

Cost Curves of the Firm

12.1 Production and Input Supply: Why Marginal Cost Is as It Is

What to Read For
What are the two fundamental messages of the idea of the production function? What are concavity and diminishing marginal returns to scale? What is the factual reason that economists assume diminishing marginal returns to scale? How do you derive a total cost curve from a production function and a price of the input? How do you derive marginal and average cost from the total cost? What is the necessary relation between marginal and average cost curves? How does a firm behave if its marginal cost curve slopes down? How can delivery costs and rising costs of inputs offset a perversely shaped marginal cost curve? What is duality? Why is it a consequence of the equality of cost and revenue?

Output Depends on Input
For some questions it is enough to know about the firm’s marginal cost curve. But for others one wants to look behind the curve to its causes. Its causes are, put briefly, the constraints imposed on the firm by the state of markets and of knowledge. The firm combines costly ingredients according to the best recipes it knows to produce a given output at minimum cost. A steel company produces outputs of bars, angles, sheets, rails, structural shapes, and so forth with inputs of coal, iron ore, marketing managers, insurance, limestone, blast furnaces, soaking pit crane operators, computers, file clerks, rolling mills, and thousands of other distinct entities, whether purchased from other people or owned by the steel firm itself. The set of recipes for combining the inputs is called the production function—an idea that was used earlier and will be used later still again. A firm might have dozens of outputs and inputs. For present purposes, however, one output and one input will do. This means adding up steel sheets and cold rolled bars into one output, in amount Q, and adding up soaking pit crane
operators and rolling mills into one input, in amount $I$. That is, some function $F$ connects output to input; thus, $Q = F(I)$. A function, for example, might be a constant elasticity one, such as $Q = 5^{0.8}$. The simple one-input production function conveys the idea that output depends on input. (A many-input production function conveys the additional message that there is more than one way to skin a cat; that is, there are different combinations of inputs that can be used to produce the same level of output.)

A feature this particular function shares with other functions that might describe a firm’s recipes is concavity. It is the shape marked Acceptable in Figure 12.1. What is acceptable about it is that it exhibits diminishing marginal returns to scale, which is to say that, as the firm attempts to produce more output by pushing in more input, each additional dose of input produces less additional output. The slope diminishes.

Unlike the man in the street, economists do not believe that bigger is always better. Early in the application of fertilizer, labor, machinery, and so forth to a given plot of land (or all these and land to a given farmer), there may well be "economies of scale" in producing corn. But eventually the firm is overwhelmed by the additional inputs. The 50 tractors crowded onto the 2 acres of Mr. Craft’s farm smash into each other and explode; the fertilizer left in tons on each square yard buries the young corn plants to a depth of 6 feet; the thousands of laborers become a mob and take to sleeping, fighting, and card playing on company time, then turn to Craft’s house and burn it down for sport. The point of Catastrophic Overcrowding in Figure 12.1, needless to say, will never be reached by a rational firm. In fact, it will never reach beyond

An empirical law of production is that at high levels of input the marginal product of input will be falling. The increase in output per unit of input is lower the greater the input. Marginal product cannot be everywhere increasing.
Maximum Output. Less colorful stories apply to the earlier parts of the curve, but the moral is the same. Along Acceptable, marginal returns diminish continually; along Also Acceptable, they at first increase, then diminish. The critical point is that the functions do not end like Unacceptable, the upward-curving dashed line. The reasons are two, the first being fact: When they can be measured well—as in the case of farming—production functions of firms do not exhibit continuously increasing returns to scale.

How to Reason from the Shape of the Production Function to the Shape of the Cost Functions

The second reason the slope declines will become clear as the argument goes forward. Suppose that the all-purpose input has a fixed per-unit price of \( w \) (for wage, the price of the input that is most often the subject of thought). The price is one of the constraints facing the firm. Given \( w \), it is easy to get from the production function to the cost curve. Look at Figure 12.2. The left panel has the production function laid on its side, with a shape like Also Acceptable in Figure 12.1. Look at it with the arrow pointing Up. You will see that it merely reverses the direction in which input is measured. If \( w \) is fixed, the total cost is of course simply \( wI \) (that is, the input price times the input amount). The vertical axis can be stretched by the factor \( w \). The stretching clearly will leave the general shape of the curve unaltered. The result is the Total Cost curve in the second panel of the figure. The third (right) and final panel simply plots the slope of the Total Cost curve against quantity. The slope is the rise in total cost per unit of a rise in quantity; that is, the slope is Marginal Cost.

The diagram contains or implies everything you will ever need to know about cost curves. So read the previous paragraph again, slowly: It’s worth the investment. Notice the way in which the little tangencies along the Total Cost curve follow as they should the falling and rising shape of the Marginal Cost curve. Notice, too, that the Average Cost (the dashed curve) is the slope of a ray.

Figure 12.2
The Shape of the Production Function (a) Determines the Shape of Total (b) and Marginal (Average) (c) Cost Curves

If inputs can be bought at a constant price, the total cost curve and the production function are identical except for scale. The slope of the total cost curve at a given quantity is of course marginal cost at that quantity. The slope of a line through the origin intersecting the Total Cost curve is average cost at that quantity.
from the origin to the total cost curve. Such a slope measures total cost divided by output, and therefore when the ray reaches a minimum in slope at **Lowest Average Cost** on the **Total** Cost curve (middle panel) it is also identical to the slope of the curve. That is, marginal cost is equal to average cost once, slicing up through the average cost at its minimum point. This amusing and surprising fact is useful for drawing self-consistent diagrams of average and marginal cost.

### A Numerical Example

It will help in driving home the diagram to work through a numerical example. Try to do it without peeking at the answer.

**Q:** Marcello de Cecco hires labor to harvest his olive trees and press out the oil at a fixed wage of $4 per hour. His production of olive oil as a function of hours of labor is:

<table>
<thead>
<tr>
<th>labor Input (hours)</th>
<th>Olive Oil Output (gallons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>30</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>300</td>
<td>200</td>
</tr>
<tr>
<td>400</td>
<td>350</td>
</tr>
<tr>
<td>500</td>
<td>800</td>
</tr>
<tr>
<td>600</td>
<td>1200</td>
</tr>
<tr>
<td>700</td>
<td>1400</td>
</tr>
<tr>
<td>800</td>
<td>1575</td>
</tr>
<tr>
<td>900</td>
<td>1700</td>
</tr>
<tr>
<td>1000</td>
<td>1800</td>
</tr>
<tr>
<td>1100</td>
<td>1850</td>
</tr>
</tbody>
</table>

1. List the total costs as a function of output from 30 to 1900 gallons. Sketch its shape.
2. Calculate the marginal cost at each output. (To be definite, use, so to speak, a forward-looking definition of marginal cost: At 1200 gallons ask what additional costs are incurred for each gallon on average from 1200 to 1400.) Sketch it.
3. Calculate the average cost at each output. What, approximately, is the output at which average cost is at a minimum? Is marginal cost approximately equal to average cost there?

