Chapter 8 THEORY OF COST

8.1 INTRODUCTION

The physical conditions of production, the price of resources, and the economically efficient conduct of an entrepreneur jointly determine the cost of production of a business firm. The production function furnishes the information necessary to trace out the isoquant map. Resource prices establish the isocost curves. Finally, efficient entrepreneurial behavior dictates production of any level of output by that combination of inputs which equates the marginal rate of technical substitution and the input-price ratio. Each position of tangency therefore determines a level of *output* and its associated *total cost.* From this information, one may construct a table, a schedule, or a mathematical function relating total cost to the level of output. This is the cost schedule or cost function which is the subject of this chapter.[1]

[1] Let C represent total cost, q output, and A the cost of (short-run) fixed inputs. The cost of fixed inputs is the same regardless of the level of output; but the cost of the variable inputs changes with their usage. Hence the cost function may be written as $C = A + g(q)$.

Alternatively, a specific cost function may be derived directly from the production function if the latter is known. As will be obvious, a direct derivation can become a messy business even for very simple production functions.

Suppose the production function takes the Cobb-Douglas form:

$$(8.1.1) \qquad q = a x_1^b x_2^c$$

where q represents physical output; x_1 and x_2 represent the physical quantities of two inputs; and a, b, and c are technologically given positive constants. Further, let w and r represent the given constant unit prices of the inputs x_1 and x_2 respectively. From equation (8.1.1) equality between the marginal rate of technical substitution and the input-price ratio requires that

$$(8.1.2) \qquad \frac{w}{r} = \frac{b x_2}{c x_1}.$$

Take the logarithms of expressions (8.1.1) and (8.1.2) and write them as simultaneous equations:

$$(8.1.3) \qquad \begin{aligned} b \log x_1 + c \log x_2 &= \log q - \log a, \\ - \log x_1 + \log x_2 &= \log c - \log b + \log w - \log r. \end{aligned}$$

Solving equations (8.1.3) simultaneously yields the following expressions:

Before turning to the mechanics of cost analysis, however, we need to pause for a somewhat broader view and to pose the question, "Just *what* constitutes the legitimate costs of production?" There are two answers to this question which, under ideal circumstances, happen to become one and the same. At present we must be content with the two; but in Chapter 16 we set out the conditions that must obtain to make the answers the same.

8.1.a—Social Cost of Production

Economists are principally interested in the social cost of production, the cost a society incurs when its resources are used to produce a given commodity. At any point in time a society possesses a pool of

---

$$(8.1.4) \qquad x_2^* = [a^{-1} q c^b b^{-b} w^b r^{-b}]^{\frac{1}{b+c}},$$

$$x_1^* = [a^{-1} q c^{-c} b^c w^{-c} r^c]^{\frac{1}{b+c}},$$

where x_1^* and x_2^* are the quantities of input required to produce q units of output at the cost-minimizing input ratio given by expression (8.1.2).

The cost (C) of producing q units is

$$(8.1.5) \qquad C = w x_1^* + r x_2^*.$$

Substituting expression (8.1.4) into (8.1.5) yields

$$(8.1.6) \qquad C = r\left[\left(\frac{cw}{br}\right)^b \frac{q}{a}\right]^{\frac{1}{b+c}} + w\left[\left(\frac{cw}{br}\right)^{-c} \frac{q}{a}\right]^{\frac{1}{b+c}}.$$

Consider

$$w\left[\left(\frac{cw}{br}\right)^{-c} \frac{q}{a}\right]^{\frac{1}{b+c}} = \frac{w\left[\left(\frac{cw}{br}\right)^{-c}\left(\frac{cw}{br}\right)^{b+c} \frac{q}{a}\right]^{\frac{1}{b+c}}}{\left[\left(\frac{cw}{br}\right)^{b+c}\right]^{\frac{1}{b+c}}}$$

$$= \frac{w\left[\left(\frac{cw}{br}\right)^b \frac{q}{a}\right]^{\frac{1}{b+c}}}{\frac{cw}{br}} = \frac{br}{c}\left[\left(\frac{cw}{br}\right)^b \frac{q}{a}\right]^{\frac{1}{b+c}}.$$

Substituting the final expression in (8.1.6) one obtains

$$(8.1.7) \qquad C = r\left[\left(\frac{cw}{br}\right)^b \frac{q}{a}\right]^{\frac{1}{b+c}} + \frac{br}{c}\left[\left(\frac{cw}{br}\right)^b \frac{q}{a}\right]^{\frac{1}{b+c}},$$

or

$$(8.1.8) \qquad C = r\left(\frac{b+c}{c}\right)\left[\left(\frac{cw}{br}\right)^b \left(\frac{q}{a}\right)\right]^{\frac{1}{b+c}}.$$

Since the technological parameters a, b, and c are given, as well as the market parameters r and w, expression (8.1.8) shows cost as a function of the level of output only.

resources either individually or collectively owned, depending upon the political organization of the society in question. From a social point of view the objective of economic activity is to get as much as possible from this existing pool of resources. What is "possible," of course, depends not only upon the efficient and full utilization of resources but upon the specific list of commodities produced as well. A society could obviously have a greater output of automobiles if only small compact cars were produced. Larger, more luxurious cars require more of almost every input. But in their private evaluation schemes some members of the society may attach much greater significance to luxury cars than to compact cars.

Balancing the relative resource cost of a commodity with its relative social desirability entails a knowledge of both social valuations and social cost. This broad problem is deferred to Chapter 16 so that our attention can now be directed exclusively to social cost.

The social cost of using a bundle of resources to produce a unit of commodity X is the number of units of commodity Y that must be sacrificed in the process. Resources are used to produce both X and Y (and all other commodities). Those resources used in X production cannot be used to produce Y or any other commodity. To use a popular wartime example, devoting more resources to the production of guns means using fewer resources to produce butter. The social cost of guns is the amount of butter foregone.

Economists speak of this as the *alternative* or *opportunity cost* of production.

Definition: the *alternative* or *opportunity cost* of producing one unit of commodity X is the amount of commodity Y that must be sacrificed in order to use resources to produce X rather than Y. This is the social cost of producing X.

8.1.b—Private Cost of Production

There is a close relationship between the opportunity cost of producing commodity X and a calculation the producer of X must make. The use of resources to produce X rather than Y entails a social cost; there is a private cost as well because the entrepreneur must pay a price to get the resources he uses.

Suppose he does. The entrepreneur pays a certain amount to purchase resources, uses them to produce a commodity, and sells the commodity. He can compare the receipts from sales with the cost of resources and, roughly speaking, determine whether he has made an accounting profit or not. But an accountant would be quick to tell the entrepreneur he should make some further calculations. He has invested his time and money in producing commodity X. If he had not undertaken this line of business, he could have invested his time and money elsewhere—in another line of business, perhaps, or by purchasing securities with his money and using his time as an employee of another entrepreneur.

The producer of X incurs certain *explicit costs* by purchasing resources. He incurs some *implicit costs* also, and a full accounting of profit or loss must take these implicit costs into consideration. The pure economic profit an entrepreneur earns by producing commodity X may be thought of as his accounting profit minus what could be earned in the best alternative use of his money and what could be earned in the best alternative use of his time. These two factors are called the implicit cost of production.

