Unit III = Theory of Production & Theory of Cost

Theory of Production and cost - Theory of Production: Meaning of Production, Factors of production, Concepts of Total Product, Average product and Marginal Product, Production Function, Short Run Production Function: Law of Variable Proportion, ISO Quant Curve, Long Run Production Function, Law of Returns to Scale.
Theory of Cost – Meaning of Cost, Cost Concepts, Cost Function, Concepts of total Cost ,Average Cost and Marginal Cost, Short Run Average Cost Curves, Long Run Cost Curve, Economies and Diseconomies of Scale

Theory of Production

Concept of Production : In economics, production means any economic activity which is directed for the satisfaction of human wants.

Meaning : Production may be called as creation of utility i.e. creating of want satisfying goods and services. For eg. Production of cars by manufacturers, selling of vegetables by traders etc.

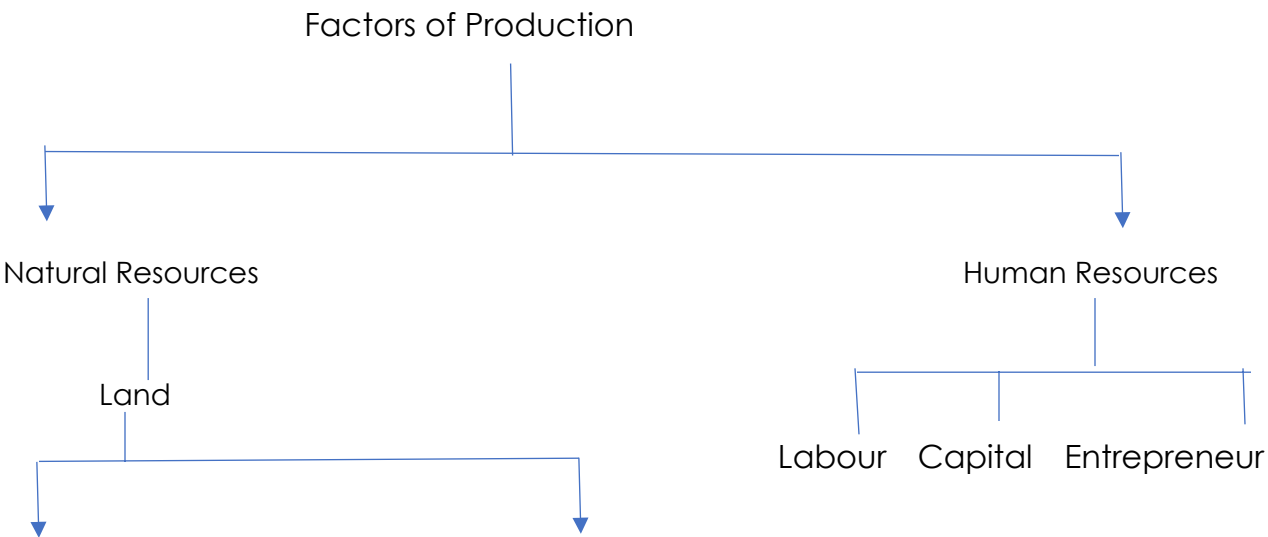
Definition by Adam Smith "Production is a creation of physical sset".

Definition by Alfred Marshall "Production is a creation of utility".

Definition by Mayers "Any action undertaken for the exchange of commodities and services is production".

Factors of Production

To produce goods, producer requires factors of production. Productive resources are classified into two parts. Natural resources and human resources. Natural resource is land and human resources are labour, capital and entrepreneur.



On the surface of earth Beneath the surface of earth

1. Water River
2. Sea Water
3. Agricultural land
4. Mountains etc….

1. Coal
2. Iron
3. Silver
4. Gold etc……

Rewards / Incentives obtained from different factors of production in separate ways.

SN	Factors of Production	Incentives / Rewards	Nature of Reward
01	Land	Rent	Fixed
02	Labour	Wages	Fixed
03	Capital	Interest	Fixed
04	Enterprenure	Profit	Variable/fluctuating

LAND – It refers to the surface of the earth. In economics land is used in a wider scope. Land is a primary, natural and original factor of production.
Land is a free gift of nature.
Land is passive factor of production.
Land is permanent. It can not be increased or decreased.
Land is fixed in quantity. It has inelastic supply means man can not increase or decrease its volume.
Land has no geographical mobility. It can not be moved from one place to another. Land is heterogenous factor, it differ in quality and grade.
Land varied in fertility and use. No 2 piece of land yield same returns. Land is natural factor. Its value depends on its location.

LABOUR – according to Marshall "Mental and physical exertion directed to produce goods and services with a view to gain an economic reward". Any activity done for personal interest or an hobby is not a labour. Labour is most active and living factor of production.
Labour involves human efforts to gain economic reward. It is a self source.
Labour is perishable in nature. If labour is absent on a work for a day, labour is lost forever. Less bargaining power. No specific prices are fixed, although labour laws are followed upto some extent.
Variation in efficiency of labour.
Restricted mobility. Labour can not be moved from one place to another easily.

CAPITAL – capital is that part of wealth, of an individual or community which is used for future production of wealth and which yield an income. For eg. Plant and machineries, raw materials, fuels, dams and canals etc.
Capital is a man made factor. Capital is a part of wealth.
Capital is produced means of production means passive factor of

production. Capital has got mobility.
It is perishable in nature.

Production Function

What is production function?

Production function is a functional relationship between physical input (factors of production) and output (total quantity of goods and services produced. In production process firms combines various inputs in different quantities / proportions to produce different levels of output. The technological-physical relationship between input and output is known as production function

Definition by Stigler "*A production function refers to the functional relationship under the given technological relationship under the given technology between physical rates of inputs and output of a firm per unit of time*".

Production function states the relationship between inputs and output i.e. maximum amount of output can be produced with given quantities of inputs in existing state of technology. Production function gives minimum quantities of various inputs that are required to yield given quantity of output.