**A:**

1. Simply multiply the 'hours of labor input by the $4 cost per hour and list the result next to the corresponding output level (ignore the last two columns):

<table>
<thead>
<tr>
<th>output (gallons)</th>
<th>Cost</th>
<th>Marginal Cost per Gallon</th>
<th>Average Cost per Gallon</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>$400</td>
<td>$5.71</td>
<td>$13.33</td>
</tr>
<tr>
<td>100</td>
<td>800</td>
<td>4.00</td>
<td>8.00</td>
</tr>
<tr>
<td>200</td>
<td>1200</td>
<td>2.67</td>
<td>6.00</td>
</tr>
<tr>
<td>350</td>
<td>1600</td>
<td>.89</td>
<td>4.57</td>
</tr>
<tr>
<td>800</td>
<td>2000</td>
<td>1.00</td>
<td>2.50</td>
</tr>
<tr>
<td>1200</td>
<td>2400</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>1400</td>
<td>2800</td>
<td>2.28</td>
<td>2.00</td>
</tr>
<tr>
<td>1575</td>
<td>3200</td>
<td>3.20</td>
<td>2.03</td>
</tr>
<tr>
<td>1700</td>
<td>3600</td>
<td>4.00</td>
<td>2.12</td>
</tr>
<tr>
<td>1800</td>
<td>4000</td>
<td>8.00</td>
<td>2.22</td>
</tr>
<tr>
<td>1850</td>
<td>4400</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sketch is just like Figure 12.2, middle panel, appropriately labeled.

2. At 30 gallons an additional 70 gallons costs an additional $400. So marginal cost is on average $400/70 = $5.71 per gallon. At 100 gallons, an additional 100 gallons costs an additional $400. So marginal cost is on average $400/100 = $4.00. Going through the table in like fashion gives the third column above, the marginal cost. The sketch is the right panel of Figure 12.2.

3. Average cost is just total cost divided by output. For instance, at 1400 gallons of oil it is $2800/1400 = $2 per gallon. The last column fills in the rest. The minimum average cost is somewhere around 1200 or 1400 gallons. Marginal cost is approximately equal to average cost there. In other words, just as Figure 12.2 claims, the marginal cost curve, starting below the average cost ($5.71 is less than $13.33, $4.00 is less than $8.00), is equal to the average cost when the average cost hits bottom. Thereafter, the marginal cost is higher ($3.20 is higher than $2.03, and so forth).
Other Things Equal, Diminishing Returns Imply Rising Marginal Cost

The crucial result is the mere upward slope of the marginal cost curve. The upward slope comes directly from the diminishing returns to scale. Put verbally, the lower increments to output forthcoming from additional units of input mean that given increments to output will require larger and larger doses of inputs. That is, they will require higher and higher costs. With marginal cost rising, the firm will arrive at a particular finite output if it applies the rule (of rational life) “to maximize profit, set output such that marginal cost equals marginal revenue.’’

T or F: If on the contrary the production function implied higher increments to output from additional units of input, the firm will expand without limit.

A: With increasing returns to scale (and a fixed factor price, w) the marginal cost curve slopes downward. A competitive firm would be minimizing, not maximizing, profit if it stopped at the output that equalizes marginal revenue and marginal cost (see Figure 12.3). The firm makes a positive marginal profit on each quantity sold beyond the point $MC = MR$ and will therefore continue moving to the right indefinitely. Therefore, it would expand indefinitely. Therefore, true.

Again, consider the following.

T or F: Although increasing returns to the scale of a perfectly competitive firm is not consistent with equilibrium, constant returns is.

A: By the reasoning just given, constant returns implies a flat marginal cost curve—by contrast with the rising one of decreasing returns and the falling one of increasing returns. If a flat demand curve—a price given to a price-taking firm—is the marginal benefit, then there are three possibilities, none of which is attractive (see Figure 12.4). Therefore, false.

The usual case is that marginal cost eventually rises:

Q: If de Cecco in the earlier problem can sell his olive oil at $2.30 a gallon, where does he produce? How much profit does he make there?

A: He produces where marginal benefit ($2.30 a gallon) equals marginal cost: that is, at an output of 1400 gallons, approximately, where marginal cost is $2.28. His

Figure 12.3 Why Marginal Cost Must Slope Upward

A firm facing a horizontal demand curve cannot have a marginal cost curve that is falling and that lies below the demand curve. The optimal size of the firm would be infinitely large.
Figure 12.4  
**Constant Returns Is Inconsistent with an Equilibrium Scale for the Firm**

A firm facing a horizontal demand curve and having a horizontal marginal cost curve would produce either nothing or an infinite quantity or would be indifferent to the amount it produced.

### Possibility

#### (a)
Marginal cost more than market price

- **Diagram**
  - Marginal Cost ($MC$) is above the demand curve ($Price$).
  - **Consequence**: No output makes money; firm closes down.

#### (b)
Marginal cost exactly equal to market price

- **Diagram**
  - Marginal Cost ($MC$) is exactly equal to market price ($Price$).
  - **Consequence**: Any output equally profitable (namely, zero) to any other: no theory of firm's output.

#### (c)
Marginal cost less than market price

- **Diagram**
  - Marginal Cost ($MC$) is below the demand curve ($Price$).
  - **Consequence**: Infinite output maximizes profit: firm expands without limit.

Profit is his revenue on 1400 gallons—namely, $(1400 \times \$2.30) = \$3220$—minus his total cost, which at 1400 gallons is $\$2800$. So he makes $\$3220 - \$2800 = \$420$.

### Transport Costs Might Not Be Equal

Such considerations demonstrate that a flat marginal cost is inconsistent with a price-taking firm—a firm that is so small in its market that it faces a flat demand curve. But a firm that is big in its market, and that faces therefore a downward-sloping demand curve (because it faces such a big part of the entire demand curve), could rationally stop at a finite size though its marginal cost of production were flat. The reason, to be elaborated in Chapter 17 on monopoly, is that selling too much could carry the firm so far down the demand curve that the fall in price would overbalance the rise in quantity. A monopolist, in
other words, faces diminishing returns \textit{in selling} that can play the same role of limiting the scale of the firm as do diminishing returns \textit{in production} or (as we shall see in a moment) in buying inputs.

Another case of a single firm facing a downward-sloping demand curve is that of a firm that can get more customers only from farther afield, with higher costs therefore of delivering the product. Costs may be constant or even declining as output rises \textit{at the plant}, but the rising cost of \textit{delivery} from the plant leads to a finite equilibrium. Big hardware stores, for example, are better, since they can stock a wider range of hardware and assure that each visit by a customer is more certain to yield the kind of paint, nail, or tool the customer wants. Should Chicago therefore have one big store? No. It has in fact over 400, because a local store is convenient; the trip to the One Big Store would be expensive.

\textbf{Q:} A careful study of economies of scale in the generation of electricity by privately owned American utilities in 1955 concluded that they were very strong: In the constant-elasticity form $Q = a^p$, the $a$ was 2 or higher for the firm (any $a$ greater than 1, recall, is an unacceptable elasticity in Figure 12.1).\footnote{The facts, the question, and the answer are taken from Marc Nerlove, “Returns to Scale in Electricity Supply,” in C. F. Christ et al., eds., \textit{Measurement in Economics} (Stanford, Calif.: Stanford University Press, 1963), reprinted in Arnold Zellner, ed., \textit{Readings in Economic Statistics and Econometrics} (Boston: Little, Brown, 1968).} Yet there were over 400 privately owned electrical utilities in the United States in 1955. Why!

\textbf{A:} Transporting electricity is expensive, and the farther it is transported the more expensive it is. Therefore a flat or downward-sloping marginal cost at the plant is consistent with a rising marginal cost at the place of use. Other \textit{natural monopolies} limited by transport costs are said to include grocery stores, schools, drugstores, banks, and the like.

