Theory of consumer Behaviour

- Utility
- Total utility
- Marginal utility
- Law of diminishing marginal utility

Harish

Theory of consumer Behaviour

- consumer theory shows how individuals make choices given their income and the prices of goods and services and helps us to understand how individuals' tastes & incomes influence the demand curve.

- The theory of consumer choice is the branch of microeconomics that relates preferences to consumption expenditures and to consumer demand curves.

UTILITY

- To satisfy our need we consume goods & service that provides benefit. this satisfaction or benefit is called ~~benefit~~ Utility.
- Utility means usefulness. In Economics utility is defined as the power of a commodity or a service to satisfy a human want. Utility is a subjective or psychological concept.

Total Utility

- Total utility refers to the sum of utilities of all units of a commod consumed.

$$TU = MU_1 + MU_2 + \cdots + MU_n$$

For example, if a consumer consumes ten biscuits, then the total utility is the sum of satisfaction of consuming all the ten biscuits.

Marginal Utility

- Marginal Utility is the addition made to the total utility by consumin one more unit of a Commodity.

For example, if a consumer consumes 10 biscuits, the Marginal utility is the utility derived from the 10th unit. It is nothing but the total utility of 10 biscuits minus the total utility of 9 biscuits.

Thus $MU_n = TU_n - TU_{n-1}$

Where, MU_n = marginal utility of nth commodit

TU_n = Total Utility of n units

TU_{n-1} = Total Utility of n-1 units.

Formula: $MU = \dfrac{\text{Change in TU}}{\text{Change in Quantity Demanded}}$

$= \dfrac{\Delta TU}{\Delta Q}$; ΔTU = Change in total utility
ΔQ = Change in Quantity of consumed.

Q: What is Law of Diminishing marginal ~~Be~~ utility (DMU):

- If a consumer takes more and more units of a commodity, the additional utility he derives from an extra unit of the commod goes on falling. Thus, according to this law, the Marginal utility decreases with the increase in the consumption of a commodity.

When marginal utility decreases, the total utility increases at a diminishing rate.

Law of Diminishing Marginal Utility:

Definition of the Law:

"Other things remaining the same when a person takes successive units of a commodity, the marginal utility diminishes constantly".

The marginal utility of a commodity diminishes at the consumer gets larger quantities of it. Marginal utility is the change in the total utility resulting from one unit change in the consumption of a commodity per unit of time.

Assumptions:

Following are the assumptions of the law of **diminishing marginal utility**.

1. The utility is measurable and a person can express the utility derived from a commodity in qualitative terms such as 2 units, 4 units and 7 units etc.
2. A rational consumer aims at the maximization of his utility.
3. It is necessary that a standard unit of measurement is constant
4. A commodity is being taken continuously. Any gap between the consumption of a commodity should be suitable.
5. There should be proper units of a good consumed by the consumer.
6. It is assumed that various units of commodity homogeneous in characteristics.
7. The taste of the consumer remains same during the consumption o the successive units of commodity.
8. Income of the consumer remains constant during the operation of the law of diminishing marginal utility.
9. It is assumed that the commodity is divisible.
10. There should be not change in fashion. For example, if there is a fashion of lifted shirts, then the consumer may have no utility in open shirts.
11. It is assumed that the prices of the substitutes do not change. For example, the demand for CNG increases due to rise in the prices of petroleum and these price changes effect the utility of CNG.

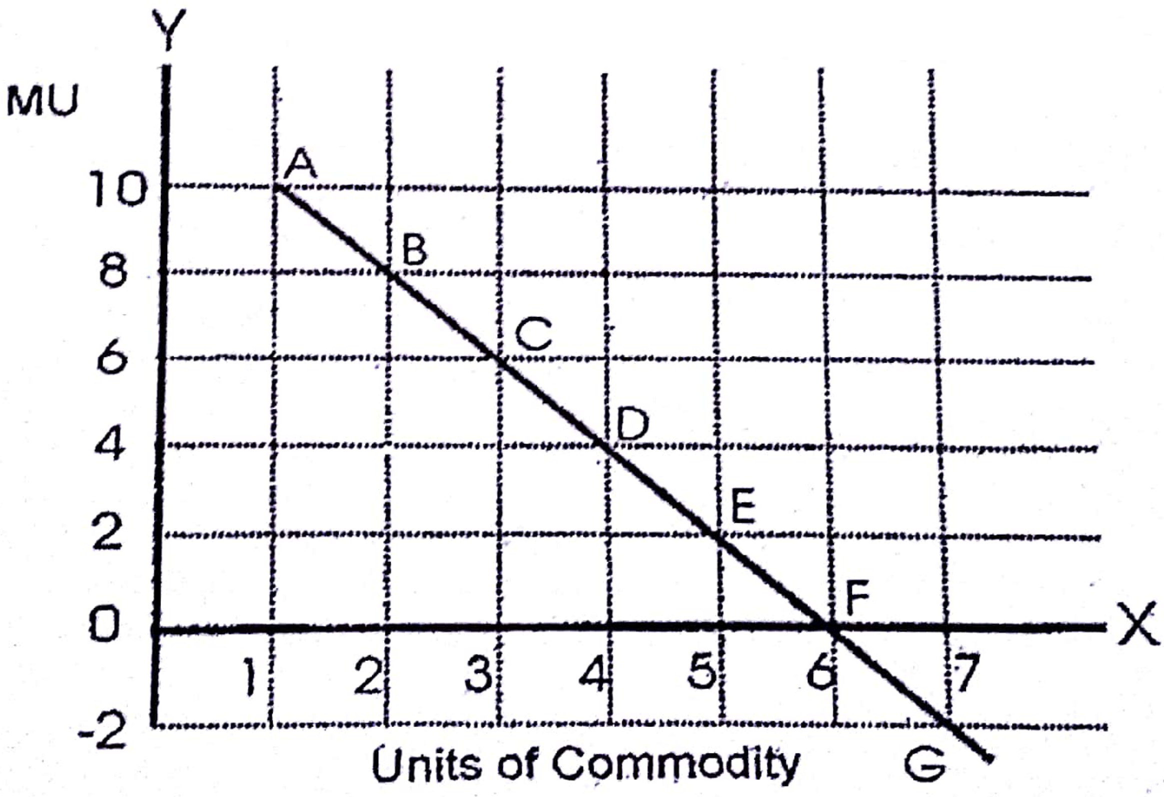
Explanation With Schedule and Diagram:

We assume that a man is very thirsty. He takes the glasses of water successively. The marginal utility of the successive glasses of water decreases, ultimately, he reaches the point of satiety. After this point the marginal utility becomes negative, if he is forced further to take a glass of water. The behavior of the consumer is indicated in the following schedule:

Units of commodity	Marginal utility	Total utility
1st glass	10	10
2nd glass	8	18
3rd glass	6	24
4th glass	4	28
5th glass	2	30
6th glass	0	30
7th glass	-2	28

On taking the 1st glass of water, the consumer gets 10 units of utility, because he is very thirsty. When he takes 2nd glass of water, his marginal utility goes down to 8 units because his thirst has been partly satisfied. This process continues until the marginal utility drops down to zero which is the saturation point. By taking the seventh glass of water, the marginal utility becomes negative because the thirst of the consumer has already been fully satisfied.

The law of diminishing marginal utility can be explained by the following diagram drawn with the help of above schedule:



In the above figure, the marginal utility of different glasses of water is measured on the y-axis and the units (glasses of water) on X-axis. With the help of the schedule, the points A, B, C, D, E, F and G are derived by the different combinations of units of the commodity (glasses of water) and the marginal utility gained by different units of commodity. By joining these points, we get the marginal utility curve. The marginal utility curve has the downward negative slope. It intersects the X-axis at the point of 6th unit of the commodity. At this point "F" the marginal utility becomes zero. When the MU curve goes beyond this point, the MU becomes negative. So there is an inverse functional relationship between the units of a commodity and the marginal utility of that commodity.

Exceptions or Limitations:

The limitations or exceptions of the law of diminishing marginal utility are as follows:

1. The law does not hold well in the rare collections. For example, collection of ancient coins, stamps etc.
2. The law is not fully applicable to money. The marginal utility of money declines with richness but never falls to zero.
3. It does not apply to the knowledge, art and innovations.
4. The law is not applicable for precious goods.
5. Historical things are also included in exceptions to the law.
6. Law does not operate if consumer behaves in irrational manner. For example, drunkard is said to enjoy each successive peg more than the previous one.
7. Man is fond of beauty and decoration. He gets more satisfaction by getting the above merits of the commodities.
8. If a dress comes in fashion, its utility goes up. On the other hand its utility goes down if it goes out of fashion.
9. The utility increases due to demonstration. It is a natural element.

Importance of the Law of Diminishing Marginal Utility:

The importance or the role of the law of diminishing marginal utility is as follows:

1. By purchasing more of a commodity the marginal utility decreases. Due to this behaviour, the consumer cuts his expenditures to that commodity.
2. In the field of public finance, this law has a practical application, imposing a heavier burden on the rich people.
3. This law is the base of some other economic laws such as law of demand, elasticity of demand, consumer surplus and the law of substitution etc.
4. The value of commodity falls by increasing the supply of a commodity. It forms a basis of the theory of value. In this way prices are determined

Theory of consumer Behaviour

- Indifference curve analysis
- Marginal rate of substitution (MRS)
- Budget line
- Utility maximization.

Nazrul

Theory of Consumer Behaviour
- Indifference curve analysis
- Marginal rate of substitution (MRS)
- Budget line
- Utility maximization

✠ Some Questions on these Topics

i) What is indifference curve? Explain the properties of IC.
ii) Why two indifference curves can not intersect each other?
iii) What is MRS? How we calculate MRS?
iv) What is Budget line? How Budget line changes?
v) With the help of IC & B.L, explain consumer's equilibrium.
vi) How the equilibrium position changes?
vii) How do we derive demand curve from consumer's equilibri with the help of IC & B.L.
viii) Construction of B.L & drawing of B.L

Chapter # Theory of Consumer Behaviour

✠ Indifference curve analysis (Please Read From sheet & Books)

i)- Defn
ii)- Properties

iii) ~~MRS~~ the Marginal rate of substitution:

- The slope of an indifference curve is known as the marginal rate of substitution (MRS). MRS measures the rate at which the consumer is just willing to substitute one good for the other.

- suppose that we take a little of good X, ΔX away from the consumer. Then we give him Δy, an amount that is just sufficient to put him back on his indifference curve, so that he is just as well off after this substitution of y for x as he was before.

We think of the ratio $\Delta Y/\Delta X$ as being the rate at which the consumer is willing to substitute good 2 (Y) for good 1 (X).

- Now think of ΔX as being a very small change - a marginal change. Then the rate $\frac{\Delta Y}{\Delta X}$ measures the marginal rate of substitution of good 2 (Y) for good (1) (X).

As ΔX gets smaller, $\frac{\Delta Y}{\Delta X}$ approaches the slope of the indifference curve.

- MRS is always a negative number and measures slope of the Ic. It measures the rate at which the consumer is just willing to substitute a little more consumption of Y for a little less consumption of good 1 (X)

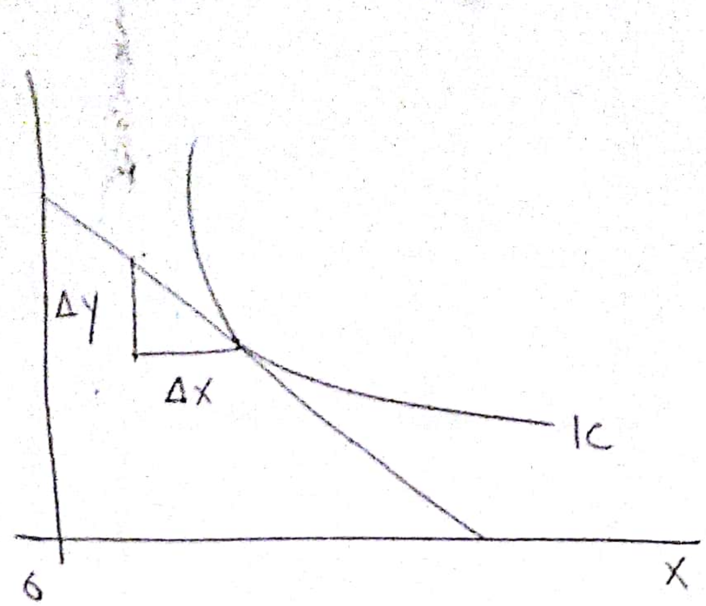


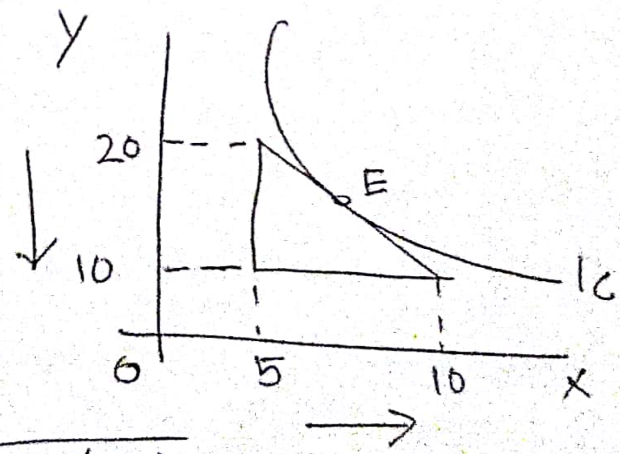
Fig – measuring MRS

– Here, In the above figure, X is shown on the horizontal axis ~~ab~~ and Y is shown on the vertical axis. Ic is the indifference curve.

– If consumer take ΔX (more of X), he has to loose ΔY (Less of Y). $\frac{\Delta Y}{\Delta X}$ is the slope of Ic. So,

$$MRS_{XY} = \frac{\Delta Y}{\Delta X}$$

Calculation of MRS :



Slope at point E = $\frac{\Delta Y}{\Delta X}$

$= \frac{\text{change in vertical axis}}{\text{change in horizontal axis}}$

$= \frac{10-20}{10-5} = -\frac{10}{5} = -2$ (MRS)

Here. MRS = 2 indicates that, if the consumer takes 1 more X, he has to give up 2 Y so that he remains on the same indifference curve.

Here, slope of indifference curve is -2.

- Indifference curves must have a negative slope, so the MRS always involves reducing the consumption of one good in order to get more of another good.

- strictly convex indifference curves means - the MRS (slope of IC) decreases as we increase X. Thus the indifference curves exhibit diminishing marginal rate of substitution. This means that the amount of good X that the person is willing to give up for an additional amount of good 2 (Y) increases the amount of good 1 (X) increases

- Stated in this way, convexity of indifference curves says that the more you have of one good, the more willing you are to give some of it up in exchange for the other good

Nimarul

2. ~~FK~~ Budget line

Definition: The Budget Line shows all the combinations of two commodities that a consumer can afford at given market prices and within the particular income level.

- We know that the higher the indifference curve, the higher is the utility, and thus, utility maximizing consumer will strive to reach the highest possible indifference curve. But, he has two strong constraints: limited income and given the market price of goods and services.

The income in hand is the main constraint that decides how high a consumer can go on the indifference map.

In a two-commodity model, the budgetary constraint can be expressed in the form of the budget equation:

$$P_x X + P_y Y = M$$

Here P_x = Price of X

P_y = Price of Y

X = amount of X good

Y = amount of Y good

M = Consumer's Income

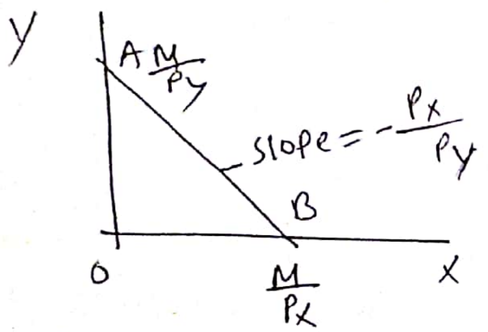
- The budget equation states that the consumer's expenditure on commodity x and y can not exceed his money income (M). Thus, the quantities of commodities x and y that a consumer can buy from his income (M) at given prices P_x and P_y can be calculated through the budget equation given below:

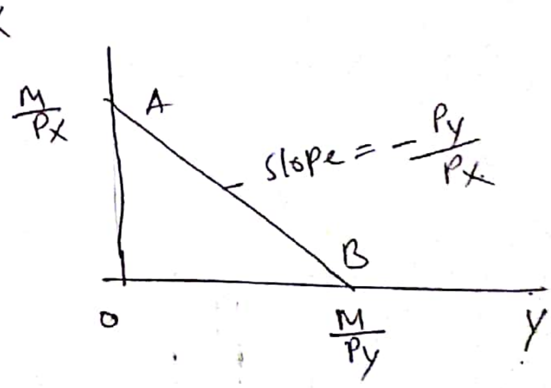
$$P_x X + P_y Y = M \qquad \text{or} \qquad P_x X = M - P_y Y$$

$$\Rightarrow P_y Y = M - P_x X \qquad \Rightarrow X = \frac{M}{P_x} - \frac{P_y}{P_x} Y$$

$$\Rightarrow Y = \frac{M}{P_y} - \frac{P_x}{P_y} X \qquad \text{Here, slope} = -\frac{P_y}{P_x}$$

Here, slope $= -\frac{P_x}{P_y}$





The values of x and y are plotted on the x & y axis, and a line with a negative slope is drawn connecting the points so obtained. This line is called the budget line or price.

⇒ Drawing of Budget line:

$$P_x X + P_y Y = M$$

if $X=0$, $P_x \cdot 0 + P_y Y = M$

$$\Rightarrow 0 + P_y Y = M$$

Scanned by CamScanner

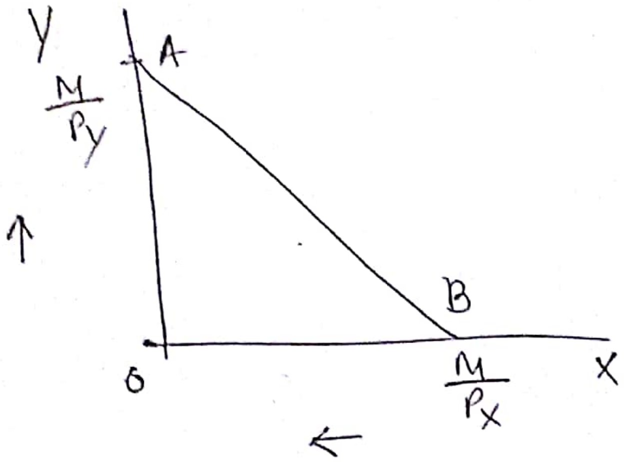
$\Rightarrow P_y Y = M$

$\Rightarrow Y = \frac{M}{P_y}$ (max amount of y when $x=0$)

And if $Y=0$, $P_x X + P_y \cdot 0 = M$

$\Rightarrow P_x X = M$

$\Rightarrow X = \frac{M}{P_x}$ (Maximum amount of x when $y=0$)



$\Rightarrow$ slope of Budget line = Rate of exchange $= \frac{\Delta Y}{\Delta X}$

$$= \frac{\frac{M}{P_y} - 0}{0 - \frac{M}{P_x}}$$

$$= - \frac{M}{P_y} \cdot \frac{P_x}{M}$$

$$= - \frac{P_x}{P_y}$$

Numerical Example:

Suppose, X & Y are two goods. Price of X is 5 and price of Y is 10. Your income level is 100. - Draw the budget line slope of Budget line.

- construction of B.L: We are given

$P_x = 5$
$P_y = 10$
$M = 100$

P.

Budget equation ⇒ $P_x X + P_y Y = M$

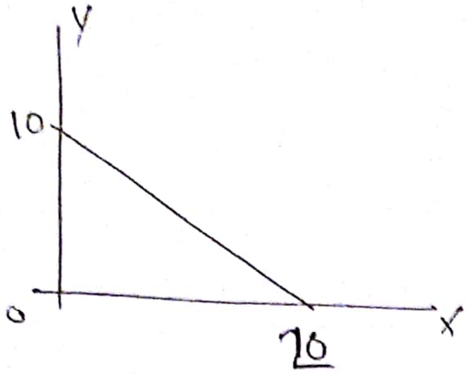
$5x + 10y = 100$

Drawing B.L : if $x = 0$, $10y = 100$

$\Rightarrow y = \frac{100}{10} = 10$

and if $y = 0$, $5x = 100$

$\Rightarrow x = \frac{100}{5} = 20$



Slope of B.L :

One method : $\frac{\Delta y}{\Delta x} = \frac{10-0}{0-20} = -\frac{10}{20} = -\frac{1}{2}$

another method : $\frac{\Delta y}{\Delta x} = -\frac{P_x}{P_y} = -\frac{5}{10} = -\frac{1}{2}$

Here slope of B.L $= -\frac{1}{2}$ means that if we take 1 more X, we have to Loose $\frac{1}{2}$ unit of Y.

