CHAPTER SIX

THEORY OF PRODUCTION and COSTS

6.1 Some Basic Concepts of Production Theory

6.1.1 Production Defined

This chapter examines the theory of producer behavior which is the supply side of the market.
In this theory of production, firms organize/combine resources or inputs such as labor, capital, land and entrepreneurship and so on, to produce final goods and services. Thus, production refers to the process of converting inputs into outputs. In other words, production is the creation of goods and services from inputs or resources, such as labor, machinery and other capital equipment, land, raw materials, and so on.

Examples: when a company such as Ford makes a truck or car or when Exxon refines a gallon of gasoline, the activity is production. But production goes much further than that. A doctor produces medical services, a teacher produces education, and a singer produces entertainment. So production involves services as well as making the goods people buy. Production is also undertaken by governments and non-profit organizations. A city police department produces protection, a public school produces education, and a hospital produces health care.

The following points are worth noting about the notion of production:

- **Production may not involve physical conversion of raw materials into tangible goods.** Some kinds of production may involve an intangible input to produce an intangible output. For example, in the production of legal, medical, social and consultancy services both input and output are intangible. Lawyers, doctors, social workers, consultants, hairdressers, musicians, orchestra players are all engaged in producing intangible goods.
- **Production process may take a variety of forms other than manufacturing.** For example, *transporting* a commodity from one place to another where it can be used is production. Such activities too are '**production**'. Storing a commodity for future sale or consumption is also '**production**'. *Wholesaling, retailing, packaging, assembling* are all productive activities. These activities are just as good examples of production as manufacturing.

6.1.2 An Input

An input is a good or service that goes into the process of production. In other words, an input is simply anything which the firm buys for use in its production or other process for sale.

Inputs can be classified into

1. Labor (including entrepreneurial talent);
2. Capital;
3. Land or natural resources;
4. Raw materials;
5. Times

Inputs are also classified as (*i*) fixed inputs (*ii*) variable inputs. A **fixed input** is one for which the level of usage cannot readily be changed. To be sure, no input is ever absolutely fixed, no matter how short the period of time under consideration. However, the cost of immediately varying the use of an input may be so great that, for all practical purposes, the input is fixed. For example, buildings, major pieces of machinery, and managerial personnel are inputs that generally cannot be rapidly augmented or diminished. A **variable input**, on the other hand, is one for which the level of usage may be changed quite readily in response to desired changes in output. Many types of labor services as well as certain raw and processed materials would be this category.

6.1.3 An Output

On the other hand, is any good or service that comes out of production process.

- The output of a firm can be a final commodity (such as home automobile) or an intermediate product, such as semiconductors (which are used in the production of computers and other goods).
- The output can be a service rather than a good. Examples of services are education, medicine, banking, communication, transportation, and many others.

6.1.4 Short-run and Long-run

The short-run refers to a period of time in which the supply of certain inputs (example, plant, building, and machines, etc.) is fixed or inelastic. In the short-run, therefore, production of a commodity can be increased by increasing the use of only variable inputs, like labor and raw materials. Long-run refers to a period of time in which the supply of all the inputs is elastic, but not enough to permit a change in technology. That is, in the long run, all the inputs are variable. Therefore, in the long-run production of a commodity can be increased by employing more or both, variable and fixed, inputs. To sum up, it can be said that the firm operates in the short-run and plans increases or reductions in its scale of operation in the long run.

6.2 Production Function

A production function is the link between levels of input usage and attainable levels of output. That is, the production formally describes the relation between physical rates of output and physical rates of input usage. A production function is a schedule (or table or mathematical equation) showing the maximum amount of output that can be produced from any specified set of inputs, given the existing technology or state of the art of production.

$Q = f(X_1, X_2, ..., X_n)$

For the sake of illustration, let's consider the simple case of production function in which only two inputs are involved in the production process (usually labor and capital).

$Q = f(L, K)$

Where, Q = Quantity produced;

L = Labor

K = Capital

However, we must stress that the principles to be developed apply to situations with more than two points and, as well, to inputs other than capital and labor.

6.2.1 Technical Efficiency and Economic Efficiency

Technical efficiency is achieved when the maximum possible amount of output is being produced with a given combination of inputs. The definition of a production function assumes that technical efficiency is being achieved because the production function gives the *maximum* output level that can be achieved for any particular combination of inputs. Thus, technical efficiency is implied by the production function.

Economic efficiency is achieved when the firm is producing a given amount of output at the lowest possible cost. One should be careful about labeling a particular production process inefficient. Certainly a process would be technically inefficient if another process can produce the same amount of output using less or one or more inputs and the same amounts of all others. If, however, the second process uses less of some inputs but more of others, the economically efficient method of producing a given level of output depends on the prices of the inputs. Even when both are technically efficient, one process might cost less- be economically efficient- under one set of input prices while the other may be economically efficient at other input prices.

6.3 Production Function in the Short Run (Optimization in the Case of one Variable Input)

The supply of the fixed inputs remained unchanged (i.e., supply inelastic) so that there are only one variable input and one fixed input. Thus, production of a commodity increased by using more of the variable inputs. In the short run, the firm faces a decision problem on how much of variable inputs should be employed with a given employment of fixed inputs. To address this issue, one needs a clear understanding of relationship among the total, average, and marginal productivity of factors.

Total Product (TP): the total amount of output produced as a result of employing all the inputs.

$TP = Q = f(L, \overline{K}) = f(L)$

Suppose a firm with a production function of the form *Q = f (L, K)* can, in the long run, choose levels of both labor and capital between 0 and 10 units. A production function giving the maximum amount of output that can be produced from every possible combination of labor and capital is shown in Table 6.1. For example, from the table, 4 units of labor combined with 3 units of capital can produce a maximum of 325 units of output; 6 labor and 6 capital can produce a maximum of 655 units of output; and so on. Note that with 0 capital, no output can be produced regardless of the level of labor usage. Likewise, with 0 labor, there can be no output.

Once the level of capital is fixed, the firm is in the short run, and output can be changed only by varying the amount of labor employed. Assume now that the capital stock is fixed at 2 units of capital. The firm is in the short run and can vary output only by varying the usage of labor (the variable input). The column in Table 6.1 under 2 units of capital gives the total output, or total product of labor, for 0 through 10 workers. This column, for which K = 2, represents the short-run production function when capital is fixed at 2 units.

