The Theory of the Firm
I. Introduction: A Schematic Comparison of the Neoclassical Approaches to the Studies Between the Theories of the Consumer and the Firm
A. The Theory of Consumer Choice:

B. The Theory of the Firm


## II. A Closer Look at the Competitive Firm and Its Objective: The Neoclassical Perspective

## 1. The Firm and its Objective:

What is a firm? A functional definition of a firm is as follow. A firm is an organizational entity engaged in production activities of goods and services for the sole purpose of attaining a stated economic objective. Examples are Farmer XYZ, family doctor, AT\&T, K-College, and so on.

One central issue in the above definition of a firm is production activity. In broad terms production activity involves the creation of values and such activities encompass virtually all phases of economic activities except consumption. At operational level, production activity entails a series of activities by which resource inputs (labor, capital, raw materials and managerial talents) are transformed using a specified production technique over some period of time into goods and services.

The above definition of the firm also conveys that firms' operate with a clearly stated economic objective. In this respect, the common approach taken is to assume that firms are primarily guided by profit motivation, hence, firms' main objective is to maximize profit.

$$
\begin{aligned}
& \text { Max.: } \Pi=T R-T C, \text { or } \\
& \text { Max.: } \Pi=P_{e} q-C\left(L, K, R, M, T^{*} ; w_{e}, r_{e}, g_{e}, \delta\right)
\end{aligned}
$$

where, $\Pi$ represents profit, $P_{e}$ market price of the output, $q$ is the output level of the firm, C is the cost function of the firm. The cost function of the firm depends on the quantity/quality of the labor (L), capital (K), raw materials (R), and managerial talents, and the prices labor $\left(\mathrm{w}_{\mathrm{e}}\right)$, capital $\left(\mathrm{r}_{\mathrm{e}}\right)$, raw materials $\left(\mathrm{g}_{\mathrm{e}}\right)$ and profit margin $(\delta)$. The variable $T^{*}$ represents the specific production technology under consideration. It is important to note that the analysis so far assumes a single product firm.

Furthermore, once the usual assumption that the firm is price-taker in both the inputs and output markets is recognized, the firm's motive to maximize profit is reduced to a purely technical consideration. This is because since a competitive firm has no control over prices, the only decision variables to the firm are the output level (q) or the inputs mix (L, K, R, M). This being the case, profit depends solely by how efficiently the inputs are used to produce the desired level of output. In other words, more than anything else, technical efficiency determines the profit of a competitive firm.

## 2. The Production Technology of the Firm

## A. Introduction

As discussed above, a competitive firm will have a strong desire to delve into the technical relationship between output (q) and inputs. The formal analysis of the physical relationship between a firm's inputs of productive resources ( $\mathrm{L}, \mathrm{K}, \mathrm{R}, \mathrm{M}$ ) and its output of goods and services is done using a concept known as production function. It is defined as function that maps the maximum attainable output that can be produced from a given combination of inputs used for production holding technology constant. Note the stress put on the phrase "maximum attainable output". This presupposes technical efficiency as being a prerequisite to a production activity of a competitive firm (which is, as discussed earlier, is consistent with the assumption of profit maximization).

$$
\mathbf{q}=f\left(\mathbf{L}, \mathbf{K}, \mathbf{R}, \mathbf{M}, \mathbf{T}^{*}\right)
$$

Typically the production analysis of a firm is done using two distinct time frameworks-the short-run and the long-run. The short-run refers to a period of time so short that the firm can not readily vary some of its inputs. That is, in the short-run some of the inputs of the firm is fixed. On the other hand, long-run refers to a span of time long enough that the firm can readily vary all of its inputs (except technology).

## B. The Long-Run Production Function

For a single-product firm employing only labor and capital (where capital assumed to be a composite input of all productive resources other than labor) as inputs, we can express the long-run production function of the firm as,

$$
q=f\left(L, K, T^{*}\right)
$$

Given this functional relationship, in the long-run what we want to investigate are the following:

- Factor substitution possibilities: to what extent is the firm able to substitute one factor input for another to produce a given level of output. This determines the degree of flexibility that a given firm has in input switching in response to changes in input prices. An important consideration in the firm's effort to minimize its production cost.
- Returns to scale: by how much will output change if the firm vary all its inputs proportionately. This is important because it determines what happens
to the cost of production as a firm attempts to modify the scale or size of its production activity.

The above two considerations suggest that, in the long-run, a competitive firm will attempt to minimize its cost of production (or maximize its profit) by giving careful attention to the input mix and/or the size of its production activity. The next sub-section provides the formal derivation of the equi-marginal condition for optimal input combinations.

## C. The Equi-Marginal Condition for Optimal Input Combination:

The search for the optimal combination of inputs is addressed by looking at the decision of the firm as a constrained optimization problem that can be formulated in the following two ways:
(1) Minimize cost of production subject to a production of a given level output, that is,

- Minimize: $\mathbf{T C}=\mathbf{w}_{\mathrm{e}} \mathbf{L}+\mathbf{r}_{\mathrm{e}} \mathrm{K}$

Subject to,

$$
\mathbf{q}_{\mathbf{0}}=\boldsymbol{f}(\mathbf{L}, \mathbf{K}), \text { where } \mathrm{q}_{\mathrm{o}} \text { is the given level of output. }
$$

(2) Maximize output subject to a given capital outlays (the amount of money available to a firm for the procurement of productive resources), that is,

- Maximize: $q=f(L, K)$

Subject to,

$$
\mathbf{T C}_{\mathbf{0}}=\mathbf{w}_{\mathbf{e}} \mathbf{L}+\mathbf{r}_{\mathbf{e}} \mathbf{K}, \text { where } \mathrm{TC}_{0} \text { is the given outlay. }
$$

Note in both of the above constrained optimization problems, the firm wants to find the input mix ( $\mathrm{L}, \mathrm{K}$ ) that would either minimize the cost of producing a given level of output (case 1) or maximize the output that can be produced from a given capital outlay. As will be evident soon, these two approaches are equivalent in the sense that they yield an identical condition for an optimal mix of inputs. Once the problem is understood this way, the optimal condition for the optimal input mix is derived using both a graphical and mathematically approaches as demonstrated below.