Production function means physical relationship between inputs and outputs used. It is a purely technological relation. It is a process of getting maximum output for given input in a particular time period(long and short run).

Economist Cobb & Douglas production function

In the year 1928 economist Cobb & Douglas gave this production function. It is based on empirical studies and was initially applicable to the whole of the manufacturing industry in USA and can be applicable to the whole economy. Cobb & Douglas production function
= $\mathbf{Q = K L^{a} C^{1-a}}$
Q = output, L = quantity of labour, C= quantity of capital, input a & k are constant factors.

In mathematical terms production function may be stated as Q = f (a,b,c,d……nT) Q is a physical quantity of output (commodity produced) per unit of time, F is functional relationship, a,b,c,d…n is various factors of production per unit of time period, T is technology, and bar above T shows that technology which is used for particular production remains constant.

But often, economists denotes simple production function assuming 2 factors and represents production function as **Q = Qx f (K, L) per unit of time.**
Qx is output of commodity, f is functional relationship, k is capital used per unit of time, L is labour units used per unit of time.

Importance of Production Function

1. **Cost reduction** – different combination of input in production function can be used to reduce production cost by producer. Every seller aims to maximise his profit with given factors of input/output quantities.
2. **Economic standards of individual rises** – with the help of technological improvements and better infrastructural facilities businessmen and other related human resources like worker/labours can improve their economic condition and earn more.
3. **Utilisation of technology** – production function enables producers to choose technology with given period of time. Various combinations of factors of production with the help of advanced technology will be utilised at maximum level.
4. **Difference in production function** – producers and industrialists must have deep knowledge and information about the production function to optimise the available resources of production process. Production functions are utilised and applied in different manners depends upon the strategies of firms and industries and are implemented in different ways and it may be agricultural or manufacturing industries.
5. **Higher earnings** – implementation, adoption of production function enables a producer to minimise wastage of production resources and earns maximum profit.

Concept of Total Product, Average Product and Marginal Product.

Total product is defined as Total quantity of goods and services produced by firm with the given inputs during a specific period of time.

Average product refers to the total product per unit of a given variable factor. Thus dividing the total product by quantity of variable factor, we get A.P. AP= TP or TP
QVF L (labour input)

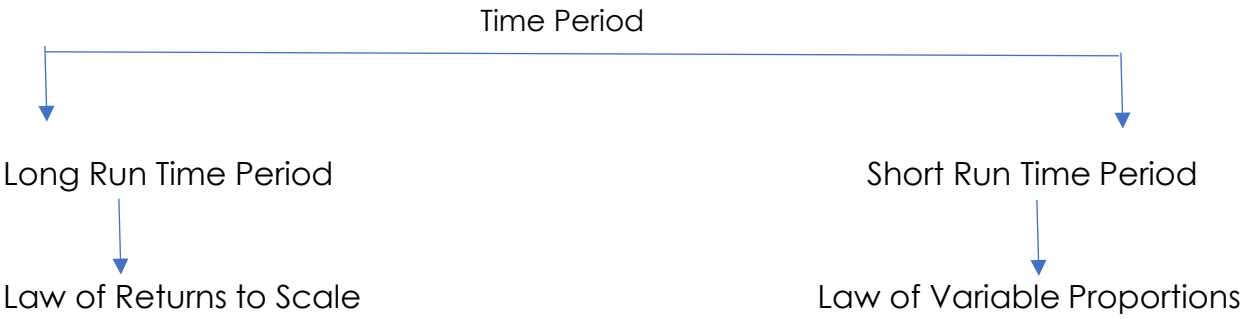
For eg. Total product produced is 400 commodities with 25 labours employed therefore

AP = 400
25 = 16 unit per labour

Marginal Product means , in addition to the unit of variable factor, all other factors are being constant, the additional factor used in total product is technically refered as marginal product. MP may be defined as

MPn = TPn – TP n-1. When 26 workers are used then total product will be increased from 400 to 440 but when 25 labours are employed then MP of 26th worker may be assumed as MP = TP 26 – TP25 = 440-400 = 40units.

Different laws using production function in specific time period.



Law of Variable Proportions Concept and meaning

Under this law it is assumed that only one factor of production is variable while other factors are fixed. It takes place in short run only. In short run period production can be increased by using more units of variable factors, but at initial stages not at final stages.

When total output or production of commodity is increased by adding units of variable inputs, while quantities of other inputs are constant, then increase in total production after some times becomes smaller and smaller.

Statement of Law given by Benham "As the proportion of one factor in a combination of factors is increased, after a point the average and marginal production of that factor will diminish".

Assumptions of Law of Variable Proportions :

1. All units of a variable factor i.e. labour are homogenous.
2. Only 1 factor of production is varying and all other factor should remain constant.
3. The scale of output is unchanged. i.e. production plant, plant size and efficiency of the firm remains constant.
4. The state of technology remains unchanged i.e. the technology used in production process remains same.

Production Schedule for Law of Variable Proportions

Units of Variable Factor (Labour) n	Total Output TP	Average Product AP=TP/QVF	Marginal Product (TPn-TPn-1)	
01	20	20	20	
02	50	25	30	Stage I
03	90	30	40	
04	120	30	30	
05	135	27	15	
06	144	24	09	Stage II
07	147	21	03	
08	148	18.5	01	
09	148	16.4	00	
10	145	14.5	-3	Stage III