These total products are reproduced in column 2 of Table 6.1 for each level of labor usage in column 1. Thus, columns 1 and 2 in Table 6.2 define a production function of the form $Q = f(L, \overline{K})$, where $\overline{K}$ = 2. In this example, total product (*Q*) rises with increases in labor up to a point (9 workers) and then declines. While total product does eventually *fall* as more workers if they knew output would fall. A manager can hire either 8 workers or 10 workers to produce 314 units of output. Obviously, the economically efficient amount of labor to hire to produce 314 units is eight workers.

Average Product (AP): the total product per unit of variable input.

$AP = TP / L = Q / L$

In our example, average product, shown in column 3, first rises, reaches a maximum at 56.7, then declines thereafter.

Marginal Product (MP): is the additional output attributable to using one additional worker with the use of all other inputs fixed (in this case, at 2 units of capital). That is,

$$MP = \Delta TP / \Delta L = \Delta Q / \Delta L$$

where Δ means "the change in." The marginal product schedule associated with the production function in Table 6.2 is shown in column 4 of the table. Because no output can be produced with 0 workers, the first worker adds 52 units of output; the second adds 60 units (i.e., increase output from 52 to 112); and so on.

Units of capital (K)

Units of labor (L)

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0	0	0	0	0	0	0	0	0	0	0
1	0	25	52	74	90	100	108	114	118	120	121
2	0	55	112	162	198	224	242	252	258	262	264
3	0	83	170	247	303	342	369	384	394	400	403
4	0	108	220	325	400	453	488	511	527	535	540
5	0	125	258	390	478	543	590	631	653	663	670
6	0	137	286	425	523	598	655	704	732	744	753
7	0	141	304	453	559	643	708	766	800	814	825
8	0	143	314	474	587	679	753	818	857	873	885
9	0	141	318	488	609	708	789	861	905	922	935
10	0	137	314	492	617	722	809	887	935	953	967

Table 6.1 A Production Function

(1) Number of workers (*L*)	(2) Total product (*Q*)	(3) Average product ($AP = Q/L$)	(4) Marginal product ($MP = \Delta Q/\Delta L$)
0	0	-	-
1	52	52	52
2	112	56	60
3	170	56.7	58
4	220	55	50
5	258	51.6	38
6	286	47.7	28
7	304	43.4	18
8	314	39.3	10
9	318	35.3	4
10	314	31.4	-4

Table 6.2 Total, Average, and Marginal Products of Labor (with capital fixed at 2 units)

Note that increasing the amount of labor from 9 to 10 actually decreases output from 318 to 314. Thus, the marginal product of the 10th worker is negative. In this example, marginal product first increases as the amount of labor increases, then decreases, and finally becomes negative. This is a pattern frequently assumed in economic analysis.

Figure 6.1 shows graphically the relations among the total, average, and marginal products set forth in Table 6.2. In Panel A, total product increases up to 9 workers, then decreases. Panel B incorporates a common assumption made in production theory: Average product first rises then falls. When average product is increasing, marginal product is greater than average product (after the first worker, at which they are equal). When average product is decreasing, marginal product is less than average product. This result is not peculiar to this production function; it occurs for any production function for which average product first increases then decreases.

An example might help demonstrate that for any average and marginal schedule, the average must increase when the marginal is above the average and decrease when the marginal is below the average. If you have taken two tests and made grades of 70 and 80, your average grade is 75. If your third test grade is higher than 75, the marginal grade is above the average the average, so your average grade increases. Conversely, if your third grade is less than 75- the marginal grade is below the average- your average falls. In production theory, if each additional worker adds more than the average, average product rises; if each additional worker adds less than the average, average product falls.

As shown in Figure 6.1, marginal product first increases then decreases, becoming negative after 9 workers. The maximum marginal product occurs before the maximum average product is attained. When marginal product is *increasing*, total product increases at an *increasing* rate. When marginal product begins to increase at a *decreasing* rate. When marginal product becomes negative (10 workers), total product declines.

We can derive the following important relationship between TP, AP and MP:

$$AP \uparrow \textit{ when } MP > AP$$
$$AP \downarrow \textit{ when } MP < AP$$
$$AP \textit{ reaches } \max imum \textit{ when } AP = MP$$
$$MP = 0 \textit{ when } TP \textit{ reaches } \max imum$$
$$MP \textit{ reaches } \max imum \textit{ before } AP$$

MATHEMATICAL RELATIONSHIP BETWEEN *APL* AND *MPL*

The mathematical relationship between the average product of labor (or any average concept) and the marginal product of labor (or any related marginal concept) may be illustrated by the use of optimization analysis.

Consider again the definition of the average product of labor

$$AP_L = \frac{TP_L}{L} = \frac{Q_L}{L} = \frac{f(K_0, L)}{L}$$

Taking the first derivative with respect to labor and setting the results equal to zero yields

$$\frac{\partial AP_L}{\partial L} = \frac{L(\partial Q_L\ \partial L) - Q_L\ (\partial L\ \partial L)}{L^2} = 0$$

$$\frac{\partial AP_L}{\partial L} = \frac{L(MP_L) - Q_L\ (1)}{L^2} = 0$$

$$\frac{MPL}{L} - \frac{QL}{L2} = \frac{MPL}{L} - \frac{QL}{L} \times \frac{1}{L} = \frac{1}{L}(MP_L - AP_L) = 0$$