## A Mathematical Approach to Cost Minimization: (Optimal Input Combinations)

Max: $Z=f(L, K)+\lambda\left(w_{e} L+r_{e} K-T_{0}\right)$
(L, K, 入)

$$
\begin{array}{lll}
\partial \mathrm{Z} / \partial \mathrm{L}=\partial \mathrm{f} / \partial \mathrm{L}+\lambda \mathbf{w}_{\mathrm{e}}=0 \quad & \Rightarrow & \mathrm{MP}_{\mathrm{L}} / \mathbf{w}_{\mathrm{e}}=-\lambda \\
\partial \mathrm{Z} / \partial \mathrm{K}=\partial \mathrm{f} / \partial \mathrm{K}+\lambda \mathbf{r}_{\mathrm{e}}=0 \quad & \Rightarrow & \mathrm{MP}_{\mathrm{K}} / \mathbf{r}_{\mathrm{e}}=-\lambda \\
\partial \mathrm{Z} / \partial \lambda=\mathbf{w}_{\mathrm{e}} \mathrm{~L}+\mathbf{r}_{\mathrm{e}} \mathrm{~K}-\mathbf{T C}_{0}=\mathbf{0} &
\end{array}
$$

Accordingly, equi-marginal principles for producing the maximum output from a given capital outlay is reached when the inputs ( $\mathrm{L}, \mathrm{K}$ ) are combined in such as way that the following two conditions are met simultaneously:

- $\mathbf{M P}_{\mathbf{L} / \mathbf{w}_{\mathbf{e}}}=\mathbf{M} \mathbf{P}_{\mathbf{K}} / \mathbf{r}_{\mathrm{e}}$, and
- $\mathbf{w}_{\mathrm{e}} \mathrm{L}+\mathbf{r}_{\mathrm{e}} \mathrm{K}-\mathrm{TC}_{\mathbf{0}}=\mathbf{0}$.

In words, the above conditions suggest that the firm is using an optimal combination of labor and capital when it operates in such a way that the marginal productivity of the last dollar spent for each input are equal. Furthermore, that this condition is met while the firm is operating with full utilization of its allotted budget or capital outlay.

## A Graphic Approach to Optimal Input Combinations

The graphic approach to the optimal input choice of a competitive firm employs a similar approach to that used in determining the optimal output choice of a consumer. In the case of the consumer choice, an indifference curve and a budget line were the two principles concepts used to graphically analyze the optimal output choice. For the firm, the analogous concepts will be an isoquant curve (a measure of the firm's technical possibilities for factor substitutions) and an isocost line (the boundary specifying resource limitations) are thoroughly discussed below.

## An Isoquant and the Measure of the Firm's Potential for Factor Substitution:

An isoquant is formally defined as a locus of all technically efficient combination of inputs of labor and capital that will produce a given level of output. Algebraically, the isoquant of a firm can be expressed as,

$$
\mathbf{q}_{\mathbf{0}}=\boldsymbol{f}(\mathbf{L}, \mathbf{K}), \text { where } \mathrm{q}_{\mathrm{o}} \text { is the given level of output. }
$$

Normally, as shown in the graph below, an isoquant is negatively sloped and convex from the origin.


The fact that an isoquant is negatively sloped indicates the existence of factor substitution possibilities. That is, conceptually, the given level of output, $q_{\mathrm{o}}$, can be produced by infinitely different combinations of labor and capital; and along a given isoquant curve an increase in the use of one input (for example labor) is accompanied by a decrease in the use of the other input (capital). The rate at which labor and capital are substituted for one another along a given isoquant curve is called the marginal rate of technical substitution, and it is measured by the slope of the isoquant curve at a point. Formally, the marginal rate of technical substitution of labor for capital ( $\mathrm{MRTS}_{\mathrm{L} / K}$ ) is defined as the amount by which the input of capital can be reduced when one (often a small) extra unit of labor is used so that output remains constant.

$$
\mathrm{MRTS}_{\mathrm{L} / \mathrm{K}}=-\mathrm{dK} / \mathrm{dL}
$$

The fact the isoquant is convex from the origin suggests that along a given isoquant, the $\mathrm{MRTS}_{\mathrm{L} / \mathrm{K}}$ tends to decrease as an increasing amount of labor is substituted for capital (see the slopes of the isoquant at points A and B). That is, with less and less capital, labor exceedingly ceases to be a good substitute of capital. Thus, convexity of an
isoquant implies that, as a general rule, the marginal rate of technical substitution of labor for capital tends to decrease as an increasing amount of labor is substituted for capital. This is known as the law of diminishing marginal rate of technical substitution.