Graph of Law of variable Proportions

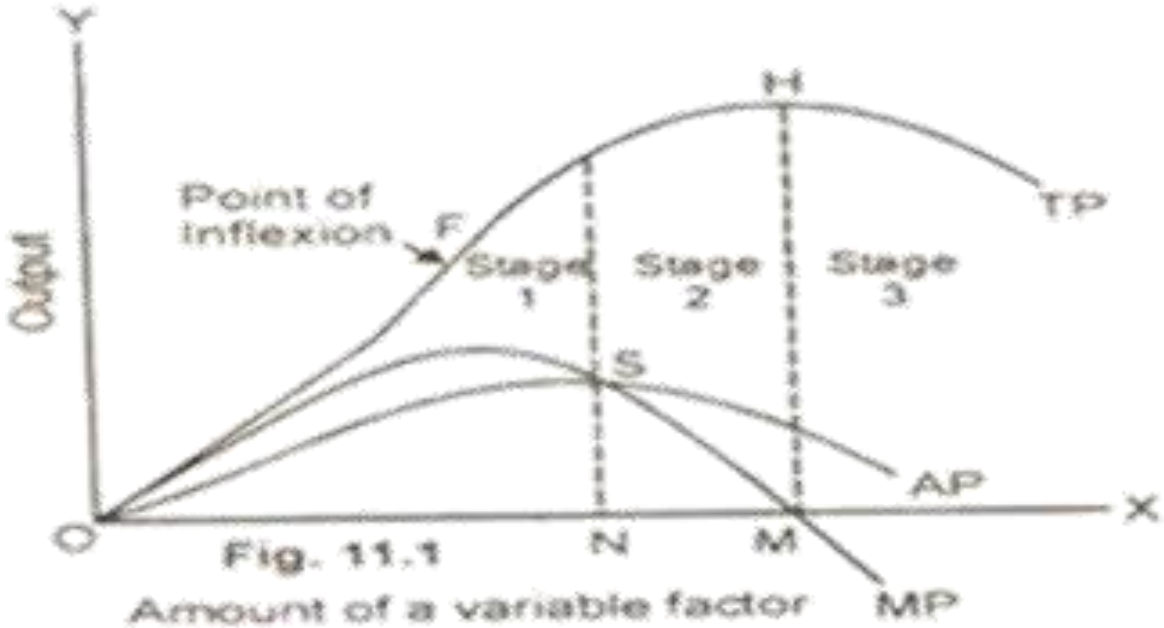


Fig. 11.1

Relationship between TP, AP & MP curves as follows :

A. Initially all 3 product are rising, but MP is greater than increase in AP.
B. When AP is maximum, MP = AP.
C. When TP is maximum MP=0.
D. When TP is falling MP is negative.
E. As long as TP is positive AP is positive.
F. Both AP & MP are inverted, U shaped.

Stage I = A rational producer will not produce in this stage because it is non economical range due to over utilisation of fixed factor land.

Stage II = The producer would like to produce more in this stage, where both AP & MP of variable factors are falling. As a rational producer may operate in this stage. Use of constant variable factors are proportionate.

Stage III = MP of variable factor is negative as its quantity is too large than fixed factor. It is non economic and inefficient range. Over utilisation of fixed factors.

Law of Returns to Scale Concept and meaning

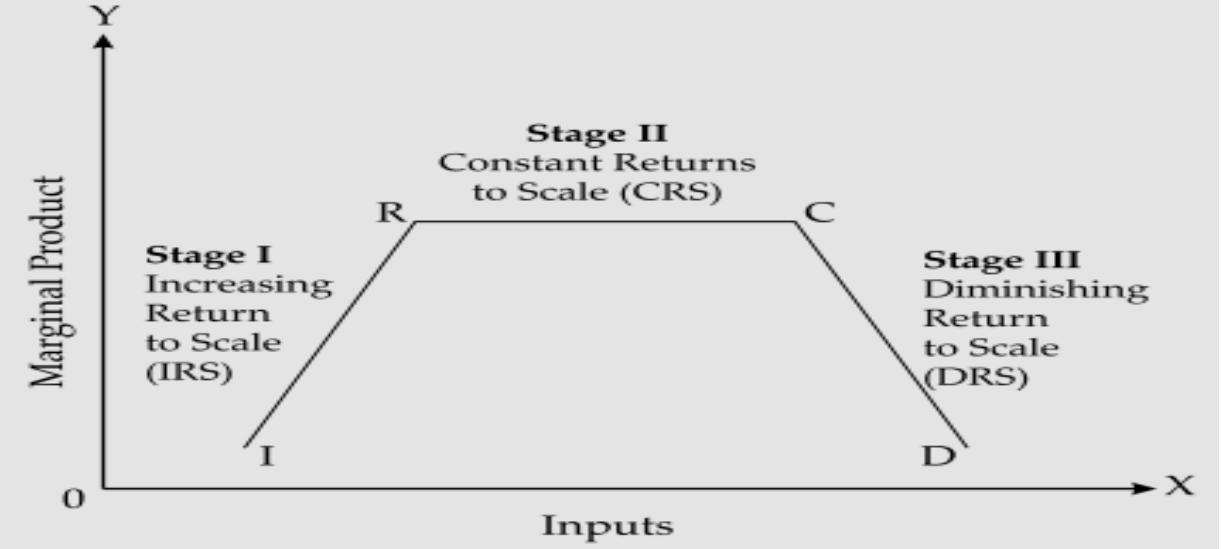
Law of returns to scale is applicable in long run, where all factors of production may changed with input factors. In long run out put can be increased by increasing all factors of production or "Scale of production." This law implies that all factors becomes variable to increase production volume.

Statement of Law "The statement of law of returns to scale states that when all factors of production are increased in the same proportion output will increase but increase may be at increasing rate, constant rate or diminishing rate."

Assumption of Law of Returns to scale

A. Applicable only in long run – This law is applicable in long run time period only. For eg. Existing plant may be dismantled or renewed to set up a new plant.
B. All factors are variable – All factors of production are assumed to be vary in long run provided that all factors must be homogenous.
C. Factors of input are varied in equal proportion to increase the production in long run, all factors are said to be varied in same proportion.
D. Technology used for production purpose remain constant, it does not change.
E. Physical units – only physical inputs and factors of output are examined, profit is ignored.