$\frac{1}{L}(MP_L-AP_L) = 0$ when AP is maximum, then $MP_L = AP_L$

$\frac{1}{L}(MP_L-AP_L) = +ve$ when AP is rising, then $MP_L = AP_L + (+ve)$

Meaning $MP_L > AP_L$

$\frac{1}{L}(MP_L-AP_L) = -ve$ when AP is declining, then $MP_L = AP_L - (+ve)$

Meaning $MP_L < AP_L$

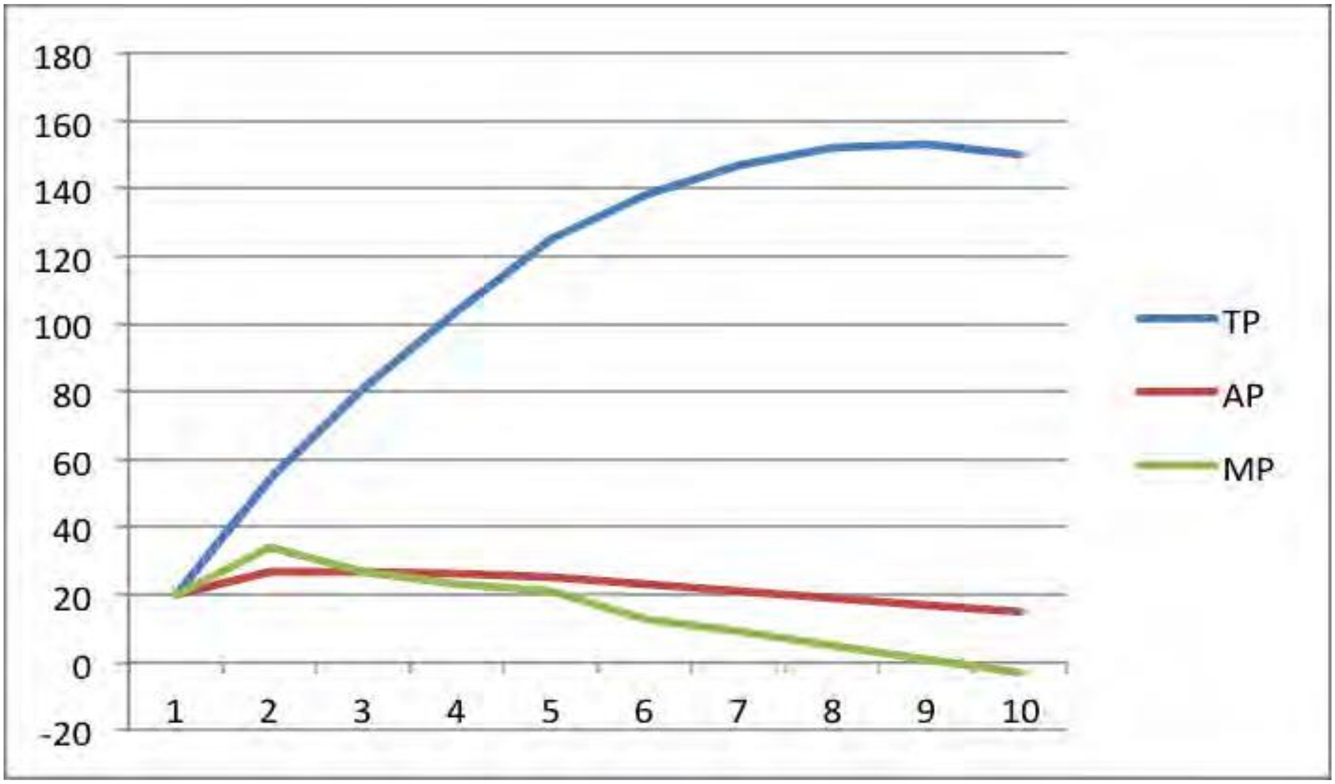
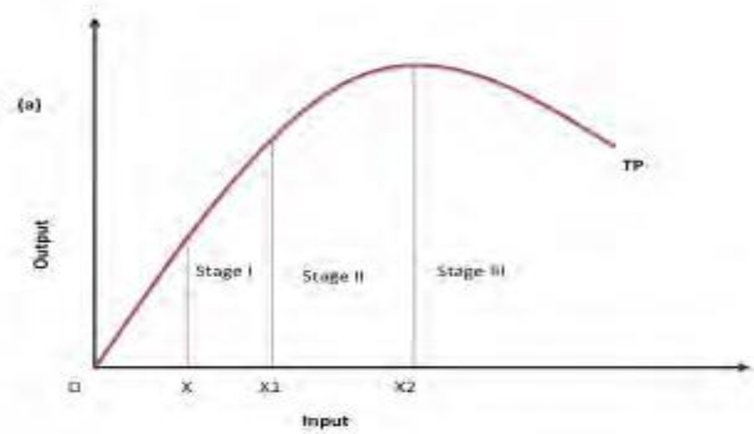

180
160
140
120
100
80
60
40
20
0
-20
1
2
3
4
5
6
7
8
9
10
TP
AP
MP

Figure 6.1 A


(a)
Output
TP
Stage I
Stage II
Stage III
X
X1
X2
Input

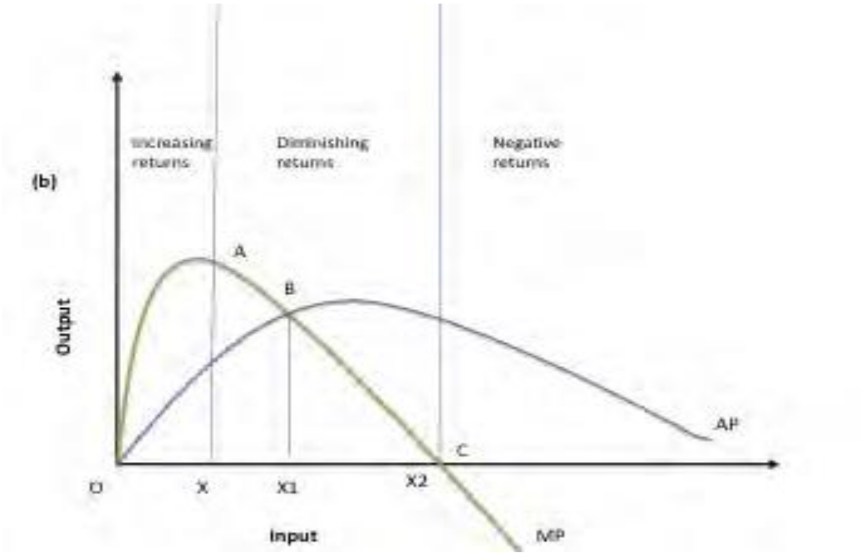


Figure 6.1 B

6.3.1 The Law of Diminishing Marginal Returns (LDMR)

When an increasing amount of variable inputs is combined with a specified amount of fixed input in the short run, the resulting increase in output diminishes. In other words, as more and more units of labor is added to a given fixed input, the marginal product of labor diminishes after some point. When the amount of the variable input is small relative to the fixed inputs, more intensive utilization of fixed inputs by variable inputs may initially increase the marginal product of the variable input as this input is increased. Nonetheless, a point is reached beyond which an increase in the use of the variable input yields progressively less additional output. Each additional units has, on average, fewer units of the fixed inputs with which to work.

Thus, the *point* here is that as a manager we need to know the levels of inputs where the *LDMR* exists and all the above stated relationship between TP, AP, and MP.