This law can be easily verified if one demonstrates that the slope of the isoquant at any point is the ratio of the marginal product of labor and capital $\left(\mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{k}}\right)$ at that point. As shown below, this can be done by taking the total derivative of the underlying isoquant function, and rearranging terms until the desired result is obtained.

$$
d q_{o}=(\partial f / \partial L) d L+(\partial f / \partial K) d K .
$$

Since $\mathrm{dq}_{\mathrm{o}}=0$ for a movement along a given isoquant curve,

$$
0=(\partial f / \partial L) d L+(\partial f / \partial K) d K .
$$

Furthermore, by rearranging terms, the above equation can be rewritten as,

$$
-d K / d L=(\partial f / \partial L) /(\partial f / \partial K)=M_{L} / M P_{K}
$$

Thus, along a given isoquant, as we move from left to right (indicating a successive increase in labor utilization) the tendency is, consistent with the law of diminishing marginal products, for the marginal product of labor to decrease and the marginal product of capital to increase. Thus, as we move along a given isoquant, the ratio of the $\mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{K}}$ (which as demonstrated above is also the measure of the slope of the isoquant at a point) monotonically declines--see the slopes of the isoquant at points A and B).

## An Isocost: The Budget Constraint Function of the Firm

Like any other economic entities, a firm operates in a world of scarcity. At a point in time firms have a limited amount of resources (money) to purchase the productive inputs they need to produce output. In a world with only two productive inputs, the resource constraint of a firm can be expressed by the following equation:

$$
\mathbf{T C}_{\mathbf{0}}=\mathbf{w}_{\mathbf{e}} \mathbf{L}+\mathbf{r}_{\mathbf{e}} \mathbf{K}, \text { where } \mathrm{TC}_{0} \text { the total budget (outlay). }
$$

Alternatively, this above equation can be written as:

$$
K=T_{0} / \mathbf{r}_{\mathrm{e}}-\left(\mathbf{w}_{\mathrm{e}} / \mathbf{r}_{\mathrm{e}}\right) \mathbf{L}
$$

The above equation, indeed, indicates the equation for the isocost line of the firm. An isocost line portrays the various alternative combination of labor and capital inputs that a
firm can purchase, given the resource prices and the limited budget of the firm. Graphically, the isocost is presented as shown below.


## The Equi-marginal Principle for Cost Minimization: A Graphical Approach

Once the concepts of isoquant and isocost are understood as discussed above, the graphical illustration of the optimal input combination is rather straightforward as shown below.


According the above graph, the firm will be using the combination of labor and capital corresponding to point $\mathrm{A}\left(\mathrm{L}_{\mathrm{A}}, \mathrm{K}_{\mathrm{A}}\right)$. At this point the equi-marginal condition for cost minimization is achieved. Any deviation away from this point will entail a higher cost for producing the same level of output or the production of a lower level of output for the same cost. For example, if the firm chooses to use the input combination indicated by point B , it will end up producing a lower level of output ( 25 instead of 30 ) and without a reduction in cost. In fact, at point $B$, the $\mathbf{M R T S}_{\mathbf{L} / \mathbf{K}}=\mathbf{M P}_{\mathbf{L}} / \mathbf{M} \mathbf{M}_{\mathbf{K}}>\mathbf{w}_{\mathbf{e}} / \mathbf{r}_{\mathbf{e}}$ indicating that the relative productivity of labor is higher than its relative cost. Hence, labor is under utilized. Thus, the firm should use more labor and less capital until equilibrium is achieved again. How does the movement from B to A achieved? Assuming input prices are not changing, as the firm moves towards A, the marginal product of labor will continue to decline while the marginal product of capital increases (by the law of diminishing marginal products). This process will cause the ratio of the marginal product of labor and capital (or the marginal rate of technical substitution of labor for capital) to decline until equilibrium is restored at point A .

One last issue we need to address is this. Suppose the firm is at equilibrium and all of a sudden the price for labor increases. Will the optimal input combination change if the firm still desires to produce the same level of output? In general, the answer to this question is affirmative; that is, holding everything else constant, a change in input price of a factor of production will cause the alteration of input mix the input would like to employ to produce a given level of output. Suppose the original equilibrium condition was,

$$
\mathbf{M} \mathbf{P}_{\mathbf{L}} / \mathbf{M} \mathbf{P}_{\mathrm{K}}=\mathbf{w}_{\mathbf{0}} / \mathbf{r}_{\mathbf{0}}
$$

Suppose, the wage rate increases from $\mathrm{w}_{\mathrm{o}}$ to $\mathrm{w}_{1}$, then at old equilibrium condition,

$$
\mathbf{M P}_{\mathbf{L}} / \mathbf{M P} \mathbf{P}_{\mathrm{K}}<\mathbf{w}_{1} / \mathbf{r}_{\mathbf{o}}
$$

This is, in fact, a disequilibrium situation. In this case, to reestablish equilibrium, the firm needs to decrease labor and increase capital. Given the firm is constrained to produce the same level of output, by how much the firm's cost will increase depends on the ease by which the firm is able to substitute capital for labor-the rate factor substitutions at the margin. The easier it is to substitute capital for labor, the lower will be the increase in the cost of production as a result of an increase in the market price for labor.