Factor combination	Scale of inputs (labour & capital in units)	TP (in quintals)	MP (in quintals)
IRS	2+1	20	20
	4+2	50	30
	6+3	90	40
CRS	8+4	140	50
	10+5	190	50
	12+6	240	50
	14+7	290	50
DRS	16+8	320	30
	18+9	340	20
	20+10	350	10



IRS (Increasing returns to scale) = With increase in all input and when output increases more than increase in input then it is called increasing returns to scale. For eg. If input of firm is increased by 15% and output is increased by 40% here law of increasing returns to scale is applicable.

CSR (Constant Returns to scale) = In case of constant returns, it refers that when there is an increase in input by 20% and output also increases by 20% then it is said constant returns to scale. Increase in input results with proportionate / same increase in output.

DRS (Diminishing Returns to scale) = When there is increase in input but output does not change or decrease with an increase in input then such situation is called diminishing returns to scale. For eg. If input is increased by 10% and output increased by 5% only then it is diminishing returns to scale.

ISO Quant Curves

The term Iso-quant or Iso-product is composed of two words,

Iso = equal, quant = quantity or product = output.

Thus it means equal quantity or equal product. Different factors are needed to produce a good. These factors may be substituted for one another.

A given quantity of output may be produced with different combinations of factors. Iso- quant curves are also known as Equal-product or Iso-product or Production Indifference curves. Since it is an extension of Indifference curve analysis from the theory of consumption to the theory of production.

Thus, an Iso-product or Iso-quant curve is that curve which shows the different combinations of two factors yielding the same total product. Like, indifference curves, Iso- quant curves also slope downward from left to right. The slope of an Iso-quant curve expresses the marginal rate of technical substitution **(MRTS).**

Definitions:

"The Iso-product curves show the different combinations of two resources with which a firm can produce equal amount of product." Bilas

"Iso-product curve shows the different input combinations that will produce a given output." Samuelson

ADVERTISEMENTS:

"An Iso-quant curve may be defined as a curve showing the possible combinations of two variable factors that can be used to produce the same total product." Peterson

"An Iso-quant is a curve showing all possible combinations of inputs physically capable of producing a given level of output." Ferguson

Assumptions:

The main assumptions of Iso-quant curves are as follows:

1. Two Factors of Production:

Only two factors are used to produce a commodity.

2. Divisible Factor:

Factors of production can be divided into small parts.

3. Constant Technique:

Technique of production is constant or is known before hand.

4. Possibility of Technical Substitution:

The substitution between the two factors is technically possible. That is, production function is of 'variable proportion' type rather than fixed proportion.

5. Efficient Combinations:

Under the given technique, factors of production can be used with maximum efficiency.

Iso-Product Schedule

Let us suppose that there are two factor inputs—labour and capital. An Iso-product schedule shows the different combination of these two inputs that yield the same level of output as shown in table 1.

Table 1. Iso-Product Schedule.

Combination	Units of labour	Units of capital	Output of cloth (metres)
A	1	15	200
B	2	11	200
C	3	8	200
D	4	6	200
E	5	5	200

The table 1 shows that the five combinations of labour units and units of capital yield the same level of output, i.e., 200 metres of cloth. Thus, 200 metre cloth can be produced by combining.

(a) 1 units of labour and 15 units of capital

(b) 2 units of labour and 11 units of capital

(c) 3 units of labour and 8 units of capital

(d) 4 units of labour and 6 units of capital

(e) 5 units of labour and 5 units of capital

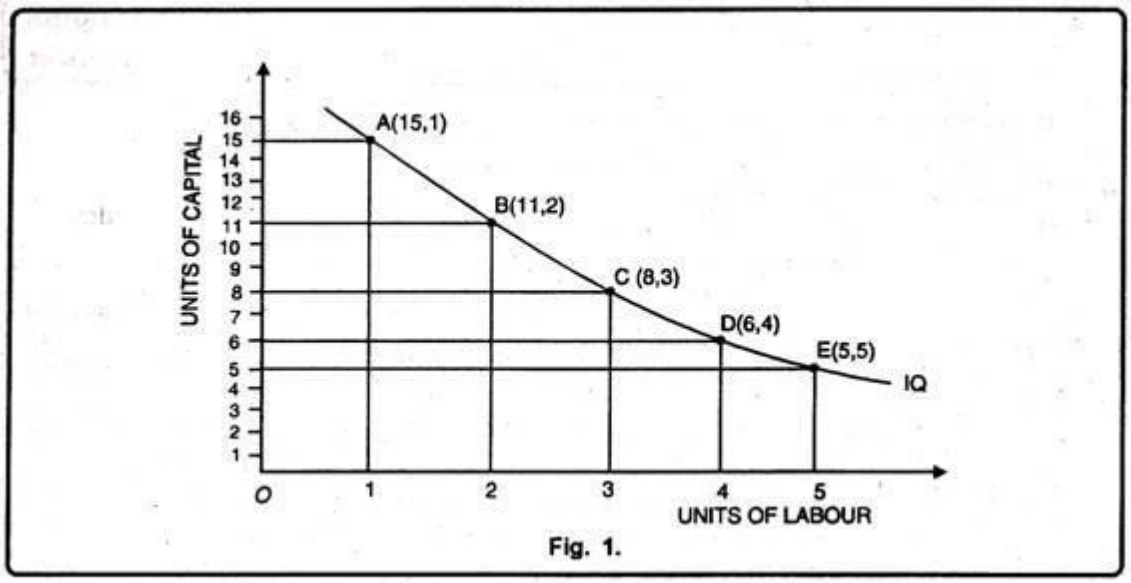


Fig. 1.

Economies and Diseconomies of Scale

Scale of production

Scale of production means the size or volume of production unit of a firm or industry. The scale of a production may be a small and large scale depending upon the quantity of output per unit of a firm. The scale of production differs with the size of the firm.