When the first order derivative of production function is zero, total output reaches maximum
When the second order derivative of production function is zero, that level of unit (labor) is the point where DMR starts to operate.
The optimization rule states that the marginal revenue product of an input (in this case labor) is equal to the price of the input. $MRP_L = P_L$

$MP_L * P_L = MRP_L$

Marginal cost of labor means price of each additional labor

$M\pi L = MRPL - MCL$

At optimum level $M\pi L = 0$ so

$0 = MRPL - MC_L$, $MRPL = MC_L = P_L$

Labor should be increased until the marginal revenue product equals to the marginal cost of labor. Increasing the labor force any further will be unprofitable.

6.4 Production Function in the Long Run (Optimization in the Case of Multiple Variable Inputs)

In the long run all inputs are variable inputs and thus the firm can produce a certain output by using more and more of these variable inputs (labor and capital).

$Q = f(L, K)$

Managers are, therefore, trying to find and exploit opportunities to change the fixed inputs in the short run into variable inputs by adding new products: changing the type and amount of inputs, restructuring the size of the firm and the production capacity.

6.4.1 Production Isoquants and Isocosts

The terms '*isoquant*' has been derived from the Greek word '*iso*' meaning '*equal*' and Latin word *quantus* meaning '*quantity*'. The '*isoquant curve*' is therefore also known as '*Equal Product Curve*'. Production in the case of two variable inputs (i.e., the long run production) is studied with the help of isoquant and isocost. An **isoquant** is a curve (or locus of points) showing all possible combinations of the inputs physically capable of producing a given (fixed) level of output. Each point on an isoquant is technically efficient; that is, for each combination on the isoquant, the maximum possible output is that associated with the given isoquant. The concept of an isoquant implies that it is possible to substitute some amount of one input for some of the other, say, labor for capital, while keeping output constant.

To understand the concept of an isoquant, return for a moment to Table 6.1 in the preceding section. This table shows the maximum output that can be produced by combining different levels of labor and capital. Now note that several levels of output in this table can be produced in two ways. For example, 108 units of output can be produced using either 6 units of capital and 1 worker or 1 unit of capital and 4 workers. Thus, these two combinations of labor and capital are two points on the isoquant associated with 108 units of output. And if we assumed that labor and capital were continuously divisible, there would be many more combinations of this isoquant.

Other input combinations in Table 6.1 that can produce the same level of output are:

Q = 285: using K = 2, L = 5, or K = 8, L = 2

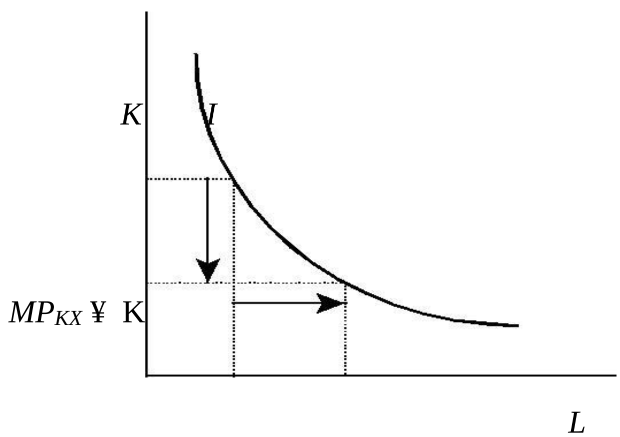
Q = 400: using K = 9, L = 3, or K = 4, L = 4

Q = 453: using K = 5, L = 4, or K = 3, L = 7

Q = 708: using K = 6, L = 7, or K = 5, L = 9

Q = 753: using K = 10, L = 6, or K = 6, L = 8

Figure 6.2 Typical Isoquants



$$\mathbf{MP_K = \frac{\Delta Q}{\Delta K} \longrightarrow \Delta Q = MP_K . \Delta K}$$

$$\mathbf{MPL = \frac{\Delta Q}{\Delta L} \longrightarrow \Delta Q = MP_L . \Delta L}$$

In isoquant output is constant. So ΔQ = ΔQ

$$\mathbf{MP_K . \Delta K = MP_L . \Delta L}$$

$\frac{\Delta K}{\Delta L} = \frac{MPL}{MPK}$ **as it is technical substitution between two products its slope is negative**

$\frac{\Delta K}{\Delta L} = - \frac{MPL}{MPK}$ **this value indicates the rate of capital given up to increase one unit of labor.**

Characteristics of Isoquants

We now set forth the typically assumed characteristics of isoquants when labor, capital, and output are continuously divisible.

- The isoquant curves never cross each other or they do not intersect.
- The higher the isoquant the more the output level in the isoquant map.
- The slope of the isoquant, that is, the marginal rate of technical substitution (MRTS) is negative.
- The isoquant curve is convex to the origin.

6.4.2 Marginal Rate of Technical Substitution (MRTS)

As depicted in Figure 6.2, isoquants slope downward over the relevant range of production. This negative slope indicates that if the firm decreases the amount of capital employed, more labor must be added in order to keep the rate of output constant. Or if labor use is decreased, capital usage must be increased to keep outside constant. Thus, the two inputs can be substituted for one another to maintain a constant level of output. This rate at which one input is substituted for another along an isoquant is called the **marginal rate of technical substitution (*MRTS*)** and is defined as:

$MRTS = -\Delta K / \Delta L$ This can clearly show the rate of substitution of any two ordered pairs of capital and labor.

The minus sign is added to make *MRTS* a positive number, since $\Delta K / \Delta L$, the slope of the isoquant, is negative.

Over the relevant range of production, the marginal rate of technical substitution diminishes. That is, as more and more labor is substituted for capital while holding output constant, the absolute value of ΔK/ΔL decreases. This can be seen in Figure 4.2. If capital is reduced from 50 to 40 (a decrease of 10 units), labor must be increased by 5 units (from 15 to 20) in order to keep the level of output at 100 units. That is, when capital is plentiful relative to labor, the firm can discharge 10 units of capital but must substitute only 5 units of labor in order to keep output at 100. The marginal rate of technical substitution in this case is - ΔK/ΔL = -(-10)/5 = 2, meaning that for every unit of labor added, 2 units of capital can be discharged in order to keep the level of output constant. However, consider a combination where capital is more scarce and labor more plentiful. For example, if capital is decreased from 20 to 10 (again a decrease of 10 units), labor must be increased by 35 units (from 40 to 75) to keep output at 100 units. In this case the *MRTS* is 10/35, indicating that for each unit of labor added, capital can be reduced by slightly more than one-quarter of a unit.