## Illustrative Problems

(a) Given that $\mathbf{q}=\mathbf{5} \mathbf{L}^{1 / 2} \mathbf{K} ; \mathbf{w}_{\mathbf{e}}=\$ 15$ and $\mathbf{r}_{\mathrm{e}}=\$ 50$; and $\mathbf{T C}_{\mathbf{0}}=$ $\$ 1125$, using the Lagrange optimization technique show that output is maximized when 25 units of labor and 15 units of capital are used.
(b) Show the result in part (a) graphically. Note that given that 25 units of labor and 15 units of capital are used, the level of output that corresponds to this level of input usage is $\mathrm{q}^{*}=5(25)^{1 / 2} 15=375$. Accordingly, the relevant isoquant equation for the question under consideration will be:

$$
K=75 L^{-1 / 2}
$$

(c) From the isoquant equation above, derive the general equation for the marginal rate of substitution. Use this equation to verify that at the optimal input combination $(25,15)$ the marginal rate of technical substitution of labor for capital is 0.3 . What does this number inform us?
(d) For the above isoquant equation, compute the marginal rate of technical substitution of labor for capital as the use of labor increases from 4 to 16,25 , and 36 ? Is the result of this exercise consistent with the law of diminishing marginal rate of substitution? Explain.
(e) What input combination of labor and capital should be used if the price for labor increases from $\$ 15$ to $\$ 25$ and the firm still wants to produce 375 units of output? Use the Lagrange method to solve this problem. (Answer: $\mathrm{L}=17.8, \mathrm{~K}=17.8$, and $\mathrm{TC}=\$ 1335$ ).

Note: Because of the increase in the relative cost of labor, the firm will end up using more capital (17.8 versus 15) and less labor ( 17.8 versus 25), and the total cost of producing the 375 units of output will increase from $\$ 1,125$ to $\$ 1,335$. The increase in the total cost depends on the ease by which labor and capital are substituted and the magnitude of the increase in the input price. To see this, note that the cost is increased by $\$ 210(\$ 1,335-\$ 1,125)$. This can be attributed to: (i) the increase in capital of 2.8 units which costs $\$ 140$ ( $2.8 \times 50$ ), and (ii) and the change in labor cost which is equal to $\$ 70[(\$ 25 \times 17.8)-(\$ 15 \times 25)]$. Or,

$$
\Delta T C=\Delta K r_{e}+w_{1} \mathbf{L}_{\mathbf{1}}-\mathbf{w}_{\mathbf{0}} \mathbf{L}_{\mathbf{0}}=\left(\mathbf{r}_{\mathbf{e}} \mathbf{K}_{\mathbf{1}}+\mathbf{w}_{\mathbf{1}} \mathbf{L}_{\mathbf{1}}\right)-\left(\mathbf{r}_{\mathrm{e}} \mathbf{K}_{\mathbf{0}}+\mathbf{w}_{\mathbf{0}} \mathbf{L}_{\mathbf{0}}\right)
$$

## Cobb-Douglas Type Production Functions and Their Properties

The production function of a firm can be expressed in several explicit algebraic forms. One of the most widely used functional forms in modern theoretical and empirical analysis for the purpose of production analysis is the Cobb-Douglas type production function. For two input cases, it takes the following form:

$$
\mathbf{q}=\mathbf{A L} \mathbf{L}^{\alpha} \mathbf{K}^{\beta}
$$

where $\mathrm{A}, \alpha$, and $\beta$ are parameters.
The popularity of the Cobb-Douglas type production function stems from the fact that it greatly simplifies the analysis of certain basic production analysis-computational ease. For example, as shown below, the equations for the marginal products of labor and capital, the marginal rate of technical substitution, and the expansion path depend on the following: the production and price parameters ( $\alpha, \beta, w_{e}, r_{e}$ ), the outputinput and the two inputs ratios.
i) Marginal product of labor:

$$
\mathrm{MP}_{\mathrm{L}}=\partial \mathrm{q} / \partial \mathrm{L}=\boldsymbol{\alpha}(\mathbf{q} / \mathrm{L})
$$

ii) Marginal Product of Capital:

$$
\mathrm{MP}_{\mathrm{K}}=\partial \mathrm{q} / \partial \mathrm{K}=\boldsymbol{\beta ( \mathbf { q } / \mathbf { K } )}
$$

iii) Marginal Rate of Technical Substitution of Labor for Capital:

$$
\mathrm{MRTS}_{\mathrm{L} / \mathrm{K}}=\mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{K}}=\boldsymbol{\alpha} / \boldsymbol{\beta}(\mathbf{K} / \mathbf{L})
$$

iv) The Equation for the Expansion Path:

$$
\begin{aligned}
& \mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{K}}=\mathrm{w}_{\mathrm{e}} / \mathrm{r}_{\mathrm{e}} \\
& \alpha / \beta(\mathrm{K} / \mathrm{L})=\mathrm{w}_{\mathrm{e}} / \mathrm{r}_{\mathrm{e}} \\
& \mathrm{~K}=\left[(\boldsymbol{\beta} / \boldsymbol{\alpha})\left(\mathbf{w}_{\mathbf{e}} / \mathbf{r}_{\mathrm{e}}\right)\right] \mathbf{L}
\end{aligned}
$$

## III. The Long-run Production and Cost Functions of A Competitive Firm

## 1. Introduction:

In this section, attempts will be made to examine how a competitive firm's decision to change the scale of its production activities (expansion or contraction its operations) would impact on the cost of production. The central issue here is to observe what happen to the productivity of resource inputs as the size of the firm increases. Of course, the changes in resource productivity will ultimately trigger corresponding changes the average and marginal costs of production.