How the position of B.L changes ?

i) ceteris paribus, only change in X's price

Old		New
$P_x = 5$	Here	$P_x' = 2.5$
$P_y = 10$	$P_x' < P_x$	$P_y = 10$
$M = 100$		$M = 100$
$P_x X + P_y Y = M$		$P_x' X + P_y Y = M$

Scanned by CamScanner

Nazmul

if $P_x' x + P_y y = M$

$\Rightarrow 2.5x + 10y = 100$

if $x=0$, $y = \frac{100}{10} = 10$

and if $y=0$, $2.5x = 100$

$\Rightarrow x = \frac{100}{2.5} = 40$ (Maximum amount of x when y=0)

and if $x=0$, $y = \frac{100}{10} = 10$ (Max amount of y when x=0)

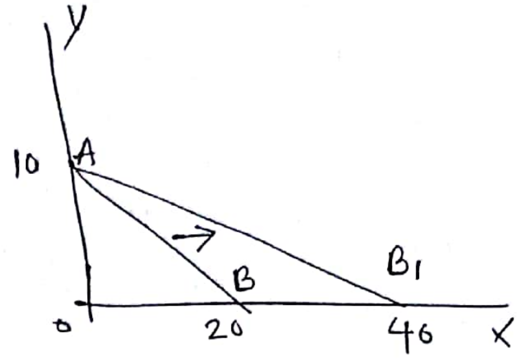


Fig- B.L Pivotes outward

Here x is shown on the horizont[al] axis and y is shown on the vertical axis.

- Initial B.L is AB. Now when Price of x decreases by 50% Price of x decreases from 5 t[o] 2.5.
- New quantity becomes $x=40$ & $y=10$

2) ceteris paribus, only change in y's price

$P_x X + P_y' Y = M$ (Price of y doubles from 10 to 20)

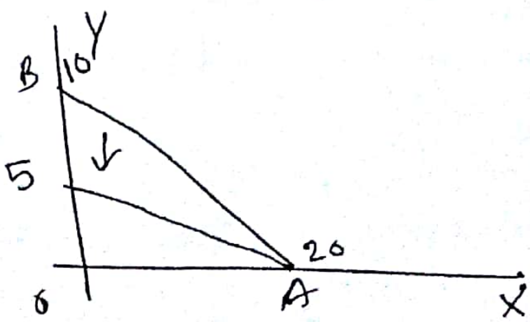
$5X + 20Y = 100$

if $x=0$, $y = \frac{100}{20} = 5$

and $y=0$, $x = \frac{100}{5} = 20$

slope $= -\frac{P_x}{P_y} = -\frac{5}{20} = -\frac{1}{4}$

$= -.25$

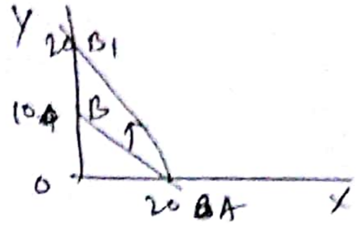


Scanned by CamScanner

- If price of y decreases by 50%. That is $P_y = 5$

$5X + 5Y = 100$

$X = 20$ & $Y = 20$



$\text{slope} = -\frac{P_x}{P_y} = -\frac{5}{5} = -1$

3) change in both x's & y's price

Initial	Changed
$P_x = 5$	$P_x' = 2.5$
$P_y = 10$	$P_y' = 20$
$M = 100$	$M = 100$

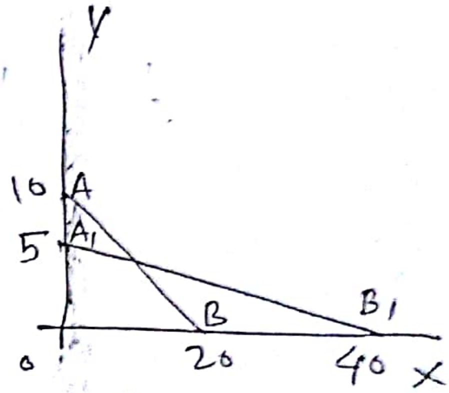
P_x decreases by 50% & P_y increases by 100%.

$5X + 10Y = 100$

$2.5X + 20Y = 100$

if $X = 0$, $Y = 5$

and $Y = 0$, $X = \frac{100}{2.5} = 40$



→ B.L. Pivotes

4) change in income

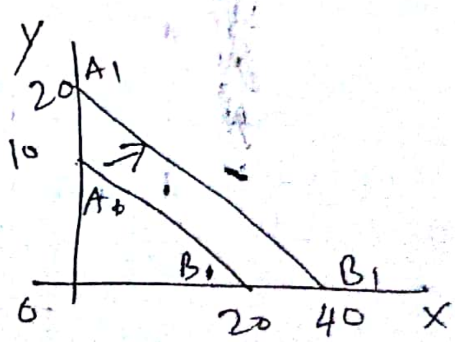
Old	Changed
$P_x = 5$	$P_x = 5$
$P_y = 10$	$P_y = 10$
$M = 100$	$M' = 200$

(Income doubles)

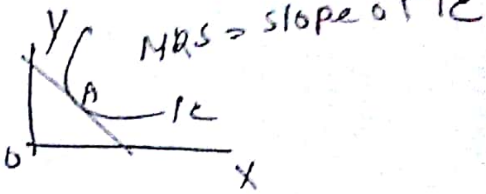
$5X + 10Y = 200$

if $X = 0$, $Y = 20$

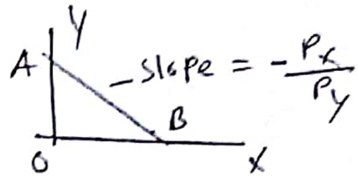
and $Y = 0$, $X = 40$



Consumer's equilibrium & utility maximization



- We know that higher IC represents higher utility. As consumer we want to stay on the higher IC. But have limitations. our budget is limited.



- We get consumer's equilibrium two indifference curve & budget line are tangent to each other.

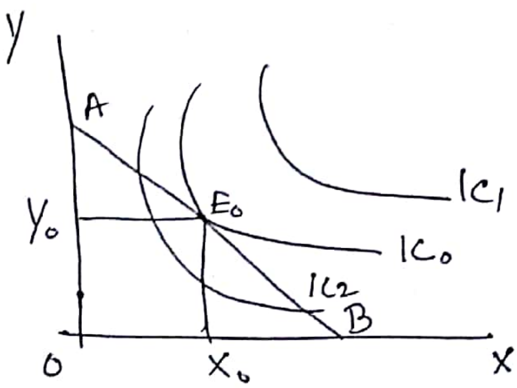


Fig- Consumer's equilibrium

Explanation: In the above graph, X is shown on the horizontal axis and Y good is shown on the vertical axis. IC_0 is the initial indifference curve and AB is the budget line.

- the consumer will not be able to be on IC_1 because it is not affordable. The consumer will not want to stay on IC_2 because lower IC_2 represents lower utility. And on IC_2 resources will be unutilized.

- Equilibrium occurs when B.L and IC are tangent. point E_0 is the Equilibrium point because here IC_0 and budget line AB are tangent with each other.

At point E_0,

$$\text{slope of IC} = \text{slope of BL}$$

$$\Rightarrow MRS_{xy} = \text{Rate of exchange}$$

$$\Rightarrow MRS_{xy} = -\frac{P_x}{P_y}$$

⇒ At point E_0, the consumer consumes X_0 amount of x good and Y_0 amount of y good. Consuming at E_0 point, the consumer is at equilibrium and get maximum utility constraint to budget line.

<u>change in equilibrium position:</u>

1) change in price of x (P_x ↓es)

(P_x ↓es)

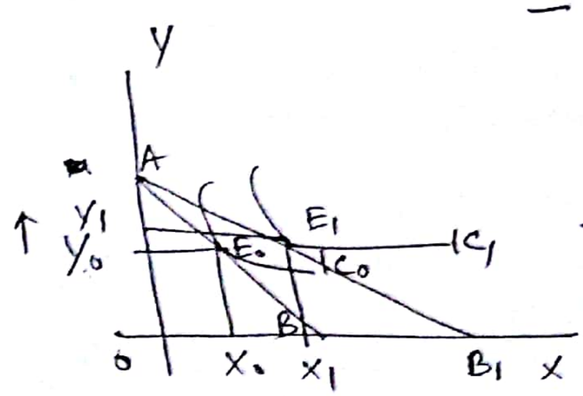


Fig-a

- Here, X is shown on X axis and Y is shown on Y axis.
- Initial budget line is AB and initial IC is IC_0.

- The consumer is initially at equilibrium point E_0 and consumes x_0 unit of x and Y_0 unit of Y.
- Now when price of x decreases, Budget line

AB changes from AB to AB_1. New equilibrium become E_1. New IC is IC_1.

- At E_1 equilibrium point consumer now consumes X_1 unit of x and Y_1 unit of y.

That is the consumption of both good increases.

- When P_x decreases, Real income of consumer increases, so the consumer can now consume more unit of both goods.

2) ~~change in~~ (Increase in P_x) ⇒ (write description)

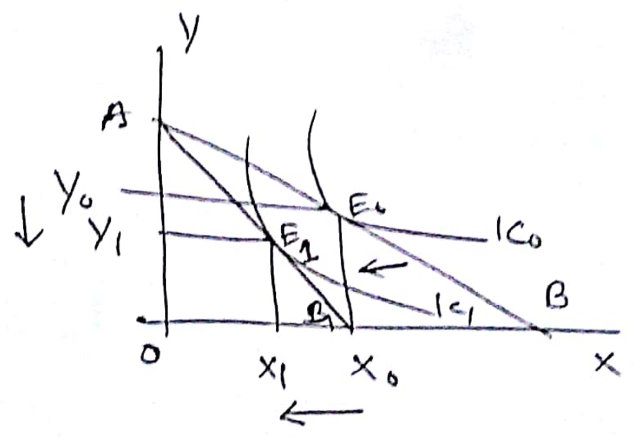
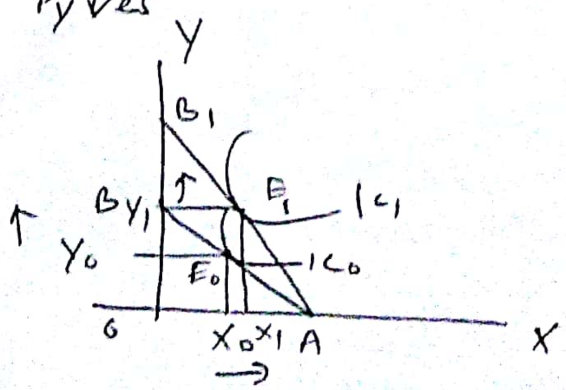


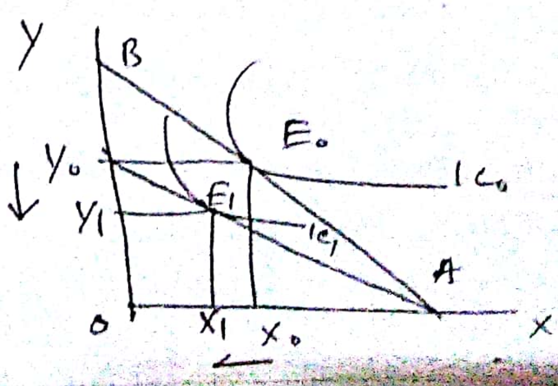
Fig-b

2) change in y's price (Price of y decreases)

a) P_y ↓es

b) P_y ↑es





3) change in income

Decreased income

a)

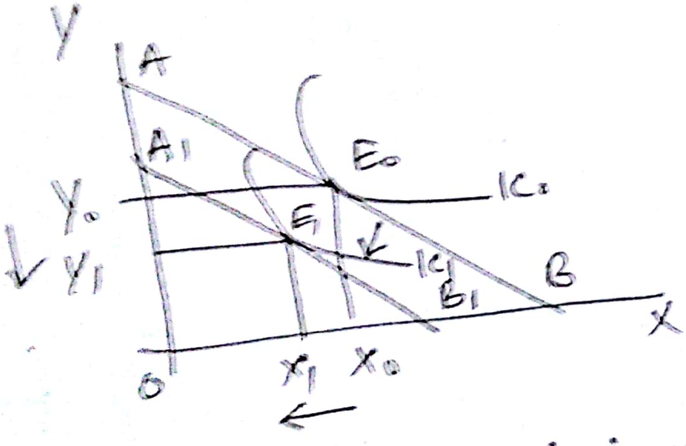


Fig-a: Decrease in income

Increase in income

b)

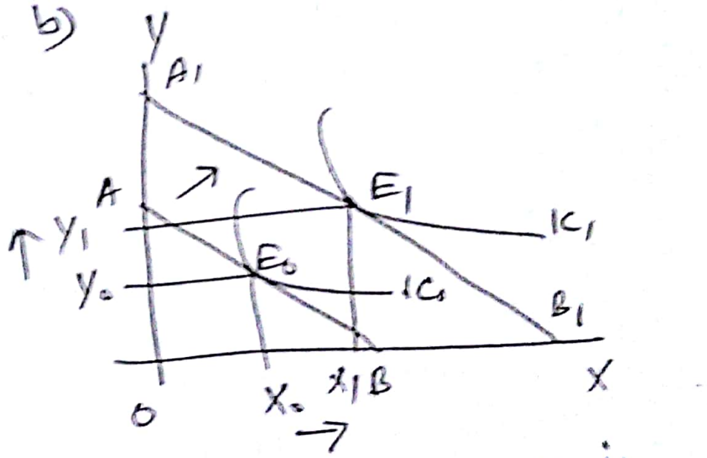


Fig-b: increase in income

5. Derivation of Demand curve from equilibrium position with the help of IC & B.L

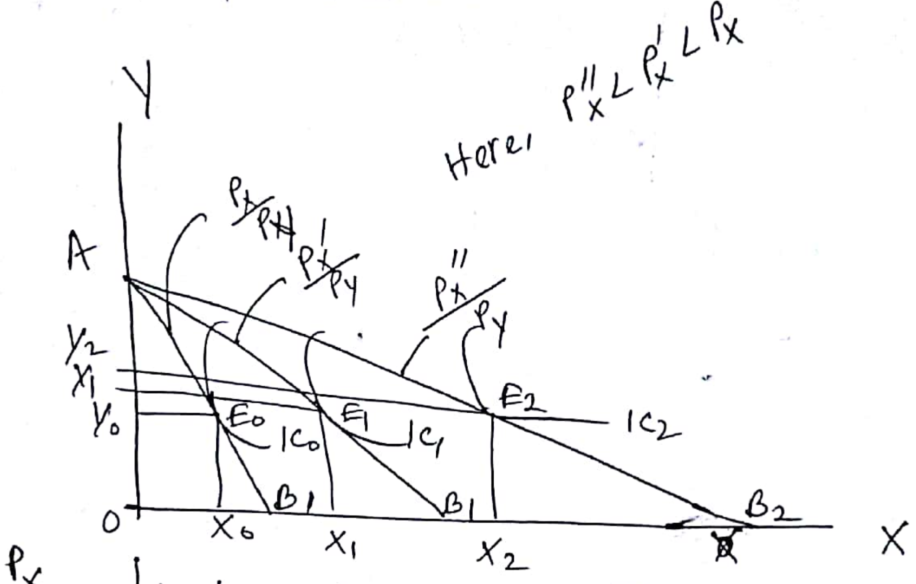


Fig-a : consumer's equilibrium

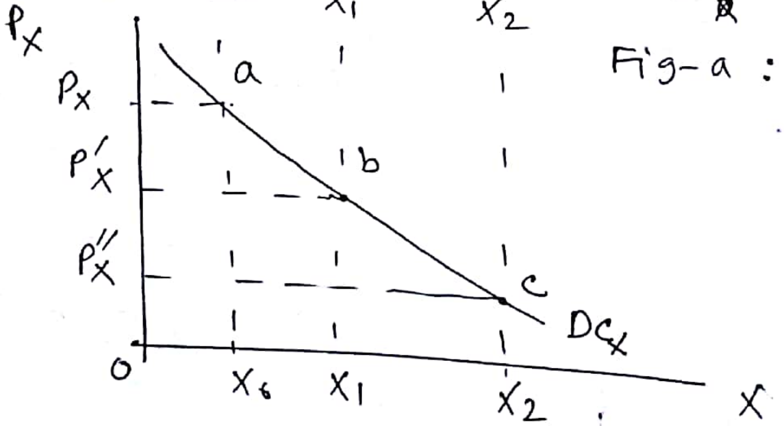


Fig- b: Demand curve

In Fig a, on the horizontal axis, X is shown and on the vertical axis good Y is shown. AB is initial budget line and IC_0 is initial indifference curve. E_0 is initial equilibrium point. At E_0 point the consumer consumes X_0 unit of X and Y_0 unit of Y. consumer's utility is maximum subject to budget line AB line.

slope of Budget line is $\frac{P_x}{P_y}$.

P Amounts of x & y at E_0 equilibrium point is transformed in figure b denoted by point a. point a shows when price is P_x, Quantity demanded for X is X_0. point a is a point of the demand curve.

— Now if price of x decreases from P_x to P_x'. Budget line pivots outward from AB to AB_1. As price of x decreases, Real income increases and the consumer now consumes more of x and y.

the consumer now moves from E_0 to a higher equilibrium point E_1. At point E_1, IC_1 & B.L AB_1 are tangent with each other.

At point E_1, the consumer consumes x_1 unit of x and y_1 of y goods. This amounts of x & y is plotted on fig-2 that is denoted by point b.

point b shows when price is P_x', Quantity demanded is X_1. point b is another point of demand curve.

Again, Price of x decreases from P'_x to P''_x. The new budget line is AB_2. New equilibrium is at point E_2 when B.L AB_2 and IC_2 are tangent to each other. At point E_2, ~~consumer~~ consumer consumes X_2 unit of X and Y_2 unit of Y.

- That is when Price of x is P''_x, the consumer ~~consumes~~ consumes X_2 unit of X.

- slope of AB_2 B.L $= \frac{P''_x}{P_y}$.

- ~~Amo~~ Equilibrium amounts of X & Y denoted at point E_2 is plotted in fig-2b that is denoted by c point.

Connecting point a, b, and c, we get demand curve for X good. Here Demand curve shows relationship between Price of X and quantity demanded of X good.

Scanned by CamScanner

Indifference Curve Analysis.

Nazrul

Indifference Curve Analysis

In Microeconomics, the Indifference Curve Analysis is an important analytical tool in the study of consumer behaviour. The indifference curve analysis was developed by the British economist Francis Ysidro Edgeworth, Italian economist Vilfredo Pareto and others in the first part of the 20th century J.R.Hicks & R.G.D. Allen in their research paper,' A Reconsideration of the Theory of Value' criticized Marshallian cardinal approach of utility and propounded Indifference curve theory of consumer's demand. It is also called as Ordinal Approach.

Concept Of Indifference Curve:

An indifference curve is a locus of combinations of goods which derive the same level of satisfaction, so that the consumer is indifferent to any of the combination he consumes.If a consumer equally prefers two product bundles, then the consumer is indifferent between the two bundles. The consumer gets the same level of satisfaction (utility) from either bundle. Graphically speaking, this is known as the indifference curve. An indifference curve shows combinations of goods between which a person is indifferent.

Symbolically,in the equation form,

An Indifference Curve =

$U=f(x_1,x_2,x_3,.....x_n)=k$ U=f(x1,x2,x3,.....xn)=k

......where, k is a constant.

In indifference curve approach only ordination of preferences is needed.It overcomes the weakness of Cardinal measurement as the satisfaction cannot be measured objectively.

The cardinal approach provides the assumption of constant utility of money, which is unrealistic.In indifference curve approach, this assumption has been dropped.

Indifference curve approach is base for the measurement of 'consumer's surplus'.In a way it contributes to the Welfare economics.

Indifference curve is a better tool to classify substitutes and complementary goods.

Properties of Indifference Curve:

The main attributes or properties or characteristics of indifference curves are as follows:

1) Indifference Curves are Negatively Sloped:

The indifference curves must slope downward from left to right. As the consumer increases the consumption of X commodity, he has to give up certain units of Y commodity in order to maintain the same level of satisfaction.