Isocost Curves

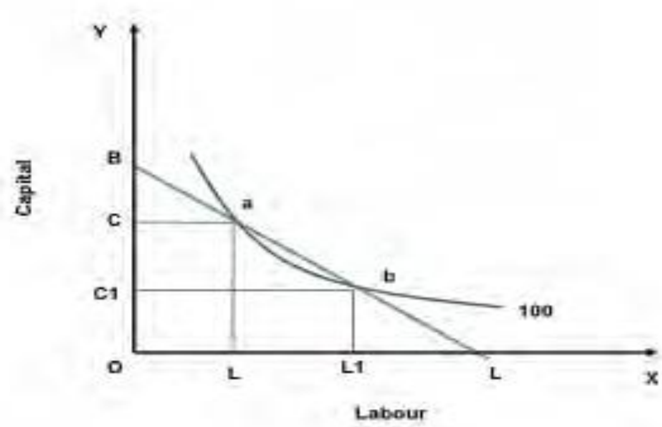
Producers must consider relative input prices in order to find the least-cost combination of inputs to produce a given level of output. An **isocost curve** shows all combinations of inputs that may be purchased for a given level of total expenditure at a given input prices. For most managers, the price of each input is determined in the market for that input by the intersection of the demand for the input and the supply of the input.

The role of managers in the concepts of isoquant and isocost is to reduce the costs. For instance, the managers may be large enough buyers of the resources that they can bargaining with the sellers of the resources to get better prices. Considering the quantities of labor and capital as L and K and their respective prices as wages (w) and rent (r), the total cost, C, is given by: $C = wL + rK$

Slope the isocost line: $\overline{C} = wL + rK \Rightarrow rK = \overline{C} - wL \Rightarrow K = \overline{C}/r - (w/r)L$

The slope of the isocost curve is equal to the negative of the relative input price ratio, - *w/r*. This ratio is important because it tells the manager how much capital must be given up if one or more unit of labor is purchased. As illustrated in Figure 4.3, - *w/r* = - $25/$50 = -1/2. If the manager wishes to purchase 1 more unit of labor at $25, 1/2 unit of capital, which costs $50, must be given up to keep the total cost of the input combination constant. If the price of labor happens to rise to $50 per unit, *r* remaining constant, the slope of the isocost curve is -$50/$50 = -1, which means the manager must give up 1 unit of capital for each additional unit of labor purchased in order to keep total cost constant.

Figure 6.3 An Isocost Curve (w = $25 and r = $50)



6.4.2.1 Determining the Optimal Combination of Inputs

A manager who wishes to maximize profit must first decide how much output to produce and then how to produce that amount at the lowest possible total cost. The optimal combination of inputs or the producer's equilibrium is achieved at point where the slope of isoquant ($MRTS_{LK}$) is equal to the slope of the isocost line ($w/r = P_L/P_K$)

$$MRTS_{LK} = -\Delta K / \Delta L = MP_L / MP_k = P_L / P_K = w / r$$

Thus, at optimization: $MP_L / w = MP_K / r$

And in the multiple input case with $X_1, X_2, ..., X_n$ inputs: $MP_{X1} / P_{X1} = MP_{X2} / P_{X2} = ... = MP_{Xn} / P_{Xn}$

Example: Production function of Q= 40L – L^2 +54K-1.5K^2

Given price of labor is 10 and capital is 15

Total cost = 120

How many combinations of labor and capital used to produce maximum output given 120 as total cost? This is because at the cost of 120 different units of output can be produced. Meaning with different combinations of labor and capital that has same cost, different levels of output can be produced. So what is that maximum level of output with this cost? It is the tangency point of isoquant curves and isocost lines.

6.5 The Law of Returns to Scale

We will now describe the effect of a proportional increase in all inputs on the level of output produced. For example, if the firm's usage of all inputs doubles, output would increase. The question is: By how much? The answer to this question depends upon the concept of returns to scale. The *law of returns to scale* shows how output is changing when inputs such as labor and capital are proportionately and simultaneously changed in the long run. This law is called the long run production function. There are three possible ways in which output may be increase when all inputs are proportionally increased.

- **Increasing returns to scale (IRS)**: increase in output (30%) is more than proportionate to increase in inputs (20%).
- **Decreasing returns to scale (DRS)**: increase in output (30%) is less than proportionate to increase in inputs (40%).
- **Constant returns to scale (CRS)**: increase in output (30%) is equal to the proportional change in inputs (30%)

6.6 THEORY OF COST

6.6.1 Concepts and Types of Costs

The theory of production shows the relationship between inputs and outputs in physical terms while the theory of cost deals the money value of inputs (i.e., costs of production) and the money value of outputs (i.e., revenue). The excess of revenue over cost yields profit which is the primary objective of firms. To achieve the profit goal of the firm a manager would make a decision on costs and outputs. Since managerial decisions are affected by the types of costs the manager should understand the different types of costs so as to focus on a relevant cost for a particular decision making. And also understand the relationship between cost and output both in the short run and long run.

6.6.1.1 Types of Costs

Actual and Opportunity Cost

Costs that are actually incurred in acquiring or producing a good or service is known as actual cost. The opportunity cost, on the other hand, refers to the foregoing of opportunities to produce an alternative good or service. Because of the fundamental economic problem of scarce resources the manager forced to choose the best out of the available alternative. Thus, the value that must be forgone in the second/next best alternative as a result of putting resources on the first best alternative. For example, a firm with Birr 100 can either make a fixed deposit with a bank and earn an interest of 10% per annum or can purchase a factors of production for producing t-shirts. Let the costs of land, labor, capital and management be Birr 20, 35, 30 and 10 respectively. Thus, the actual cost is Birr 95 while the opportunity cost is Birr 10.

Explicit and Implicit Costs

Explicit costs are monetary payments, that is out of pocket or cash expenditures, which a firm make to those “out-sider” who supply labor and other raw materials. The costs related to the firm use of certain resources which they own are called **explicit or imputed costs**. While the wages and other costs of raw materials constitute explicit costs, depreciation, salary of owner manager and other costs of self owned/self sponsored resources are **implicit costs**.

Incremental and Sunk Costs

Costs which depend on decision and relevant for decision making are incremental costs while costs which do not depend on decision and irrelevant on decision making are sunk costs. For illustration, consider a university as a firm which going to start an evening program on commercial basis beyond its regular program which runs in the day time.