Definition: the implicit costs incurred by an entrepreneur in producing a specific commodity consist of the amounts he could earn in the best alternative use of his time and money. He earns a pure economic profit from producing X if, and only if, his total receipts exceed the sum of his explicit and implicit cost.

Implicit costs are thus a fixed amount (in the short run) that must be added to explicit costs in a reckoning of pure economic profit.

8.2 SHORT AND LONG RUNS

In Chapter 6 a convenient analytical fiction was introduced, namely the *short run,* defined as a period of time in which certain types of inputs could not be increased or decreased. That is, in the short run there are certain inputs whose usage cannot be changed regardless of the level of output. Similarly, there are other inputs, variable inputs, whose usage can be changed. In the long run, on the other hand, all inputs are variable—the quantity of all inputs can be varied so as to obtain the most efficient input combination.

Corresponding to fixed inputs are short-run fixed costs. The various fixed inputs have unit prices; the fixed explicit cost is simply the sum of unit prices multiplied by the fixed number of units used. In the short run, implicit costs are also fixed; thus it is an element of fixed cost.

Definition: total fixed cost is the sum of the short-run explicit fixed costs and the implicit cost incurred by an entrepreneur.

Inputs variable in the short run give rise to short-run variable cost. Since input usage can be varied in accordance with the level of output, variable costs also vary with output. If there is zero output no units of variable input need be employed. Variable cost is accordingly zero and total cost is the same as total fixed cost. When there is a positive level of output, however, variable inputs must be used. This gives rise to variable costs, and total cost is then the sum of total variable and total fixed cost.

Definition: total variable cost is the sum of the amounts spent for each of the variable inputs used.

Definition: total cost in the short run is the sum of total variable and total fixed cost.

TABLE 8.3.1

FIXED, VARIABLE, AND TOTAL COST

Quantity of Output	*Total Fixed Cost*	*Total Variable Cost*	*Total Cost*
1	$100	$ 10.00	$ 110.00
2	100	16.00	116.00
3	100	21.00	121.00
4	100	26.00	126.00
5	100	30.00	130.00
6	100	36.00	136.00
7	100	45.50	145.00
8	100	56.00	156.00
9	100	72.00	172.00
10	100	90.00	190.00
11	100	109.00	209.00
12	100	130.40	230.40
13	100	160.00	260.00
14	100	198.20	298.20
15	100	249.50	349.50
16	100	324.00	424.00
17	100	418.50	518.50
18	100	539.00	639.00
19	100	698.00	798.00
20	100	900.00	1000.00

8.3 THEORY OF COST IN THE SHORT RUN

Our analysis of cost begins with the theory of short-run cost; we then move to the "planning horizon" in which all inputs are variable and study the theory of cost in the long run.

8.3.a—Total Short-Run Cost

Analysis of total short-run cost depends upon two propositions already discussed in this chapter: (a) the physical conditions of production, the unit prices of inputs, and efficient operation determine the cost of production associated with each possible level of output; and (b) total cost may be divided into two components, fixed cost and variable cost.

Suppose an entrepreneur has a fixed *plant* that can be used to produce a certain commodity. Further suppose this plant costs $100.

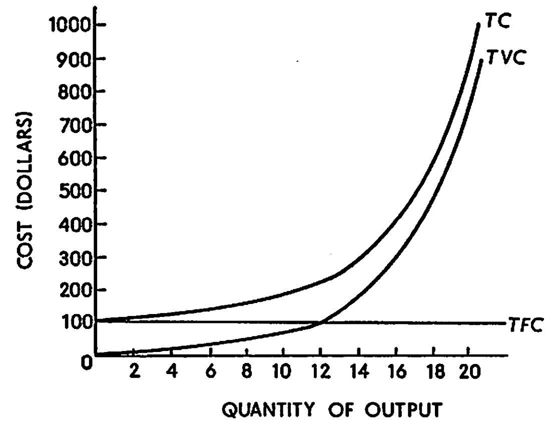

FIGURE 8.3.1

FIXED, VARIABLE, AND TOTAL COST CURVES

Total fixed cost is, therefore, $100—it does not change in magnitude irrespective of the level of output. This is reflected in Table 8.3.1 by the column of $100 entries labeled "Total Fixed Cost." It is furthermore shown by the horizontal line labeled *TFC* in Figure 8.3.1. Both table and graph emphasize that fixed cost is indeed fixed.

Variable inputs must also be used if production exceeds zero. In the spirit of Chapter 6, you might suppose there is only one variable input; alternatively, the multiple-input approach of Chapter 7 may be adopted. The choice is really not material, because an increase in the level of output requires an increase in the usage of inputs—whether this be one variable input or many variable inputs used in the optimal combination. In either case, the greater the level of variable input the greater the variable cost of production. This is shown in column 3 of Table 8.3.1 and by the curve labeled *TVC* in Figure 8.3.1.

Summing total fixed and total variable cost gives total cost, the entries in the last column of Table 8.3.1 and the curve labeled *TC* in Figure 8.3.1. From the figure one may see that *TC* and *TVC* move together and are, in a sense, parallel. That is to say, the slopes of the two curves are the same at every output point; and at each point, the two curves are separated by a vertical distance of $100, the total fixed cost.

8.3.b—Average and Marginal Cost

The total cost of production, including implicit cost, is very important to an entrepreneur. However, one may obtain a deeper understanding of total cost by analyzing the behavior of various average costs and of marginal cost.

The illustration of Table 8.3.1 is continued in Table 8.3.2. Indeed

TABLE 8.3.2

AVERAGE AND MARGINAL COST CALCULATIONS

Quantity of Output	*Total Fixed Cost*	*Total Variable Cost*	*Total Cost*	*Average Fixed Cost*	*Average Variable Cost*	*Average Total Cost*	*Marginal Cost*
1........	$100	$ 10.00	$ 110.00	$100.00	$10.00	$110.00	$—
2........	100	16.00	116.00	50.00	8.00	58.00	6.00
3........	100	21.00	121.00	33.33	7.00	40.33	5.00
4........	100	26.00	126.00	25.00	6.50	31.50	5.00
5........	100	30.00	130.00	20.00	6.00	26.00	4.00
6........	100	36.00	136.00	16.67	6.00	22.67	6.00
7........	100	45.50	145.50	14.29	6.50	20.78	9.50
8........	100	56.00	156.00	12.50	7.00	19.50	10.50
9........	100	72.00	172.00	11.11	8.00	19.10	16.00
10........	100	90.00	190.00	10.00	9.00	19.00	18.00
11........	100	109.00	209.00	9.09	9.90	19.00	19.00
12........	100	130.40	230.40	8.33	10.87	19.20	21.40
13........	100	160.00	260.00	7.69	12.30	20.00	29.60
14........	100	198.20	298.20	7.14	14.16	21.30	38.20
15........	100	249.50	349.50	6.67	16.63	23.30	51.30
16........	100	324.00	424.00	6.25	20.25	26.50	74.50
17........	100	418.50	518.50	5.88	24.38	30.50	94.50
18........	100	539.00	639.00	5.55	29.94	35.50	120.50
19........	100	698.00	798.00	5.26	36.74	42.00	159.00
20........	100	900.00	1000.00	5.00	45.00	50.00	202.00

the first four columns of the latter exactly reproduce Table 8.3.1. The remaining four columns show the new concepts to be introduced.