\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Diminishing Returns Can Limit the Size of the Firm} & \\
\hline
\end{tabular}
\end{table}

Unless it is offset by transport costs, then, a flat or downward-sloping marginal cost curve for a competitive firm is a great theoretical nuisance. For many firms the costs of transport are insignificant. Clothing manufacturers on Seventh Avenue or automobile dealers in Los Angeles or wheat farmers in Kansas are not isolated from each other’s competition by economically significant distances, yet they are finite in size. To have a theory of such finite firms actually observed in the world, one is forced to abandon the common notion that firms always reap endless economies of scale or even the less common notion that costs are constant.\footnote{The various attempts to actually observe the cost curves of firms, however, show with embarrassing frequency that cost curves do appear to be flat. The observations can often be rationalized as mismeasurements; compare Milton Friedman, in National Bureau of Economic Research, \textit{Business Concentration and Price Policy} (Princeton, N.J.: Princeton University Press, 1955), pp. 230–238. The best case for constancy of marginal cost and against the conventional theory of the firm as presented here and in other texts is J. Johnston, \textit{Statistical Cost Analysis} (New York: McGraw-Hill, 1960), especially Chapters 5 and 6. His view is that many firms are monopolists.} The argument is one of survival, in the style of biology since Darwin. A firm that is “fit” in the size it has chosen will survive by being profitable. Unfit species will vanish. If sizes of firms in the retail women’s clothing trade cluster around one size, the economist seeks reasons for the apparent optimality of the size. She uses the results of selection to guide what would otherwise be an impossibly difficult inquiry into the cost curves of firms.\footnote{The idea was made explicit by George Stigler, “The Economics of Scale,” \textit{Journal of Law and Economics} 1 (October 1958): 54–71.} If there were no presumption that the sizes and other characteristics of firms that survive are in fact the least-cost characteristics, economic studies would be crippled. For
the same reason, so would ecological studies that could make no presumption that the cowardice of wolves in the hunt or the falling of leaves in autumn are in fact valuable for the survival of wolves or of broadleaf plants.  

Chapter 14 will exploit the argument from survival to its limit. For the present the argument serves merely to buttress the belief that marginal cost curves rise. It is the second reason for the upward curvature: Not only is the upward curvature in fact observed when observation is possible, but it would have to be observed in a rational and finite-sized firm.

**An Inelastic Supply of Inputs (Like a Downward-Sloping Demand) Can Also Limit the Size of the Firm**

> **T or F:** The production function of Bethlehem Steel might exhibit constant or increasing returns to scale, yet the resulting cost curve could have the normal shape if the supply curve of iron ore facing the firm were inelastic instead of perfectly elastic.

> **A:** The reasoning is simply that as Bethlehem produces more, it buys more iron ore, which raises the price of the ore (if, as assumed, the ore is supplied inelastically to the firm). The rise in the price of the inputs can offset the advantage gained from increasing returns in needing less and less inputs per unit of output. Therefore, true.

**The Dual and the Primal Problems**

In short, the production function and the cost function are linked. The link is called duality, a thought so powerful in economics that in the higher reaches of abstraction in the field it dominates much thinking. The word comes from the jargon of programming, that is, the branch of applied mathematics that deals with maximizing something under many constraints. In the present context, the thought is that the production function is the answer to one question—what is the most quantity, $Q$, the firm can attain for a given input, $I$? And the cost function is the answer to a related question—what is the least cost, $wI$, the firm can attain for a given $Q$? And that the two questions are two sides of one question—what is the best thing the firm can do? They are dual to each other; that is, each is one of a pair. The two solutions are solutions to the same problem, but one solution is to maximize and the other is to minimize. In doing well an executive for General Motors can either start with a work force, a plant, and some raw materials and produce as many cars as possible or, equivalently, start with a certain number of cars and produce them with the given work force, plant, and raw materials as cheaply as possible. The two are not really two solutions, but the same, single solution looked at two ways. In other words, maximizing output for a given input is the same as minimizing input (and therefore input costs) for a given output; that is, $Q/I$, or the physical productivity of the inputs, is maximized when $w(I/Q)$, or the input cost per unit of output, is minimized. Elementary though it is, it is easy to get confused.

---

about the matter, and confusion about it is a reliable indicator of economic ignorance.

T or F: In a nonexpansive market, managers will attempt to minimize costs rather than maximize output.

A: Taken literally, “minimizing costs” would occur at an output of zero, and “maximizing output” would occur at an output of infinity. Read more sympathetically, the assertion must refer to minimizing (input) costs for a given output and maximizing output for a given input. But these are merely alternative ways of expressing the same thing. Staying on the cost curve (minimizing costs) is identical to staying on the production function (maximizing output). Therefore, false or meaningless.

Similarly, consider the following.

T or F: The link between new farming methods and productivity is weak. The new methods may show up in higher output per unit of input (which is higher productivity) but alternatively may be used merely to pay higher wages, with no increase in output.

A: The second sentence is true, but it is false evidence for the first. The confusion is again between the amount produced and the efficiency with which it is produced. New, higher productivity methods will raise output per unit of input (which is what higher productivity means). They will therefore reduce input per unit of output. They will therefore permit higher pay for each unit of input, for less input is wanted per unit of sellable output. The argument illustrates the strength of the link between new methods and productivity, not its weakness.

The link is so very strong because it is so very obvious. A firm must earn as much as it pays. That is, the price of its product times the quantity produced must equal the price per unit of inputs times the quantities of them used. The assertion that \( PQ = wI \) implies, by the magic of elementary algebra, that \( Q/I = w/p \). The new equation says that output per physical unit of input equals the price of the input divided by (or “deflated by”) the price of the output. In other words, there are two ways of measuring productivity, one with quantities and another—the price “dual” of the quantity “primal” (primal means “the original problem”)—with prices. The two are equal. The productivity of inputs is equal to their real price (that is, their money price deflated by the price of output). The higher real wages, rents, profits, and so forth that Americans get compared with Ethiopians are not something apart from the higher physical productivity of the American economy but identical to it, because \( Q/I = w/p \). The failure of incomes (which are input prices) to “keep up” with inflation (which is the output price) is not a consequence of an economic footrace but, rather, the necessary consequence of falling productivity unrelated to inflation, because the rate of change of \( w/p \) must equal the rate of change of \( Q/I \). A similarity between a measure of productivity change based on quantities and one based on prices is not a happy accident but, rather, a necessary correspondence, because again \( Q/I = w/p \).

Summary

Behind the marginal cost curve is the production function. Either diminishing returns to scale in the function or rising delivery costs or sufficiently inelastic supply curves of inputs can cause the marginal cost curve to slope upward. If it slopes downward, the firm will expand without limit, which suggests that it must slope upward. Somewhere in the constraints facing the firm, then, more gets you less.
The relationship between costs and production is one of *duality*. As the production function is the most output for a given input, so the cost function (the “dual” of the production function) is the least cost for a given output. The link between the two is the budget equation, \( PQ = wI \), which is to say that revenues equal expenditures. The equation implies that \( Q/I = w/p \), which is to say that physical productivity equals the real reward of the input.

**EXERCISES FOR SECTION 12.1**

1. Which if any of the following production functions have acceptable shapes?

   a. When labor is: Quantity is:

<table>
<thead>
<tr>
<th>Labor (hours)</th>
<th>Quantity (gallons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>200</td>
</tr>
<tr>
<td>200</td>
<td>500</td>
</tr>
<tr>
<td>1,000</td>
<td>3,000</td>
</tr>
<tr>
<td>10,000</td>
<td>40,000</td>
</tr>
</tbody>
</table>

   b. \( Q = L^2 \).

   c. “As labor input rises by 1 percent, output always rises by more than 1 percent.”