When we talk about the scale of production of a firm, we often hear about the fact that large-scale production, usually, helps in reducing the cost of production. Economies of scale refer to these reduced costs per unit arising due to an increase in the total output. Diseconomies of scale, on the other hand, occur when the output increases to such a great extent that the cost per unit starts increasing.

Internal and External Economies

When a firm opts for large-scale production, the economies arising out of it are grouped into two categories:

1. Internal economies – economies of production that the firm accrues when it increases the output leading to a drop in the cost of production. These arise due to endogenous factors like entrepreneurial efficiency, talents of the management team, type

of machinery, etc. These economies arise within the firm and help the firm only.

Internal Diseconomies and Economies of Scale

While studying returns to scale, we observed that they increase during the initial stages, remain constant for a while, and then start decreasing. The reason is simple – initially, the firm enjoys internal economies of scale and after a certain limit, it suffers from internal diseconomies of scale. Let's look at the types of economies and diseconomies:

Technical

Large-scale production is linked to technical economies. When a firm increases its scale of operations, it needs to use a more specialized and efficient form of capital equipment and machinery. Such machinery helps to produce larger outputs at a lower unit cost.

Further, as the scale of production increases and the amount of labor and other factors becomes larger, the firm manages to reduce costs by introducing a degree of division of labor and specialization.

However, beyond a certain point, the firm experiences diseconomies of scale. This happens because after reaching a large enough output, the firm utilizes almost all possibilities of the division of labor and employment of efficient machinery.

Post this, any increase in the size of the plant causes the costs to rise. When the scale of operations becomes too large, the management finds it more difficult to control and coordinate the operations.

Managerial

As the output increases, the firm can apply the division of labor to the management as well. For example, the production manager can look after production, the sales manager can look after sales, etc. When the scale of production increases further, the firm divides each department into sub-departments like sales is divided into advertising, exports, and service.

Thus helps in increasing the efficiency and productivity of the management team since a specialist manages each sub-department. Further, the firm has the option to decentralize decision-making authority enhancing the efficiency further. Therefore, specialized management allows the firm to reduce managerial costs.

However, as the firm increases its scale of operations beyond a certain limit, the management finds it difficult to control and coordinate between departments. This leads to managerial diseconomies.

Commercial

As a firm increases its volume of production, it requires large amounts of raw material and components. Hence, it places a bulk order for such material and components and enjoys discounted pricing for them.

Economies are also achieved during sales. If the sales staff is working under-capacity, then the firm can sell additional output at little extra cost.

Further, as the scale of production increases, the advertising cost per unit fall. Hence, the firm benefits from economies of advertising too. After an optimum level, these economies start becoming diseconomies though.

Financial

When a firm wants to raise finance, a large-scale firm has many benefits like:

- Better security to bankers
- Well-known
- Can raise finance at lower costs, etc.

However, after the optimum scale of production, the financial costs rise faster due to the increased dependence on external finances.

Risk-bearing

A firm enjoys the economies of risk-bearing if it has a large-scale operation with diverse and multi-production capabilities. However, if the

diversification increases the economic disturbances rather than covering them, then the risk increases.

2. External economies – these are the benefits that each member firm of the industry accrues due to the expansion of the entire industry.

External Diseconomies and Economies of Scale

External diseconomies and economies of scale are very important to a firm. These are a result of the expansion of output of the entire industry and not limited to an individual firm. They are available to one or more firms in the following forms:

Cheaper Raw materials and Capital Equipment

At times, the expansion of an industry results in new and cheaper sources of raw material, machinery, and other capital equipment. It also results in an increased demand for the various types of materials and equipment required by the industry.

Hence, such materials/equipment can be purchased from other industries on a large scale. This, eventually, leads to a lower cost of production and lower price. Therefore, firms using these materials/equipment get them at lower prices.

Technological External Economies

Usually, when an entire industry expands, new technical knowledge is discovered leading to new and improved machinery for the said industry. This changes the technological coefficient of production and enhances the productivity of the firms in the industry. Hence, the cost of production reduces.

Development of Skilled Labor

As the industry expands, the labor gets accustomed to managing various production processes and learns from the experience. This increases the number of skilled workers which in turn has a favorable effect on the levels of productivity.

Growth of Ancillary Industries

When a certain industry expands, many ancillary industries start specializing in the production of raw materials, tools, machinery, etc. These ancillary industries offer the materials/machinery at a low price.

Similarly, some ancillary industries also start processing industrial waste and create a useful product out of it. Overall, it leads to a lower cost of production.

Better Transportation and Marketing Facilities

An expanding industry, usually, results in better transportation and marketing networks. These aspects help reduce the cost of production in the firms from the industry.

It is important to note that, certain disadvantages can neutralize the advantages of the expansion of industry and cease the external economies of scale. These are external diseconomies. When an industry expands, the demand for certain materials and skilled labor increases.

If these factors are in short supply, then their prices can increase. Further, the geographical concentration of firms from the industry can lead to higher transportation costs, marketing costs, pollution control costs, etc.

Meaning of Cost

Cost has a key role to play in business economics. All business decisions involves a comparison between cost and returns. From seller and business owner point of view payment of rent, wages, interest are the costs. The cost of production of goods depends on the factors of production necessary to produce a given level of output and current prices of factors units. So the cost is jointly determined by techniques of production adopted and the organizational efficiency of entrepreneurs and factors of production and their prices along with the rate of output and size of plant.
Cost analysis is concerned with financial aspect of the production as against the physical aspect which is considered in production analysis. Cost analysis is a study of behavior of cost in relation to one or more production functions.

Factors determinants of cost

1. Law of returns – while determining the cost laws of returns are considered. In law of variable proportions and diminishing returns cost of production will rise and in case of production in long run law of increasing returns is applicable.
2. Size of plant – when the size of plan t is big at initial stage fixed costs are higher but variable costs are low as compared to small plant size.
3. Time period – in short run time period the cost curve will rise sharply and in long run cost curve will not increase sharply as involvement of input factors are at slow and gradual level.
4. Capacity utilisation - when the plant size is considerable for longer period more time with higher capacity utilization fixed cost per unit of output is low.
5. Prices of factors of production – the cost of a product is affected by the prices of factors of production. If the prices of factors are higher the cost of producing gods will be higher.
6. Technology – technology is a big influence on cost of production. Innovations and improvements in the technology are made to reduce the cost of production.
7. Stability – from stability point of view in long run time when the firm is involved in long run production the costs are generally lower.