The following pedagogical steps are used to formally analyze the causal relationships between the size of the firm and the productivity and, therefore, the unit costs of its operation.

- First, since what is involved is a change in the size or the scale of the firm, the issues of interest typically deal with long-run production and cost analysis.
- Second, a change in the scale or the size of the firm is identified by a proportionate change in all inputs.
- Third, from a purely physical viewpoint, one of the central analyses in this section deals with what is known as returns to scale--what happen to output as a firm's resource inputs are allowed to vary proportionately.
- Fourth, since it is in the interest of any firm to minimize its cost, what specific path should a firm take in order to assure the attainment of this particular end-the expansion path.
- Fifth, the last issue deals with what happen to unit costs of output as a firm expands its operation-economies or diseconomies of scale.


## 2. Returns Scale:

As a general concept, returns to scale refers to how output level of a firm changes (behaves) when the inputs (all inputs) are changed proportionately. For example, what happen to output if all inputs are doubled? Will output be less than or more than doubled? Or, will output remain unchanged? In general, three possible outcomes could be observed:

- Constant Returns to Scale: Multiplying all inputs by some positive constant results in a multiplication of output by the same constant. For example, doubling all inputs will double output.
- Decreasing Returns to Scale: If all inputs are multiplied by a constant greater than one, output is multiplied by a smaller positive constant. For example, if all inputs are doubled, output is less than doubled.
- Increasing Returns to Scale: Multiplying all inputs by a constant greater than one results in a multiplication of output by a larger positive constant. For instance, doubling all inputs will more than double output.


## Formal Analysis of Returns to Scale:

Suppose the long-run production function of a firm is expressed functionally as follows:

$$
\mathbf{Q}=f(\mathbf{L}, \mathbf{K}) .
$$

This production function is said to be a homogeneous function if,

$$
\mathbf{Q}=f(\lambda \mathbf{L}, \lambda \mathbf{K})=\lambda^{\mathrm{t}} f(\mathbf{L}, \mathbf{K})=\lambda^{\mathrm{t}} \mathbf{Q}
$$

where " $\lambda$ " is a positive constant and " $t$ " is the degree of homogeneity.
Thus, if that the production function is homogeneous as described above, if:

- $\mathbf{t}=\mathbf{1 . 0}$, the production function is said to be exhibiting constant return to scale.
- $\mathbf{t}<\mathbf{1 . 0}$, the production function is said to exhibit decreasing returns to scale.
- $\mathbf{t}>\mathbf{1 . 0}$, the production function is said to exhibit increasing returns to scale.


## Exercise:

i) Using the above method, verify that for a Cobb-Douglas type production function the returns to scale is the sum of the exponents of labor and capital, $\alpha$ $+\beta$.
ii) Evaluate the returns to scale for the production function:

$$
\mathrm{q}=18 \mathrm{LK}^{2}+7.5 \mathrm{~L}^{2} \mathrm{~K}-\mathrm{L}^{3} .
$$

What would the proportionate increase in output if the inputs (L, K) are allowed to double?

## 3. The Expansion Path of the Firm

In this section what want to examine is the firm's decision on resource utilization (how much more of labor and capital to use) should market conditions warrant a longterm expansion or contraction of output. Since the firm is assumed to be free to vary all its inputs as needed (long-run), it is natural to expect that the decision to either expand or reduce output will be done in such a way that the least-cost combination of resource inputs are utilized. In other word, a competitive firm in its decision to increase or decrease output in the long-run follows a particular path where the equi-marginal condition for least cost combination of resource utilization is always met. This road map for output expansion in the long-run is called the expansion path.

Formally, the expansion path is defined as the locus of optimal input combinations for each possible rate of output given that the prices for inputs (labor and capital) are held constant. In other words, it shows the optimal combinations of inputs (labor and capital) which the firm should use as it expands its long-run rate of output.


## Exercise:

Demonstrate that, for a Cobb-Douglas type production function, the expansion path is linear and it slope depends on four parameters $\left(\alpha, \beta, w_{e}, r_{e}\right)$.

## 4. Derivation of the Long-run Cost Function of a Competitive Firm:

In this section, given market prices for inputs ( $\mathrm{w}_{e}, \mathrm{r}_{\mathrm{e}}$ ) and the firms desire to minimize cost, an attempt will be made to show the link between the production technology (returns to scale) and the long-run cost function of a competitive firm. This involves the interaction of three key variables: the underlying production function of the firm, the market prices of labor and capital, and the equation of the expansion path. This will be formally demonstrated using a Cobb-Douglas type production function.

Case I: Constant Returns to Scale and Constant Cost


| $=\mathbf{4 L}$ |  |  |  |  |  |
| :--- | ---: | ---: | :--- | ---: | ---: |
| $\underline{\mathrm{L}}$ | $\underline{\mathrm{K}}$ | $\underline{\mathbf{q}}$ | $\underline{\text { TC }}$ | $\underline{\mathbf{A C}}$ | $\underline{\text { MC }}$ |
| 1 | 4 | $\mathbf{1 0}$ | $\mathbf{\$ 8}$ | $\mathbf{\$ 0 . 8}$ | $\mathbf{\$ 0 . 8}$ |
| 2 | 8 | $\mathbf{2 0}$ | $\mathbf{1 6}$ | $\mathbf{0 . 8}$ | $\mathbf{0 . 8}$ |
| 4 | 16 | $\mathbf{4 0}$ | $\mathbf{3 2}$ | $\mathbf{0 . 8}$ | $\mathbf{0 . 8}$ |
| 6 | 24 | $\mathbf{6 0}$ | $\mathbf{4 8}$ | $\mathbf{0 . 8}$ | $\mathbf{0 . 8}$ |
| 8 | 32 | $\mathbf{8 0}$ | $\mathbf{6 4}$ | $\mathbf{0 . 8}$ | $\mathbf{0 . 8}$ |
| 10 | 40 | $\mathbf{1 0 0}$ | $\mathbf{8 0}$ | $\mathbf{0 . 8}$ | $\mathbf{0 . 8}$ |
| 12 | 48 | $\mathbf{1 2 0}$ | $\mathbf{9 6}$ | $\mathbf{0 . 8}$ | $\mathbf{0 . 8}$ |