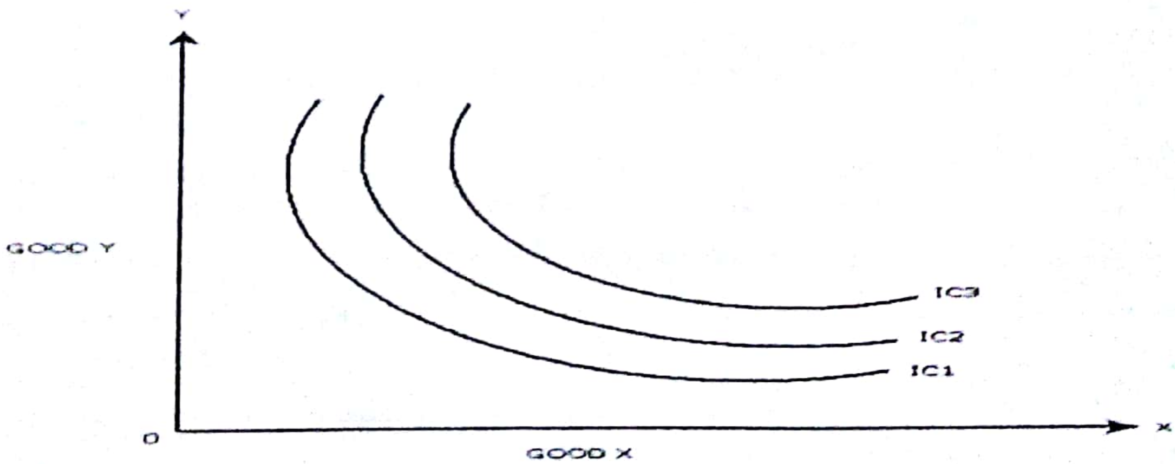
DIAGRAM:

In the above diagram, two combinations of commodity cooking oil and commodity wheat is shown by the points a and b on the same indifference curve. The consumer is indifferent towards points a and b as they represent equal level of satisfaction.

(2) Higher Indifference Curve Represents Higher Level of Satisfaction:

Indifference curve that lies above and to the right of another indifference curve represents a higher level of satisfaction. The combination of goods which lies on a higher indifference curve will be preferred by a consumer to the combination which lies on a lower indifference curve.

Diagram:

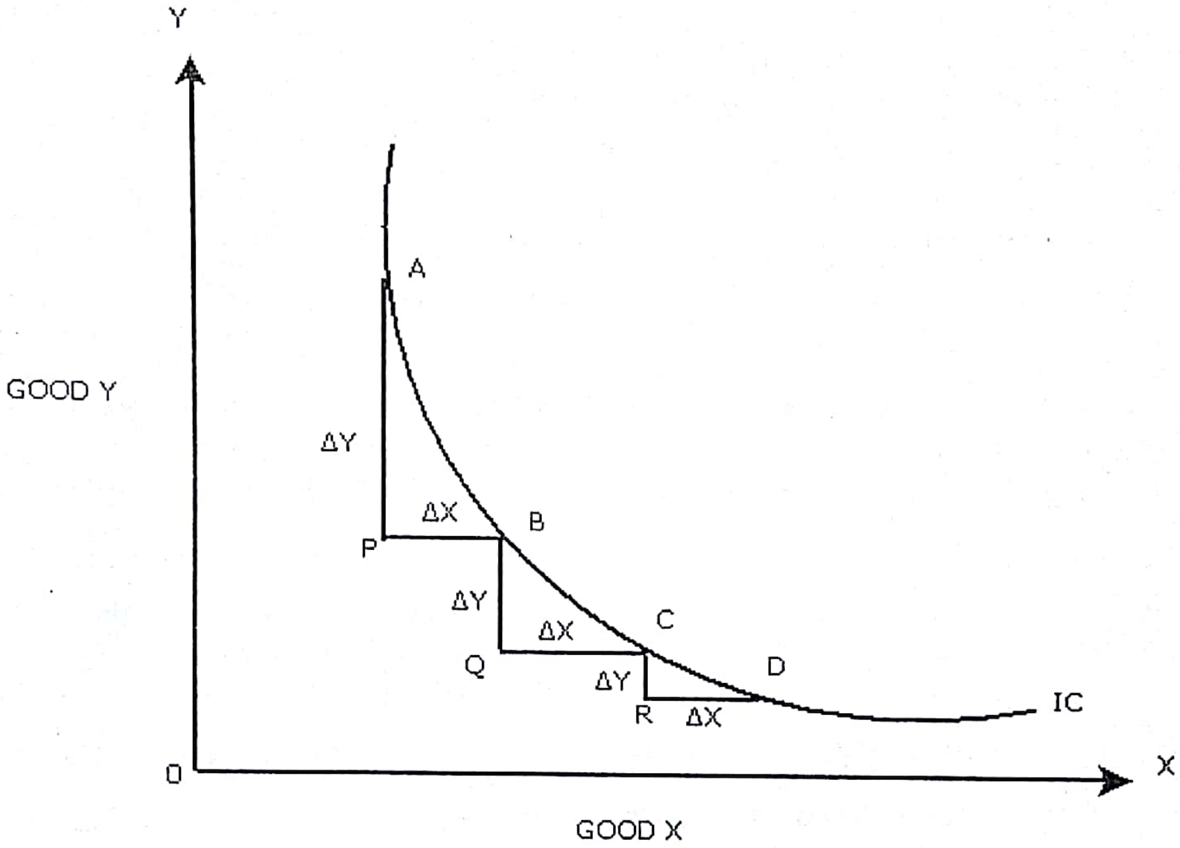


In this diagram, there are three indifference curves, IC1, IC2 and IC3 which represents different levels of satisfaction. The indifference curve IC3 shows greater amount of satisfaction and it contains more of both goods than IC2 and IC1. IC3 > IC2> IC1.

(3) Indifference Curves are Convex to the Origin:

This is an important property of indifference curves. They are convex to the origin. As the consumer substitutes commodity X for commodity Y, the marginal rate of substitution diminishes as X for Y along an indifference curve. The Slope of the curve is referred as the Marginal Rate of Substitution. The Marginal Rate of Substitution is the rate at which the consumer must sacrifice units of one commodity to obtain one more unit of another commodity.

Diagram:



In the above diagram, as the consumer moves from A to B to C to D, the willingness to substitute good X for good Y diminishes. The slope of IC is negative.In the above diagram, diminishing MRSxy is depicted as the consumer is giving AP>BQ>CR units of Y for PB=QC=RD units of X. Thus indifference curve is steeper towards the Y axis and gradual towards the X axis. It is convex to the origin.

If the indifference curve is concave, MRSxy increases. It violets the fundamental feature of consumer behaviour.

If commodities are almost perfect substitutes then MRSxy remains constant. In such cases the indifference curve is a straight line at an angle of 45 degree with either axis.

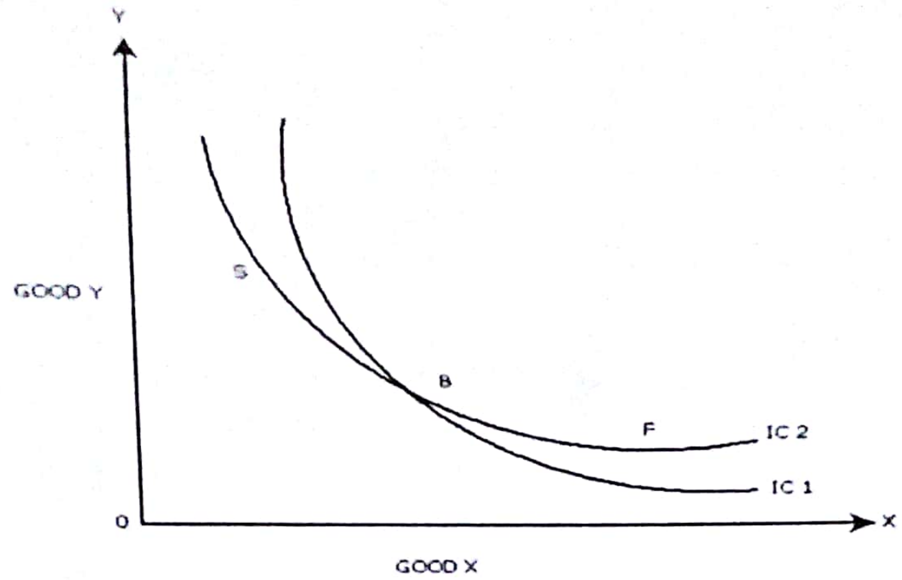
If two commodities are perfect complements, the indifference curve will have a right angle.

In reality, commodities are not perfect substitutes or perfect complements to each other.Therefore MRSxy usually diminishes.

(4) Indifference Curves cannot Intersect Each Other:

The indifference curves cannot intersect each other. It is because at the point of tangency, the higher curve will give as much as of the two commodities as is given by the lower indifference curve. This is absurd and impossible.

Diagram:



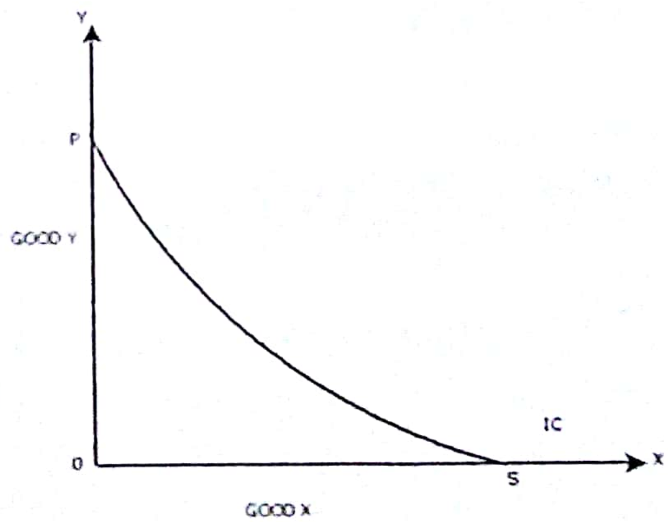
In the above diagram, two indifference curves are showing cutting each other at point B. The combinations represented by points B and F given equal satisfaction to the consumer because both lie on the same indifference curve IC2. Similarly the combinations shows by points B and E on indifference curve IC1 give equal satisfaction top the consumer.

If combination F is equal to combination B in terms of satisfaction and combination E is equal to combination B in satisfaction. It follows that the combination F will be equivalent to E in terms of satisfaction. This conclusion looks quite funny because combination F on IC2 contains more of good Y (wheat) than combination which gives more satisfaction to the consumer. We, therefore, conclude that indifference curves cannot cut each other.

(5) Indifference Curves do not Touch the Horizontal or Vertical Axis:

One of the basic assumptions of indifference curves is that the consumer purchases combinations of different commodities. He is not supposed to purchase only one commodity. In that case indifference curve will touch one axis. This violates the basic assumption of indifference curves.

Diagram:



In the above diagram, it is shown that the indifference IC touches Y axis at point P and X axis at point S. At point C, the consumer purchase only OP commodity of Y good and no commodity of X

- Production
- Production function
- Factors of production
- IQ (Iso-Quant)
- RTS

ECO101
(ST D)

1) Q: What is Production?

The processes and methods used to transform tangible inputs (raw materials, semi-finished goods, subassemblies) and intangible inputs (ideas, information, knowledge) into goods or services. Resources are used in this process to create an output that is suitable for use or has exchange value. Production is a process of workers combining various material inputs and immaterial inputs (plans, know-how) in order to make something for consumption (the output). It is the act of creating output, a good or service which has value and contributes to the utility of individuals.

Production is a process of transforming (converting) inputs (Raw materials) into outputs (finished goods). So, production means the creation of goods and services. It is done to satisfy human wants. Thus, production is a process of transformation.

2) Q: What is production function?

In economics, a production function relates physical output of a production process to physical inputs or factors of production. It is a mathematical function that relates the maximum amount of output that can be obtained from a given number of inputs - generally capital and labor. The production function, therefore, describes a boundary or frontier representing the limit of output obtainable from each feasible combination of inputs.

Firms use the production function to determine how much output they should produce given the price of a good, and what combination of inputs they should use to produce given the price of capital and labor (). When firms are deciding how much to produce they typically find that at high levels of production, their marginal costs begin increasing. This is also known as diminishing returns to scale - increasing the quantity of inputs creates a less-than-proportional increase in the quantity of output. If it weren't for diminishing returns to scale, supply could expand without limits without increasing the price of a good.

In economics, a production function relates physical output of a production process to physical inputs or factors of production.

Generally, production is the transformation of raw material into the finished goods. These raw materials are classified as land, labor, capital or natural resources. These may be fixed or variable depending upon the nature of the business.

This function establishes the physical relationship between these inputs and the output. The efficiency of this relationship depends on the different quantities used in the production process, the quantities of output and the productivity at each point.

$$Short\ Run:\ q = f(L)$$

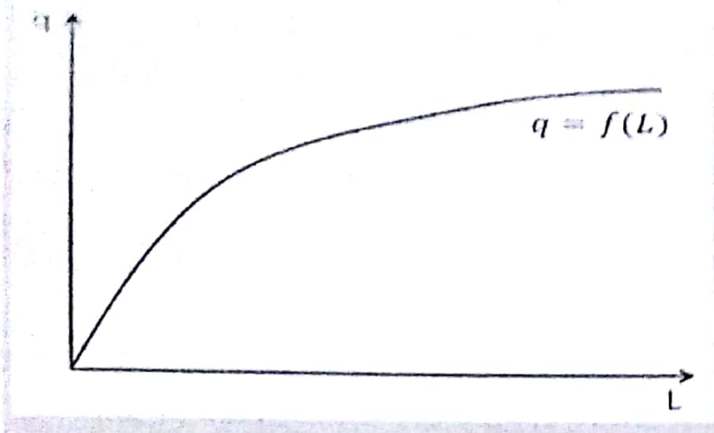
$$Long\ Run:\ q = f(K, L)$$

Here, L=Labor and K=Capital

The production function simply states the quantity of output (q) that a firm can produce as a function of the quantity of inputs to production, or . There can be a number of different inputs to production, i.e. "factors of production," but they are generally designated as either capital or labor. (Technically, land is a third category of factors of production). The particular functional form of the production function depends on the specific technology and production processes that a firm uses.

In the short run, the amount of capital that a factory uses is generally thought to be fixed. (The reasoning is that firms must commit to a particular size of factory, office, etc. and can't easily change these decisions without a long planning period.) Therefore, the quantity of labor (L) is the only input in the short-run production function. In the long run, on the other hand, a firm has the planning horizon necessary to change not only the number of workers but the amount of capital as well, since it can move to a different size factory, office, etc. Therefore, the long-run production function has two inputs that be changed- capital (K) and labor (L). Both cases are shown in the diagram above.

The Production Function in the Short Run

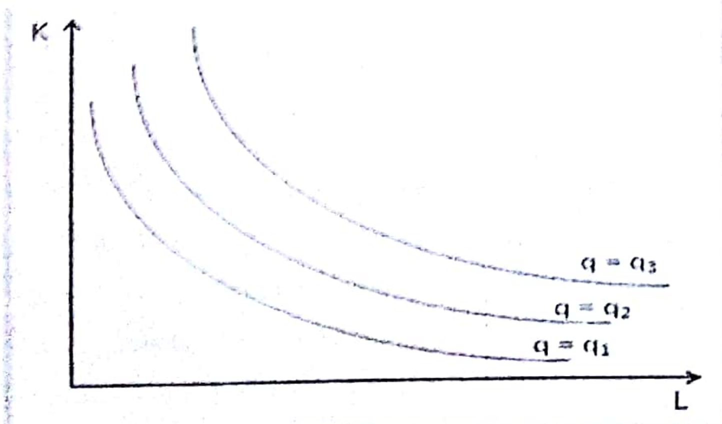


As shown in the above diagram, the short-run production function puts the quantity of labor (L) on the horizontal axis (since it's the independent variable) and the quantity of output (q) on the vertical axis (since it's the dependent variable).

The short-run production function has two notable features. First, the curve starts at the origin, which represents the observation that the quantity of output pretty much has to be zero if the firm hires zero workers.

Second, the production function gets flatter as the amount of labor increases, resulting in a shape that is curved downward. Short-run production functions typically exhibit a shape like this due to the phenomenon of diminishing marginal product of labor.

The Production Function in the Long Run



Because it has two inputs, the long-run production function is a bit more challenging to draw. One mathematical solution would be to construct a three-dimensional graph, but that is actually more complicated than is necessary. Instead, economists visualize the long-run production function on a 2-dimensional diagram by making the inputs to the production function the axes of the graph, as shown above. Technically, it doesn't matter which input goes on which axis, but it is typical to put capital (K) on the vertical axis and labor (L) on the horizontal axis.

With each line on the graph representing a particular quantity of output. In fact, each line on this graph is called an "isoquant" curve, so even the term itself has its roots in "same" and "quantity."

In general, curves that are further away from the origin correspond to larger quantities of output. (In the diagram above, this implies that q_3 is greater than q_2, which is greater than q_1.) This is simply because curves that are further away from the origin are using more of both capital and labor in each production configuration. It is typical (but not necessary) for the curves to be shaped like the ones above, as this shape reflects the tradeoffs between capital and labor that are present in many production processes.

3) Q: What are the factors of production?

Factors of production is an economic term that describes the inputs that are used in the production of goods or services in order to make an economic profit. The factors of production include land, labor, capital and entrepreneurship. These production factors are also known as management, machines, materials and labor, and knowledge has recently been talked about as a potential new factor of production.

Resources required for generation of goods or services, generally classified into four major groups:
Land (including all natural resources),
Labor (including all human resources),
Capital (including all man-made resources), and
Enterprise (which brings all the previous resources together for production).
These factors are classified also as management, machines, materials, and money (this, the 4 Ms), or other such nomenclature. More recently, knowledge has come to be recognized as distinct from labor, and as a factor of production in its own right.

Q: What is Iso-Quant (IQ)?

The isoquant curve is a graph, used in the study of microeconomics, that charts all inputs that produce a specified level of output. This graph is used as a metric for the influence that the inputs have on the level of output or production that can be obtained. The isoquant curve assists firms in making adjustments to inputs in order to maximize outputs, and thus profits.

An isoquant is a firm's counterpart of the consumer's indifference curve. An isoquant is a curve that shows all the combinations of inputs that yield the same level of output. 'Iso' means equal and 'quant' means quantity. Therefore, an isoquant represents a constant quantity of output. The isoquant curve is also known as an "Equal Product Curve" or "Production Indifference Curve" or Iso-Product Curve."

The concept of isoquants can be easily explained with the help of the table given below:

Table : An Isoquant Schedule

Combinations of Labor and Capital	Units of Labor (L)	Units of Capital (K)	Output of Cloth (meters)
A	5	9	100
B	10	6	100
C	15	4	100
D	20	3	100

The above table is based on the assumption that only two factors of production, namely, Labor and Capital are used for producing 100 meters of cloth.

Combination A = 5L + 9K = 100 meters of cloth

Combination B = 10L + 6K = 100 meters of cloth

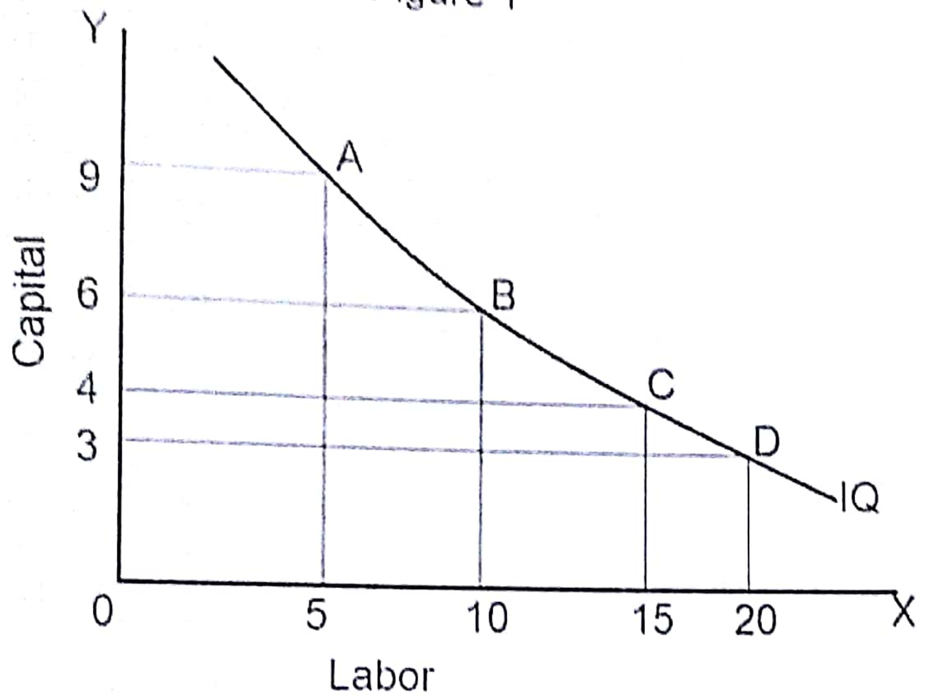
Combination C = 15L + 4K = 100 meters of cloth

Combination D = 20L + 3K = 100 meters of cloth

The combinations A, B, C and D show the possibility of producing 100 meters of cloth by applying various combinations of labor and capital. Thus, an isoquant schedule is a schedule of different combinations of factors of production yielding the same quantity of output.

An iso-product curve is the graphic representation of an iso-product schedule.