First consider the column labeled "average fixed cost."

Definition: average fixed cost is total fixed cost divided by output.

The calculation is very simple. When one unit of output is produced, *AFC* is $\$100/1 = \$100$. When two units are produced, *AFC* $= \$100/2 = \$50$; and so on. Graphically, average fixed cost is shown

by the curve designated *AFC* in Figure 8.3.2. Cost in dollars is plotted on the vertical axis and output on the horizontal axis. The *AFC* curve is negatively sloped throughout because as output increases the ratio of fixed cost to output must decline.[2] Mathematically, the *AFC* curve is a rectangular hyperbola.

Next move to column 6, Table 8.3.2. This column is labeled "average variable cost," a concept that is entirely analogous to average fixed cost.

Definition: average variable cost is total variable cost divided by output.

Again the calculation is simple and gives rise to the curve labeled *AVC* in Figure 8.3.2.[3] But now there is a great difference between *AVC* and *AFC:* the former does not have a negative slope throughout its entire range. Indeed, in this illustration *AVC* first declines, reaches a minimum, and rises thereafter.

The reason for this curvature lies in the theory of production. Total variable cost equals the number of units of variable input used (V) multiplied by the unit price of the input (P). Thus in the one variable-input case, $TVC = PV$.

Average variable cost is *TVC* divided by output O, or

$$AVC = \frac{TVC}{O} = P\frac{V}{O}.$$

Consider the term V/O, the number of units of input divided by the number of units of output. In Chapter 6, average product (AP) was defined as total output (O) divided by the number of units of input (V). Thus

$$AVC = P\left(\frac{1}{AP}\right),$$

or price per unit of input multiplied by the reciprocal of average product. Since average product normally rises, reaches a maximum, and then declines, average variable cost normally falls, reaches a minimum, and rises thereafter.

Relationship: a production function such as that shown in Figure 6.2.1 gives rise to the average product curve shown in Figure 6.2.2. This type of production function also determines a total variable cost curve such as that in Figure 8.3.1 and the average variable cost curve shown in Figure 8.3.2.

[2] As in footnote 1, let the cost function be $C = A + g(q)$, where A is total fixed cost and $g(q)$ gives the total variable cost associated with each level of output. Thus average fixed cost is A/q and its slope is $-A/q^2$.

[3] Using footnote 2, average variable cost is $g(q)/q$ and its slope is

$$\frac{qg'(q) - g(q)}{q^2}.$$

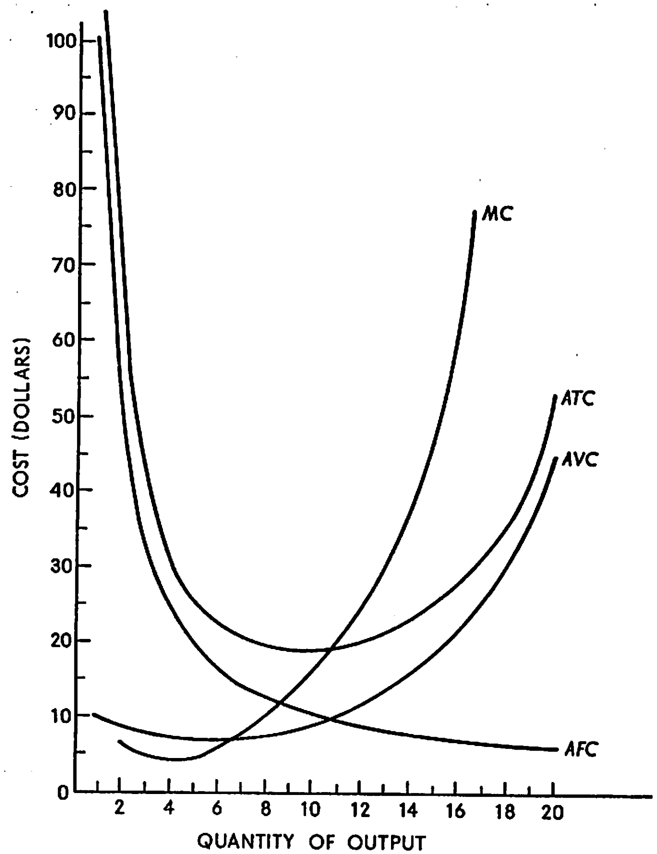

FIGURE 8.3.2

AVERAGE AND MARGINAL COST CURVES

Column 7, Table 8.3.2, contains the entries for average total cost, which may also be called average cost or unit cost.

Definition: average total cost is total cost divided by output.

In light of this definition, ATC may be computed by dividing the entries in column four by the corresponding entries in column one.[4] However, since $TC = TFC + TVC$,

[4] In the notations of the previous footnotes,

$$ATC = \frac{C}{q} = \frac{A}{q} + \frac{g(q)}{q}$$

and its slope is

$$-\frac{A}{q^2} + \frac{qg'(q) - g(q)}{q^2} = \frac{qg'(q) - A - g(q)}{q^2} = \frac{qg'(q) - C}{q^2}.$$

$$ATC = \frac{TC}{O} = \frac{TFC}{O} + \frac{TVC}{O} = AFC + AVC.$$

Thus one may calculate average cost as the sum of average fixed and average variable cost.

This method of calculation also explains the shape of the average total cost curve in Figure 8.3.2. Over the range of values for which both AFC and AVC decline, ATC must obviously decline as well. But even after AVC turns up, the marked decline in AFC causes ATC to continue to decline. Finally, however, the increase in AVC more than offsets the decline in AFC; ATC therefore reaches its minimum and increases thereafter.

Finally, column 8 of Table 8.3.2 contains the entries for marginal cost.

Definition: marginal cost is the addition to total cost attributable to the addition of one unit to output.

Marginal cost is thus calculated by subtracting successively the entries in the total-cost column.[5] For example, the marginal cost of the second unit produced is $MC_2 = TC_2 - TC_1$. Since only variable cost changes in the short run, however, marginal cost may be computed by successive subtraction of the entries in the total variable cost column. Thus the marginal cost of the second unit is also $MC_2 = TVC_2 - TVC_1$.

As shown in Figure 8.3.2, MC—just as AVC—first declines, reaches a minimum, and rises thereafter. The explanation for this curvature also lies in the theory of production. Let Δ denote "the change in." As shown just above, $MC = \Delta\ (TVC)$ for a unit change in output. More generally, if output does not change by precisely one unit, $MC = \Delta(TVC)/\Delta O$. In our previous notation, $TVC = PV$. Thus $\Delta TVC = P(\Delta V)$, for an entrepreneur who is a perfect competitor in the input market (input price is given by market demand and supply and changes in his purchases do not affect the price).

Using the two relations,

$$MC = P\frac{\Delta V}{\Delta O}.$$

In Chapter 6, marginal product (MP) was defined as the change in output attributable to a change in input, or $MP = \Delta O/\Delta V$. Thus

$$MC = P(1/MP).$$

Since marginal product normally rises, reaches a maximum, and

[5] For infinitesimally small changes in output, $MC = dC/dq = g'(q)$ and its slope is $g''(q)$.

declines, marginal cost normally declines, reaches a minimum, and rises thereafter.