2. Suppose that de Cecco’s olive grove produced output this way:

<table>
<thead>
<tr>
<th>Labor Input (hours)</th>
<th>Olive Oil Output (gallons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>30</td>
</tr>
<tr>
<td>200</td>
<td>120</td>
</tr>
<tr>
<td>300</td>
<td>250</td>
</tr>
<tr>
<td>400</td>
<td>370</td>
</tr>
<tr>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>600</td>
<td>600</td>
</tr>
<tr>
<td>700</td>
<td>675</td>
</tr>
<tr>
<td>800</td>
<td>700</td>
</tr>
<tr>
<td>900</td>
<td>710</td>
</tr>
<tr>
<td>1000</td>
<td>715</td>
</tr>
</tbody>
</table>

   He still pays $4 an hour for labor. Calculate and list against the various outputs from 30 to 715 gallons the total cost, marginal cost, and average cost.

3. In Exercise 2, what approximately is the output with minimum average cost? If de Cecco sells oil at $1.6 a gallon, where does he produce?

4. Contrast the cost curves of Exercise 2 with those in the text. Which represents a small-scale firm? Which has more sharply diminishing returns to scale?

5. What would de Cecco in the text produce if the price of olive oil were $1 a gallon? $3.20 a gallon? $8? What curve, then, is de Cecco’s supply curve, that is, the curve showing how much output he supplies at various different prices?
PROBLEMS FOR SECTION 12.1

1. The McGouldrick Textile Mill of Manchester, New Hampshire, is planned as a cube of height (and width and depth), $h$. Suppose that the construction and maintenance costs of the mill itself are proportional to the surface area of the mill, because the cost is determined by the extent of brick and glass ("Factory windows are always broken / Somebody’s always throwing bricks."). Suppose too that the output, $Q$, of the mill is proportional to its cubic volume, because the larger the volume the more spindles, looms, and other producers of output can be crammed in. **True or false**: Total construction and maintenance costs will vary as $cQ^{2/3}$, where $c$ is some constant; that is, marginal construction and maintenance costs will decline as planned output rises, two thirds being less than one. (**Hint:** Use the output equation to solve for $h$ as a function of $Q$, then substitute into the cost equation.)

2. A lone dealer in the London Times newspaper living at Isca at the end of the Fosse Way can get as many copies as he wants at 10 pence a copy (that is, his marginal cost at Isca is constant). All his customers will pay only 15 pence a copy. At each mile marker north of Isca (the Fosse Way runs north from Isca to Lindum, you see, and is the only road), there are 1000 potential buyers of the Times, beginning at mile 1. Each mile of transport cost a quarter of a penny per copy. How many copies will he sell in this case? Draw a diagram illustrating the general case, with distance in miles along the horizontal axis (distance being equivalent to numbers of copies) and cost and selling price along the vertical.

**True or False**

3. If the production function for the airframe industry is characterized by constant returns to scale, the supply curve of the industry will be infinitely elastic in the long run.

4. If total costs $= cQ^a$ and $a$ is less than 1.0, marginal cost is declining (the firm is experiencing increasing returns to scale).

12.2 Cost Curves in Use: The Long Run and the Short Run

*What to Read For*

What is the principle that adding constraints hurts? How does the short-run cost curve of a firm illustrate the principle? What are the two reasons a firm will tolerate being on a higher-cost, short-run curve? Would it tolerate it if there was a secondhand market in equipment? Why might a firm prefer an expensive but flexible plan to a cheap but inflexible one? What is the envelope of all short-run cost curves?

*The Production Function or Cost Curve Is the Best*

That the production and cost curves measure the most output and the least cost implies that the firm can do worse. Indeed it can. It can take up positions anywhere in the upper area of Figure 12.5, such as the point Error, using excess inputs in amount $AI$ to attain $Q_0$, which will therefore cost more than it should, by the amount $AI$ multiplied by the cost of each unit of input.

A firm having full knowledge of its costs and managed by a genius would never be in error. Knowledge and genius, however, are expensive, more expensive for some firms than for others, with the result that actual firms depart more
The production and cost functions show the least costly methods of production. If the profit-maximizing firm had full information, it would always choose points such as Efficient. In practice, ignorance, which is costly to remove, makes the firm choose points like Error.

Figure 12.5
The Firm Can Do No Better than the Production (a) and Cost Curves (b), but It Can Do Worse

The key point is that real firms spend much of their time above "the" cost curve. The point can be put in the form of another one of those stunningly obvious but frequently misunderstood principles of economics: namely, adding constraints hurts. The principle was first stated in the chapter on utility functions. A firm that is constrained to use dull managers because of the cost of finding better ones will have higher costs than one free to use the best. Likewise
a runner constrained to leap over hurdles while running 100 yards will have worse times than one free to run straight. The women’s 100-meter freestyle record in swimming is never slower than the 100-meter backstroke, since freestyle means that the swimmer could choose the backstroke if it were the fastest way of covering 100 meters.

**T or F:** The costs of operating the World Trade Center building in New York will be higher with a new federal rule that air conditioning must be no cooler than 78°F than without it.

**A:** If constrained to shift from 72°F to 78°F, the building might save on energy costs (as a matter of fact, because existing equipment is designed to operate best at the old, low temperature, it might not) but will lose on other costs, such as the efficiency of the office workers who are using the building or the life expectancy of stored paper. The conclusion would be inevitable if the building were being operated at minimum cost before the federal rule was imposed, because if the higher 78°F temperature were the optimum, it would have been chosen anyway, in the absence of the federal rule. Therefore, true.

---

**Constraints Put the Firm on the Short-Run Cost Curve**

The most important class of examples of how firms can do worse and how constraints make them do so is that short-run costs are always above long-run costs. By short run, the economist always means “constrained to use the equipment in place, that is, unwilling to adjust fully and immediately to a new scale of operation.” Since the short run is a situation with more constraints than is the long run, it is clearly more costly than the long run. Constraints always hurt.

There are two reasons why a firm might suffer the constraints of the short run. The first is that, if a change in output is known to be short lived, the firm will not want to buy a long-lived piece of equipment to service it. For example, suppose (as was the case) that the output of skateboards was very large during the time that children were just discovering them, but fell to the replacement output as soon as every house in the land had a couple of skateboards in it. An intelligent firm wishing to cash in on the temporary boom would not buy specialized factories, retail outlets, and so forth suited to a long persistence of the boom output, because it knows the boom will not last. The specialized factories will last for 20 years. If the boom lasts as long, the per-skateboard cost of the factories will be low. But if the boom lasts only two years, the cost per skateboard will be high. At such costs it may well be better to rent a factory than to buy, to get a factory suited to general use rather than one specialized in the production of skateboards, and to pay premiums to old workers to work overtime rather than to hire and train new ones. The first principle, then, is that one does not use a cannon to kill a fly.

