**THEORY OF PRODUCTION**

The **theory of production** explains the market supply of goods and services. A firm produces the goods and services supplied in a market and thus, it is an economic agent that produces goods and services (output) using inputs. It is assumed that the firm's primary objective is to maximize profits. In maximizing profits, firms are subject to two constraints: the consumers' demand for their product and the costs of production.

**Some basic concepts in production**

**Production**

It is the process through which factors of production (inputs) are transformed into output.

**Inputs**

These are resources/factors of production used in the production process e.g. land, raw materials, labor, capital etc.

**Output (also called total product)**

Refers to the amount of product a firm can produce with a given set of inputs.

Although many inputs are used in most production processes, we usually abstract from the complexity and talk about just two: capital and labor.

**A Production function**

This is a relationship between the quantity of inputs and the quantity of the product produced. It can be in form of a table, graph, or equation showing the maximum output of the product that can be achieved from a given set of inputs.

Factors of production can be fixed or variable.

**Variable and fixed factors of production**

**Variable factors**

**These are factors of production whose quantity change with changes in output.** In order to increase quantity produced of a commodity, you can increase the quantity of these inputs. Examples of variable factors include; labor, raw materials, energy inputs etc.

**Fixed factors**

These are factors of production whose quantity does not change with changes in output. For example, land, buildings, capital such as the size of the factory, the machinery, computers and other capital equipment.

**The short-run and long-run periods of production**

1. **The short-run period of production**

The short-run period of production is that period over which some factors of production are fixed and others variable.

We normally assume that the quantity of capital inputs (e.g. plant and machinery) is fixed and that production can be changed through changing the quantities for variable inputs such as labor, raw materials, energy inputs etc.

The length of the short run varies from one business to another depending on the time scale that permits a business to alter all of the inputs.

The short run period for a restaurant differs from that of a power generating plant. The point is that for some businesses the short run can be a matter of weeks! Whereas for businesses that require very expensive capital equipment which may take several months or perhaps years to become available, it can be a sizeable period of time.

**Measurement of production and productivity in the short-run**

1. **Total product (TP) or Output (Q)**

This is the amount of a product a firm can produce with a given quantity of inputs. At first output increases at a faster rate with the increase in the variable factor, then at a decreasing rate, reaches a maximum and later it falls when the fixed factor gets over utilized.

**Illustration**

1. **Average product** (AP)

$$AP=\frac{Q}{L}$$

This is output produced per unit of a variable factor. It is obtained when total output is divided by the number of units of the variable factor (e.g. output per worker employed). AP first increases, reaches a maximum and later falls but it will never become zero since output cannot be zero for as long as factors of production are employed. The point at which AP reaches its maximum is the point of diminishing average productivity.

**Illustration**

1. **Marginal product (MP)**

This is the additional output produced by an additional unit of a factor of production. It is the change in total product when an additional unit of the variable factor is employed holding all other factors constant. MP will also first increase, reach a maximum and then diminishing returns set it causing it to fall.
$$MP=\frac{∆Q}{∆L}$$

NOTE: MP is the slope of the TP curve and the maximum point of the MP curve is the point of diminishing marginal productivity.

**Illustration**

The standard economic assumption which affects the shape of the production function or the TP curve and other curves (AP, MP) is the; “law of diminishing marginal returns” Or “The law of variable proportions. This law operates in the short-run period of production and it states that;

**“Holding other factors constant, as more and more units of a variable factor are employed on a fixed factor, output will first increase but beyond a certain point, additional units will result into a decrease in output”.**

At low levels of labor input, the fixed factors of production - land and capital, is still under**utilized** which means that each additional worker will have plenty of capital to use and, as a result, marginal product will rise. Beyond a certain point however, the fixed factors of production become scarce and new workers will not have as much capital to work with. As a result, the marginal productivity of each worker tends to fall.

**Assumptions of the Law:**

1. The state of technology is assumed constant. If there is improvement in the technology, then marginal and average products may rise instead of diminishing.