Relationship: a production function such as that shown in Figure 6.2.1 gives rise to the marginal product curve shown in Figure 6.2.2. This type of production function also determines a total cost curve such as that in Figure 8.3.1 and the marginal cost curve shown in Figure 8.3.2.

8.3.c—Geometry of Average and Marginal Cost Curves

In Chapter 6, the average and marginal product curves were derived geometrically from the total product curve. In like manner, the

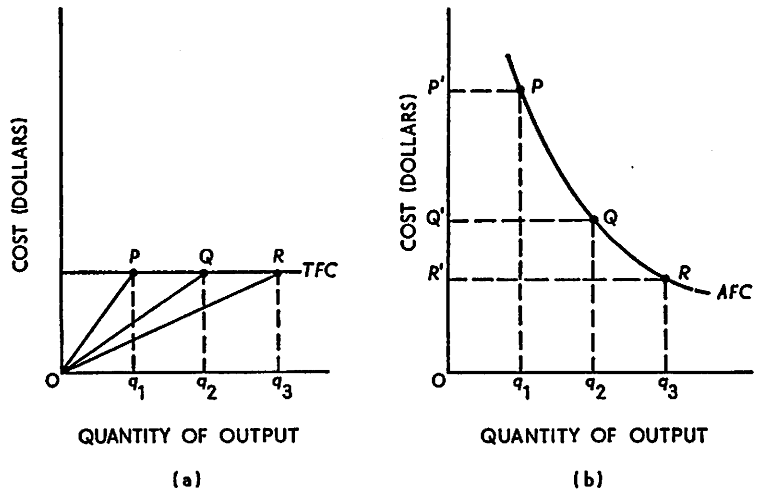

FIGURE 8.3.3

DERIVATION OF THE AVERAGE FIXED COST CURVE

average and marginal cost curves can be derived from the corresponding total cost curves.

Figure 8.3.3 illustrates the derivation of average fixed cost. (Note: vertical axes of panels a and b have different scales.) In panel a, total fixed cost is plotted and the outputs Oq_1, Oq_2, and Oq_3 are measured so that $Oq_1 = q_1q_2 = q_2q_3$. Since $AFC = TFC/O$, average fixed cost is given by the slope of a ray from the origin to a point on the TFC curve. For output Oq_1, AFC is the slope of the ray OP, or q_1P/Oq_1. Similarly, for output Oq_2, AFC is q_2Q/Oq_2, and so on. Since TFC is always the same, $q_1P = q_2Q = q_3R$. By construction, $Oq_2 = 2Oq_1$, and $Oq_3 = 3Oq_1$. Thus AFC for output Oq_2 is $q_2Q/Oq_2 = q_1P/2Oq_1 = 1/2\ (q_1P/Oq_1) = 1/2\ AFC$ for output Oq_1. This is shown in panel b by the difference in OP' and OQ'—more specifically, $OQ' = 1/2\ OP'$.

Similarly, as you can demonstrate for yourself, $OR' = 1/3OP'$. The remaining points on AFC are determined in the same way.

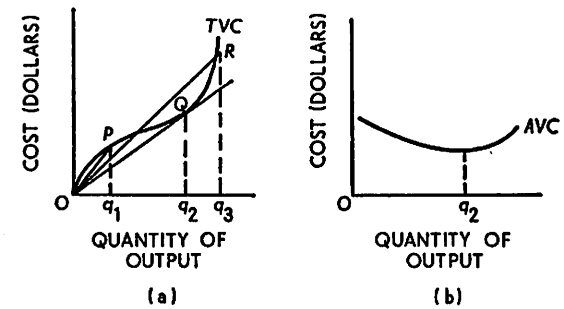

FIGURE 8.3.4

DERIVATION OF THE AVERAGE VARIABLE COST CURVE

Figure 8.3.4 shows how AVC is derived from TVC. As is true of all "average" curves, the average variable cost associated with any level of output is given by the slope of a ray from the origin to the corresponding point on the TVC curve. As may easily be seen from panel a, the slope of a ray from the origin to the curve steadily diminishes as one passes through points such as P; and it diminishes until the ray is just tangent to the TVC curve at point Q, associated with output Oq_2. Thereafter the slope increases as one moves from Q toward points such as R. This is reflected in panel b by constructing AVC with a negative slope until output Oq_2 is attained. After that point, the slope becomes positive and remains positive thereafter.

Exactly the same argument holds for panels a and b of Figure 8.3.5, which show the derivation of ATC from TC. The slope of the ray

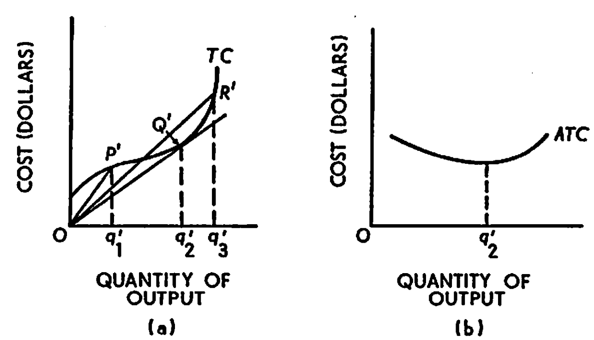

FIGURE 8.3.5

DERIVATION OF THE AVERAGE TOTAL COST OR UNIT COST CURVE diminishes as one moves along TC until the point Q' is reached. At Q' the slope of the ray is least, so minimum ATC is attained at the output level Oq'_2. Thereafter the slope of the ray increases continuously, and the ATC curve has a positive slope. (Note: the output level Oq'_2 does not represent the same quantity in Figure 8.3.3–8.3.6.)

Finally, the derivation of marginal cost is illustrated in Figure 8.3.6. Panel a contains the total cost curve TC. As output increases from Oq_1 to Oq_2, one moves from point P to point Q, and total cost increases from TC_1 to TC_2. Marginal cost is thus

$$MC = \frac{TC_2 - TC_1}{Oq_2 - Oq_1} = \frac{QR}{PR}.$$

Now let the point P move along TC toward point Q. As the distance between P and Q becomes smaller and smaller the slope of the tangent

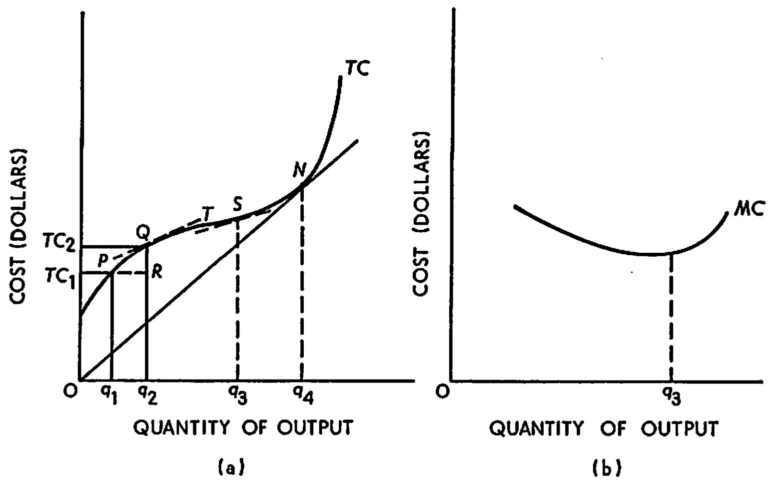

FIGURE 8.3.6
DERIVATION OF THE MARGINAL COST CURVE

T at point Q becomes a progressively better estimate of QR/PR. And in the limit, for movements in a tiny neighborhood around point Q the slope of the tangent is marginal cost.