<u>Cost function</u>

Cost function is a combination of various determinants of factor units of costs. These determinants are identifiable in cost behavior of the firm. Cost function refers to the mathematical relationship between cost of production and other different determinants of cost. Cost is a function of prices of input and the rate of output with the size of plant and state of technology used.

Therefore C = f (f, o, p, t, u)

C= cost , F=functional relationship

P= size of plant

F=factor input prices

O=rate of output

T= technology used

U=utilization of production capacity

Opportunity cost is the forgone benefit that would have been derived by an option not chosen. To properly evaluate opportunity costs, the costs and benefits of every option available must be considered and weighed against the others.
the opportunity cost of a particular activity option is the loss of value or benefit that would be incurred (the cost) by engaging in that activity, relative to engaging in an alternative activity offering a higher return in value or benefit.

Opportunity Cost = FO (returns on best forgone option)
Sacrifice is a given measurement in opportunity cost of which the decision maker forgoes the opportunity of the next best alternative.In other words, to disregard the equivalent utility of the best alternative choice to gain the utility of the best perceived option. If there were decisions to be made that require no sacrifice then these would be cost free decisions with zero opportunity cost. Only through the analysis of opportunity cost, the company can choose the most beneficial project, based on when the actual benefits are greater than the opportunity cost, so that the limited resources can be optimally allocated to achieve maximum efficiency.
An explicit cost is a direct payment made to others in the course of running a business, such as wage, rent and materials as opposed to *implicit costs*, where no actual payment is made.[2] It is possible still to underestimate these costs, however: for example, pension contributions and other "perks" must be taken into account when considering the cost of labour. Explicit costs are taken into account along with implicit ones when considering economic profit. Accounting profit only takes explicit costs into account an implicit cost, also called an imputed cost, implied cost, or notional cost, is the opportunity cost equal to what a firm must give up in order to use a factor of production for which it already owns and thus does not pay rent. It is the opposite of an explicit cost, which is borne directly.[1] In other words, an implicit cost is any cost that results from using an asset instead of renting it out or selling it. The term also applies to foregone income from choosing not to work.

Implicit costs also represent the divergence between economic profit (total revenues minus total costs, where total costs are the sum of implicit and explicit costs) and accounting profit (total revenues minus only explicit costs). Since economic profit includes these extra

opportunity costs, it will always be less than or equal to accounting profit.

Calculation of Short-Run Average Total Cost

Let us now have a look at the various short-run average cost functions.

1. Short-run average variable cost - It is the variable cost of production per unit product. The formula for short-run average variable cost can be written as -

$$AVC = TVC / Q$$

Where AVC is the average variable cost and TVC is the total variable cost.

Where AVC is the average variable cost and TVC is the total variable cost.

2. Short-run average fixed cost - It is defined as the fixed cost for production per unit of output. It is calculated as -

$$AFC = TFC / Q$$

Where AFC is the average fixed cost and TFC is the total fixed cost. Where AFC is the average fixed cost and TFC is the total fixed cost.

3. Short-run average total cost - It refers to the total cost of production per unit product. The formula for the short-run average total cost is as follows-

$$ATC = TC / Q$$

Where TC is the average total cost, TC is the total cost. Where ATC is the average total cost,TC is the total cost.

The short-run average total cost can also be calculated as the sum of short-run average variable cost and average fixed cost.

$$ATC = AVC + AFC$$

All these functions are important for plotting the cost curves in the short-run.

The Short-Run Average Cost Curve

After having talked about the short-run average cost definition and a thorough understanding of its components, we will now discuss the average cost curve in the short-run.

On the X-axis is the cost of production (in rupees) and on the Y-axis is the quantity of output.

The graph of the average fixed cost goes on decreasing because it is a fixed number and as we keep dividing it by the increasing number of products, it keeps getting smaller. The marginal cost curve goes down and up because of the law of diminishing marginal returns. It goes down at first due to the additional output produced by the workers as they specialize, but eventually, it starts rising because the resources become limited after a certain period.

The short-run average total cost curve and the short-run average variable cost curve also go down first, intersect the curve of marginal cost at their minimum, and then go on rising to form a U-shape. This graph could also be used to calculate total costs by finding out the area under a particular curve.

In 1998, Alan Blinder, former vice president of the American Economics Association, conducted a survey in which 200 US firms were shown different cost curves and asked to specify which one of those curves represented the US economy the best. He found that about 88.4% of firms reported cost curves with constant or marginal cost.

In other words, it is the total cost divided by the number of units produced.

The diagram below shows the AFC, AVC, ATC, and Marginal Costs (MC) curves:

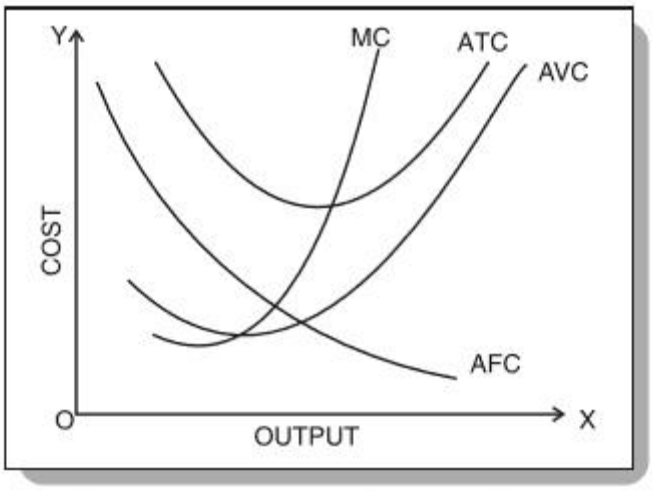


Fig. 1 : Short run Average and Marginal Cost Curves

It is important to note that the behaviour of the ATC curve depends upon that of the AVC and AFC curves. Observe that:

- In the beginning, both AVC and AFC curves fall. Hence, the ATC curve falls as well.
- Next, the AVC curve starts rising, but the AFC curve is still falling. Hence, the ATC curve continues to fall. This is because, during this phase, the fall in the AFC curve is greater than the rise in the AVC curve.
- As the output rises further, the AVC curve rises sharply. This offsets the fall in the AFC curve. Hence, the ATC curve falls initially and then rises.