The second reason why a firm might suffer the short run is that, even if the change in output is known to be permanent, it may be cheaper to adjust to it slowly. For example, if the college of the University of Chicago decides to expand from 2500 to 3500 students, it will eventually want to have more housing, more classrooms, and more teachers to accommodate the increased number of students. But only “eventually.” It will not immediately knock down the old buildings and rebuild them to better suit the higher scale; it will not immediately hire new teachers. The costs of adjustment are high. Haste makes waste. It will, instead, house the students in rented hotels and hire visiting teachers on temporary assignment. These are expensive and inconvenient devices but cheaper than excess haste in adjusting. Only over the long run will it take
advantage of the wearing out of buildings and the wider pool of good teachers available from a leisurely and thorough search to remake the plant and faculty to suit the higher scale.

The long-run costs will be lower than the short-run costs, which are in turn lower than the Hasty costs. See Figure 12.6. A rational university wanting to increase its size will move from point Start to Crowded to Suitable. The dashed short-run cost curve is exactly suitable only to the output of Start. Everywhere else—at student numbers below as well as above 2500—it is higher than long-run cost. The second principle then is that haste makes waste.

The two principles are widely applicable. A sudden rise in the number of students in grade school should not inspire an immediate remaking of whole schools, even if the rise itself is permanent and especially if it is not. A rise in the demand for automobiles should inspire Ford to hire people on overtime at old plants before constructing new plants. And it applies to falls in output as well as to rises. A fall in attendance at Boston Red Sox baseball games should not inspire the management to tear down Fenway Park and build a new, smaller stadium at once. Firms spend some time on their short-run cost curves.

Both principles would be false, however, if there was a frictionless secondhand market in university buildings, grade schools, auto plants, and baseball stadiums. Unless getting into and out of owning such things were expensive, the distinction between long- and short-run costs would be empty, because the firm would never be stuck with a short-run expedient.
**T or F:** A shirtmaker who owns easily sellable space on Seventh Avenue, rents his sewing machines, and buys labor by the week is never off his long-run cost curve.

**A:** He can resell all these things to the market at a moment's notice, repurchasing them at the next, with no significant cost of making and unmaking the deals. He will therefore never have to suffer the inconvenience of an unsuitable set of equipment. That is, true.

**Figure 12.7**

*Flexibility Is Valuable in Terms of Total (a) and Average (b) Cost*

Flexible methods of production are less costly at very high or very low levels of output than are inflexible methods. If the demand facing the firm is sufficiently variable—if there is often a boom or a bust—the firm will choose **flexible**.
The distinction between the long- and short-run cost curve, then, is a crude way of allowing for the costs imposed by inconstancy of output, by ignorance of its duration, by haste in adjusting. In a word, the distinction allows for the cost of transacting. The costs of transacting are not given to the firm by God, for a firm has some choice of the set of transaction costs it will face. For example, it can choose either to rent or to buy equipment, such as sewing machines for shirtmaking. Renting has different sizes and types of transaction costs from ownership, and the balance of advantage of one over the other is not obvious. In particular, to mention an elementary point, owning one’s machines does not make them costless.

**T or F:** Friedman, the sweatshop shirtmaker, who owns his sewing machines outright, has a cost advantage over Schwartz, who must pay rent on her machines.

**A:** One way of seeing the point is to note that Friedman had to borrow $500 (or whatever) per machine to become an owner and therefore must pay to the banker interest equal to the annual rental Schwartz pays to an owner. Another and deeper way is to note that the real “cost” of a machine to Friedman is what it could earn outside his own factory, which is precisely (in a rental market without transaction costs) the rental that Schwartz would pay. By either argument, false.

A firm must often make a choice between one or another short-run cost curve (that is, between one or another set of transaction costs keeping the curve above the long-run curve). Stueland Electric in St. Joseph, Michigan, for example, must choose between arrangements such as owning outright its own trucks or supplies of cable that give the least cost when the firm is running at normal output and arrangements such as depending on a truck-renting firm in Baroda or a supplier of cable in South Bend that give higher costs at normal output but lower costs at very high or low outputs. If Stueland faces a predictable, stable demand, it will choose the inflexible but (perhaps) cheap arrangement of outright ownership. If it faces a highly variable demand, it will choose the flexible arrangement, for this will give the lowest cost over high and low outputs taken together. The choice is between the dashed total cost curve and the solid total cost curve in the top panel of Figure 12.7. The bottom panel displays the same choice in terms of average cost (that is, cost per unit), bearing in mind that if one cost curve is higher than another in total cost for a given output, it must also be higher in average cost for that output. If the firm faces outputs such as boom and bust frequently enough, it will prefer the curve Flexible to Inflexible.

**The Long Run Is the Envelope of All the Short Runs**

The engine of analysis can at the end be thrown into reverse, deriving the long-run cost curve from the short instead of viewing the short-run cost curves as deviations from the long. There are infinitely many short-run cost curves, some high, some low, some flexible, some inflexible. What is true of all of them is that the long-run cost curve is by definition below them and in fact consists of their combined lowest borders. The long-run total cost curve shows the lowest cost at which the firm can produce any given level of output. Speaking technically, the long-run curve is the envelope of all the short-run curves. In mathematics a curve that just touches all of a number of curves is known as the envelope of the curves, a term you can remember by thinking of the curve’s
The long-run total and long-run average cost curves are lower bounds, or envelopes, of all the points on the short-run curves. Evidently, the short-run cost curves are always above their envelope, since being always below is how the envelope is defined (see Figure 12.8). The diagram simply restates the opening theme of the section: The best that a firm can do is its long-run cost curve, but it can do worse.

**Summary**

The production function can be viewed as best practice, in which case the corresponding cost curve is lowest cost. Adding constraints hurts, a principle applicable in the present case to cost curves. When additional constraints drive firms off
the lowest curve, the cost observed is not a good estimate of the lowest cost. An important example of additional constraints is the transaction costs that force firms to operate along short-run cost curves. Firms do not fall blindly into such curves. On the contrary, they choose them, choosing, for example, a flexible plant that has lower costs calculated over episodes of boom and bust relative to an inflexible one that has low costs only if output is steady. In any event the long-run cost curve is the lowest attainable for a given output, which is to say that it is the envelope of all the short-run curves.

EXERCISES FOR SECTION 12.2

1. Suppose that you had a set of observations of cost and output such as the scattered dots in Figure 12.5. If you wanted to know the Cost Function, would you run a line through the points?

2. Explain in each case the relevance of the adding-constraints-hurts principle:
   a. The 55-mph speed limit on balance hurts.
   b. Energy conservation cannot make the nation better off.
   c. If without railroads a shipment of wheat had to follow the pattern it in fact followed with railroads, the shipping cost would be higher than if the pattern was allowed to adjust to the absence of railroads.

3. True or false: Sears, Roebuck is better off owning the land under its stores than renting it, because then it doesn't pay rent.

PROBLEMS FOR SECTION 12.2

1. You are Jan deVries, the fabulously wealthy owner of a fleet of oil tankers. As a common carrier you are required by law to accept any shipment of oil demanded of you; that is, you have no control over your output (contrary to the assumption in the last few chapters that output is the one thing a firm can control). Initially your output of oil shipped is Old Output, and you have chosen a collection of ships that minimizes the cost of producing Old Output (see Figure 12.9). Output and total cost are those that will persist into the indefinite future. You are on the Old short-run cost curve.
   a. Look at the vertical distance called Transaction Cost: It is the cost (all costs that persist into the indefinite future) of shifting from the Old to the (dashed) New short-run cost curve. Assume that it is the same length no matter where on the diagram it is moved. If output changes permanently to New Output, will you find it worth your while to adjust your fleet to get the New cost curve?
   b. Suppose, however, that the old fleet deteriorates a little each year, driving the Old cost curve up. When will investment in an adjusted fleet take place?
   c. Suppose that the New cost curve falls a little each year, reflecting the improvements in technology that permit a fleet built to embody the technology cheaper to operate. When will investment in an adjusted fleet take place?
   d. Would you expect an industry with rapidly growing output, equipment that wore out quickly, and experiencing rapid technological change to have few or many occasions to invest in changing from one cost curve to another?