As one moves along TC through points such as P and Q the slope of TC diminishes. The slope continues to diminish until point S is reached at output Oq_3. Thereafter the slope increases. Therefore, the MC curve is constructed in panel b so that it decreases until output Oq_3 is attained and increases thereafter.

One final point should be noted about Figures 8.3.4 and 8.3.6. As already shown, TC and TVC have the same slope at each output point; TC is simply TVC displaced upward by the constant amount TFC. Since the slopes are the same, MC is given by the slope of either curve.

In panel a, Figure 8.3.4, the slope of the ray OQ gives minimum AVC. But at this point the ray OQ is just tangent to TVC; hence it also gives MC at this point. Thus $MC = AVC$ when the latter attains its minimum value.[6] Similarly, in panel a, Figure 8.3.6, the slope of the ray ON gives minimum ATC. But at this point the ray is tangent to TC; thus its slope also gives MC. Consequently $MC = ATC$ when the latter attains its minimum value.[7]

8.3.d—Short-Run Cost Curves

The properties of the average and marginal cost curves, as derived in subsection 8.3.c, are illustrated by the "typical" set of short-run cost curves shown in Figure 8.3.7. The properties may be summarized as follows.

Relationships: (i) AFC declines continuously, approaching both axes asymptotically, as shown by points one and two in the figure. AFC is a rectangular hyperbola. (ii) AVC first declines, reaches a minimum at point four, and rises thereafter. When AVC attains its minimum at point four, MC equals AVC. As AFC approaches asymptotically close to the horizontal axis, AVC approaches ATC asymptotically, as shown by point five. (iii) ATC first declines, reaches a minimum at point three, and rises thereafter. When ATC attains its minimum at point three, MC equals ATC. (iv) MC first declines, reaches a minimum at point six, and rises thereafter. MC equals both AVC and ATC when these curves attain their minimum values. Furthermore, MC lies below both AVC and ATC over the range in which the curves decline; it lies above them when they are rising.

8.4 LONG-RUN THEORY OF COST

The conventional definition of the long run given in Chapter 6 and elsewhere is "a period of time of such length that all inputs are

[6] The equality of MC and AVC when the latter is a minimum follows from the relationship of MC and MP, AVC and AP. *Explain why.*

[7] Using the notation of the previous footnotes, these two points can easily be proved mathematically. $AVC = g(q)/q$; hence AVC reaches its minimum when $d(AVC)/dq = 0$. Performing this operation, we have

$$\frac{d\left(\frac{g(q)}{q}\right)}{dq} = \frac{qg'(q) - g(q)}{q^2} = \frac{1}{q}\left(g'(q) - \frac{g(q)}{q}\right) = 0.$$

Since $q > 0$, the expression in parentheses must be zero, or $g'(q) = g(q)/q$ at the point of minimum AVC.

Similarly, $ATC = A/q + g(q)/q$. Equating its first derivative with zero, one obtains

$$\frac{d\left(\frac{A}{q} + \frac{g(q)}{q}\right)}{dq} = -\frac{A}{q^2} + \frac{qg'(q) - g(q)}{q^2} = 0.$$

Thus $qg'(q) = A + g(q)$, or $g'(q) = A/q + g(q)/q$ at the point of minimum ATC.

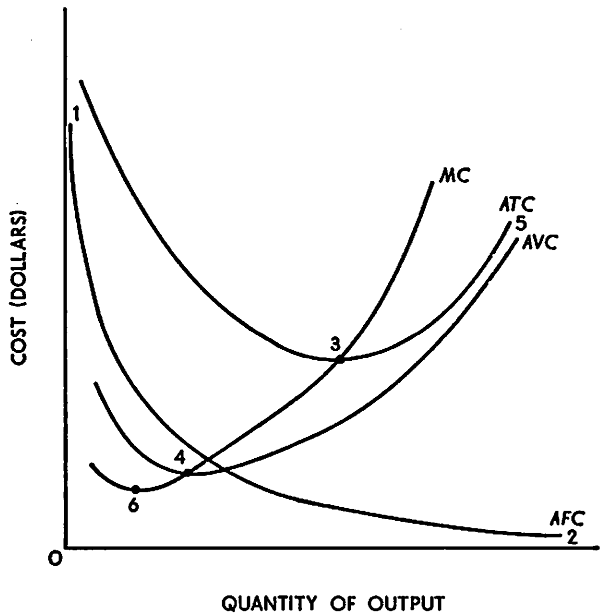

FIGURE 8.3.7
TYPICAL SET OF COST CURVES

variable." Another aspect of the long run has also been stressed—an aspect that is, perhaps, the most important of all. The long run is a *planning horizon.* All production, indeed all economic activity, takes place in the short run. The "long run" refers to the fact that economic agents—consumers and entrepreneurs—can plan ahead and choose many aspects of the "short run" in which they will operate in the future. Thus in a sense, the long run consists of all possible short-run situations among which an economic agent may choose.

As an example, *before* an investment is made an entrepreneur is in a long-run situation. He may select any one of a wide variety of different investments. After the investment decision is made and funds congealed in fixed capital equipment, the entrepreneur operates under short-run conditions. Thus perhaps the best distinction is to say that an economic agent *operates* in the short run and *plans* in the long run.

8.4.a—Short Run and the Long

To begin with a highly simplified situation, suppose technology is such that plants in a certain industry can have only three different sizes. That is, the fixed capital equipment comprising the "plant" is available in only three sizes—small, medium, and large.

The plant of smallest size gives rise to the short-run average cost curve labeled SAC_1 in Figure 8.4.1. The medium-size plant has short-run average cost given by SAC_2; the large plant has an average cost given by SAC_3. In the long run, an entrepreneur has to choose among the three investment alternatives represented by the three short-run average cost curves. If he expects his most profitable output to be Ox_1, he will select the smallest plant. If he expects Ox_2 to be most profitable, he will select the medium plant; and so forth. Such decisions would be made because the entrepreneur chooses the plant capable of producing the expected output at the lowest unit cost.

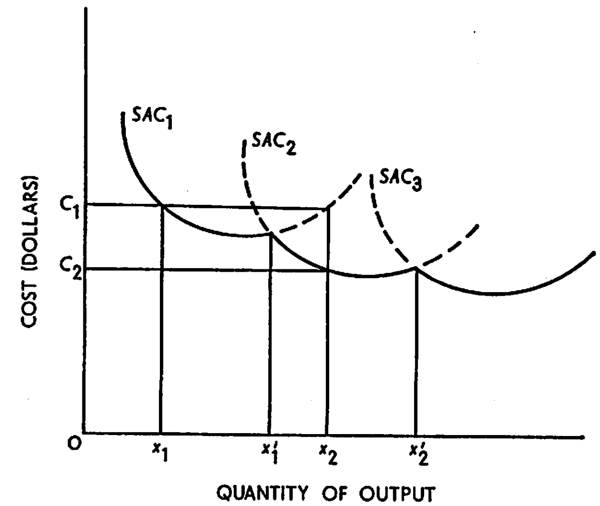

FIGURE 8.4.1
SHORT-RUN AVERAGE COST CURVES FOR PLANTS OF DIFFERENT SIZE

If he expected to produce either Ox_1' or Ox_2', his decision would be more difficult. At each of these points, two plants give rise to the same average cost. An entrepreneur might choose the smaller plant because it requires a smaller investment. On the other hand, he might select the larger plant in order to meet a possible expansion of demand. In these two examples the entrepreneur's decision would be based upon considerations other than least-cost output.