Figure 1

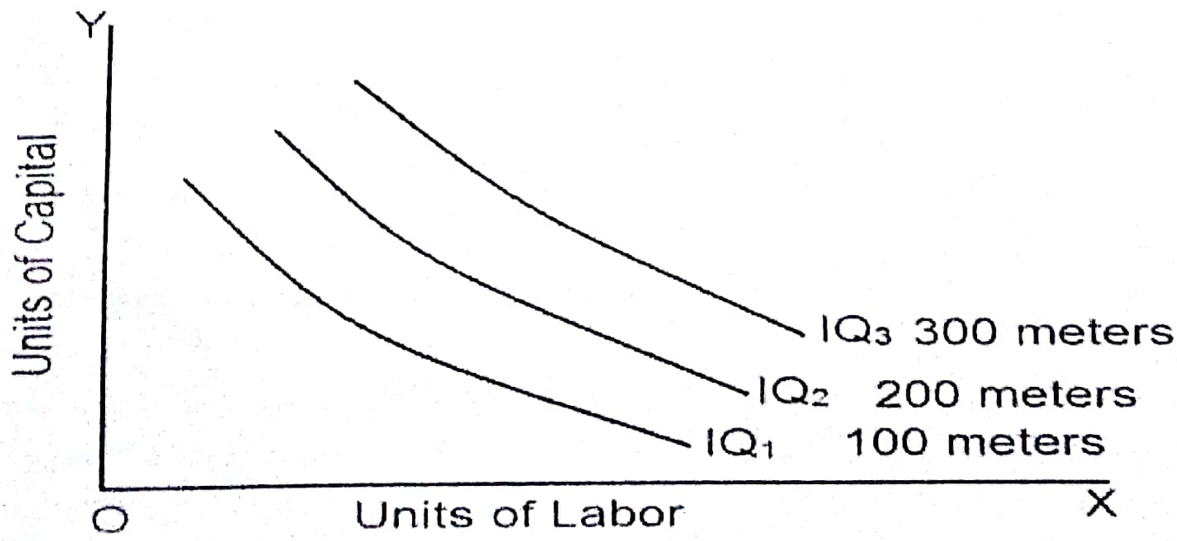


Thus, an isoquant is a curve showing all combinations of labor and capital that can be used to produce a given quantity of output.

Isoquant Map

An isoquant map is a set of isoquants that shows the maximum attainable output from any given combination inputs.

Figure 2



5) Returns to Scale (RTS).

Law of Returns to Scale

In the long- run the dichotomy between fixed factor and variable factor ceases. In other words, in the long-run all factors are variable. The law of returns to scale examines the relationship between output and the scale of inputs in the long-run when all the inputs are increased in the same proportion.

Assumptions

This law is based on the following assumptions:

1. All the factors of production (such as land, labor and capital) but organization are variable
2. The law assumes constant technological state. It means that there is no change in technology during the time considered.
3. The market is perfectly competitive.
4. Outputs or returns are measured in physical terms.

Three phases of returns to scale

There are three phases of returns in the long-run which may be separately described as (1) the law of increasing returns (2) the law of constant returns and (3) the law of decreasing returns.

Depending on whether the proportionate change in output equals, exceeds, or falls short of the proportionate change in both the inputs, a production function is classified as showing constant, increasing or decreasing returns to scale.

a) Increasing Returns to Scale(IRS)

Increasing returns to scale or diminishing cost refers to a situation when all factors of production are increased, output increases at a higher rate. It means if all inputs are doubled, output will also increase at the faster rate than double. Hence, it is said to be increasing returns to scale. This increase is due to many reasons like division external economies of scale. Increasing returns to scale can be illustrated with the help of a diagram 8.

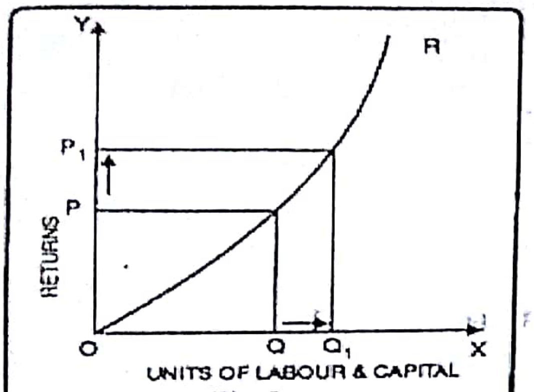


Fig 8.

In figure 8, OX axis represents increase in labour and capital while OY axis shows increase in output. When labour and capital increases from Q to Q_1, output also increases from P to P_1 which is higher than the factors of production i.e. labour and capital.

The firm enjoys various internal and external economies such as dimensional economies, economies flowing from indivisibility, economies of specialization, technical economies, managerial economies and marketing economies when it has IRS. Economies simply mean advantages for the firm. Due to these economies, the firm realizes increasing returns to scale. Marshall explains increasing returns in terms of "increased efficiency" of labor and capital in the improved organization with the expanding scale of output and employment factor unit. It is referred to as the economy of organization in the earlier stages of production.

b) Constant Returns to Scale

Constant returns to scale or constant cost refers to the production situation in which output increases exactly in the same proportion in which factors of production are increased. In simple terms, if factors of production are doubled output will also be doubled.

In this case internal and external economies are exactly equal to internal and external diseconomies. This situation arises when after reaching a certain level of production, economies of scale are balanced by diseconomies of scale. This is known as homogeneous production function. Cobb-Douglas linear homogenous production function is a good example of this kind. This is shown in diagram 10. In figure 10, we see that increase in factors of production i.e. labour and capital are equal to the proportion of output increase. Therefore, the result is constant returns to scale.

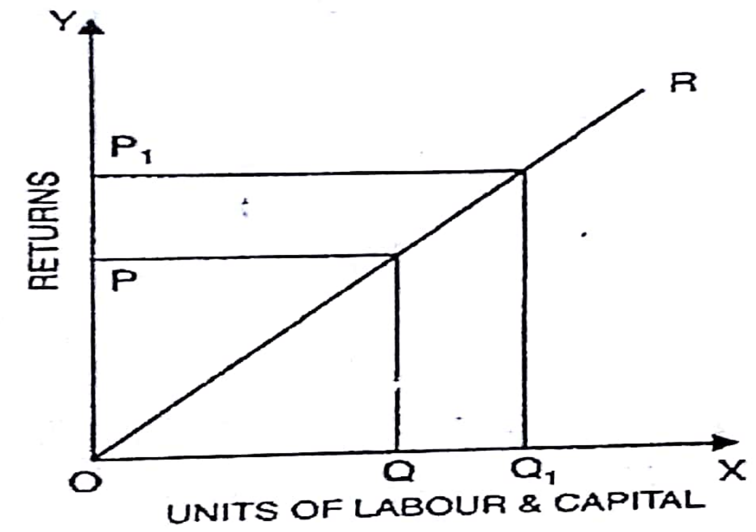


Fig. 10.

When a firm is at constant returns to scale, an increase in all inputs leads to a proportionate increase in output but to an extent.

A production function showing constant returns to scale is often called 'linear and homogeneous' or 'homogeneous of the first degree.' For example, the Cobb-Douglas production function is a linear and homogeneous production function.

c) Diminishing Returns to Scale

2. Diminishing Returns to Scale:

Diminishing returns or increasing costs refer to that production situation, where if all the factors of production are increased in a given proportion, output increases in a smaller proportion. It means, if inputs are doubled, output will be less than doubled. If 20 percent increase in labour and capital is followed by 10 percent increase in output, then it is an instance of diminishing returns to scale.

The main cause of the operation of diminishing returns to scale is that internal and external economies are less than internal and external diseconomies. It is clear from diagram 9.

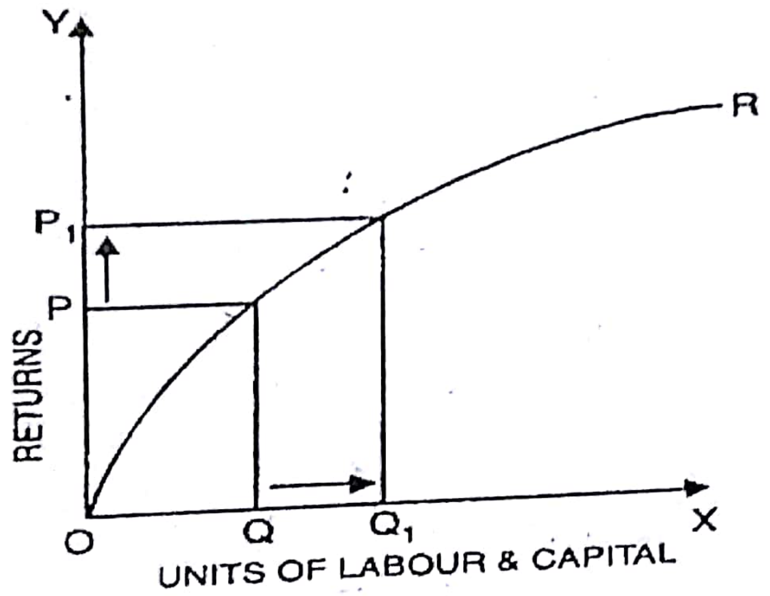
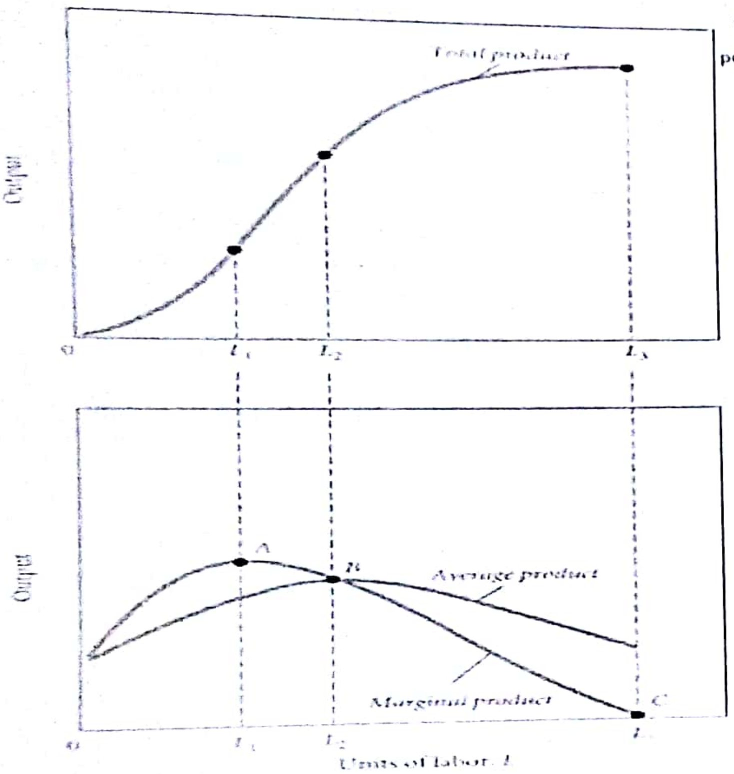
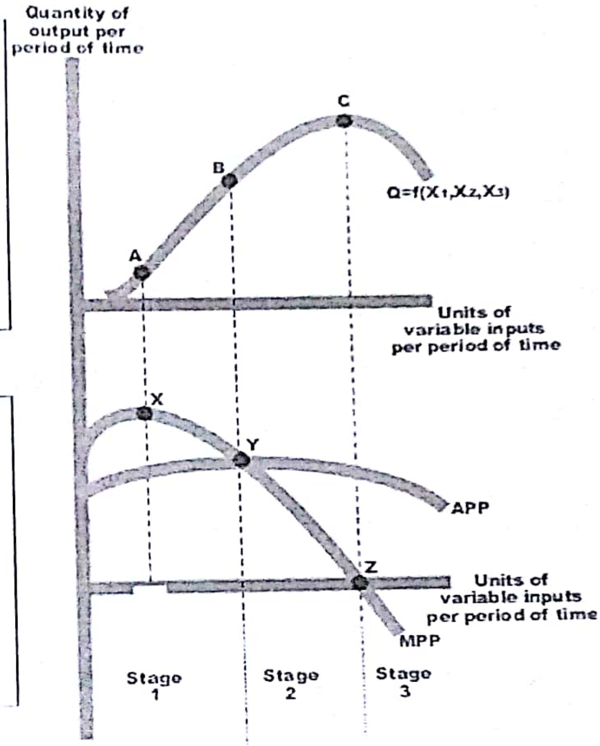


Fig. 9.

In this diagram 9, diminishing returns to scale has been shown. On OX axis, labour and capital are given while on OY axis, output. When factors of production increase from Q to Q_1 (more quantity) but as a result increase in output, i.e. P to P_1 is less. We see that increase in factors of production is more and increase in production is comparatively less, thus diminishing returns to scale apply.

This situation arises when a firm expands its operation even after the point of constant returns. Decreasing returns mean that increase in the total output is not proportionate according to the increase in the input. Because of this, the marginal output starts decreasing . Important factors that determine diminishing returns are managerial inefficiency and technical constraints.





See Samuelson

Industrial Economics

- firm, Industry, plant

- Market Structure

- - - - - - -

Perfect competition

- Short Run
- Long Run
- Monopoly

- hand notes
+ extra Sheet

Industrial → 01

Industrial Economics

Definition:

- Industrial Economics is defined as the application of micro-economic theory to the analysis of firms, market and industries;
- Industrial economics is a distinctive branch of economics which deals with the economic problems of firms and industries, and their interrelationship with society.

Plant, Firm and Industry Its Interrelationship: Stages of Industries

Plant: Plant is a unit of an Enterprise where production is carried out with the help of Machine and man. For Example-Steel plant.

Firm

- Firm refers to a unit of management which is operated under a trade name organized either to extract minerals, produce or manufacture goods or to sale gods or services, or to engage two or three of these activities simultaneously.
- Firm is a unit of organization within which the combination and deployment of the factors of production which are governed by a single, or at least a unified will. For example, sugar producing firm or a research center.

Industry

- Industry within the frame work of a perfectly competitive market structure can be defined as a large number firms competing each other in the production of homogeneous product. It refers to all manifold activities of country which offers employment but more narrowly those activities which are concerned with the productions of goods rather than services.
- Industry is considered as an aggregate of firms producing identical of substitutable products.

Stages of Industries

- Industries are divided into three stages:
  - Primary
  - Secondary
  - Tertiary

Contd.

- Primary: Industries are describe as primary because they carry out the first stages of production, produce raw material either by extracting or growing them.
- Secondary: Secondary is called processing and manufacturing. They change the raw materials into finished and semi- finished goods.
- Tertiary: Firm at this state provide serves to the primary and secondary Industries. They supplied materials or services directly to the customers.

Characteristics of Different Market Structures, Price and Output Determination under Perfectly Competitive Markets, Monopoly, Oligopoly and Monopolistic Competition

Market Structure

- Market structure is a particular situation which influences the firm's pricing and output decision.
- Markets are of different types:

1. Perfectly competitive market, and
2. Imperfectly competitive market.

- Imperfectly competitive markets are:

1. Monopoly market
2. Oligopoly market
3. Monopolistic competitive market.

A Comparative Structure of Perfect and Imperfect Markets

	Perfect Competition	Imperfect Competition	Monopoly
1	Large number of buyers and sellers	Large number of buyers, few sellers	Single seller, large number of buyers
2	Free entry and exit	Entry and exit are limited	No entry or exit
3	Products are homogenous	Product differentiation and price variation	Possibility or product differentiation
4	Normal price	High price	Abnormal price
5	Normal profit	High profit	Abnormal profit

A Comparative Structure of Perfect and Imperfect Markets (contd.)

	Perfect Competition	Imperfect Competition	Monopoly
6	Perfect knowledge	Imperfect knowledge	Imperfect knowledge
7	Seller price taker	Seller price setter	Seller price setter
8	No advertisement costs	High advertisement cost	Low advertisement cost
9	Buyers are not betrayed (well informed)	Buyers can be betrayed (Ignorance)	Buyers can be betrayed
10	Equilibrium MC=MR	Equilibrium MC=AR	Equilibrium MC=AR
11	Non-existence in real life	High existence in real life	Little existence in real life

Characteristics of a Perfectly Competitive Market

- There is large number of firms in the industry and the product is heterogeneous.
- Competition is perfect in the sense that every seller or buyer will sell or purchase an insignificant amount of output which cannot influence the price.
- Although market is perfect there is no rivalry among the individual firms. Each one acts autonomously disregarding the others.
- Price is determined in the market.
- Seller is price taker.
- Products of the firm are perfect substitutes for one another so that the price elasticity of demand curve of individual firm is infinite.

Characteristics of a Perfectly Competitive Market (contd.)

- The demand curve for perfectly competitive firm is horizontal.
- Together all buyers form the market demand curve.
- Equilibrium price is at the intersection point of the market demand and supply.
- Each firm takes the equilibrium point as a given.
- Each firm sells all quantities of output it produces at this point.
- Therefore, the demand curve of perfectly competitive market is horizontal.

Monopoly

- Monopoly is a market structure, where exists a single firm in the industry, there are no substitutes for commodities it produces.
- The demand of the monopolist coincides with the industry demand, which has finite price elasticity.

Characteristics of Monopoly

- Only one seller- but many buyers.
- No close substitute of the product.
- The total supply of the product is concentrated in a single firm.
- Entry is blocked.
- In monopoly, the demand of the firm is also the demand of the industry and hence is negative sloping.
- There is no supply curve.

Oligopoly

- In an oligopolistic market there is a small number of firms, so that sellers are conscious of their interdependence.
- Each firm takes into account the rival's reaction.
- The competition is not perfect, yet the rivalry among firms is high, unless they make a collusive agreement.
- The products are homogenous (pure oligopoly) or differentiated (differentiated oligopoly).

Collusive Oligopoly

➤ Cartel:

- Group of firms come to an agreement aiming at joint profit maximization by controlling a significant proportion of market share.

➤ Merger:

- A merger involves the decision of a number of firms to form a single corporation.
- The new firm may act as a cartel: it may decide to change the output quota of each plant so as to maximize the overall profit.
- Each plant or firm will be allocated a quota defined by the equality of marginal cost with the common marginal revenue.
- Cartel agreements are illegal in Britain and USA but merger is legal.

Three different types of price leadership

- Dominant: This price leadership occurs when one large firm sets price, other firms follow.
- Barometric: When a firm sets a price after consciously evaluating the market behavior and if it is proved to be correct, others follow
- Collusive: When a few large firms having a similar market share as well as market demand sets a price.

Monopolistic Competition

- Monopolistic Competition is one type of imperfect competition in the market which lies between monopoly and perfect competition.

Principal Characteristics of Monopolistic Competition

- The industry comprises so many firms that each ignores the possible reaction of its competitors when it makes its decision on price and output.
- The firm produces one specific variety or brand of the industry's differentiated product.

Principal Characteristics of Monopolistic Competition (contd.)

- There exists freedom of entry.
- There is a symmetry i.e , when a new firm enters the industry selling a new version of a product, the consumers take it equally from all existing firms.
- Large number of buyers.
- Each firm faces a demand curve, although negatively sloped, is highly elastic, because of product variation.

Literature

- Douglas Needham: Economic Analysis and Industrial Structure;
- J.F. Pickering: Industrial Structure and market conduct;
- James Kocls: Industrial organization and price;
- R.R. Borthwal: Industrial Economics;
- Roser Clarker: Industrial Economics.

Perfect competition

Question-1

Suppose the market demand for a good produced under perfect competition is given by $P = 100 - Q$. The total cost of the firm is $TC = 16 + 4q + q^2$ and price of the good is TK. 20 per unit. Calculate the equilibrium output, profit of the firm, number of existing firm both in short run and long run. Show the scenario graphically.