Shapes of Long-Run Average Cost Curves

While in the short run firms are limited to operating on a single average cost

curve (corresponding to the level of fixed costs they have chosen), in the long run when all costs are variable, they can choose to operate on any average cost curve. Thus, the long-run average cost (LRAC) curve is actually based on a group of short-run average cost (SRAC) curves, each of which represents one specific level of fixed costs. More precisely, the long-run average cost curve will be the least expensive average cost curve for any level of output. Figure 3 shows how the long-run average cost curve is built from a group of short-run average

cost curves.

Five short-run-average cost curves appear on the diagram. Each SRAC curve represents a different level of fixed costs. For example, you can imagine $SRAC_1$ as a small factory, $SRAC_2$ as a medium factory, $SRAC_3$ as a large factory, and $SRAC_4$ and $SRAC_5$ as very large and ultra-large. Although this diagram shows only five SRAC curves, presumably there are an infinite number of other SRAC curves between the ones that we show. Think of this family of short-run average cost curves as representing different choices for a firm that is planning its level of investment in fixed cost physical capital—knowing that different choices about capital investment in the present will cause it to end up with different short-run average cost curves in the future.

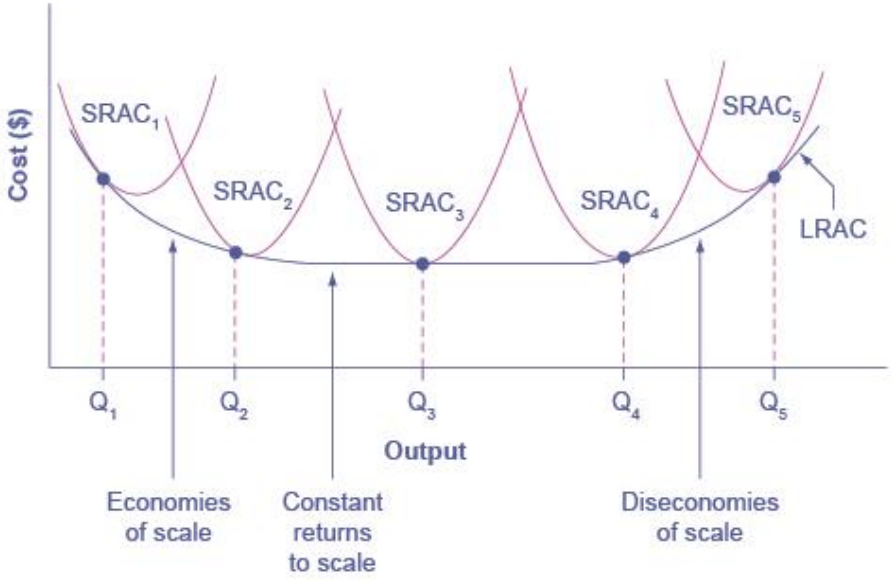


Figure 3. From Short-Run Average Cost Curves to Long-Run Average Cost Curves The five different short-run average cost (SRAC) curves each represents a different level of fixed costs, from the low level of fixed costs at $SRAC_1$ to the high level of fixed costs at $SRAC_5$. Other SRAC curves, not shown in the diagram, lie between the ones that are shown here. The long-run average cost (LRAC) curve shows the lowest cost for producing each quantity of output when fixed costs can vary, and so it is formed by the bottom edge of the family of SRAC curves. If a firm wished to produce quantity Q_3, it would choose the fixed costs associated with $SRAC_3$.

The long-run average cost curve shows the cost of producing each

quantity in the long run, when the firm can choose its level of fixed costs and thus choose

which short-run average costs it desires. If the firm plans to produce in the long run at an output of Q_3, it should make the set of investments that will lead it to locate on $SRAC_3$, which allows producing q_3 at the lowest cost. A firm that intends to produce Q_3 would be foolish to choose the level of fixed costs at $SRAC_2$ or $SRAC_4$. At $SRAC_2$ the level of fixed costs is too low for producing Q_3 at lowest possible cost, and producing q_3 would require adding a very high level of variable costs and make the average cost very high. At $SRAC_4$, the level of fixed costs is too high for producing q_3 at lowest possible cost, and again average costs would be very high as a result.

The shape of the long-run cost curve, in Figure 3, is fairly common for many industries. The left-hand portion of the long-run average cost curve, where it is downward- sloping from output levels Q_1 to Q_2 to Q_3, illustrates the case of economies of scale. In this portion of the long-run average cost curve, larger scale leads to lower average costs. We illustrated this pattern earlier in Figure 2.

In the middle portion of the long-run average cost curve, the flat portion of the curve around Q_3, economies of scale have been exhausted. In this situation, allowing all inputs to expand does not much change the average cost of production. We call this constant returns to scale. In this LRAC curve range, the average cost of production does not change much as scale rises or falls.