2. Marie and Brownie contemplate running an all-season skiing and camping resort near Montpelier, Vermont. They have two possible designs, one a general design that serves both for skiing in the winter and camping in the summer and the other a specialized design that makes skiing cheaper but camping more expensive. The costs expected into the indefinite future are as follows:
Figure 12.9  
Transaction Cost Is a Vertical Distance, Placeable Anywhere

The vertical distance transaction cost tells the cost of using the market to change the method of production. The change shifts the firm from the Old cost curve to the New. The firm that must increase output from Old to New output will make the shift if the gain from changing over at New output is greater than the transaction cost.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Summer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of guests</td>
<td>10,000</td>
<td>5,000</td>
</tr>
<tr>
<td>Cost per guest, general</td>
<td>$10</td>
<td>$10</td>
</tr>
<tr>
<td>Cost per guest, special ski</td>
<td>$5</td>
<td>$30</td>
</tr>
</tbody>
</table>

Suppose that the number of guests is given (the common carrier assumption) and that it is invariant to the choice of design. What is the unit of time over which costs should be measured? Which design has the lowest cost?

**True or False**

3. Compulsory conservation of energy will save the nation's resources.
4. Since the long-run cost is the lowest-cost way of producing each output, the long-run cost curve runs through the minimum points of short-run cost curves.

### 12.3 Cost Curves in Use: Fixed and Variable Costs

**What to Read For**  
How do the total cost curve and the total revenue curve determine how much a firm wants to produce? Are past, sunk costs relevant? Why not? Are fixed costs relevant? What are the two meanings of "fixed costs"?
After so much attention to the domestic affairs of the firm, it is now time to return to its foreign affairs, that is, its decision of how much to produce for sale. The story is really just a review, because you already know how a firm arranges its foreign affairs—it sets output so that marginal benefit equals marginal cost. The goal of making profit is served by two steps. First, plan how to produce any output at least cost. Second, choose the output that maximizes the excess of revenues over costs. Lars Sandberg, a farmer in Ohio producing soybeans for sale to a market that will buy all he produces at a fixed dollar price, is in such a situation. His total cost has the usual stretched shape. His total revenue is a straight line out of the origin. This is because each additional ton brings him a fixed additional number of dollars (that is, the total revenue curve has a constant slope).

The top panel of Figure 12.10 shows what he does. He produces the output
Chapter 12  COST CURVES OF THE FIRM

that makes the gap between the total revenue curve and the total cost curve as large as possible, because the gap is profit.

The bottom panel gives the usual correspondences between the totals on the top and the averages and marginals on the bottom. As usual, in the bottom panel the Maximum Profit point is at the output that equalizes marginal cost and price. And as usual the amount of Profit is the area of revenue under the Price line (equal here to the two shaded areas) minus the area of Cost under the Marginal Cost curve. The heavy line in the top panel is an alternative picture of the same Profit. It will not come as a complete surprise that the Maximum Profit is at the output at which the slope of the total cost curve (look at the dashed tangent) is equal to the slope of the total revenue: These slopes are exactly the marginal cost and the price, and by a well-known argument equalizing them brings bliss. You can overcome any surprise by translating the argument into the terms of the total cost diagram. At Too Much Output, for example, total revenue has risen above what it is at Maximum Profit by the amount \( AR \), but total cost has risen even farther, by \( AC \). Clearly it will be better to move back to Maximum Profit.\(^5\)

Past Costs Are Irrelevant

The central message of the diagram—and of the last two chapters—is that in the pursuit of profit a rational firm is influenced by its costs. What costs? Future costs that can be altered by the amount of output.

Q: A jeweler in Harper Court leaves the prices of her gold necklaces at their price when made even though the price of gold has since doubled. She says to an amazed and incredulous but grateful customer that “I make money even at the old prices.” Does she?

A: The “cost” relevant to a fully rational firm is not yesterday’s outlay of money but the cost from the moment the sale is made into the future. To take an even more extreme case, if gold prices were going to triple tomorrow and if God had informed the jeweler, then it would be madness to make any sales today. Likewise in the present case: The true “cost” is the opportunity cost, that is, what the necklaces would bring now or later, at another time or place. A doubling of the price of gold sharply increases the cost (though by less than doubling, for the labor and capital in fabrication and in marketing would not double in price and are also part of the cost). Selling a necklace worth $70 for $50 is to incur on this account a $20 loss of opportunity, not a gain.

The point is that historical costs are not directly relevant to the forward-looking decisions of a rational firm. Historical costs are, as the vivid word in accounting has it, “sunk.” The proverb is “let bygones be bygones.” Forget about the past and look to the future, for purposes of moneymaking if not of other sorts of wisdom. That Manhattan once cost $24.00 does not mean that for present purposes owners should care. If the owner of the land under the World Trade Center foolishly sells the land for $24.50 he cannot comfort himself by thinking, “Well, at least I made some profit: Once the whole island sold

\(^{5}\) First-year calculus reduces the matter to routine. Remember, though, calculus is not necessary. Suppose that the cost function were some function \( f(q) \), where \( q \) is the output chosen. Total revenue is of course \( pq \), where \( p \) is the given price. Profit is then revenue minus cost, \( pq - f(q) \), and to find the condition for maximizing this profit function, one sets its first derivative with respect to \( q \) equal to zero. Well, its derivative is \( p - f'(q) \), which when set equal to zero implies \( p = f'(q) \). This says that the maximum gap of profit occurs when the slope of the total revenue function (in this case the slope is just \( p \) ) is equal to the slope of the cost curve, \( f'(q) \). Profit is maximized at the output level for which marginal revenue is equal to marginal cost.
firms in regulated industries, to be sure, must often pay close attention to historical costs, because the accounting principles demanded by the regulations do. A regulated telephone company constrained to earn no more than 5% on the acquisition costs of its assets must calculate the acquisition costs historically, not at the cost in future that is relevant to their economic use. Furthermore, the past acquisition cost of even a nonregulated firm, such as the Evans Slipper factory, may be a useful estimate of its present cost of replacement if not much has happened since it was built to alter the cost of building or if what has happened is measurable. But these uses aside, historical costs do not determine present and future costs and are therefore useless for measuring profit.

Unavoidable Fixed Costs Are Past

A closely related point is that fixed costs do not affect the output of the firm. There are two distinct types of fixed costs. One type is fixed costs that are “fixed and unavoidable.” These have as much effect on present and future decisions as do historical costs: none. Fixed costs in this sense, for example, are often identified with the repayment of debts incurred to invest in equipment. If you borrow $10,000,000 from your bank to invest in a ship, you must pay it back with interest. Suppose that your monthly payment to the bank is $150,000. You would be pleased to earn more than this amount from the ship, and surely you expected to when you made the investment. But whether or not you actually do so is irrelevant to your behavior as a shipowner. Whatever you do, the bank each month presents you with a bill for $150,000. The bank could care less if you are still a shipping magnate or if business has been good. For your part, that you face a bill for $150,000 affects your business no differently than would a bill for $150,000 for something unrelated to your business, such as child support. The way in which you run the shipping business is unaffected by the burden of debt “on” it. The fixed and unavoidable cost has no effect on your business decisions—whether or not you keep the ship and how much a year you run it.