Short run

In the short run, the profit maximizing condition of the firm under perfect competition is $P = MC$

Here, we are given,

$$TC = 16 + 4q + 2q^2$$

$$MC = \frac{\Delta TC}{\Delta q} = 4 + 2q \quad \text{and} \quad P = 20$$

So, profit is maximized when, $P = MC$

$$\Rightarrow 20 = 4 + 2q$$

$$\Rightarrow -2q = 4 - 20$$

$$\Rightarrow 2q = 16 \quad \Rightarrow q = 8$$

(equilibrium quantity of a firm in short run under per. comp is $q = 8$

Profit in short run

$$\text{Profit}(\pi) = \cancel{TR}\ \text{Total revenue (TR)} - \text{Total cost (TC)}$$
$$= \cancel{PX}\ \text{Price}(P) \times \text{quantity}(q) - \cancel{TC}\ 16 + 4q + 2q^2$$
$$= 20 \times 8 - (16 + 4 \times 8 + 8^2)$$
$$= (20 \times 8) - (16 + 32 + 64)$$
$$= 160 - 112$$
$$= 48$$

Total market supply of the Industry in equilibrium:

$$P = \text{market Demand}$$
$$\Rightarrow 20 = 100 - Q$$
$$\Rightarrow Q = 100 - 20$$
$$\Rightarrow Q = 80$$

Number of firms in the industry in short run:

$$n = \frac{\text{total market supply}}{\text{firm's equilibrium output}}$$
$$= \frac{Q}{q} = \frac{80}{8} = 10$$

So, there will be 10 firms in the Industry supplying total output of 80 at the price 20. [Here we get price 20, do not need to find eq^m price because firm is a price taker, not setter here)

Marginal cost :

$MC = P$

$\Rightarrow 20 = 4 + 2q$

$\Rightarrow 20 = 4 + 2 \times 8$

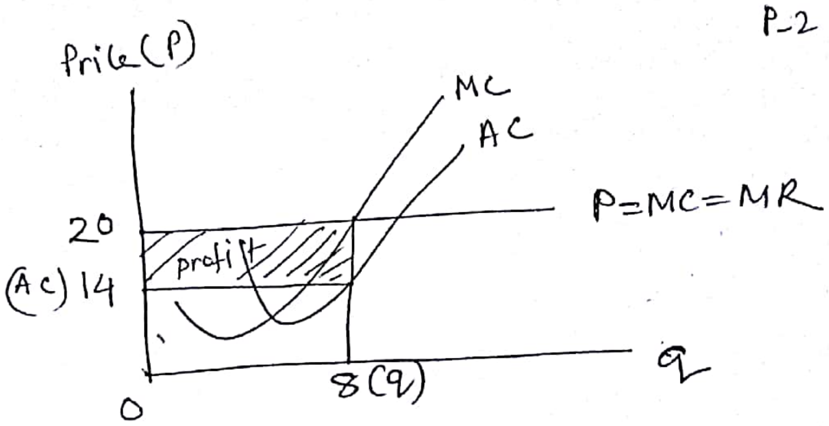


Fig – Short run equilibrium

Average Cost $= \frac{TC}{q}$

$= \frac{16 + 4q + q^2}{q}$

$= \cancel{16+4q}$ $\frac{16}{q} + 4 + q$

$= \frac{16}{8} + 4 + 8$

$= 2 + 4 + 8$

$= 14$

$\boxed{\text{Long run}}$

– In the Long run, profit maximizing condition for perfect compet is

Marginal Cost = Average cost = Price

$MC = AC = P$

– Given, $P = 100 - Q$

$TC = 16 + 4q + q^2$

$\frac{dTC}{dq} = 4 + 2q$

$AC = \frac{TC}{q} = \frac{16}{q} + \frac{4q}{q} + \frac{q^2}{q}$

$= \frac{16}{q} + 4 + q$

so, In Long run,

$$MC = AC$$

$$4+2q = \frac{16}{q} + 4 + q$$

$$\Rightarrow 2q = \frac{16}{q} + q$$

$$\Rightarrow 2q - q = \frac{16}{q}$$

$$\Rightarrow q = \frac{16}{q}$$

$$\Rightarrow q^2 = 16$$

$$\Rightarrow q = \sqrt{16}$$

$$\Rightarrow q = \pm 4$$ [output can not be negative]

$$\Rightarrow q = +4, -4$$

$\therefore$ equilibrium output of a firm in Long run $q=4$

$\therefore$ equilibrium price in Long run :

$$P = MC$$

$$= 4+2q$$

$$= 4 + 2\times 4$$

$$\Rightarrow 4+8$$

$$\Rightarrow 12$$

- Equilibrium supply of the industry in the long run

~~Q = 100 - Q~~ $P = 100 - Q$

$$\Rightarrow 12 = 100 - Q$$

$$\Rightarrow Q = 100 - 12$$

$$\Rightarrow Q = 88$$

Number of firms in the long run,

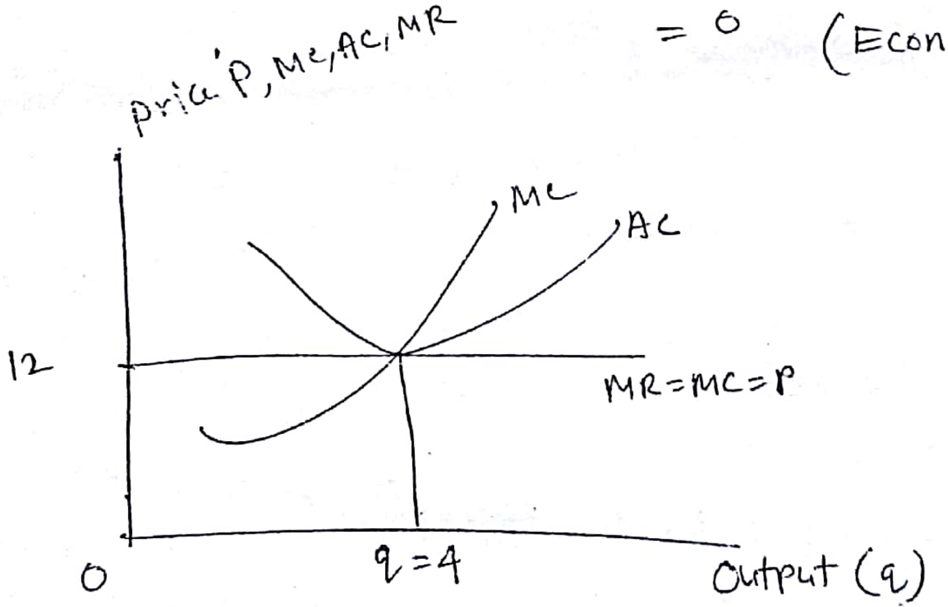
$$N = \frac{Q}{q} = \frac{\text{Market output}}{\text{firm's output}}$$

$$= \frac{88}{4} = 22$$

there will be 22 firms in the long run under perfect competition in the industry.

— profit of a firm in the long run:

$$\pi = TR - TC$$
$$= P \times q - (16 + 4q + q^2)$$
$$= 12 \times 4 - (16 + 4 \times 4 + 4^2)$$
$$= 48 - (16 + 16 + 16)$$
$$= 48 - 48$$
$$= 0$$ (Economic or normal profit



Here, $MC = P$

$\Rightarrow 4 + 2q = P$

$\Rightarrow 4 + 4 \times 4 = P$

$\Rightarrow P = 12$

$AC = \frac{16}{4} + 4 + 4$

$= 4 + 8$

$= 12$

Monopoly

-Question: A monopolist faces the following demand curve $P = 100 - Q$. The total cost is given by $TC = 10 + 4Q^2$. How much output will the monopolist produce? What price will it charge? What will be profit of the monopolist?

Here, $P = 100 - Q$

$TR = P \times Q = 100Q - Q^2$

$\frac{\Delta TR}{\Delta Q} = MR = 100 - 2Q$

And $TC = 10 + 4Q^2$

$\frac{\Delta TC}{\Delta Q} = 8Q$

Profit is maximized in monopoly market when

$MC = MR$

$\Rightarrow 8Q = 100 - 2Q$

$\Rightarrow 10Q = 100$

$\Rightarrow Q = 10$

So, output produced by the monopolist is $Q = 10$

~~Profit~~ Price charged by the monopolist

$P = 100 - Q$

$= 100 - 10$

$= 90$

Profit of the monopolist: $\pi = TR - TC$

$= \cancel{10}\ P \times Q - (10 + 4Q^2)$

$= 10 \times 90 - (10 + 4 \times 10^2)$

$= 900 - 410$

$= 490$

$MC = 8Q$
$= 8 \times 10$
$= 80$

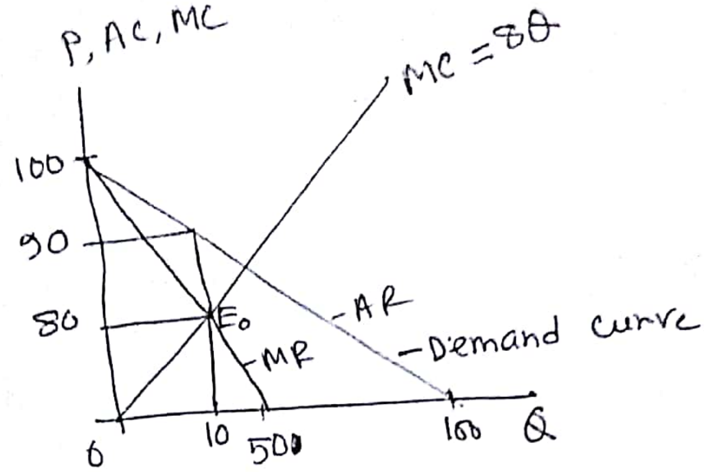


Fig – Equilibrium of a monopoly firm/market

Question

A monopolist faces the following demand curve $P = 100 - Q$ and average cost is given by $AC = 4 + Q$. How much output will it produce? What price will be charged? Suppose that the industry above were perfectly competitive. What will be the equilibrium price and quantity in such a situation? Calculate the size of the welfare loss of the society resulting from the presence of the monopolist?

Solution: We are given,

$$P = 100 - Q$$

$$\therefore TR = P \times Q = (100 - Q)Q = 100Q - Q^2$$

$$\frac{\Delta TR}{\Delta Q} = MR = 100 - 2Q$$

$AC = 4 + Q$

$\therefore TC = AC \times Q = \cancel{4Q} \div (4+Q)Q = 4Q + Q^2$

$\therefore \frac{\Delta TC}{\Delta Q} = MC = 4 + 2Q$

— A monopolist maximizes profit when MC = MR

so, $MC = MR$

$4 + 2Q = 100 - 2Q$

$\Rightarrow 4Q = 100 - \cancel{2Q}\,4$

$\Rightarrow 4Q = 96$

$\Rightarrow Q_m = \frac{96}{4} = 24$

Equilibrium output produced by monopolist is $Q = 24$

— Price charged by the monopolist,

$P_m = 100 - Q_m$ $(Q = 24)$

$= 100 - 24$

$= 76$

— Profit of the monopolist

$\pi_M = TR - TC$

$= P \times Q - (4Q + Q^2)$

$= (76 \times 24) - (4 \times 24 + 24^2)$

$= 1824 - 672$

$= 1152$

so, The monopolist sets price at $P = 76$ and sells output $Q = 24$ and makes profit of 1152.

Scanned by CamScanner

If the Industry would be under perfect competition :-

- under perfect competitive market, firm maximizes profit at $P = MC$.

So, $P = MC$

$100 - Q = 4 + 2Q$

$\Rightarrow -Q - 2Q = 4 - 100$

$\Rightarrow -3Q = -96$

$\therefore Q_c = 32$

Here, $P = 100 - Q$

$AC = 4 + 2Q$

$TC = 4Q + Q^2$

In case of perfect competitive market, equilibrium output will be $Q_c = 32$

- Equilibrium price under perfect competition :

$P_c = 100 - Q$

$= 100 - 32$

$= 68$

- profit under perfect competition:

$\pi_c = P_c \times Q_c - TC$

$= 68 \times 32 - (4 \times 32 + 32^2)$

$= 2176 - (128 + 1024)$

$= 2176 - 1152$

$= 1024$

Here, $\pi_m = 1152 > \pi_c = 1024$

profit under monopoly > profit under perfect competition.

Dead Weight loss (Welfare loss):

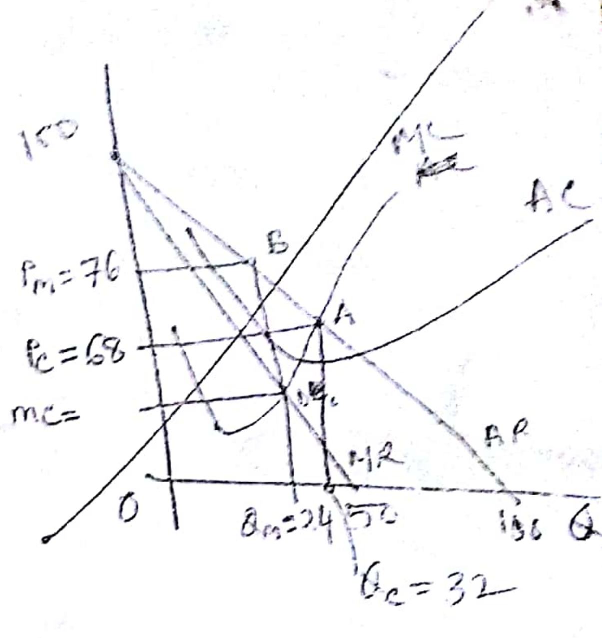
Welfare loss (BAD) =

$\frac{1}{2} \times (P_m - mc) \times (Q_c - Q_m)$

$= \frac{1}{2} \times (76 - \quad) \times (32 - 24)$

$= \frac{1}{2} \times 24 \times 8$

$= 96$



Here area BAD is Deadweight loss.

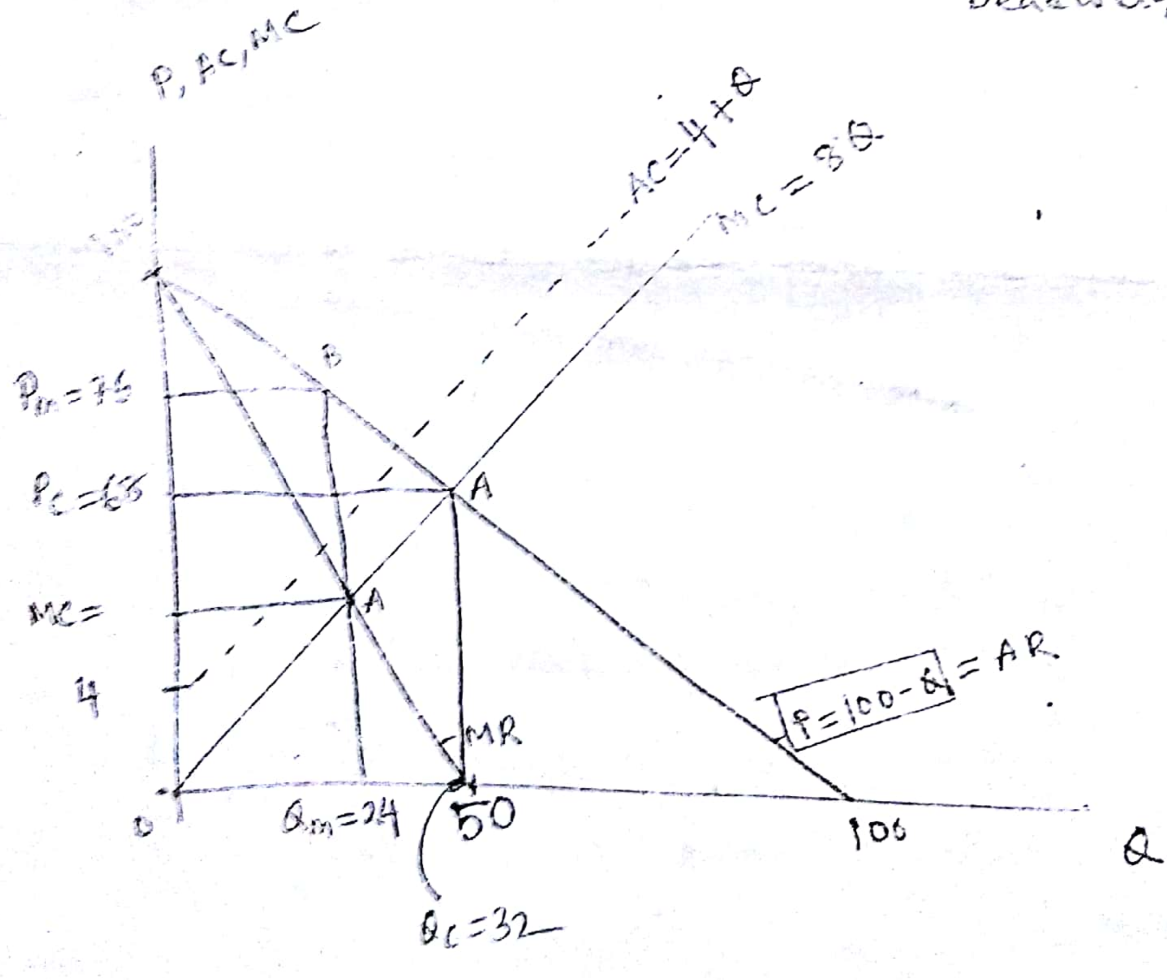


Fig. Measuring Welfare ~~loss~~ loss arising from monopoly.

Nazrul

Decision Time Frames

People who operate firms make many decisions, and all of their decisions are aimed at achieving one overriding goal: maximum attainable profit. But not all decisions are equally critical. Some decisions are big ones. Once made, they are costly (or impossible) to reverse. If such a decision turns out to be incorrect, it might lead to the failure of the firm. Other decisions are small. They are easily changed. If one of these decisions turns out to be incorrect, the firm can change its actions and survive.

The biggest decision that an entrepreneur makes is in what industry to establish a firm. For most entrepreneurs, their background knowledge and interests drive this decision. But the decision also depends on profit prospects—on the expectation that total revenue will exceed total cost.

Cindy has decided to set up Campus Sweaters. She has also decided the most effective method of organizing the firm. But she has not decided the quantity to produce, the factors of production to hire, or the price to charge for sweaters.

Decisions about the quantity to produce and the price to charge depend on the type of market in which the firm operates. Perfect competition, monopolistic competition, oligopoly, and monopoly all confront the firm with *different* problems. Decisions about *how* to produce a given output do not depend on the type of market in which the firm operates. *All* types of firms in *all* types of markets make similar decisions about how to produce.

The actions that a firm can take to influence the relationship between output and cost depend on how soon the firm wants to act. A firm that plans to change its output rate tomorrow has fewer options than one that plans to change its output rate six months or six years in the future.

To study the relationship between a firm's output decision and its costs, we distinguish between two decision time frames:

- The short run
- The long run

The Short Run

The **short run** is a time frame in which the quantity of at least one factor of production is fixed. For most firms, capital, land, and entrepreneurship are fixed factors of production and labor is the variable factor of production. We call the fixed factors of production the firm's *plant*: In the short run, a firm's plant is fixed.

For Campus Sweaters, the fixed plant is its factory building and its knitting machines. For an electric power utility, the fixed plant is its buildings, generators, computers, and control systems.

To increase output in the short run, a firm must increase the quantity of a variable factor of production, which is usually labor. So to produce more output, Campus Sweaters must hire more labor and operate its knitting machines for more hours a day. Similarly, an electric power utility must hire more labor and operate its generators for more hours a day.

Short-run decisions are easily reversed. The firm can increase or decrease its output in the short run by increasing or decreasing the amount of labor it hires.

The Long Run

The **long run** is a time frame in which the quantities of *all* factors of production can be varied. That is, the long run is a period in which the firm can change its *plant*.

To increase output in the long run, a firm can change its plant as well as the quantity of labor it hires. Campus Sweaters can decide whether to install more knitting machines, use a new type of machine, reorganize its management, or hire more labor. Long-run decisions are *not* easily reversed. Once a plant decision is made, the firm usually must live with it for some time. To emphasize this fact, we call the past expenditure on a plant that has no resale value a **sunk cost**. A sunk cost is irrelevant to the firm's current decisions. The only costs that influence its current decisions are the short-run cost of changing its labor inputs and the long-run cost of changing its plant.

REVIEW QUIZ

1. Distinguish between the short run and the long run.
2. Why is a sunk cost irrelevant to a firm's current decisions?

You can work these questions in Study Plan 11.1 and get instant feedback. MyEconLab

We're going to study costs in the short run and the long run. We begin with the short run and describe a firm's technology constraint.

Short-Run Technology Constraint

To increase output in the short run, a firm must increase the quantity of labor employed. We describe the relationship between output and the quantity of labor employed by using three related concepts:

1. Total product
2. Marginal product
3. Average product

These product concepts can be illustrated either by product schedules or by product curves. Let's look first at the product schedules.

Product Schedules

Table 11.1 shows some data that describe Campus Sweaters' total product, marginal product, and average product. The numbers tell us how the quantity of sweaters produced increases as Campus Sweaters employs more workers. The numbers also tell us about the productivity of the labor that Campus Sweaters employs.

Focus first on the columns headed "Labor" and "Total product." **Total product** is the maximum output that a given quantity of labor can produce. You can see from the numbers in these columns that as Campus Sweaters employs more labor, total product increases. For example, when 1 worker is employed, total product is 4 sweaters a day, and when 2 workers are employed, total product is 10 sweaters a day. Each increase in employment increases total product.

The **marginal product** of labor is the increase in total product that results from a one-unit increase in the quantity of labor employed, with all other inputs remaining the same. For example, in Table 11.1, when Campus Sweaters increases employment from 2 to 3 workers and does not change its capital, the marginal product of the third worker is 3 sweaters—total product increases from 10 to 13 sweaters.

Average product tells how productive workers are on average. The **average product** of labor is equal to total product divided by the quantity of labor employed. For example, in Table 11.1, the average product of 3 workers is 4.33 sweaters per worker—13 sweaters a day divided by 3 workers.