Finally, the right-hand portion of the long-run average cost curve, running from output level Q4 to Q5, shows a situation where, as the level of output and the scale rises, average costs rise as well. This situation is called diseconomies of scale. A firm or a factory can grow so large that it becomes very difficult to manage, resulting in unnecessarily high costs as many layers of management try to communicate with workers and with each other, and as failures to communicate lead to disruptions in the flow of work and materials. Not many overly large factories exist in the real world, because with their very high production costs, they are unable to compete for long against plants with lower average costs of production. However, in some planned economies, like the economy of the old Soviet Union, plants that were so large as to be grossly inefficient were able to continue operating for a long time because government economic

planners protected them from competition and ensured that they would not make losses.

Diseconomies of scale can also be present across an entire firm, not just a large factory. The leviathan effect can hit firms that become too large to run efficiently, across the entirety of the enterprise. Firms that shrink their operations are often responding to finding itself in the diseconomies region, thus moving back to a lower average cost at a lower output level.

Economies and Diseconomies of Scale

Scale of production

Scale of production means the size or volume of production unit of a firm or industry. The scale of a production may be a small and large scale depending upon the quantity of output per unit of a firm. The scale of production differs with the size of the firm.

When we talk about the scale of production of a firm, we often hear about the fact that large-scale production, usually, helps in reducing the cost of production. Economies of scale refer to these reduced costs per unit arising due to an increase in the total output. Diseconomies of scale, on the other hand, occur when the output increases to such a great extent that the cost per unit starts increasing.

Internal and External Economies

When a firm opts for large-scale production, the economies arising out of it are grouped into two categories:

3. Internal economies – economies of production that the firm accrues when it increases the output leading to a drop in the cost of production. These arise due to endogenous factors like entrepreneurial efficiency, talents of the management team, type of machinery, etc. These economies arise within the firm and help the firm only.

Internal Diseconomies and Economies of Scale

While studying returns to scale, we observed that they increase during the initial stages, remain constant for a while, and then start decreasing. The reason is simple – initially, the firm enjoys internal economies of scale

and after a certain limit, it suffers from internal diseconomies of scale. Let's look at the types of economies and diseconomies:

Technical

Large-scale production is linked to technical economies. When a firm increases its scale of operations, it needs to use a more specialized and efficient form of capital equipment and machinery. Such machinery helps to produce larger outputs at a lower unit cost.

Further, as the scale of production increases and the amount of labor and other factors becomes larger, the firm manages to reduce costs by introducing a degree of division of labor and specialization.

However, beyond a certain point, the firm experiences diseconomies of scale. This happens because after reaching a large enough output, the firm utilizes almost all possibilities of the division of labor and employment of efficient machinery.

Post this, any increase in the size of the plant causes the costs to rise. When the scale of operations becomes too large, the management finds it more difficult to control and coordinate the operations.

Managerial

As the output increases, the firm can apply the division of labor to the management as well. For example, the production manager can look after production, the sales manager can look after sales, etc. When the scale of production increases further, the firm divides each department into sub-departments like sales is divided into advertising, exports, and service.

Thus helps in increasing the efficiency and productivity of the management team since a specialist manages each sub-department. Further, the firm has the option to decentralize decision-making authority enhancing the efficiency further. Therefore, specialized management allows the firm to reduce managerial costs.

However, as the firm increases its scale of operations beyond a certain limit, the management finds it difficult to control and coordinate between departments. This leads to managerial diseconomies.

Commercial

As a firm increases its volume of production, it requires large amounts of raw material and components. Hence, it places a bulk order for such material and components and enjoys discounted pricing for them.

Economies are also achieved during sales. If the sales staff is working under-capacity, then the firm can sell additional output at little extra cost.

Further, as the scale of production increases, the advertising cost per unit fall. Hence, the firm benefits from economies of advertising too. After an optimum level, these economies start becoming diseconomies though.

Financial

When a firm wants to raise finance, a large-scale firm has many benefits like:

- Better security to bankers
- Well-known
- Can raise finance at lower costs, etc.

However, after the optimum scale of production, the financial costs rise faster due to the increased dependence on external finances.

Risk-bearing

A firm enjoys the economies of risk-bearing if it has a large-scale operation with diverse and multi-production capabilities. However, if the diversification increases the economic disturbances rather than covering them, then the risk increases.

4. External economies – these are the benefits that each member firm of the industry accrues due to the expansion of the entire industry.

External Diseconomies and Economies of Scale

External diseconomies and economies of scale are very important to a firm. These are a result of the expansion of output of the entire industry and not limited to an individual firm. They are available to one or more firms in the following forms:

Cheaper Raw materials and Capital Equipment

At times, the expansion of an industry results in new and cheaper sources of raw material, machinery, and other capital equipment. It also results in an increased demand for the various types of materials and equipment required by the industry.

Hence, such materials/equipment can be purchased from other industries on a large scale. This, eventually, leads to a lower cost of production and lower price. Therefore, firms using these materials/equipment get them at lower prices.

Technological External Economies

Usually, when an entire industry expands, new technical knowledge is discovered leading to new and improved machinery for the said industry. This changes the technological coefficient of production and enhances the productivity of the firms in the industry. Hence, the cost of production reduces.

Development of Skilled Labor

As the industry expands, the labor gets accustomed to managing various production processes and learns from the experience. This increases the number of skilled workers which in turn has a favorable effect on the levels of productivity.

Growth of Ancillary Industries

When a certain industry expands, many ancillary industries start specializing in the production of raw materials, tools, machinery, etc. These ancillary industries offer the materials/machinery at a low price.

Similarly, some ancillary industries also start processing industrial waste and create a useful product out of it. Overall, it leads to a lower cost of production.

Better Transportation and Marketing Facilities

An expanding industry, usually, results in better transportation and marketing networks. These aspects help reduce the cost of production in the firms from the industry.

It is important to note that, certain disadvantages can neutralize the advantages of the expansion of industry and cease the external economies of scale. These are external diseconomies. When an industry expands, the demand for certain materials and skilled labor increases.

If these factors are in short supply, then their prices can increase. Further, the geographical concentration of firms from the industry can lead to higher transportation costs, marketing costs, pollution control costs, etc.