Clearly, as indicated by the entry of average and marginal costs in the above table, constant cost industry results in a long-run constant and average costs.



Case II: The Association of Decreasing Returns to Scale and Increasing Long-run Average and Marginal Costs:

Let us modify the production function so that it exhibits decreasing returns to scale. The new production function is given to be:

$$
\mathrm{q}=5 \mathrm{~L}^{1 / 4} \mathrm{~K}^{1 / 4}
$$

| $\underline{\mathrm{K}=\mathbf{4 L}}$ | $\underline{\mathrm{L}}$ | $\underline{\mathbf{( q )}}$ | $\underline{\mathbf{T C}}$ | $\underline{\mathbf{A C}}$ | $\underline{\mathbf{M C}}$ |
| :--- | :--- | :--- | :--- | ---: | ---: |
| 2 | 8 | $\mathbf{1 0 . 0}$ | $\mathbf{\$ 1 6}$ | $\mathbf{\$ 1 . 6 0}$ | -- |
| 4 | 16 | $\mathbf{1 4 . 1}$ | $\mathbf{3 2}$ | $\mathbf{2 . 2 9}$ | $\mathbf{\$ 4 . 0 0}$ |
| 6 | 24 | $\mathbf{1 7 . 3}$ | $\mathbf{4 8}$ | $\mathbf{2 . 7 7}$ | $\mathbf{4 . 8 5}$ |
| 8 | 32 | $\mathbf{2 0 . 0}$ | $\mathbf{6 4}$ | $\mathbf{3 . 2 0}$ | $\mathbf{5 . 9 0}$ |
| 10 | 40 | $\mathbf{2 2 . 6}$ | $\mathbf{8 0}$ | $\mathbf{3 . 5 4}$ | $\mathbf{6 . 1 5}$ |
| 12 | 48 | $\mathbf{2 4 . 5}$ | $\mathbf{9 6}$ | $\mathbf{3 . 9 2}$ | $\mathbf{8 . 4 2}$ |

Again, by looking the entries of the average and marginal costs in the above table it is evident that decreasing returns to scale is associated with increasing unit costs.



Case III: The Association of Increasing Returns to Scale and Decreasing Longrun Average and Marginal Costs

In this case let the modified production function be represented by:

$$
\mathrm{q}=5 \mathrm{~L}^{1 / 2} \mathrm{~K}
$$

| $=8 \mathrm{~L}$ |  |  |  |  |  |
| :--- | :---: | :---: | :--- | :---: | :---: |
| $\underline{\mathrm{~L}}$ | $\underline{\mathrm{~K}}$ | $\mathbf{9}$ | $\underline{\mathbf{T C}}$ | $\underline{\mathbf{A C}}$ | $\underline{\mathbf{M C}}$ |
| 2 | 16 | 113.0 | 24 | $\$ 0.21$ | -- |
| 4 | 32 | 320.0 | 48 | 0.15 | $\$ 0.12$ |
| 6 | 48 | 588.0 | 72 | 0.12 | 0.09 |
| 8 | 64 | 919.0 | 96 | 0.10 | 0.073 |
| 10 | 80 | 1265 | 120 | 0.095 | 0.069 |
| 12 | 96 | 1663 | 144 | 0.087 | 0.06 |

Evidently, when the production function exhibits increasing returns to scale, long-run average and marginal costs are declining monotonically-decreasing unit costs (see below)

q


## General Features of Long-run Cost Function: Economies and Diseconomies

 of ScaleEconomists hypothesize that, in general, firms initially encounter increasing returns to scale as they attempt to expand their operation. That is, until certain level of output is reached ( $\mathrm{q}_{0}$ in the Figure below), expansion is associated with increased productivity of resources or decreasing average cost. This phenomenon in economics is called economics of scale; that is, there is a benefit to be gained by increasing plant size (scale) of operation. The reason for this are several, and chief among them are: division of labor, capacity utilization of previously under used resources (space, machines, etc.), and better use of byproducts (which reduces waste in resources).

This process will not continue indefinitely. When a firm reaches to a certain size ( $\mathrm{q}^{*}$ in the Figure below), economies of scale would seize to materialize. This situation may be followed by a steady-state situation where unit cost remains constant for a certain range of output ( $\mathrm{q}^{*}$ to $\mathrm{q}^{* *}$ in the Figure below). After a certain level of output ( $\mathrm{q}^{* *}$ in the Figure below) is reached, the cost per unit of output starts to increase as the firm attempts to further expand. This situation is called diseconomies of scale and it reflect to the loss of efficiency or productivity associated with getting big. The primary sources for this perceived inefficiency are: bureaucratic mess - communication lines will be disrupted, and the problem with managing big organization.