In all other cases his decision is determined by unit cost. Suppose he expects to produce output Ox_1. He accordingly builds the plant represented by SAC_1. Now suppose he actually finds it desirable to produce Ox_2 units. He can do this with his plant, at an average cost of Oc_1 per unit. In the short run this is all he can do; he has no option. But he can plan for the future. Once his old plant has "worn out" he can replace it with a new one—and it will be a medium-size plant because the output Ox_2 can be produced for an average cost of Oc_2 per unit, substantially less than with the small plant.

In the short run, an entrepreneur must operate with SAC_1, SAC_2, or SAC_3. But in the long run, he can plan to build the plant whose size leads to the least average cost for any given output. Thus as a planning device he regards the heavily shaded curve as his long-run average cost curve because this curve shows the least unit cost of producing each possible output.

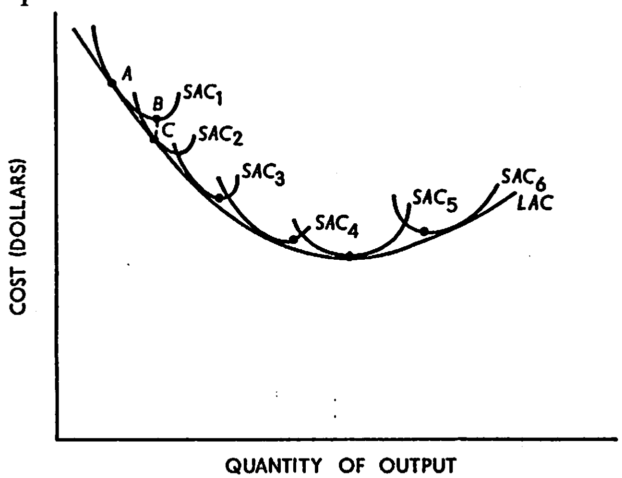

FIGURE 8.4.2

LONG-RUN AVERAGE COST CURVE

8.4.b—Long-Run Average Cost Curve

This illustration is, as we said, highly simplified. An entrepreneur is normally faced with a choice among quite a wide variety of plants. In Figure 8.4.2, six short-run average cost curves are shown; but this is really far from enough. Many curves could be drawn between each of those shown. These six plants are only representative of the wide variety that could be constructed.

These many curves, just as the three in subsection 8.4.a, generate *LAC* as a planning device. Suppose an entrepreneur thinks the output associated with point *A* will be most profitable. He will build the plant represented by SAC_1 because it will enable him to produce this output at the least possible cost per unit. With the plant whose short-run average cost is given by SAC_1, unit cost could be reduced by expanding output to the amount associated with point *B*, the minimum point on SAC_1. If demand conditions were suddenly changed so this larger

output were desirable, the entrepreneur could easily expand—and he would add to his profitability by reducing unit cost. Nevertheless, when setting his future plans the entrepreneur would decide to construct the plant represented by SAC_2 because he could reduce unit costs even more. He would operate at point *C*, thereby lowering his unit cost from the level at point *B* on SAC_1.

The long-run planning curve, *LAC*, is a locus of points representing the least unit cost of producing the corresponding output. The entrepreneur determines the size of plant by reference to this curve. He selects that short-run plant which yields the least unit cost of producing the volume of output he anticipates.

8.4.c—Long-Run Marginal Cost

A marginal cost curve may be constructed for the planning curve or the long-run average cost curve. This is illustrated in Figure 8.4.3.

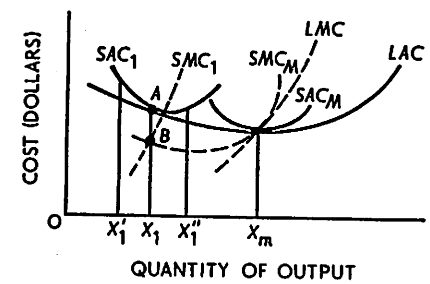

FIGURE 8.4.3

LONG-RUN AND SHORT-RUN MARGINAL COST

Consider the plant represented by the short-run average cost curve SAC_1 with the associated short-run marginal cost curve SMC_1. At point *A*, corresponding to output Ox_1, *SAC* and *LAC* are equal. Hence short-run and long-run total cost are also equal.

For smaller outputs, such as Ox_1', SAC_1 exceeds *LAC*, so that short-run total cost is greater than long-run total cost. Thus for an expansion of output toward Ox_1, long-run marginal cost—whatever it may be—must exceed the known short-run marginal cost. That is, we have moved from a point where short-run total cost exceeds long-run total cost to a point where they are equal. The addition to total cost, or marginal cost, must consequently be smaller for the short-run curve than for the long-run curve. Therefore *LMC* is greater than *SMC* to the left of point *A*.

For an expansion of output from Ox_1 to Ox_1'', the opposite situation holds. SAC_1 is greater than *LAC* at Ox_1'', so short-run total cost exceeds long-run total cost at this point. Now we have moved from a point where short-run and long-run total cost are equal (Ox_1) to a point where short-run total cost exceeds long-run total cost (Ox_1''). Therefore, the addition to total cost, or marginal cost, must be greater for the short-run curve than for the long-run curve. Whatever *LMC* might be, it must be less than SMC_1 to the right of Ox_1.

But now we have the information to find one point on the *LMC* curve. *LMC* must exceed SMC_1 to the left of Ox_1 and it must be less than SMC_1 to the right of Ox_1. Therefore, *LMC* must equal SMC_1 at output Ox_1. This gives us point *B* on the *LMC* curve. To find all the other points, this process is repeated. Take the next short-run average cost curve, together with its known short-run marginal cost. *LMC* must equal this *SMC* for the output at which the *SAC* curve is tangent to *LAC*. Performing this process for all plant sizes generates the *LMC* curve.

There is one important point to notice. *LMC* intersects *LAC* when the latter is at its minimum point. There will be one, and only one, short-run plant size whose minimum short-run average cost coincides with minimum long-run average cost. This plant is represented by SAC_M and SMC_M in Figure 8.4.3. SMC_M equals SAC_M at the minimum point on the latter curve. SAC_M is tangent to *LAC* at their common minimum; and as we have shown, *LMC* equals *SMC* at the point where *SAC* and *LAC* are tangent. Therefore, *LMC* must pass through the minimum point on *LAC*.

8.5 SHAPE OF LAC

The long-run average cost curve and the short-run average cost curve are alike in that each has been drawn with a U shape. The reasons for this shape, however, are quite different. *SAC* is U-shaped because the decline in average fixed cost is ultimately more than offset by the rise in average variable cost—the latter occurring because average product reaches a maximum and declines. But this has nothing at all to do with the curvature of *LAC*. Economies and diseconomies of scale are the factors governing the shape of *LAC*.[8]

8.5.a—Economies of Scale

As the size of plant and the scale of operation become larger, considering expansion from the smallest possible plant, certain economies of scale are usually realized. That is, after adjusting *all* inputs optimally the unit cost of production can be reduced by increasing the size of plant.