The evening program is intended to use the time of academic as well as administrative staff for whom extra payments would have to be made for their services. In addition, there will be some costs on electricity, chalk etc. Besides, the classroom and blackboard of the university would be utilized for the purpose. With this regard, the costs of the time of the staff and the amount spent on electricity bill and chalk etc are incremental costs while the cost of the use of classroom and blackboard are the sunk cost. Other types of costs, but not part of this chapter, includes: economic and accounting costs, private and social costs, separable and common cost, historical and replacement cost.

6.6.1.2 Short Run and Long Run Costs

Short-run and long-run cost concepts are related to variable and fixed costs, respectively, and often figure in economic analysis interchangeably.

Short-run costs are the costs which vary with the variation in output, the size of the firm remaining the same. In other words, short-run costs are the same as variable costs. **Long-run costs,** on the other hand, are the costs which are incurred on the fixed assets like plant, building, machinery, etc. Such costs have long-run implication in the sense that these are not used up in the single 'batch of production'.

Long-run costs are, by implication, the same as fixed costs. In the long-run, however, even the fixed costs become variable costs as the size of the firm or sale of production increases. Broadly speaking, 'the short-run costs are those associated with variables in the utilization of fixed plant or other facilities whereas long-run costs are associated with the changes in the size and kind of plant.

Production Cost in the Short-Run

In the short run, the relationship between the TC and per unit costs (average & marginal costs) can be displayed using tabular and graphical illustration.

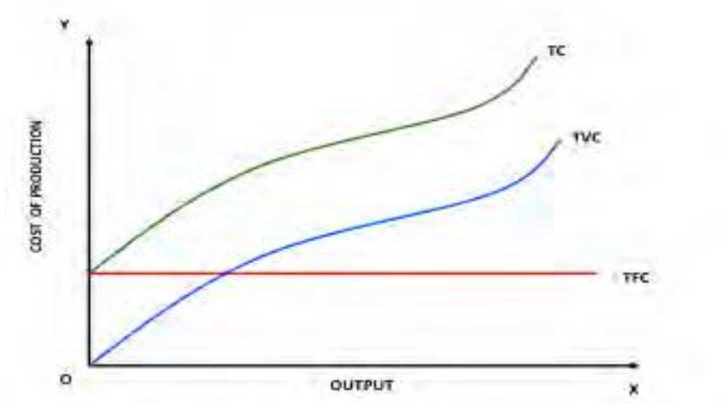
Fixed and Variable Costs

Total Fixed Costs (*TFC*) are costs of a firm that do not vary with the change in output so that firm's might incur costs even if no output is produced. Example, cost of firm's plant and machinery, rent of building and factory etc. **Total Variable Cost (*TVC*)** is the sum of the amounts spent for each of the variable inputs used. Total variable cost increases as output increases. Example, costs of raw materials and wages. **Total Cost (*TC*)** is the sum of total fixed cost and total variable cost. Total cost increases with increases in output (TC = TVC + TFC).

Table 6.3 Short-Run Total Cost Schedules

(1) Output (*Q*)	(2) Total fixed cost (*TFC*)	(3) Total variable cost (*TVC*)	(4) Total cost (*TC*) *TC = TFC + TVC*
0	$6,000	$ 0	$6,000
100	6,000	4,000	10,000
200	6,000	6,000	12,000
300	6,000	9,000	15,000
400	6,000	14,000	20,000
500	6,000	22,000	28,000
600	6,000	34,000	40,000

Figure 6.4 Total Cost Curves



Average and Marginal Costs

Average Fixed Cost (*AFC*): is total fixed cost divided by output:

AFC = TFC/Q

Average fixed cost is obtained by dividing the fixed cost (in this case $6,000) by output. Thus, *AFC* is high at relatively low levels of output; since the denominator increases as output increases, *AFC* decreases over the entire range of output. If output were to continue increasing, *AFC* would approach 0 as output became extremely large.

Average Variable Cost (*AVC*): is total variable cost divided by output:

$$AVC = TVC/Q$$

The average variable cost first falls to $30, then increases thereafter.

Average Total Cost (*ATC*): it is the per unit cost of the total cost and computed as the sum of average variable cost (AVC) and average fixed cost (AFC).

$$ATC = TC/Q = (TVC + TFC)/Q = AVC + AFC$$

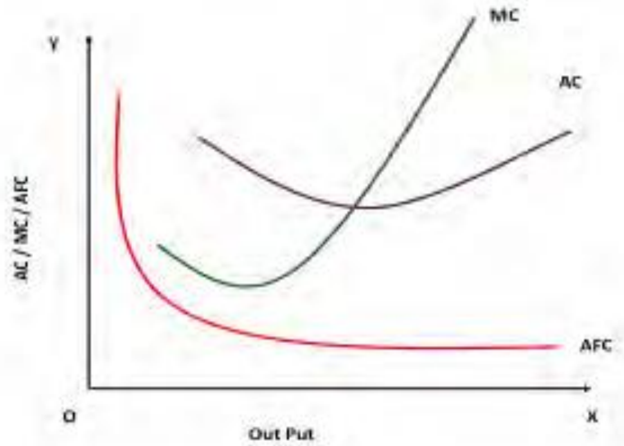
Marginal Cost (*MC*): it is the change in TC due to the production of one additional unit of output. In other words, the change in either total variable cost or total cost per unit change in output.

$$MC = \Delta TC/\Delta Q$$

Table 6.4 Average and Marginal Cost Schedules

(1) Output (*Q*)	**(2) Average fixed cost (AFC) *AFC = TFC/Q***	**(3) Average variable cost (AVC) *AVC = TVC/Q***	**(4) Average total cost (ATC) *ATC = TC/Q***	**(5) Marginal cost (MC) *MC = ΔTC/ΔQ***
0	-	-	-	$40
100	$60	$40	$100	20
200	30	30	60	30
300	20	30	50	50
400	15	35	50	80
500	12	44	56	120
600	10	56.7	66.7	

Figure 6.5 Average and Marginal Cost Curves



The average total cost first declines, reaches a minimum at $50, then increases thereafter. The minimum *AFC* is attained at a larger output (between 300 and 400) than that at which *AVC* attains its minimum (between 200 and 300). This result is not peculiar to the cost schedules given above.

It can be seen that *MC* first declines, reaches a minimum of $20, then rises. Note that minimum marginal cost is attained at an output (between 100 and 200) below that at which either *AVC* or *ATC* attains its minimum. Marginal cost equals *AVC* and *ATC* at their respective minimum levels.