## Short-run Production and Cost Functions:

## An Outline

## 1. The Link of Long-run to Short-run Cost Functions

Analyze this using the following graph:


What the above graph shows is that when capital is assumed to be fixed at certain level, $\mathrm{K}^{*}$, a firm can not any more be able to follow its expansion path as it attempts to expand output due to changes in demand for its product. For example, for output levels below $\mathrm{q}_{2}$, for any level of output it wishes to produce, the firm will be put in bind to use input combinations that under utilize labor and over utilize capital (that is, $\mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{K}}>\mathrm{w}_{\mathrm{e}} / \mathrm{r}_{\mathrm{e}}$ ). Similarly, for output level above $\mathrm{q}_{2}$, the firm will be forced to use input combinations that increasingly indicates over an over utilization of labor and under utilization of capital (that is, $\mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{K}}$ $<\mathrm{w}_{\mathrm{e}} / \mathrm{r}_{\mathrm{e}}$. Given that the capital is fixed a level of $\mathrm{K}^{*}$, only at $\mathrm{q}_{2}$ level of output is the firm producing using input combination that is at its expansion path. The implication of this is that, for any given size of capital that is held constant (plant size), there is a single point where the short-run and long-run costs of the firm are the same. In any other situations, the production cost of the firm is higher in the short-run than in the long-run. This conclusion should not be surprising given the apparent loss of flexibility in input usage in the short-run.

## 2. Basic Features of Short-run Cost Function

To begin with it is important to note that in the short-run, firms have two categories of costs-fixed and variable costs. Furthermore, short-run costs are analyzed in terms of both total and unit costs. In this section we will closely examine the salient features of both the total and unit short-run cost functions.

Families of Short-run Total Costs:
The fixed costs arise from the fixed inputs of a firm. The amount of the fixed cost depends on the quantity of each of the various fixed inputs and the respective price paid for them. Example may include top management salaries, property taxes, interest on borrowed money, depreciation charges, rent on office space, the implicit costs of equity capital, and insurance premiums.

Formally, total fixed cost (TFC) may be defined as the sum total of the explicit costs of all the fixed inputs plus the implicit costs associated with the firm's equity capital. In our two inputs world, the fixed cost is associated with the cost of capital. The following some of the salient features of total fixed cost:

- Total fixed cost is constant unless the prices of the fixed inputs change.
- Total fixed cost does not depend on output. It will remain the same regardless of the size of the firm's output.
- Thus, total fixed cost is a cost that can not be avoided by a firm even when production of output is reduced to zero. Graphically, this is depicted as follows:

- On the other hand, while the total fixed cost does not vary with output, it is important to not that the average fixed cost (AFC) declines monotonically with increases in output. Furthermore, the decline of
the average fixed cost is more pronounced at the early stage in production and becomes insignificant as output continues to increase. (see the graph below)


Output (q)
On the other hand, the variable costs arise from the short-run variable costs. The short-run total variable cost (TVC) is the sum of the amount a firm spends for variable inputs employed in the production process. Examples include payroll expenses, raw materials outlays, power and fuel charges, and transportation costs. In our two inputs world, the variable cost is primarily identified with the cost for labor. Since in the short-run a firm modifies its output rate by changing the use of its variable inputs, variable costs depend on and vary with the quantity of output. As will be explained shortly, the shape of the conventional short-run total variable cost is expected to be similar with the one that we have examined for the long-run cost-sigmoid shape. Accordingly, it indicates that initially the total variable cost increases at a decreasing rate, and beyond certain level of output is attained (see q${ }^{*}$ in the Figure below) it starts to increase at an increasing rate. The explanation for this will be evident when the link between the short-run cost and production functions are discussed later.


Finally, the total cost in the short-run is composed of the total variable and fixed costs (TC = TVC + TFC $)$. Since the fixed is invariant to output, the shape of the total cost and total variable cost are identical. In contrast to the total variable cost, the total cost is simply scaled upward by the amount of the fixed cost.

## Families of Short-run Unit Costs:

- Families of short-run unit costs and their pertinent features: definitions and geometry
- What explains the structure of short-run cost function?
- The link between the short-run cost and production functions.
- Mathematical specification of the cost functions: The quadratic form

The Equi-Marginal Condition for Profit Maximization and the Supply Curve of a Competitive Firm

- The nature of the revenue function of a competitive firm; $\mathrm{MR}=\mathrm{P}_{\mathrm{e}}$.
- Condition for profit maximizing (loss minimizing) output, $\mathrm{MR}=\mathrm{MC}$ and MC is rising.
- The decision rule for a competitive firm to shut-down in short-run: $P_{e}$ < AVC.
- The formal derivation of the short-run supply curve of a competitive firm

- Factors affecting the supply curve of a competitive firm
- An explanation for why the short-run supply curve is upward sloping
- The market supply curve of a competitive firm


## The Demand for Factors of Production: The Case of the Single Variable Input

- Profit maximization and the input demand of :
- $\mathbf{M R}=\mathbf{P}_{\mathrm{e}}=\mathbf{M C}$, but we know that $\mathrm{MC}=\mathrm{w}_{\mathrm{e}} / \mathrm{MP}_{\mathrm{L}}$

Thus, at a profit maximizing level of output,

- $P_{e}=w_{e} / M P_{L}$, or
- $\mathbf{P}_{\mathbf{e}} \mathbf{M P}_{\mathbf{L}}=\mathbf{w}_{\mathrm{e}}$, or $\mathbf{M R P}_{\mathbf{L}}=\mathbf{w}_{\mathbf{e}}$. (Note for this condition to make sense the $M P_{L}$ has to be positive and a declining function of labor.)
- Show that for a single input, the demand curve of labor for a firm is the negatively sloped part of the marginal revenue product (use data).