Adam Smith gave one of the outstanding reasons for this: specialization and division of labor. When the number of workers is expanded, fixed inputs remaining fixed, the opportunities for specialization and division of labor are rapidly exhausted. The marginal product curve rises, to be sure; but not for long. It very quickly reaches its maximum and declines thereafter. When workers and equipment are expanded together, however, very substantial gains may be reaped by division of jobs and the specialization of workers in one job or another.

Proficiency is gained by concentration of effort. If a plant is very small and employs only a small number of workers, each worker will usually have to perform several different jobs in the production process. In doing so he is likely to have to move about the plant, change tools, and so on. Not only are workers not highly specialized but a part of their worktime is consumed in moving about and changing tools. Thus important savings may be realized by expanding the scale of operation. A larger plant with a larger work force may permit each worker to specialize in one job, gaining proficiency and obviating time-consuming interchanges of location and equipment. There naturally will be corresponding reductions in the unit cost of production.

Technological factors constitute a second force contributing to economies of scale. If several different machines, each with a different rate of output, are required in a production process, the operation may have to be quite sizeable to permit proper "meshing" of equipment. Suppose only two types of machines are required, one which produces and one which packages the product. If the first machine can produce 30,000 units per day and the second can package 45,000, output will have to be 90,000 per day in order to utilize fully the capacity of each machine.

[8] As explained in detail in the first footnote, the cost function, and hence the cost curves, are derived from the production function. Yet the shape of the average cost curve depends, among other things, upon the nature and cost of the fixed inputs. On the other hand, the shape of the long-run average cost curve depends exclusively upon the production function. The purpose of this footnote is to dispel any misunderstanding that might arise in regard to production functions homogeneous of degree one.

First, note that the short-run average cost curve is U-shaped, regardless of the form of the production function. By our analytical fiction one or more inputs are *fixed*, which gives rise to fixed or "overhead" cost; and the act of "spreading overhead," other things equal, must impart a downward slope to the short-run average cost curve, at least over a portion of its range.

Second, in the long run no inputs are fixed. If all input prices are constant and if the production function is homogeneous of degree one, the long-run average cost curve is a straight line parallel to the abscissa. If the production function is homogeneous of degree greater (less) than one, the long-run average cost curve is a monotonically decreasing (increasing) curve. These statements do not hold if input prices vary with their rates of utilization.

Another technological element is the fact that the cost of purchasing and installing larger machines is usually proportionately less than the cost of smaller machines. For example, a printing press that can run 200,000 papers per day does not cost ten times as much as one that can run 20,000 per day—nor does it require ten times as much building space, ten times as many men to work it, and so forth. Again, expanding size tends to reduce the unit cost of production.

A final technological element is perhaps the most important of all: as the scale of operation expands there is usually a *qualitative*, as well as a quantitative, change in equipment. Consider ditch digging. The smallest scale of operation is one man and one shovel. But as the scale expands beyond a certain point one does not simply continue to add men and shovels. Shovels and most workers are replaced by a modern ditch-digging machine. In like manner, expansion of scale normally permits the introduction of various types of automation devices, all of which tend to reduce the unit cost of production.

Thus two broad forces—specialization and division of labor and technological factors—enable producers to reduce unit cost by expanding the scale of operation.[9] These forces give rise to the negatively sloped portion of the long-run average cost curve.

But why should it ever rise? After all possible economies of scale have been realized, why doesn't the curve become horizontal?

8.5.b—Diseconomies of Scale

The rising portion of *LAC* is usually attributed to "diseconomies of scale," which means limitations to efficient management. Managing any business entails controlling and coordinating a wide variety of activities—production, transportation, finance, sales, etc. To perform these managerial functions efficiently the manager must have accurate information; otherwise the essential decision making is done in ignorance.

As the scale of plant expands beyond a certain point, top management necessarily has to delegate responsibility and authority to

[9] This discussion of economies of scale has concentrated upon physical and technological forces. There are financial reasons for economies of scale as well. Large-scale purchasing of raw and processed materials may enable the buyer to obtain more favorable prices (quantity discounts). The same is frequently true of advertising. As another example, financing of large-scale business is normally easier and less expensive; a nationally known business has access to organized security markets, so it may place its bonds and stocks on a more favorable basis. Bank loans also usually come easier and at lower interest rates to large, well-known corporations.

These are but examples of many potential economies of scale attributable to financial considerations. For a more detailed discussion, see William G. Husband and James C. Dockeray, *Modern Corporation Finance*, 6th ed. (Homewood, Ill.: R. D. Irwin, Inc., 1966.)

lower echelon employees. Contact with the daily routine of operation tends to be lost and efficiency of operation to decline. Red tape and paper work expand; management is generally not as efficient. This increases the cost of performing the managerial function and, of course, the unit cost of production.

It is very difficult to determine just when diseconomies of scale set in and when they become strong enough to outweigh the economies of scale. In businesses where economies of scale are negligible, diseconomies may soon become of paramount importance, causing *LAC* to turn up at a relatively small volume of output. Panel a, Figure 8.5.1, shows a long-run average cost curve for a firm of this type. In other cases, economies of scale are extremely important. Even after the efficiency of management begins to decline technological economies of scale may offset the diseconomies over a wide range of output. Thus the *LAC* curve may not turn upward until a very large volume of output is attained. This case, typified by the so-called natural monopolies, is illustrated in panel b, Figure 8.5.1.

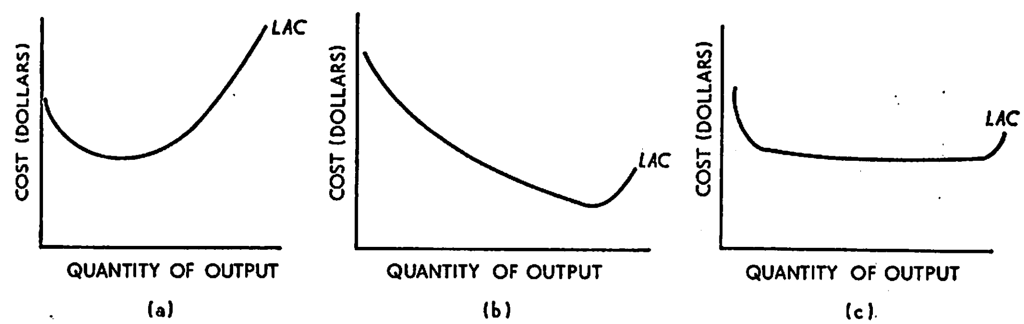

FIGURE 8.5.1
VARIOUS SHAPES OF LAC

In many actual situations, however, neither of these extremes describes the behavior of *LAC*. A very modest scale of operation may enable a firm to capture all of the economies of scale; however, diseconomies may not be incurred until the volume of output is very great. In this case, *LAC* would have a long horizontal section, as shown in panel c. Many economists and businessmen feel that this type of *LAC* curve describes most production processes in the American economy.