2. There is presence of some fixed factors. This law does not apply in case all factors are proportionately varied.

4. All units of the variable factor are assumed to be equally productive.

The table below provides a simple numerical example showing the relationship between TP, AP and MP in the short-run reflecting the law of diminishing marginal returns, given two factors of production (K, L).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **K** | **L** | **TP** | **MP** | **AP** |
|  10 | 0 | 0 |   |   |
|  10 | 1 | 2 |   |   |
|  10 | 2 | 10 |   |   |
|  10 | 3 | 30 |   |   |
|  10 | 4 | 40 |   |   |
|  10 | 5 | 45 |   |   |
|  10 | 6 | 48 |   |   |
|  10 | 7 | 49 |   |   |
|  10 | 8 | 49 |   |   |
|  10 | 9 | 48 |   |   |

.

***A graphical relationship of the three curves is illustrated as below;***

TP first rises at an increasing rate, then at a decreasing rate reaches maximum when MP=0 and it falls (MP becomes negative).

When TP is rising, MP is positive.

MP rises until maximum and it falls to zero and negative.

AP rises, reaches maximum when it is intersected by the falling MP. It starts falling but never becomes zero.

**Stages or regions of production**

**First region**

This begins from zero to where MP =AP.

In this region TP, MP, and AP are positive

The fixed factor is not fully utilized and thus every additional unit of a variable factor increases TP. It is irrational to operate in this region.

MP reaches its maximum in this region.

**Second region**

 This starts where MP = AP up to where TP is at maximum and MP=0

 TP is still rising but at decreasing rate.

 MP, AP are still positive but falling

 The fixed factor is fully and efficiently utilized

 This is the economic stage and production should take place in this region.

**Third region**

 This region starts from where TP is at maximum (and MP =0) on wards.

 TP, AP are declining and MP is negative, the fixed factor is over utilized.

 Every additional unit of a variable factor leads to decline in total output

It is also irrational to operate in this region

**LONG RUN PRODUCTION**

In the long run, **all of the factors of production can change** giving a business the opportunity to increase the **scale of its operations.**

The long run production function is generally expressed using an isoquant curve.

**Isoquant**

An isoquant is a curve showing all possible combinations of two inputs that produce the same level of output.

**Illustration of an isoquant**

**Properties of an isoquant**

1. The curve is negatively sloped or downward- sloping. In order to maintain a given level of output, a reduction in the use of one input requires an increase in the use of another input.
2. It is convex to the origin. This means that additional units of X result in smaller and smaller reduction in Y and it is because of diminishing returns.
3. An isoquant that lies above the other shows a higher level of output. This is because it has more quantity of one or both inputs.
4. Between two isoquants are many more other isoquants forming an isoquant map
5. An isoquant does not touch either axes, otherwise this would lead to a corner solution where a particular level of output is produced using one input.
6. Between any two points on an isoquant are many other points.
7. Isoquants do not intersect.

**The optimal combination of inputs in the long-run**

A firm chooses the method of production (a mixture of inputs) that produces a given level of output at the lowest possible cost. To determine the least cost combination of inputs, we use the isocost line.

**An isocost**

An isocost is a curve that shows the various combinations of two inputs that can be purchased with a firm’s given cost outlay or sum of money. All combinations of inputs along the same isocost cost the same.

If we assume that a firm has a cost outlay C to purchase both L and K; and their prices are PL, PK  respectively, the isocost line is presented as;

$C=P\_{L}+P\_{K}$

**Illustration of an isocost**

The isocost curve is downward-sloping which implies that if a firm has a fixed cost outlay, to increase the amount of one factor it has to reduce the amount of the other.

**Equilibrium of a firm in the long-run (profit maximization)**

The traditional goal of a firm is to maximize profit. To achieve this, a firm must produce at a point that maximizes output and minimizes costs. This is a point where a firm’s isocost is tangent to the highest possible isoquant.

**Illustration of the equilibrium condition**

Points on the lowest isoquant are affordable but undesirable because they indicate low levels of output. On the other hand, those on the highest isoquant are desirable since they indicate higher levels of output; however, these points are unaffordable. Points on the isoquant tangent to the isocost show the highest possible and affordable output, particularly, that specific point where the slope of the isocost is tangent to that of the isoquant **(**$ point e^{0}$**).** This is the point where the firm maximizes its profit and thus it is the equilibrium point.