Q: British owners of old coal-powered (as distinct from oil-powered) ships after World War I earned enough in revenue to cover the captains, crew, supplies, fuel, and the like for operating the ships, but not enough to also cover the interest and repayment costs on the loans taken out before the war to buy the ships. True or false: The shipowners were probably acting irrationally when they bought the ships and were certainly acting irrationally when they continued to operate the ships despite the losses.

A: What matters to a judgment on the rationality of the initial investment is the expectation of future profits. The shipowners could well have made investments in 1910 in expectation of making money yet could find later that their expectation was mistaken, as it in fact was. And having made the initial mistake, the decision is bygone, irrelevant to the decision a shipowner faces in 1930 of whether or not to keep the ship afloat. So long as the ship earns enough to pay all the variable costs (the costs that do vary with output), the rational decision is to carry on. Therefore, false.

Avoidable Fixed Costs Determine Whether or Not to Produce Anything

There are two distinct types of fixed costs. The first has just been discussed: fixed and unavoidable. The second is fixed and avoidable if one closes down entirely. Payment on a loan used to open a restaurant is fixed and unavoidable even if the restaurant closes down. Payment of a license to operate the restaurant is fixed (that is, it does not vary with the output of the restaurant) but avoidable if it closes down and no longer begs permission of the sovereign power to exist.
To look at it from the other side, there are two kinds of variable costs: costs that are zero when the output is zero and costs that are not zero when the output is zero. The significance of the distinction arises from problems of the following sort.

**T or F:** Since a tax of $1000 per firm imposed on the existence of firms does not affect marginal cost, it does not affect the output chosen by each firm (unlike a sales tax or an employment tax) and therefore does not affect the output of the industry.

**A:** The successive assertions travel from plain truth to probable error by small steps. The first assertion is plain truth and deserves a lot of discussion.

It is true that a tax of $1000 per firm does not affect marginal cost. The $1000 is a fixed cost, invariant with output, given that some output is to be produced. The total cost (including here both fixed and variable cost) will move up by $1000 when the tax is imposed, thereby cutting profits by $1000. But because a uniform rise of $1000 in the cost curve does not alter its slope, the corresponding marginal cost curve is not moved (see Figure 12.11). That the shaded area of the Profit on Variable Cost in the bottom diagram does not change after the tax even though the Profit Gap in the top diagram shrinks is due to its exact definition, now made explicit for the first time. It is profit on variable cost alone, because it is calculated by subtracting from revenue the area of the costs under the marginal cost curve (that is, all and only those costs that do vary with output). Because the marginal cost does not move from its old position, the optimal output does not change. The economic common sense here is that the firm always wants to have as large a gap as possible between revenue and cost, whether the gap is big or small. The tax makes the gap smaller, and may even make it a loss, yet the firm still does as well as it can under the circumstances by setting marginal cost equal to marginal revenue. It is often said that low profits spur the firm on to greater effort and larger outputs; and it is also often said that they discourage the firm, causing less effort and smaller outputs. The middle ground claimed here, by contrast, says that the size of the profit earned has no effect whatever on the output chosen, given that some output is to be produced. The assertion is not self-evidently true, which is to say that it is a hypothesis about how people behave that might in some cases be false. In any event, it is the assumption made in the usual theory of the firm.

The later assertions in the question, however, do not follow from it. True, the $1000 tax per firm does not affect output given that some output is to be produced. But there’s the rub. That some output is to be produced is not in the long run given but is a choice made by the firm. If the tax is so large that profits on variable costs are offset entirely, then the firm will close down. And in the long run all costs are variable; that is, all costs are avoidable. In the long run, then, the tax does affect output, for it affects whether or not the firm produces at all. Therefore it can affect the output of the industry. As always, an extreme example helps. Suppose that the tax were $100,000 instead of $1000 and that it were imposed on each hot dog stand. Clearly, none would exist at such a high price of existence. To take the other extreme, suppose that the tax were $1. Clearly, the effect on the decision to exist as a hot dog stand would be trivial. The effect on the industry running hot dog stands is the more powerful the larger the effect on profit. In short, the answer is false; a $1000 tax can affect the output of the industry.

---

6 Fixed cost, then, is a constant of integration in the calculus problem to find \( \int MC(q) \, dq \), in which \( MC \) is the marginal cost function. Or, to look at it the other way, the total cost is variable cost (a function of \( q \)) plus fixed cost (a constant), meaning that marginal cost is the derivative of variable cost alone,

\[
\frac{dTC}{dq} = \frac{d[VC(q) + FC]}{dq} = \frac{dVC(q)}{dq}
\]

since FC is a constant.
Figure 12.11
A Fixed Tax per Firm (a) Does Not Move Marginal Cost (b)

A tax that does not vary with output increases total costs by the same amount at all levels of output and therefore does not change marginal cost. As long as the firm produces some output the tax will have no effect on how much the firm produces.

Average Cost Is Only Relevant to the Decision Whether to Produce at All

The logic of fixed and average costs, then, though poorly suited for questions of what output should be, is well suited for questions of whether any output should be produced at all. For example, the investment in equipment that a firm makes is analogous to a tax. Profit over variable cost must cover the cost.
A farmer with at least 100 acres would break even on a harvesting machine if the savings per acre were 0.84 shilling. If the savings per acre were higher, the savings line would be steeper, the machine would be profitable for 100 or even fewer acres, and a profit-maximizing farmer would buy the machine.

Q: By the 1870s a reaping machine to replace hand harvesting of wheat and barley had been available for decades, yet in England (unlike the United States) less than half the harvest was done by machine. It has been argued that this was a result of the small size of English farms, the high cost of machines, and the inconvenient nature of the English landscape (plowed for purposes of drainage into a "ridge-and-furrow" configuration, which made the cumbersome reaping machines less effective).

1. If a farmer bought a machine for 660 shillings on which he had to earn 12.7% a year to pay back the banker, what in shillings does the machine have to pay back every year to justify its purchase? If the labor saving on flat land not plowed into ridges and furrows were 3.3 shillings per acre harvested per year, what would be the lowest annual acreage harvested (the "threshold" acreage) at which a machine would become profitable?

2. The labor saving due to the introduction of the reaping machine was lower on ridge-and-furrow land. Illustrate in a diagram of total cost against acres harvested per machine per year how the threshold acreage varies with the per acre labor saving. If the maximum annual capacity of a machine were 100 acres harvested, what is the minimum labor saving that will make the machine profitable?

3. If the labor saving on a ridge-and-furrow farm were

2 shillings an acre, what would you conclude about the rationality of not adopting the reaper in view of the fact that the wheat and barley acreage on a majority of English farms (mostly ridge-and-furrow) was less than 50 acres (with a good deal of variation around this average)?

4. How does this conclusion change if farmers can share a machine, if a market in machine time exists, if the size of farms can be altered, or if the percentage of a farm devoted to wheat and barley can be altered?

A: 1. The annual opportunity cost of the investment is $0.127(660) = 84$ shillings. This is a fixed cost with respect to the number of acres harvested. You pay it regardless. So if the machine saves 3.3 shillings an acre, you had better have at least $84/3.3 = 25$ acres to harvest with your machine.