If you look closely at the numbers in Table 11.1, you can see some patterns. As Campus Sweaters hires more labor, marginal product increases initially, and then begins to decrease. For example, marginal product increases from 4 sweaters a day for the first worker to 6 sweaters a day for the second worker and then decreases to 3 sweaters a day for the third worker. Average product also increases at first and then decreases. You can see the relationships between the quantity of labor hired and the three product concepts more clearly by looking at the product curves.

TABLE 11.1 Total Product, Marginal Product, and Average Product

	Labor (workers per day)	Total product (sweaters per day)	Marginal product (sweaters per additional worker)	Average product (sweaters per worker)
A	0	0		
			4	
B	1	4		4.00
			6	
C	2	10		5.00
			3	
D	3	13		4.33
			2	
E	4	15		3.75
			1	
F	5	16		3.20

Total product is the total amount produced. Marginal product is the change in total product that results from a one-unit increase in labor. For example, when labor increases from 2 to 3 workers a day (row *C* to row *D*), total product increases from 10 to 13 sweaters a day. The marginal product of going from 2 to 3 workers is 3 sweaters. Average product is total product divided by the quantity of labor employed. For example, the average product of 3 workers is 4.33 sweaters per worker (13 sweaters a day divided by 3 workers).

Product Curves

The product curves are graphs of the relationships between employment and the three product concepts you've just studied. They show how total product, marginal product, and average product change as employment changes. They also show the relationships among the three concepts. Let's look at the product curves.

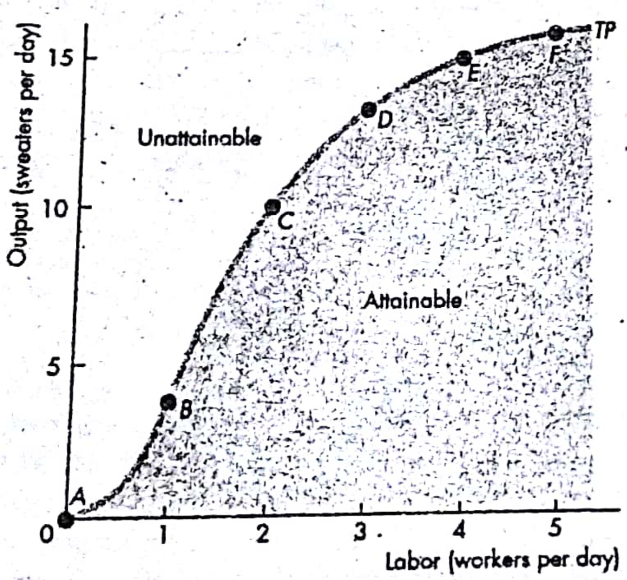
Total Product Curve

Figure 11.1 shows Campus Sweaters' total product curve, *TP*, which is a graph of the total product schedule. Points *A* through *F* correspond to rows *A* through *F* in Table 11.1. To graph the entire total product curve, we vary labor by hours rather than whole days.

Notice the shape of the total product curve. As employment increases from zero to 1 worker a day, the curve becomes steeper. Then, as employment increases to 3, 4, and 5 workers a day, the curve becomes less steep.

The total product curve is similar to the *production possibilities frontier* (explained in Chapter 2). It separates the attainable output levels from those that are unattainable. All the points that lie above the curve are unattainable. Points that lie below the curve, in the orange area, are attainable, but they are inefficient—they use more labor than is necessary to produce a given output. Only the points *on* the total product curve are technologically efficient.

FIGURE 11.1 Total Product Curve



The total product curve, *TP*, is based on the data in Table 11.1. The total product curve shows how the quantity of sweaters produced changes as the quantity of labor employed changes. For example, 2 workers can produce 10 sweaters a day (point *C*). Points *A* through *F* on the curve correspond to the rows of Table 11.1. The total product curve separates attainable outputs from unattainable outputs. Points below the *TP* curve are inefficient.

MyEconLab animation

Marginal Product Curve

Figure 11.2 shows Campus Sweaters' marginal product of labor. Part (a) reproduces the total product curve from Fig. 11.1 and part (b) shows the marginal product curve, *MP*.

In part (a), the orange bars illustrate the marginal product of labor. The height of a bar measures marginal product. Marginal product is also measured by the slope of the total product curve. Recall that the slope of a curve is the change in the value of the variable measured on the *y*-axis—output—divided by the change in the variable measured on the *x*-axis—labor—as we move along the curve. A one-unit increase in labor, from 2 to 3 workers, increases output from 10 to 13 sweaters, so the slope from point *C* to point *D* is 3 sweaters per additional worker, the same as the marginal product we've just calculated.

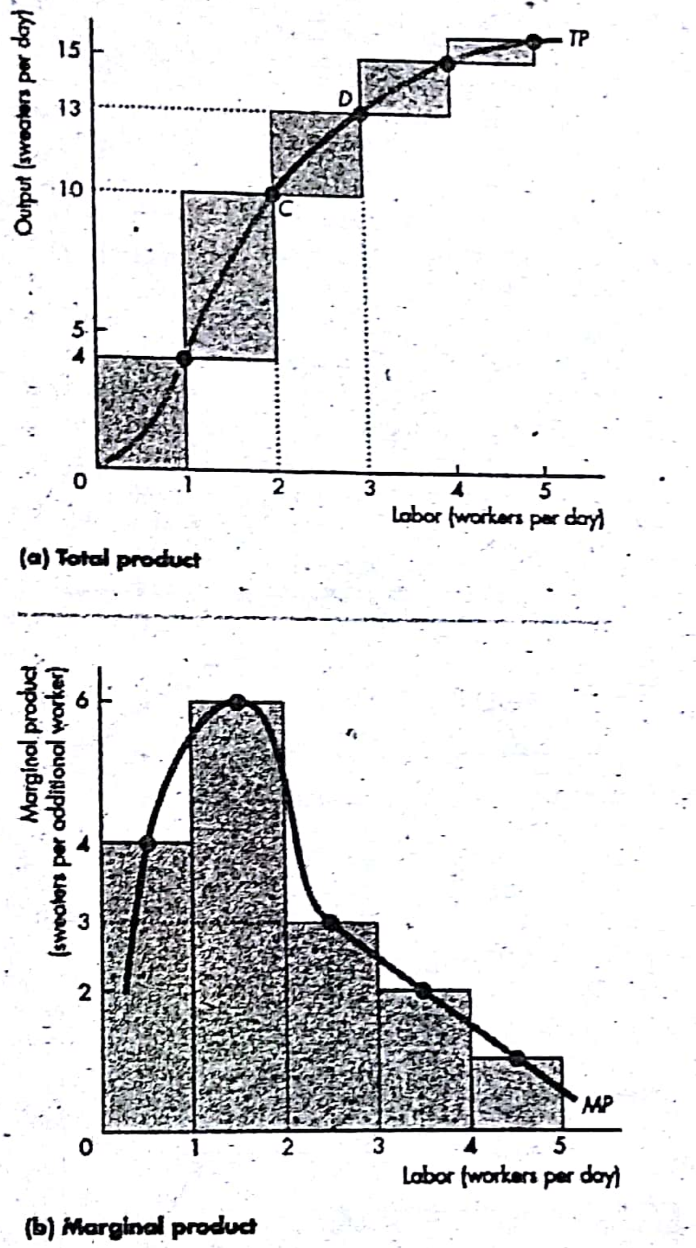
Again varying the amount of labor in the smallest units possible, we can draw the marginal product curve shown in Fig. 11.2(b). The *height* of this curve measures the *slope* of the total product curve at a point. Part (a) shows that an increase in employment from 2 to 3 workers increases output from 10 to 13 sweaters (an increase of 3). The increase in output of 3 sweaters appears on the *y*-axis of part (b) as the marginal product of going from 2 to 3 workers. We plot that marginal product at the midpoint between 2 and 3 workers. Notice that the marginal product shown in Fig. 11.2(b) reaches a peak at 1.5 workers, and at that point, marginal product is 6 sweaters per additional worker. The peak occurs at 1.5 workers because the total product curve is steepest when employment increases from 1 worker to 2 workers.

The total product and marginal product curves differ across firms and types of goods. GM's product curves are different from those of PennPower, whose curves in turn are different from those of Campus Sweaters. But the shapes of the product curves are similar because almost every production process has two features:

- Increasing marginal returns initially
- Diminishing marginal returns eventually

Increasing Marginal Returns Increasing marginal returns occur when the marginal product of an additional worker exceeds the marginal product of the previous worker. Increasing marginal returns arise from increased specialization and division of labor in the production process.

FIGURE 11.2 Total Product and Marginal Product



Marginal product is illustrated by the orange bars. For example, when labor increases from 2 to 3 workers a day, marginal product is the orange bar whose height is 3 sweaters. (Marginal product is shown midway between the quantities of labor to emphasize that marginal product results from *changing* the quantity of labor.) The steeper the slope of the total product curve (*TP*) in part (a), the larger is marginal product (*MP*) in part (b). Marginal product increases to a maximum (in this example when 1.5 workers a day are employed) and then declines—diminishing marginal product.

MyEconLab animation

For example, if Campus Sweaters employs one worker, that person must learn all the aspects of sweater production: running the knitting machines, fixing breakdowns, packaging and mailing sweaters, buying and checking the type and color of the wool. All these tasks must be performed by that one person.

If Campus Sweaters hires a second person, the two workers can specialize in different parts of the production process and can produce more than twice as much as one worker. The marginal product of the second worker is greater than the marginal product of the first worker. Marginal returns are increasing.

Diminishing Marginal Returns Most production processes experience increasing marginal returns initially, but all production processes eventually reach a point of *diminishing* marginal returns. **Diminishing marginal returns** occur when the marginal product of an additional worker is less than the marginal product of the previous worker.

Diminishing marginal returns arise from the fact that more and more workers are using the same capital and working in the same space. As more workers are added, there is less and less for the additional workers to do that is productive. For example, if Campus Sweaters hires a third worker, output increases but not by as much as it did when it hired the second worker. In this case, after two workers are hired, all the gains from specialization and the division of labor have been exhausted. By hiring a third worker, the factory produces more sweaters, but the equipment is being operated closer to its limits. There are even times when the third worker has nothing to do because the machines are running without the need for further attention. Hiring more and more workers continues to increase output but by successively smaller amounts. Marginal returns are diminishing. This phenomenon is such a pervasive one that it is called a "law"—the law of diminishing returns. The **law of diminishing returns** states that

> As a firm uses more of a variable factor of production with a given quantity of the fixed factor of production, the marginal product of the variable factor eventually diminishes.

You are going to return to the law of diminishing returns when we study a firm's costs, but before we do that, let's look at the average product of labor and the average product curve.

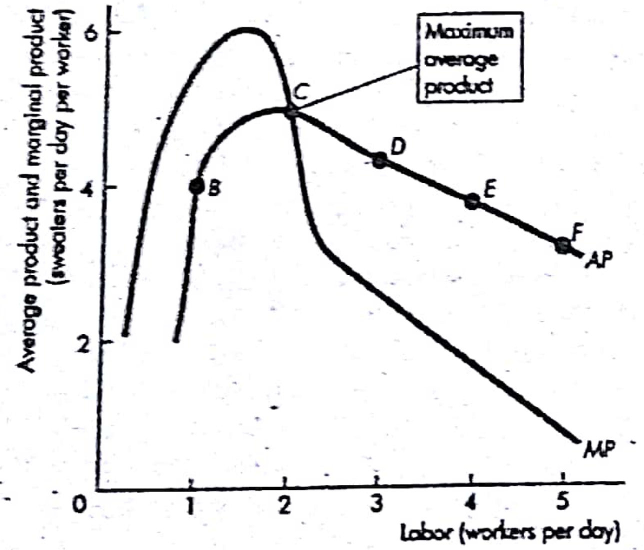
Scanned by CamScanner

Average Product Curve

Figure 11.3 illustrates Campus Sweaters' average product of labor and shows the relationship between average product and marginal product. Points *B* through *F* on the average product curve *AP* correspond to those same rows in Table 11.1. Average product increases from 1 to 2 workers (its maximum value at point *C*) but then decreases as yet more workers are employed. Notice also that average product is largest when average product and marginal product are equal. That is, the marginal product curve cuts the average product curve at the point of maximum average product. For the number of workers at which marginal product exceeds average product, average product is *increasing*. For the number of workers at which marginal product is less than average product, average product is *decreasing*.

The relationship between the average product and marginal product is a general feature of the relationship between the average and marginal values of any variable—even your grades.

FIGURE 11.3 Average Product



The figure shows the average product of labor and the connection between average product and marginal product. With 1 worker, marginal product exceeds average product, so average product is increasing. With 2 workers, marginal product equals average product, so average product is at its maximum. With more than 2 workers, marginal product is less than average product, so average product is decreasing.

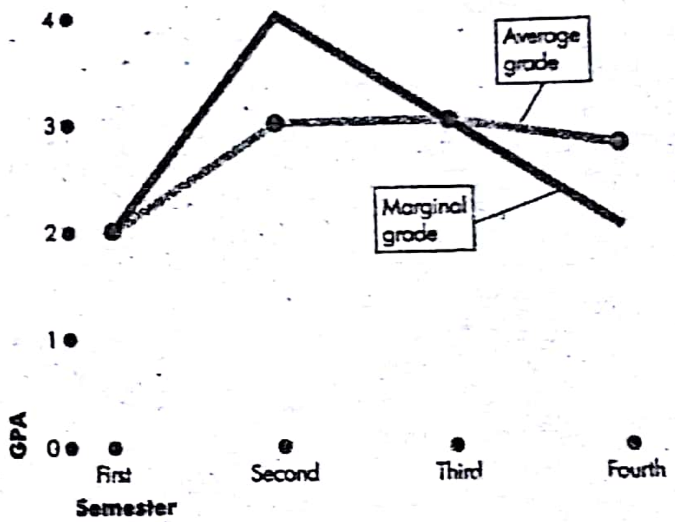
MyEconLab animation

Economics in Action

How to Pull Up Your Average

Do you want to pull up your average grade? Then make sure that your grade this semester is better than your current average! This semester is your marginal semester. If your marginal grade exceeds your average grade (like the second semester in the figure), your average will rise. If your marginal grade equals your average grade (like the third semester in the figure), your average won't change. If your marginal grade is below your average grade (like the fourth semester in the figure), your average will fall.

The relationship between your marginal and average grades is exactly the same as that between marginal product and average product.



Marginal and Average Grade Curves

REVIEW QUIZ

1 Explain how the marginal product and average product of labor change as the labor employed increases (a) initially and (b) eventually.
2 What is the law of diminishing returns? Why does marginal product eventually diminish?
3 Explain the relationship between marginal product and average product.

You can work these questions in Study Plan 11.2 and get instant feedback. MyEconLab

Campus Sweaters' product curves influence its costs, as you are now going to see.

Short-Run Cost

To produce more output in the short run, a firm must employ more labor, which means that it must increase its costs. We describe the relationship between output and cost by using three cost concepts:

- Total cost
- Marginal cost
- Average cost

Total Cost

A firm's **total cost** (*TC*) is the cost of *all* the factors of production it uses. We separate total cost into total *fixed* cost and total *variable* cost.

Total fixed cost (*TFC*) is the cost of the firm's fixed factors. For Campus Sweaters, total fixed cost includes the cost of renting knitting machines and *normal profit*, which is the opportunity cost of Cindy's entrepreneurship (see Chapter 10, p. 225). The quantities of fixed factors don't change as output changes, so total fixed cost is the same at all outputs.

Total variable cost (*TVC*) is the cost of the firm's variable factors. For Campus Sweaters, labor is the variable factor, so this component of cost is its wage bill. Total variable cost changes as output changes.

Total cost is the sum of total fixed cost and total variable cost. That is,

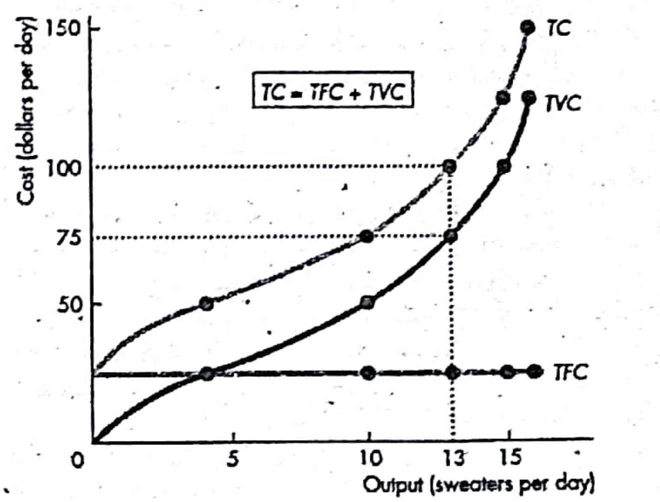
$$TC = TFC + TVC.$$

The table in Fig. 11.4 shows total costs. Campus Sweaters rents one knitting machine for \$25 a day, so its *TFC* is \$25. To produce sweaters, the firm hires labor, which costs \$25 a day. *TVC* is the number of workers multiplied by \$25. For example, to produce 13 sweaters a day, in row *D*, the firm hires 3 workers and *TVC* is \$75. *TC* is the sum of *TFC* and *TVC*, so to produce 13 sweaters a day, *TC* is \$100. Check the calculations in the other rows of the table.

Figure 11.4 shows Campus Sweaters' total cost curves, which graph total cost against output. The green *TFC* curve is horizontal because total fixed cost (\$25 a day) does not change when output changes. The purple *TVC* curve and the blue *TC* curve both slope upward because to increase output, more labor must be employed, which increases total variable cost. Total fixed cost equals the vertical distance between the *TVC* and *TC* curves.

Let's now look at a firm's marginal cost.

FIGURE 11.4 Total Cost Curves



	Labor (workers per day)	Output (sweaters per day)	Total fixed cost (*TFC*) (dollars per day)	Total variable cost (*TVC*) (dollars per day)	Total cost (*TC*) (dollars per day)
A	0	0	25	0	25
B	1	4	25	25	50
C	2	10	25	50	75
D	**3**	**13**	**25**	**75**	**100**
E	4	15	25	100	125
F	5	16	25	125	150

Campus Sweaters rents a knitting machine for \$25 a day, so this cost is the firm's total fixed cost. The firm hires workers at a wage rate of \$25 a day, and this cost is its total variable cost. For example, in row *D*, Campus Sweaters employs 3 workers and its total variable cost is 3 × \$25, which equals \$75. Total cost is the sum of total fixed cost and total variable cost. For example, when Campus Sweaters employs 3 workers, total cost is \$100—total fixed cost of \$25 plus total variable cost of \$75.

The graph shows Campus Sweaters' total cost curves. Total fixed cost is constant—the *TFC* curve is a horizontal line. Total variable cost increases as output increases, so the *TVC* curve and the *TC* curve increase as output increases. The vertical distance between the *TC* curve and the *TVC* curve equals total fixed cost, as illustrated by the two arrows.

MyEconLab animation

Marginal Cost

Figure 11.4 shows that total variable cost and total cost increase at a decreasing rate at small outputs but eventually, as output increases, total variable cost and total cost increase at an increasing rate. To understand this pattern in the change in total cost as output increases, we need to use the concept of *marginal cost.*

A firm's **marginal cost** is the increase in total cost that results from a one-unit increase in output. We calculate marginal cost as the increase in total cost divided by the increase in output. The table in Fig. 11.5 shows this calculation. When, for example, output increases from 10 sweaters to 13 sweaters, total cost increases from $75 to $100. The change in output is 3 sweaters, and the change in total cost is $25. The marginal cost of one of those 3 sweaters is ($25 ÷ 3), which equals $8.33.

Figure 11.5 graphs the marginal cost data in the table as the red marginal cost curve, *MC.* This curve is U-shaped because when Campus Sweaters hires a second worker, marginal cost decreases, but when it hires a third, a fourth, and a fifth worker, marginal cost successively increases.

At small outputs, marginal cost decreases as output increases because of greater specialization and the division of labor. But as output increases further, marginal cost eventually increases because of the *law of diminishing returns.* The law of diminishing returns means that the output produced by each additional worker is successively smaller. To produce an additional unit of output, ever more workers are required, and the cost of producing the additional unit of output—marginal cost—must eventually increase.

Marginal cost tells us how total cost changes as output increases. The final cost concept tells us what it costs, on average, to produce a unit of output. Let's now look at Campus Sweaters' average costs.