The average and marginal cost schedules in columns 3, 4, and 5 are shown graphically in Figure 6.5. Average fixed cost is not graphed because it is a curve that simply declines over the entire range of output and because, as you will see, it is irrelevant for decision making. All three curves decline at first and then rise. Marginal cost equals *AVC* and *ATC* when they are declining and above them when they are increasing. Since *AFC* decreases over the entire range of output and since *ATC* = *AVC* + *AFC*, *ATC* becomes increasingly close to *AVC* as output increases. These are the general properties of typically assumed average and marginal cost curves.

- **Relations**: *(1) AFC declines continuously, approaching both axes asymptotically (as shown by the decreasing distance between ATC and AVC). (2) AVC first declines, reaches a minimum at* Q_2*, and then rises thereafter. When AVC is at its minimum, short-run marginal cost equals AVC. (3) ATC first declines, reaches a minimum at* Q_3*, and then rises thereafter. When ATC is at its* *minimum, short-run marginal cost equals ATC. (4) Short-run marginal cost first declines, reaches a minimum at Q_1, and rises thereafter. Short-run marginal cost equals both AVC and ATC when these curves are at their minimum values. Furthermore, short-run marginal cost lies below both AVC and ATC over the range for which these curves declines; Short-run marginal cost lies above them when they are rising.*

In general, the reason marginal cost crosses *AVC* and *ATC* at their respective minimum points follows from the definitions of the cost curves. If marginal cost is below average variable cost, each additional unit of output adds less to cost than the average variable cost of that unit. Thus, average variable cost must decline over this range. When short-run marginal cost is above *AVC*, each additional unit f output adds more to cost than *AVC*. In this case *AVC must rise.*

<u>Mathematical relationships between AC and MC</u>

AC= TC/Q =VC/Q + FC/Q (by using derivative of a quotient)

$$\frac{\delta AC}{\delta Q} = \frac{Q(\delta VC/\delta Q) - VC(\delta Q/\delta Q}{Q2} + \frac{Q(\delta FC/\delta Q) - FC(\delta Q/\delta Q)}{Q2}$$

$$\frac{\delta AC}{\delta Q} = \frac{Q(MC) - VC}{Q2} + \frac{0 - FC}{Q2}$$

$$\frac{\delta AC}{\delta Q} = \frac{Q(MC)\ - VC}{Q2} + \frac{0 - FC}{Q2}$$

$$\frac{\delta AC}{\delta Q} = \frac{MC}{Q} - \frac{VC}{Q}\ \frac{1}{Q} + \frac{-FC}{Q2}$$

$$\frac{\delta AC}{\delta Q} = \frac{1}{Q}(MC - AVC) - (FC/Q2)$$

$$\frac{\delta AC}{\delta Q} = \frac{1}{Q}\ (MC\text{-}AVC\text{-}AFC)$$

1. **If $\frac{\delta AC}{\delta Q}$ = 0 when AC is minimum, then**

$$\boxed{\frac{1}{Q}\ (MC\text{-}AVC\text{-}AFC) = 0}$$

MC = AVC+AFC, MC= AC

2. If $\frac{\delta AC}{\delta Q}$ = + ve (when AC is rising)

$\frac{1}{Q}$ **(MC-AVC-AFC) = + ve**

(MC-AVC-AFC) = + ve(Q)

MC = AVC+AFC (+ ve)(Q)