2. The diagram is given in Figure 12.12. The minimum labor saving is simply $84 = 100S$, that is, $S = 0.84$.

3. The threshold acreage for 2 shillings an acre saving is $84/2 = 42$, so a good many farms in England harvested acreages below the threshold (some harvested only, say, 20 acres). These farmers were not being irrational to ignore the charms of mechanized harvesting. One could build from these elements a full theory of the adoption of the reaper: As the benefits varied over time or as the size of farms varied, the percentage of the distribution of farm sizes that adopted the reaper would vary.

4. The conclusion of (3), however, changes radically if machines can be rented, loaned, or shared. Farmers could club together to buy a machine, for example, and harvest 100 acres with it each year, even if they all have pitiful 15-acre stands of wheat and barley. Or a farmer with a reaper on 70 acres could notice that his machine stood idle for the last week or so of the harvest and could rent it out to his reaperless neighbor. Or a company could be formed specializing in reaping. The threshold idea no longer has force, and the failure to adopt reapers remains a puzzle, as long as the labor saving on ridge-and-furrow land was more than 0.84 shilling an acre. Furthermore, the threshold could be changed if the size of farms or the percentage of a farm devoted to wheat and barley could be raised. The acreage to be harvested, in other words, is not given by God but is alterable, in which case it might be altered to take advantage of the machine.

**Summary**

A firm achieves the most profit by equalizing marginal revenue and marginal cost or (equivalently) finding as large a gap as possible between total revenue and total cost.\(^8\) The total cost in question is future, variable cost. It is not past, fixed costs. Firms, like all rational folk, let bygones be bygones, do not cry over spilt milk, and look to the future. A tax levied in a lump or a machinery cost that can be avoided by refraining from any output, of course, is subject to choice by the firm. Costs of doing business do not affect marginal cost and therefore do not affect output if some is forthcoming. But if the cost imposed is too large, the firm will shut down entirely. The question—taken up in detail in Chapter 14—becomes whether or not profit on variable cost (such as the revenue over cost in a restaurant or the savings on labor cost in buying a machine) is large enough to offset the cost of licenses, buildings, or reaping machines. To be or not to be, that is the question.

**EXERCISES FOR SECTION 12.3**

1. This exercise is meant to convince you that where marginal cost equals marginal benefit is where the gap between total cost and total revenue is greatest. The diagram you construct is used later, so do it with care. Get a piece of graph paper (or simply use lined notebook paper, imagining where the vertical lines would come). Draw some axes without units, labeling the vertical one “total cost and total revenue” and the horizontal “quantity.“ Leave

---

\(^8\) It might as well be admitted here that the phrase “total cost” is ambiguous in economics. It sometimes means “all cost, including fixed and variable costs,” but sometimes, as here, it means “cost in total, as distinct from marginal or average cost.”
room below them for other axes. Now draw freehand a stretched Z curve of Total Cost in the style of the Total Cost curve in the top panel of Figure 12.10. Make sure it is a Z, not an S. Draw a straight-line Total Revenue curve cutting through the cost curve. Now mark off the distances on your axes, dollars on the vertical, and, say, tons on the horizontal (choose easy-to-work-with numbers, like 100, 200, 300, and so forth). No tricks, right? You’ve supplied all the data so far, except the general shape of the curves. So if it turns out that the output where marginal cost in your diagram equals marginal revenue is the output where the gap between total cost and revenue is greatest, you’ll have to believe that it’s a property of such curves in general, not some accident or some sleight of hand by economists.

All right. Now locate and mark where the vertical gap between revenue and cost is greatest. If your curve is curvy enough you can do it with your eye; otherwise you can move a straight edge of paper around until the greatest gap is marked off on it.

Draw axes exactly below these, again in the style of Figure 12.10. Mark off the (horizontal) quantity axis just as it is in your top panel. Calculate the slope you have implicitly given the Total Revenue curve. It will be, of course, the rise over the run, so many dollars per ton. In fact it will be the price; and since the line is straight it will be one price. Choose a convenient point on the (vertical) average-and-marginal-cost axis of your bottom panel to represent this level of price, and draw the Price line (consult Figure 12.10).

Now the delicate part. Using a straightedge pushed up against various points on the Total Cost curve, find the numerical values of the slopes of the various tangents to the curve. Remember: rise over run, so many dollars per ton. These will be, of course, marginal costs: Plot them in your bottom panel, making sure to line them up with the same tonnages and making sure that their heights are exactly as they should be relative to the Price line. Do enough points to sketch in the Marginal Cost curve. At what tonnage does Marginal Cost intersect Price? That’s right: the same tonnage as the one that makes the gap between total revenues and costs a maximum. Q.E.D.

2. By way of review of cost curves, use the diagram constructed for Exercise 1 to prove that Marginal Cost intersects Average Cost at the minimum of Average Cost. Average Cost at some output, you recall, will be the slope of a ray from the origin out to that point on the Total Cost curve.

3. Again by way of review of cost curves, use your diagrams to show that the size of the area analogous to that marked Cost in the bottom panel of Figure 12.10 is the same as the height of the total cost curve. To do this you will need to divide up the area under Marginal Cost into ten or so little rectangles, each with some easy-to-reckon base marked off along the tonnage axis and each with a height running through that part of the Marginal Cost curve. The area of all the rectangles will not be exactly equal to the total cost because the Marginal Cost curve is not exactly a series of little steps like the tops of your rectangles. But it will be pretty close. The equality proves the crucial point about marginal diagrams: Areas in marginal diagrams correspond to some vertical distance in the corresponding total diagram.

4. Draw a third diagram with tonnage on the horizontal axis and Total Profits on the vertical. This can be done by measuring the vertical gaps between revenue and cost and then plotting the result (if revenue is below cost, of course, the profit is negative). Where is profit a maximum? Define the idea of marginal profit. (Hint: Look at your second, marginal-cost-and-revenue diagram. What is the profit earned from the first unit produced? From the second?) True or false: When marginal profit is zero, total profit is at a maximum.

5. Which of the following are fixed costs? Distinguish those that are fixed and avoidable from those that are fixed and unavoidable:
   a. Heating expenses of the college buildings.
b. Payments by the Wretched of the Earth Mining Company of Pretoria to the college (which loaned Wretched money by buying its bonds).

c. Payments by the college to tenured faculty. ["Tenure" is alleged to mean that one cannot be dismissed from one's job except for gross moral turpitude or gross dereliction of duty (Look those up).]

d. Payments to untenured faculty.

e. Fees paid to the state by the college to license its dining facilities.

f. Fees paid by the college to be a college.

g. Purchase of exam booklets by the college.
PROBLEMS FOR SECTION 12.3

1. Label and explain two alternative and equal areas of profit in Figure 12.13.
2. Do the same for Figure 12.14.
3. In the diagrams of total revenue drawn in the text, it is assumed that price (average revenue being the slope of a line from the origin out to the place on the total revenue curve) does not fall with larger quantities sold. If Saudi Arabia produces substantially more oil, however, the price will fall. Only oil at a lower price will find new demanders.
   a. In such a situation, what is the shape of the total revenue curve?
   b. Suppose that Saudi Arabia has the usual stretched Z for a total cost curve and faces the total revenue curve described in (a). Where would Saudi Arabia set its output to maximize profit? Would the resulting price be equal to marginal cost?
4. An Iowa City councilman argued that "We spent $30 million refurbishing the downtown: It would be silly not to build a highway now to make it easy for people to get to the downtown." Criticize.