Average Cost

Three average costs of production are

1. Average fixed cost
2. Average variable cost
3. Average total cost

Average fixed cost (*AFC*) is total fixed cost per unit of output. **Average variable cost** (*AVC*) is total variable cost per unit of output. **Average total cost** (*ATC*) is total cost per unit of output. The average cost concepts are calculated from the total cost concepts as follows:

$$TC = TFC + TVC.$$

Divide each total cost term by the quantity produced, Q, to get

$$\frac{TC}{Q} = \frac{TFC}{Q} + \frac{TVC}{Q},$$

or

$$ATC = AFC + AVC.$$

The table in Fig. 11.5 shows the calculation of average total cost. For example, in row *C*, output is 10 sweaters. Average fixed cost is ($25 ÷ 10), which equals $2.50, average variable cost is ($50 ÷ 10), which equals $5.00, and average total cost is ($75 ÷ 10), which equals $7.50. Note that average total cost is equal to average fixed cost ($2.50) plus average variable cost ($5.00).

Figure 11.5 shows the average cost curves. The green average fixed cost curve (*AFC*) slopes downward. As output increases, the same constant total fixed cost is spread over a larger output. The blue average total cost curve (*ATC*) and the purple average variable cost curve (*AVC*) are U-shaped. The vertical distance between the average total cost and average variable cost curves is equal to average fixed cost—as indicated by the two arrows. That distance shrinks as output increases because average fixed cost declines with increasing output.

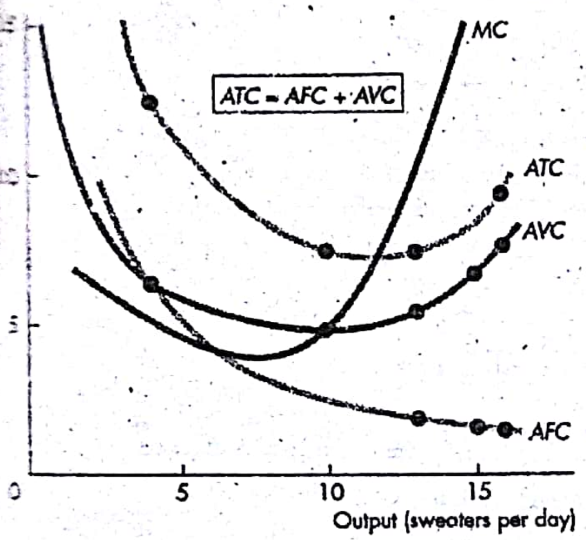
Marginal Cost and Average Cost

The marginal cost curve (*MC*) intersects the average variable cost curve and the average total cost curve *at their minimum points.* When marginal cost is less than average cost, average cost is decreasing, and when marginal cost exceeds average cost, average cost is increasing. This relationship holds for both the *ATC* curve and the *AVC* curve. It is another example of the relationship you saw in Fig. 11.3 for average product and marginal product and in your average and marginal grades.

Why the Average Total Cost Curve Is U-Shaped

Average total cost is the sum of average fixed cost an average variable cost, so the shape of the *ATC* curve

FIGURE 11.5 Marginal Cost and Average Costs



Marginal cost is calculated as the change in total cost divided by the change in output. When output increases from 4 to 10 sweaters, an increase of 6 sweaters, total cost increases by \$25. Marginal cost is \$25 ÷ 6, which is \$4.17.

Each average cost concept is calculated by dividing the related total cost by output. When 10 sweaters are produced, *AFC* is \$2.50 (\$25 ÷ 10), *AVC* is \$5 (\$50 ÷ 10), and *ATC* is \$7.50 (\$75 ÷ 10).

The graph shows that the *MC* curve is U-shaped and intersects the *AVC* curve and the *ATC* curve at their minimum points. The average fixed cost curve (*AFC*) is downward sloping. The *ATC* curve and *AVC* curve are U-shaped. The vertical distance between the *ATC* curve and the *AVC* curve is equal to average fixed cost, as illustrated by the two arrows.

	Labor (workers per day)	Output (sweaters per day)	Total fixed cost (*TFC*)	Total variable cost (*TVC*)	Total cost (*TC*)	Marginal cost (*MC*)	Average fixed cost (*AFC*)	Average variable cost (*AVC*)	Average total cost (*ATC*)
			(dollars per day)			(dollars per additional sweater)	(dollars per sweater)		
A	0	0	25	0	25		—	—	—
						6.25			
B	1	4	25	25	50		6.25	6.25	12.50
						4.17			
C	2	10	25	50	75		2.50	5.00	7.50
						8.33			
D	3	13	25	75	100		1.92	5.77	7.69
						12.50			
E	4	15	25	100	125		1.67	6.67	8.33
						25.00			
F	5	16	25	125	150		1.56	7.81	9.38

MyEconLab animation

combines the shapes of the *AFC* and *AVC* curves. The U shape of the *ATC* curve arises from the influence of two opposing forces:

1. Spreading total fixed cost over a larger output
2. Eventually diminishing returns

When output increases, the firm spreads its total fixed cost over a larger output and so its average fixed cost decreases—its *AFC* curve slopes downward.

Diminishing returns means that as output increases, ever-larger amounts of labor are needed to produce an additional unit of output. So as output increases, average variable cost decreases initially but eventually increases, and the *AVC* curve slopes upward. The *AVC* curve is U-shaped.

The shape of the *ATC* curve combines these two effects. Initially, as output increases, both average fixed cost and average variable cost decrease, so average total cost decreases. The *ATC* curve slopes downward.

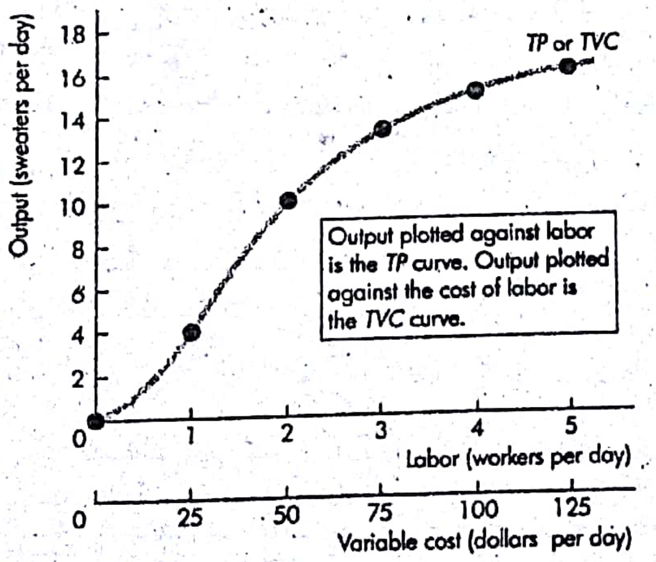
But as output increases further and diminishing returns set in, average variable cost starts to increase. With average fixed cost decreasing more quickly than average variable cost is increasing, the *ATC* curve continues to slope downward. Eventually, average variable cost starts to increase more quickly than average fixed cost decreases, so average total cost starts to increase. The *ATC* curve slopes upward.

Cost Curves and Product Curves

The technology that a firm uses determines its costs. A firm's cost curves come directly from its product curves. You've used this link in the tables in which we have calculated total cost from the total product schedule and information about the prices of the factors of production. We're now going to get a clearer view of the link between the product curves and the cost curves. We'll look first at the link between total cost and total product and then at the links between the average and marginal product and cost curves.

Total Product and Total Variable Cost Figure 11.6 shows the links between the firm's total product curve, *TP*, and its total *variable* cost curve, *TVC*. The graph is a bit unusual in two ways. First, it measures two variables on the *x*-axis—labor and variable cost. Second, it graphs the *TVC* curve but with variable cost on the *x*-axis and output on the *y*-axis. The graph can show labor and cost on the *x*-axis because variable cost is proportional to labor. One worker costs $25 a day. Graphing output against labor gives the *TP* curve and graphing variable cost against output gives the *TVC* curve.

FIGURE 11.6 Total Product and Total Variable Cost



The figure shows the total product curve, *TP*, as a graph of output (sweaters per day) plotted against labor (workers per day). It also shows the total variable cost curve, *TVC*, as a graph of total variable cost (dollars per day) against output. The only difference between the *TVC* curve here and that in Fig. 11.4 is that we've switched the *x*-axis and *y*-axis.

MyEconLab animation

ECONOMICS IN THE NE

Checkout Cost Curves

Welcome to More Self-Checkouts

Self-checkout is expanding. One estimate is that it will grow by 84 percent over the next five years. And new technologies will enable customers armed with smart phones to use apps that scan and pay for items in the aisles.

Source: MSNBC.com, July 22, 2011

DATA AND ASSUMPTIONS

A grocery store paid $20,000 to install 5 worker-operated checkout lines. With a life of 9 years and operating for 10 hours a day, these machines have an *implicit rental rate* of $1.00 an hour. Checkout clerks can be hired for $10 an hour. The total product schedule (checkouts per hour) for this store is:

Checkout clerks	1	2	3	4	5
Checkouts per hour	12	22	30	36	40

Another grocery store has converted to all self-checkout. It paid $100,000 to install a 5-line self-operated system. With a 5-year life and operating for 10 hours a day, the system has an *implicit rental rate* of $7.00 an hour. It hires checkout assistants to help customers at $10 an hour—the same wage as paid to checkout clerks. The total product schedule for this store is:

Checkout assistants	1	1	1	2
Checkouts per hour	12	22	30	36

That is, one checkout assistant can help shoppers check out up to a rate of 30 an hour and a second assistant can boost output to 36 an hour. (Shoppers using self-checkout aren't as quick as clerks, so the fastest rate at which this store can check out customers is 36 an hour.)

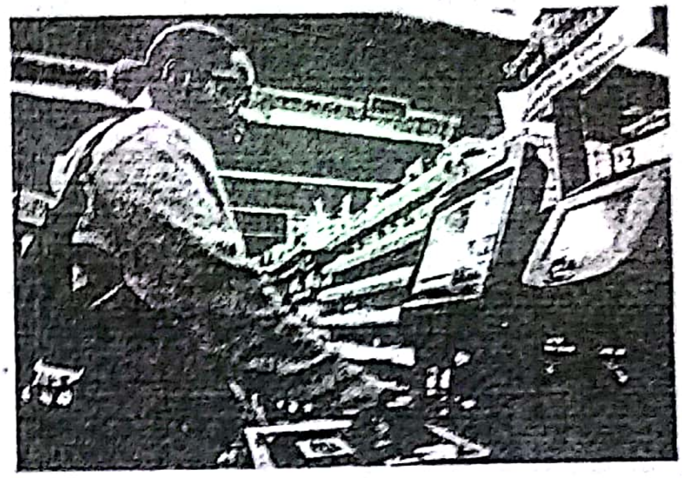
THE PROBLEM

- Which checkout system has the lower average total cost (*ATC*)? Which system has the lower marginal cost (*MC*)? Sketch the *ATC* and *MC* curves for the two systems.

THE SOLUTION

- Start with the worker-operated checkout system. Fixed cost is $1.00 per hour and variable cost is $10.00 per clerk. So the total cost schedule is:

Checkout clerks	1	2	3	4	5
Checkouts per hour	12	22	30	36	40
Total cost (*TC*) per hour	11	21	31	41	51



- Calculate MC as the change in TC divided by the change in output (change in number of checkouts) and calculate ATC as TC divided by output to get:

Checkouts per hour	12	22	30	36	40
Marginal cost (*MC*)	0.83	1.00	1.25	1.67	2.50
Average total cost (*ATC*)	0.92	0.95	1.03	1.14	1.28

- Figure 1 graphs the MC and ATC values at each output rate.

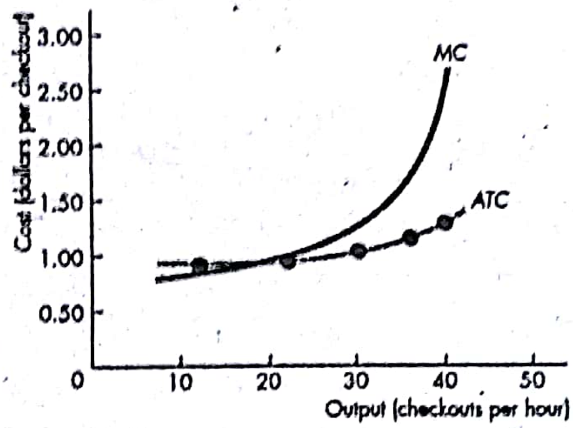


Figure 1 Operator Checkout

- Now do similar calculations for the self-checkout system. Fixed cost is \$7.00 per hour and variable cost is \$10.00 per clerk hour. So the total cost schedule is:

Checkout assistants	1	1	1	2
Checkouts per hour	12	22	30	36
Total cost (*TC*) per hour	17	17	17	27

- Calculate MC and ATC in the same way as before to get

Checkouts per hour	12	22	30	36
Marginal cost (*MC*)	0.83	0	0	1.67
Average total cost (*ATC*)	1.42	0.77	0.57	0.75

- Figure 2 graphs the MC and ATC values at each output rate.
- Figure 3 compares the ATC of the two systems. You can see that the self-checkout system has higher ATC at low output rates and lower ATC at higher output rates. The reason is that self-checkout has a higher fixed cost and lower variable cost than the worker-operated system.

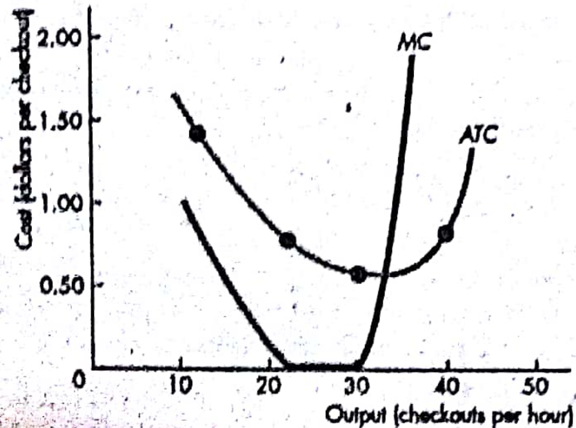


Figure 2 Self-Checkout

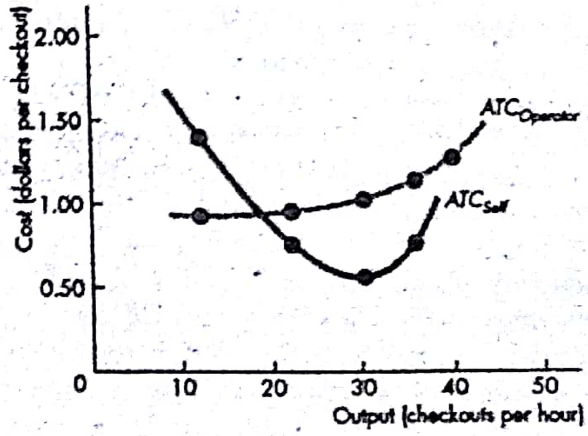


Figure 3 ATC Compared

Average and Marginal Product and Cost Figure 11.7 shows the links between the firm's average and marginal product curves and its average and marginal cost curves. The upper graph shows the average product curve, *AP*, and the marginal product curve, *MP*—like those in Fig. 11.3. The lower graph shows the average variable cost curve, *AVC*, and the marginal cost curve, *MC*—like those in Fig. 11.5.

As labor increases up to 1.5 workers a day (upper graph), output increases to 6.5 sweaters a day (lower graph). Marginal product and average product rise and marginal cost and average variable cost fall. At the point of maximum marginal product, marginal cost is at a minimum.

As labor increases from 1.5 workers to 2 workers a day, (upper graph) output increases from 6.5 sweaters to 10 sweaters a day (lower graph). Marginal product falls and marginal cost rises, but average product continues to rise and average variable cost continues to fall. At the point of maximum average product, average variable cost is at a minimum. As labor increases further, output increases. Average product diminishes and average variable cost increases.

Shifts in the Cost Curves

The position of a firm's short-run cost curves depends on two factors:

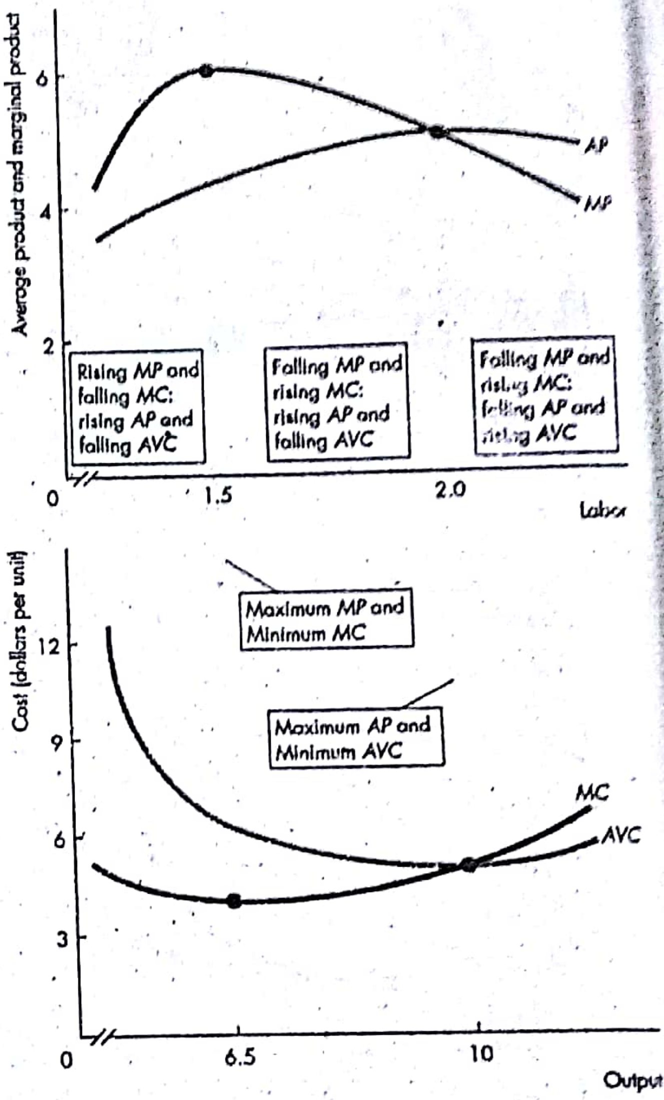
- Technology
- Prices of factors of production

Technology A technological change that increases productivity increases the marginal product and average product of labor. With a better technology, the same factors of production can produce more output, so the technological advance lowers the costs of production and shifts the cost curves downward.

For example, advances in robot production techniques have increased productivity in the automobile industry. As a result, the product curves of Chrysler, Ford, and GM have shifted upward and their cost curves have shifted downward. But the relationships between their product curves and cost curves have not changed. The curves are still linked in the way shown in Figs. 11.6 and 11.7.

Often, as in the case of robots producing cars, a technological advance results in a firm using more capital, a fixed factor, and less labor, a variable factor.

FIGURE 11.7 Average and Marginal Product Curves and Cost Curves



A firm's *MP* curve is linked to its *MC* curve. If, as the firm increases its labor from 0 to 1.5 workers a day, the firm's marginal product rises, its marginal cost falls. If marginal product is at a maximum, marginal cost is at a minimum. If, as the firm hires more labor, its marginal product diminishes, its marginal cost rises.

A firm's *AP* curve is linked to its *AVC* curve. If, as the firm increases its labor to 2 workers a day, its average product rises, its average variable cost falls. If average product is at a maximum, average variable cost is at a minimum. If, as the firm hires more labor, its average product diminishes, its average variable cost rises.

MyEconLab animation

TABLE 11.2 A Compact Glossary of Costs

Term	Symbol	Definition	Equation
Fixed cost		Cost that is independent of the output level; cost of a fixed factor of production	
Variable cost		Cost that varies with the output level; cost of a variable factor of production	
Total fixed cost	*TFC*	Cost of the fixed factors of production	
Total variable cost	*TVC*	Cost of the variable factors of production	
Total cost	*TC*	Cost of all factors of production	$TC = TFC + TVC$
Output (total product)	*TP*	Total quantity produced (output *Q*)	
Marginal cost	*MC*	Change in total cost resulting from a one-unit increase in total product	$MC = \Delta TC \div \Delta Q$
Average fixed cost	*AFC*	Total fixed cost per unit of output	$AFC = TFC \div Q$
Average variable cost	*AVC*	Total variable cost per unit of output	$AVC = TVC \div Q$
Average total cost	*ATC*	Total cost per unit of output	$ATC = AFC + AVC$

Another example is the use of ATMs by banks to dispense cash. ATMs, which are fixed capital, have replaced tellers, which are variable labor. Such a technological change decreases total cost but increases fixed costs and decreases variable cost. This change in the mix of fixed cost and variable cost means that at small outputs, average total cost might increase, while at large outputs, average total cost decreases.

Prices of Factors of Production An increase in the price of a factor of production increases the firm's costs and shifts its cost curves. How the curves shift depends on which factor price changes.

An increase in rent or some other component of *fixed* cost shifts the *TFC* and *AFC* curves upward and shifts the *TC* curve upward but leaves the *AVC* and *TVC* curves and the *MC* curve unchanged. For example, if the interest expense paid by a trucking company increases, the fixed cost of transportation services increases.