- Factors affecting the demand for a variable input:
- The price of the output, and, therefore, the demand for the product.
- The price of other inputs (substitutes and complements).
- The price of the input under consideration, for example, the wage rate for labor.
- The productivity of the input under consideration
- What happens when we have more than one variable inputs to consider?


## Long-run Equilibrium of a Competitive Industry:

Basic Facts or Assumptions:

- Economic Concept of costs- implicit, explicit and opportunity costs
- Cost is minimized in the long-run; the firm operates along its expansion path
- Freedom of entry and exit
- Firms are price takers
- Long-run equilibrium is achieved when each firm attains zero economic profit (where economic profit = total revenue - total implicit and explicit costs). This is attained when the following condition is met:

$$
\mathbf{P}_{\mathrm{e}}=\mathrm{LRAC}=\mathrm{LRMC} .
$$



- Why zero economic profit is a normal profit?
- Normal profit covers only the opportunity costs of owners' equity.
- Why competitive markets in long-run equilibrium are efficient!
- Outputs are produced using the optimal plant size or at a level where the average cost is at its minimum (see the long-run equilibrium condition).
- Firms are allowed to make a normal profit and nothing more.
- Factors of productions are compensated or rewarded according their marginal productivity.

$$
\left.\begin{array}{l}
\mathrm{P}_{\mathrm{e}} \mathrm{MP}_{\mathrm{L}}=\mathrm{w}_{\mathrm{e}} \\
\mathrm{P}_{\mathrm{e}} \mathrm{MP}_{\mathrm{K}}=\mathrm{r}_{\mathrm{e}}
\end{array}\right\} \Rightarrow \quad \mathrm{MP}_{\mathrm{L}} / \mathrm{MP}_{\mathrm{K}}=\mathrm{w}_{\mathrm{e}} / \mathrm{r}_{\mathrm{e}} .
$$

- The market provides output to consumers at the lowest possible price.

$$
\mathrm{P}_{\mathrm{e}}=\mathrm{LAC}(\text { at its minimum })
$$

- Therefore, consumers' surplus is maximized.
- Other Outcomes: The powerful role of market prices

1. Efficiency in Consumption:

- $\quad \mathrm{MRS}_{\mathrm{x} / \mathrm{y}}=\mathrm{P}_{\mathrm{x}} / \mathrm{P}_{\mathrm{y}}$
- $\left.\mathrm{MRS}_{\mathrm{x} / \mathrm{y}}=\mathrm{P}_{\mathrm{x}} / \mathrm{P}_{\mathrm{y}}\right\} \quad \mathbf{M R S}_{\mathrm{x} / \mathrm{y}}=\mathbf{M R S}_{\mathrm{x} / \mathrm{y}}=\mathbf{M R S}_{\mathrm{x} / \mathrm{y}}=\ldots=\mathbf{P}_{\mathrm{x}} / \mathbf{P}_{\mathbf{y}}$.
- $\mathrm{MRS}_{\mathrm{x} / \mathrm{y}}=\mathrm{P}_{\mathrm{x}} / \mathrm{P}_{\mathrm{y}}$

Markets assure that the relative values of any pair of goods are the same for all consumers.
2. Efficiency in Production (use of resources among firms)

- $\operatorname{MRTS}_{L / K}=w_{e} / r_{e}$
- MRTS $_{L / K}=w_{e} / r_{e}$

MRTS $_{\mathrm{L} / \mathrm{K}}=\mathrm{MRTS}_{\mathrm{L} / \mathrm{K}}=\mathrm{MRTS}_{\mathrm{L} / \mathrm{K}}=\mathbf{w}_{\mathrm{e}} / \mathbf{r}_{\mathrm{e}}$

- MRTS $_{L / K}=w_{e} / r_{e}$」

Markets assure that factors of production (labor and capital) are used with equal degree of efficiency by all firms who are engaged in the production of goods and services.

## 3. Efficiency in Output

$\left.\begin{array}{ll}\text { - } & \mathrm{P}_{\mathrm{x}}=\mathrm{MC}_{\mathrm{x}} \\ \text { - } & \mathrm{P}_{\mathrm{y}}=\mathrm{MC}_{\mathrm{y}}\end{array}\right\} \quad \mathrm{P}_{\mathrm{x}} / \mathrm{P}_{\mathrm{y}}=\mathrm{MC}_{\mathrm{x}} / \mathrm{MC}_{\mathrm{y}}$
But from (1) we know that

$$
\mathrm{P}_{\mathrm{x}} / \mathrm{P}_{\mathrm{y}}=\mathrm{MRS}_{\mathrm{x} / \mathrm{y}}
$$

Therefore, at equilibrium,

$$
\mathbf{M R S}_{\mathbf{x} / \mathbf{y}}=\mathbf{M C}_{\mathbf{x}} / \mathbf{M C}_{\mathbf{y}}=\mathbf{P}_{\mathbf{x}} / \mathbf{P}_{\mathbf{y}} .
$$

In a competitive market economy, the relative value of any pair of goods (say $X$ and Y) to society (or consumers) reflects the relative cost of producing the goods under consideration. That is, relative prices are good measures of relative scarcity.