MC = AC(+ ve)(Q), So MC > AC

3. If $\frac{\delta AC}{\delta Q}$ = (- ve)(when AC is decling)

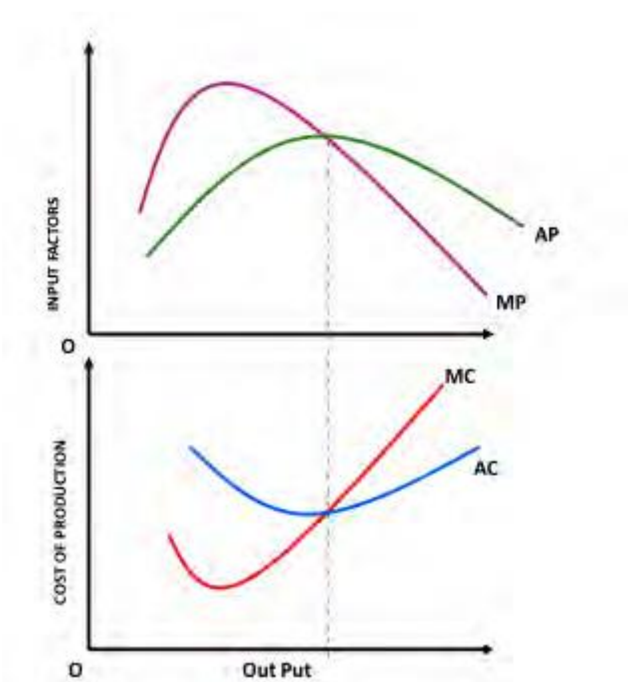
$\frac{1}{Q}$ **(MC-AVC-AFC) = (- ve)**

(MC-AVC-AFC) = (- ve)Q

MC = AVC+AFC +(- ve)Q

MC = AC+(- ve)Q, So MC < AC

Figure 6.6 Graph – Optimum Cost And Output

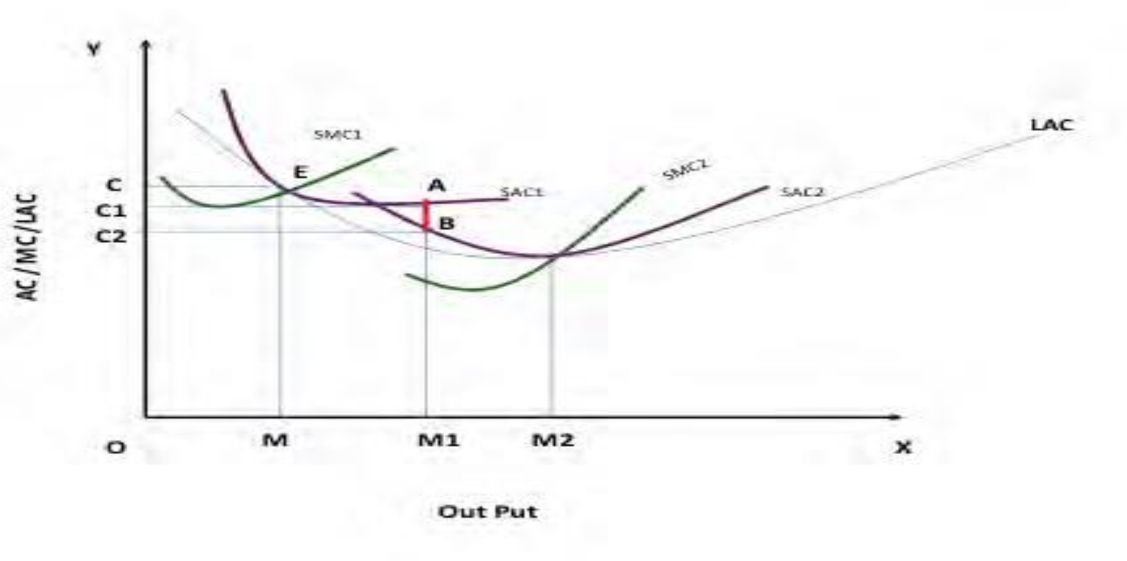


Production Cost in the Long-Run

To analyze the production costs in the long run it is good to deal above the long run cost curves which are a guide for an entrepreneur in his/her decision to plan future expansion of the firm's output. Due to this circumstance, the long run curves are known as a 'planning curve' or sometimes called the 'planning horizon'. The most important long run costs are the long run average cost (LAC) and long run marginal cost (LMC) and these curves are derived from their respective short run cost curves as depicted in the following graphs.

The LAC is derived from the short run average costs SAC_1, SAC_2 & SAC_3 which are associated with the three different plant sizes as small, medium and large plants respectively. The firm produces OX_1 at the minimum point of its short run average cost, $SAC_{1.}$ By installing a medium size plant, the firm can increase its output to OX_2 at a reduced cost of BX_2 at the minimum level of its short run average cost, $SAC_{2,}$ as well as the minimum of LAC. This production of more output at lower cost is due to economics of scale which could be achieved through specialization and factor productivity. If the firm wishes to increase its size by installing a large plant the output might increase to OX_3 but the cost, $CX_{3,}$ rising compared to that of small plant size. This production of more output at higher cost resulted from diseconomies of scale in this situation the firm becomes less efficient as its size gets large and large. Because the system of management becomes complex so that the managers are overworked and thus less efficient.

Figure 6.7 The LAC Curve



The LAC is, therefore, derived by joining then points on falling part of SAC_1 (which is on the left of its minimum indicating underutilization of the firm), the minimum of SAC_2, and the rising part of SAC_3 (which is on the right of its minimum indicating overutilization of the firm). In traditional theory of firms the LAC is U-shaped and it is often called the envelope Curve because it envelops the SAC curves.

Note: The SAC curve of the medium plant size is an optimal plant size. This is because all possible economies of scale are fully exploited. At this point more output is produced at lower cost in which both the SAC and LAC are at their minimum levels. An optimal scale of the firm (or the firm's long run equilibrium) is achieved at the point where $SAC_2 = LAC = LMC = SMC_2$ and the equilibrium level of output is X_2.

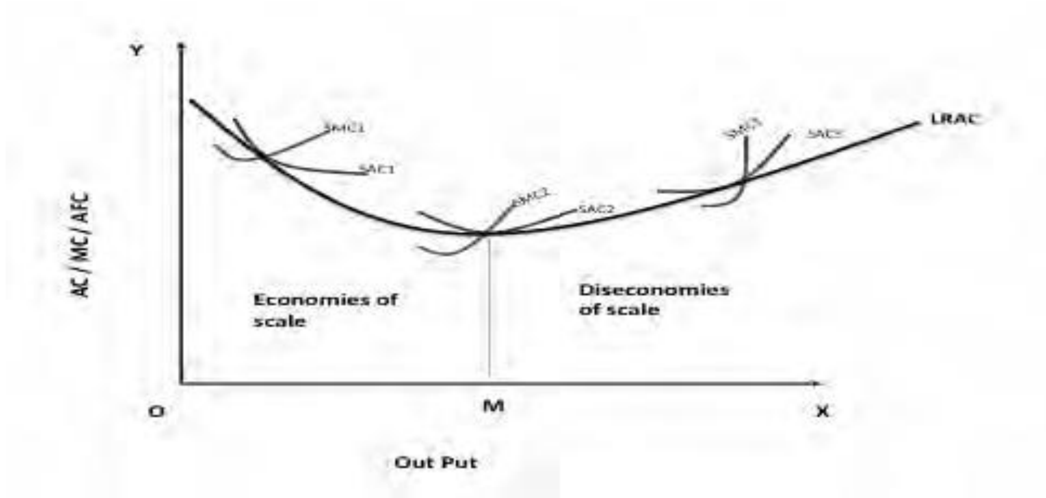
6.2 Scale and Scope Economies

Economies and diseconomies of scale are a phenomenon relating to the long run cost-output relationship while economies of scope is a phenomenon relating to the comparison of the long run costs between the joint and separate production process.

Economies of scale: it refers to range of output over which LAC falls as output increases by increasing the plant size. It is represented by the range of output between point 0 and M_2 on the LAC curve in Figure 6.7. In other words, economies of scale are referred to the production of more outputs at lower costs by increasing the size of the plant. Better facilities operative system, training and R & D are main factors causing the economies of scale.

Diseconomies of scale: it refers to a range of output over which LAC rises as output rises by further increasing the plant size. It is represented by the range of output on the right of M_2 on the LAC curve in figure 6.7. In other words, diseconomies of scale are referred to the production of more output at higher costs by further increasing the size of the plant. Difficulties in rising funds and difficulties in team-work and coordination are main factor causing diseconomies of scale.

Figure 6.8 – Economies of Scale and Diseconomies of scale



Economies of scope: it is a situation in which the joint cost of producing two or more goods by a single firm is less than the sum of a separate costs of producing the same level of output for each goods by a separate firms.

6.6.3 Cost Function

The cost of production (C) is a function of three main factors the total output produced (q), the price of inputs (p) and the efficiency of inputs (e). I.e., C = *f* (*q*, *p*, *e*).

There are various functional forms of cost function. The most important function forms are linear, quadratic and cubic forms of cost functions.

Linear: $C = a + bQ$

Quadratic: $C = a + bQ + cQ^2$

Cubic: $C = a + bQ + cQ^2 + dQ^3$

Exercise

1. The short run total cost function of Siket enterprise has been estimated as:

   $TC = 100 + 6Q + 0.25Q^2$, where TC is total cost; Q is output.

A. Determine the AFC, AVC, ATC & MC of the enterprise.

B. If you were a business advisor of the enterprise which aims to minimize its cost, how much would you recommend it to produce?

C. Determine the output level at which the ATC is at its minimum.