8.6 CONCLUSION

The physical conditions of production and resource prices jointly establish the cost of production. This is very important to individual business firms and to the economy as a whole. But it is only half the story. Cost gives one aspect of economic activity: to the individual businessman it comprises his obligations to pay out funds; to the society as a whole it represents the resources that must be sacrificed to obtain a given commodity. The other aspect is revenue or demand. To the individual businessman revenue constitutes the flow of funds from which his obligations may be met. To society, demand represents the social valuation placed on a commodity.

Thus both demand and cost must be taken into consideration. It is to the demand side that we turn in Part III.

PROBLEM

Return to the problem at the end of Chapter 6. Total product is given; and you have computed average and marginal product. You are also given the following information:

1. Total fixed cost (total price of fixed inputs) is $220 per period.
2. Units of the variable input cost $100 per unit per period.

Using this information, complete the following table:

Units of Variable Input	*Total Product*	*Average Product*	*Marginal Product*	*Total Fixed Cost*	*Total Variable Cost*	*Total Cost*	*Average Fixed Cost*	*Average Variable Cost*	*Average Total Cost*	*Marginal Cost*
1	100									
2	250									
3	410									
4	560									
5	700									
6	830									
7	945									
8	1050									
9	1146									
10	1234									
11	1314									
12	1384									
13	1444									
14	1494									
15	1534									
16	1564									
17	1584									
18	1594									

I. Graph the total cost curves on one sheet and the average and marginal curves on another.

II. By reference to table and graph, answer the following questions.

1. When marginal product is increasing, what is happening to:
   a) Marginal cost?
   b) Average variable cost?
2. When marginal cost first begins to fall, does average variable cost begin to rise?
3. What is the relation between marginal cost and average variable cost when marginal and average product are equal?
4. What is happening to average variable cost while average product is increasing?
5. Where is average variable cost when average product is at its maximum? What happens to average variable cost after this point?
6. What happens to marginal cost after the point where it equals average variable cost?
   a) How does it compare with average variable cost thereafter?
   b) What is happening to marginal product thereafter?
   c) How does marginal product compare with average product thereafter?
7. What happens to total fixed cost as output is increased?
8. What happens to average fixed cost as:
   a) Marginal product increases?
   b) Marginal cost decreases?
   c) Marginal product decreases?
   d) Marginal cost increases?
   e) Average variable cost increases?
9. How long does average fixed cost decrease?
10. What happens to average total cost as:
    a) Marginal product increases?
    b) Marginal cost decreases?
    c) Average product increases?
    d) Average variable cost decreases?
11. Does average cost increase:
    a) As soon as the point of diminishing marginal returns is passed?
    b) As soon as the point of diminishing average returns is passed?
12. When does average cost increase? Answer this in terms of:
    a) The relation of average cost to marginal cost.
    b) The relation between the increase in average variable cost and the decrease in average fixed cost.

SUGGESTED READING

1. CLARK, J. M. *The Economics of Overhead Cost*, chaps. 4–6. Chicago: University of Chicago Press, 1923.
2. VINER, JACOB. "Cost Curves and Supply Curves," *Zeitschrift für National-*

ökonomie, III (1931), pp. 23–46. Reprinted in AEA, *Readings in Price Theory,* pp. 198–232. Homewood: Richard D. Irwin, Inc., 1952.

3. HENDERSON, JAMES M. AND QUANDT, RICHARD E. *Microeconomic Theory: A Mathematical Approach,* pp. 55–62. New York: McGraw-Hill Book Co., Inc., 1958. [Elementary math required.]

ADVANCED READING, PART II

I. *The Theory of Production, General*

1. ARROW, KENNETH J.; CHENERY, HOLLIS B.; MINHAS, BAGICHA; AND SOLOW, ROBERT M. "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics,* XLIII (1961), pp. 225–50.
2. BORTS, GEORGE H. AND MISHAN, E. J. "Exploring the 'Uneconomic Region' of the Production Function," *Review of Economic Studies,* XXIX (1962), pp. 300–12.
3. CARLSON, SUNE. *A Study on the Pure Theory of Production,* Stockholm Economic Studies, No. 9. London: P. S. King & Sons, Ltd., 1939.
4. CASSELS, J. M. "On the Law of Variable Proportions," *Explorations in Economics,* pp. 223–36. New York: McGraw-Hill Book Co., 1936.
5. FERGUSON, C. E. "Transformation Curve in Production Theory: A Pedagogical Note," *Southern Economic Journal,* XXIX (1962), pp. 96–102.
6. JOHANSEN, LEIF. "Substitution versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis," *Econometrica,* XXVII (1959), pp. 157–76.
7. MACHLUP, FRITZ. "On the Meaning of the Marginal Product," *Explorations in Economics,* pp. 250–63. New York: McGraw-Hill Book Co., 1936.
8. ROBINSON, JOAN, "The Production Function," *Economic Journal,* LXV (1955), pp. 67–71.
9. SAMUELSON, PAUL A. *Foundations of Economic Analysis,* pp. 57–89. Cambridge: Harvard University Press, 1947.
10. SAMUELSON, PAUL A. "Parable and Realism in Capital Theory: The Surrogate Production Function," *Review of Economic Studies,* XXIX (1962), pp. 193–206.
11. SHEPHARD, RONALD W. *Cost and Production Functions.* Princeton: Princeton University Press, 1953.
12. SOLOW, ROBERT M. "The Production Function and the Theory of Capital," *Review of Economic Studies,* XXIII (1955–56), pp. 101–108.
13. SOLOW, ROBERT M. "Substitution and Fixed Proportions in the Theory of Capital," *Review of Economic Studies,* XXIX (1962), pp. 207–18.
14. STIGLER, GEORGE J. *Production and Distribution Theories.* New York: Macmillan Co., Inc., 1946.

15. WALTERS, A. A. "Production and Cost Functions: An Econometric Survey," *Econometrica,* XXXI (1963), pp. 1–66, with extensive bibliography.

II. *The Classification of Inventions*

1. HICKS, JOHN R. *The Theory of Wages.* London: Macmillan & Co., Ltd., 1932.
2. ROBINSON, JOAN. "The Classification of Inventions," *Review of Economic Studies,* V (1937–38), pp. 139–42.
3. SEEBER, NORTON C. "On the Classification of Inventions," *Southern Economic Journal,* XXVIII (1962), pp. 365–71.

III. *The Elasticity of Substitution*

1. HICKS, JOHN R. *The Theory of Wages.* London: Macmillan & Co., Ltd., 1932.
2. LERNER, A. P. "The Diagrammatical Representation of the Elasticity of Substitution," *Review of Economic Studies,* I (1933–34), pp. 68–71.
3. MACHLUP, FRITZ. "The Commonsense of the Elasticity of Substitution," *Review of Economic Studies,* II (1934–35), pp. 202–13; with notes by MILTON FRIEDMAN, JOAN ROBINSON, A. P. LERNER, AND FRITZ MACHLUP, *Review of Economic Studies,* III (1935–36), pp. 147–52.
4. PIGOU, A. C. "The Elasticity of Substitution," *Economic Journal,* XLIV (1934), pp. 232–41.

IV. *The Theory of Cost*

The classic reference is JACOB VINER. "Cost Curves and Supply Curves," *Zeitschrift für Nationalökonomie und Statistik,* III (1931), pp. 23–46. In addition, see citations 9, 11, and 15 under I above.