An increase in wages, gasoline, or another component of *variable* cost shifts the *TVC* and *AVC* curves upward and shifts the *MC* curve upward but leaves the *AFC* and *TFC* curves unchanged. For example, if truck drivers' wages or the price of gasoline increases, the variable cost and marginal cost of transportation services increase.

You've now completed your study of short-run costs. All the concepts that you've met are summarized in a compact glossary in Table 11.2.

REVIEW QUIZ

1. What relationships do a firm's short-run cost curves show?
2. How does marginal cost change as output increases (a) initially and (b) eventually?
3. What does the law of diminishing returns imply for the shape of the marginal cost curve?
4. What is the shape of the *AFC* curve and why does it have this shape?
5. What are the shapes of the *AVC* curve and the *ATC* curve and why do they have these shapes?

You can work these questions in Study Plan 11.3 and get instant feedback. MyEconLab

Long-Run Cost

We are now going to study the firm's long-run costs. In the long run, a firm can vary both the quantity of labor and the quantity of capital, so in the long run, all the firm's costs are variable.

The behavior of long-run cost depends on the firm's *production function*, which is the relationship between the maximum output attainable and the quantities of both labor and capital.

The Production Function

Table 11.3 shows Campus Sweaters' production function. The table lists total product schedules for four different quantities of capital. The quantity of capital identifies the plant size. The numbers for plant 1 are for a factory with 1 knitting machine—the case we've just studied. The other three plants have 2, 3, and 4 machines. If Campus Sweaters uses plant 2 with 2 knitting machines, the various amounts of labor can produce the outputs shown in the second column of the table. The other two columns show the outputs of yet larger quantities of capital. Each column of the table could be graphed as a total product curve for each plant.

TABLE 11.3 The Production Function

Labor (workers per day)	Output (sweaters per day) Plant 1	Plant 2	Plant 3	Plant 4
1	4	10	13	15
2	10	15	18	20
3	13	18	22	24
4	15	20	24	26
5	16	21	25	27
Knitting machines (number)	1	2	3	4

The table shows the total product data for four quantities of capital (plant sizes). The greater the plant size, the larger is the output produced by any given quantity of labor. For a given plant size, the marginal product of labor diminishes as more labor is employed. For a given quantity of labor, the marginal product of capital diminishes as the quantity of capital used increases.

Diminishing Returns Diminishing returns occur with each of the four plant sizes as the quantity of labor increases. You can check that fact by calculating the marginal product of labor in each of the plants with 2, 3, and 4 machines. With each plant size, as the firm increases the quantity of labor employed, the marginal product of labor (eventually) diminishes.

Diminishing Marginal Product of Capital Diminishing returns also occur with each quantity of labor as the quantity of capital increases. You can check that fact by calculating the marginal product of capital at a given quantity of labor. The *marginal product of capital* is the change in total product divided by the change in capital when the quantity of labor is constant—equivalently, the change in output resulting from a one-unit increase in the quantity of capital. For example, if Campus Sweaters has 3 workers and increases its capital from 1 machine to 2 machines, output increases from 13 to 18 sweaters a day. The marginal product of the second machine is 5 sweaters a day. If Campus Sweaters continues to employ 3 workers and increases the number of machines from 2 to 3, output increases from 18 to 22 sweaters a day. The marginal product of the third machine is 4 sweaters a day, down from 5 sweaters a day for the second machine.

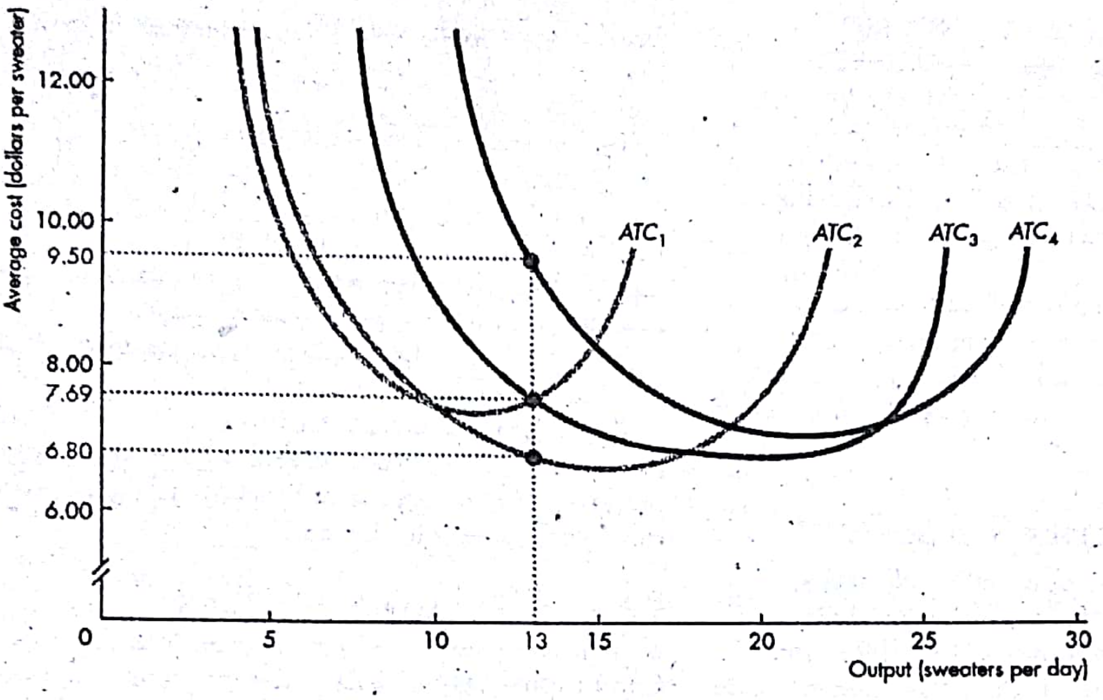
Let's now see what the production function implies for long-run costs.

Short-Run Cost and Long-Run Cost

As before, Campus Sweaters can hire workers for $25 a day and rent knitting machines for $25 a day. Using these factor prices and the data in Table 11.3, we can calculate the average total cost and graph the *ATC* curves for factories with 1, 2, 3, and 4 knitting machines. We've already studied the costs of a factory with 1 machine in Figs. 11.4 and 11.5. In Fig. 11.8, the average total cost curve for that case is ATC_1. Figure 11.8 also shows the average total cost curve for a factory with 2 machines, ATC_2, with 3 machines, ATC_3, and with 4 machines, ATC_4.

You can see, in Fig. 11.8, that the plant size has a big effect on the firm's average total cost.

FIGURE 11.8 Short-Run Costs of Four Different Plants



The figure shows short-run average total cost curves for four different quantities of capital at Campus Sweaters. The firm can produce 13 sweaters a day with 1 knitting machine on ATC_1 or with 3 knitting machines on ATC_3 for an average cost of \$7.69 a sweater. The firm can produce 13 sweaters a day by using 2 machines on ATC_2 for \$6.80 a sweater or by using 4 machines on ATC_4 for \$9.50 a sweater.

If the firm produces 13 sweaters a day, the least-cost method of production, *the long-run method*, is with 2 machines on ATC_2.

MyEconLab animation

In Fig. 11.8, two things stand out:

1. Each short-run *ATC* curve is U-shaped.
2. For each short-run *ATC* curve, the larger the plant, the greater is the output at which average total cost is at a minimum.

Each short-run *ATC* curve is U-shaped because, as the quantity of labor increases, its marginal product initially increases and then diminishes. This pattern in the marginal product of labor, which we examined in some detail for the plant with 1 knitting machine on pp. 254–255, occurs at all plant sizes.

The minimum average total cost for a larger plant occurs at a greater output than it does for a smaller plant because the larger plant has a higher total fixed cost and therefore, for any given output, a higher average fixed cost.

Which short-run *ATC* curve a firm operates on depends on the plant it has. In the long run, the firm can choose its plant and the plant it chooses is the one that enables it to produce its planned output at the lowest average total cost.

To see why, suppose that Campus Sweaters plans to produce 13 sweaters a day. In Fig. 11.8, with 1 machine, the average total cost curve is ATC_1 and the average total cost of 13 sweaters a day is \$7.69 a sweater. With 2 machines, on ATC_2, average total cost is \$6.80 a sweater. With 3 machines, on ATC_3, average total cost is \$7.69 a sweater, the same as with 1 machine. Finally, with 4 machines, on ATC_4, average total cost is \$9.50 a sweater.

The economically efficient plant for producing a given output is the one that has the lowest average total cost. For Campus Sweaters, the economically efficient plant to use to produce 13 sweaters a day is the one with 2 machines.

In the long run, Cindy chooses the plant that minimizes average total cost. When a firm is producing a given output at the least possible cost, it is operating on its *long-run average cost curve*.

The **long-run average cost curve** is the relationship between the lowest attainable average total cost and output when the firm can change both the plant it uses and the quantity of labor it employs.

The long-run average cost curve is a planning curve. It tells the firm the plant and the quantity of labor to use at each output to minimize average cost. Once the firm chooses a plant, the firm operates on the short-run cost curves that apply to that plant.

The Long-Run Average Cost Curve

Figure 11.9 shows how a long-run average cost curve is derived. The long-run average cost curve *LRAC* consists of pieces of the four short-run *ATC* curves. For outputs up to 10 sweaters a day, average-total cost is the lowest on ATC_1. For outputs between 10 and 18 sweaters a day, average total cost is the lowest on ATC_2. For outputs between 18 and 24 sweaters a day, average total cost is the lowest on ATC_3. And for outputs in excess of 24 sweaters a day, average total cost is the lowest on ATC_4. The piece of each *ATC* curve with the lowest average total cost is highlighted in dark blue in Fig. 11.9. This dark blue scallop-shaped curve made up of the pieces of the four *ATC* curves is the *LRAC* curve.

Economies and Diseconomies of Scale

Economies of scale are features of a firm's technology that make average total cost *fall* as output increases. When economies of scale are present, the *LRAC* curve slopes downward. In Fig. 11.9, Campus Sweaters has economies of scale for outputs up to 15 sweaters a day.

Greater specialization of both labor and capital is the main source of economies of scale. For example, if GM produces 100 cars a week, each worker must perform many different tasks and the capital must be general-purpose machines and tools. But if GM produces 10,000 cars a week, each worker specializes in a small number of tasks, uses task-specific tools, and becomes highly proficient.

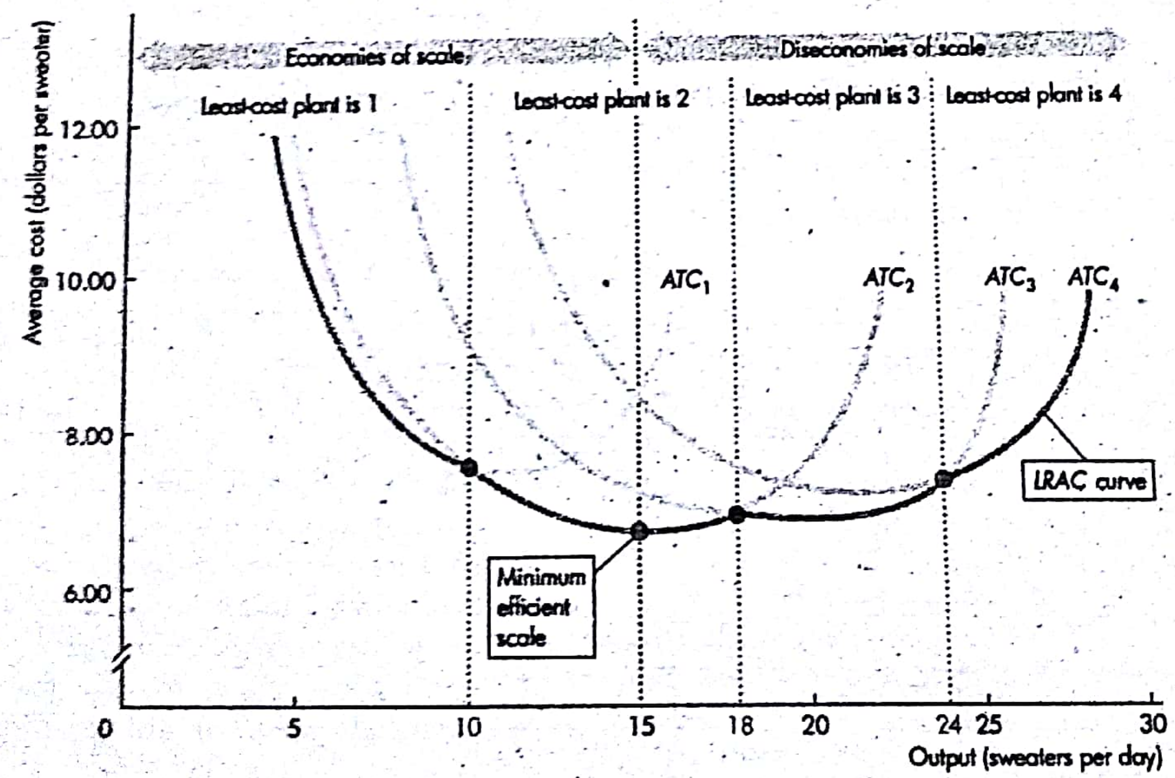
Diseconomies of scale are features of a firm's technology that make average total cost *rise* as output increases. When diseconomies of scale are present, the *LRAC* curve slopes upward. In Fig. 11.9, Campus Sweaters experiences diseconomies of scale at outputs greater than 15 sweaters a day.

The challenge of managing a large enterprise is the main source of diseconomies of scale.

Constant returns to scale are features of a firm's technology that keep average total cost constant as output increases. When constant returns to scale are present, the *LRAC* curve is horizontal.

Economies of Scale at Campus Sweaters The economies of scale and diseconomies of scale at Campus Sweaters arise from the firm's production function in Table 11.3. With 1 machine and 1 worker, the firm produces 4 sweaters a day. With 2 machines and 2 workers, total cost doubles but out-

FIGURE 11.9 Long-Run Average Cost Curve



The long-run average cost curve traces the lowest attainable *ATC* when both labor and capital change. The green arrows highlight the output range over which each plant achieves the lowest *ATC*. Within each range, to change the quantity produced, the firm changes the quantity of labor it employs.

Along the *LRAC* curve, economies of scale occur if average cost falls as output increases; diseconomies of scale occur if average cost rises as output increases. Minimum efficient scale is the output at which average cost is lowest, 15 sweaters a day.

MyEconLab animatio

Economics in Action

Produce More to Cut Cost

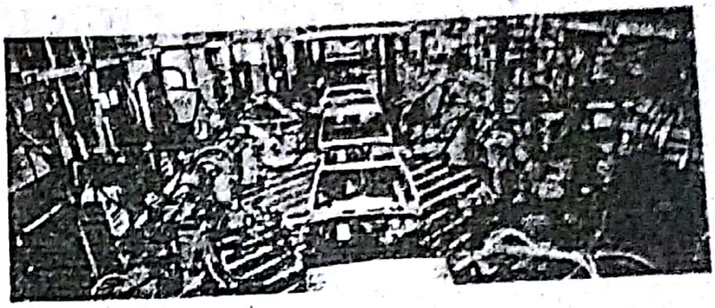
Why do GM, Ford, and the other automakers have expensive equipment lying around that isn't fully used? You can answer this question with what you've learned in this chapter.

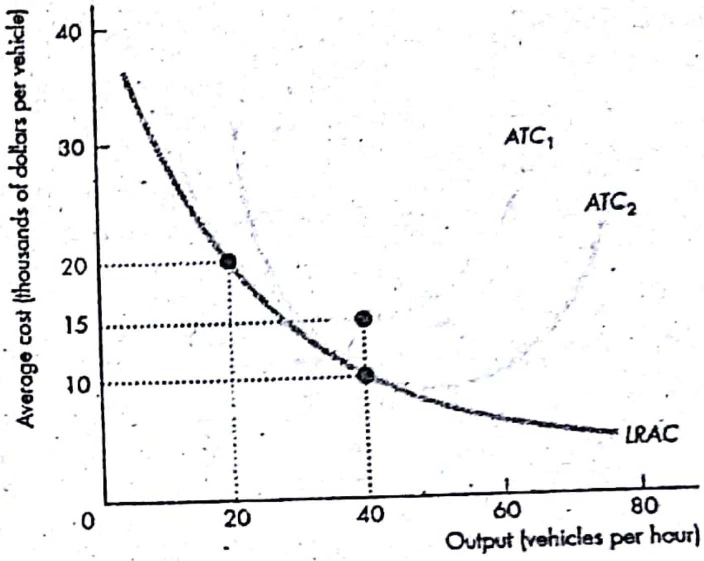
The basic answer is that auto production enjoys economies of scale. A larger output rate brings a lower long-run average cost—the firm's *LRAC* curve slopes downward.

An auto producer's average total cost curves look like those in the figure. To produce 20 vehicles an hour, the firm installs the plant with the short-run average total cost curve ATC_1. The average cost of producing a vehicle is $20,000.

Producing 20 vehicles an hour doesn't use the plant at its lowest possible average total cost. If the firm could sell enough cars for it to produce 40 vehicles an hour, the firm could use its current plant and produce at an average cost of $15,000 a vehicle.

But if the firm planned to produce 40 vehicles an hour, it would not stick with its current plant. The firm would install a bigger plant with the short-run average total cost curve ATC_2, and produce 40 vehicles an hour for $10,000 a car.





Automobile Plant Average Cost Curves

put more than doubles to 15 sweaters a day, so average cost decreases and Campus Sweaters experiences economies of scale. With 4 machines and 4 workers, total cost doubles again but output less than doubles to 26 sweaters a day, so average cost increases and the firm experiences diseconomies of scale.

Minimum Efficient Scale A firm's **minimum efficient scale** is the *smallest* output at which long-run average cost reaches its lowest level. At Campus Sweaters, the minimum efficient scale is 15 sweaters a day.

The minimum efficient scale plays a role in determining market structure. In a market in which the minimum efficient scale is small relative to market demand, the market has room for many firms, and the market is competitive. In a market in which the minimum efficient scale is large relative to market demand, only a small number of firms, and possibly only one firm, can make a profit and the market is either an oligopoly or monopoly. We will return to this idea in the next three chapters.

REVIEW QUIZ

1. What does a firm's production function show and how is it related to a total product curve?
2. Does the law of diminishing returns apply to capital as well as labor? Explain why or why not.
3. What does a firm's *LRAC* curve show? How is it related to the firm's short-run *ATC* curves?
4. What are economies of scale and diseconomies of scale? How do they arise? What do they impl for the shape of the *LRAC* curve?
5. What is a firm's minimum efficient scale?

You can work these questions in Study Plan 11.4 and get instant feedback. MyEconL

Reading Between the Lines on pp. 264–265 app what you've learned about a firm's cost curves. It lo at McDonald's cost curves and explains how increa plant size can lower average total cost.

Nazrul

Monopoly and How It Arises

A **monopoly** is a market with a single firm that produces a good or service with no close substitutes and that is protected by a barrier that prevents other firms from entering that market.

How Monopoly Arises

Monopoly arises for two key reasons:

- No close substitutes
- Barrier to entry

No Close Substitutes If a good has a close substitute, even though only one firm produces it, that firm effectively faces competition from the producers of the substitute. A monopoly sells a good or service that has no good substitutes. Tap water and bottled water are close substitutes for drinking, but tap water has no effective substitutes for showering or washing a car and a local public utility that supplies tap water is a monopoly.

Barrier to Entry A constraint that protects a firm from potential competitors is called a **barrier to entry.** The three types of barrier to entry are

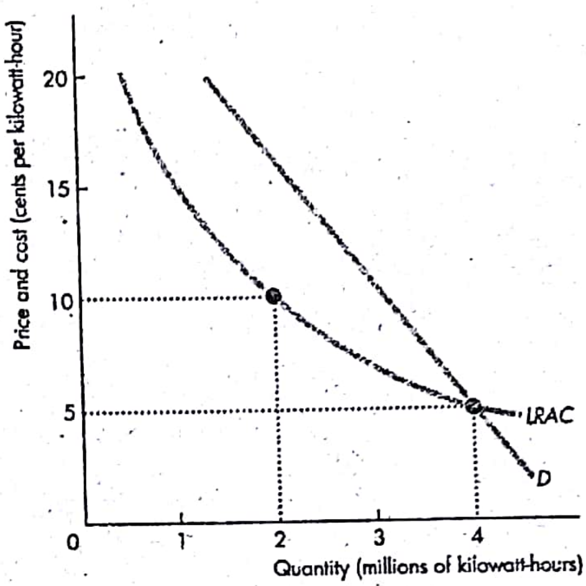
- Natural
- Ownership
- Legal

Natural Barrier to Entry A natural barrier to entry creates a **natural monopoly:** a market in which economies of scale enable one firm to supply the entire market at the lowest possible cost. The firms that deliver gas, water, and electricity to our homes are examples of natural monopoly.

Figure 13.1 illustrates a natural monopoly. The market demand curve for electric power is *D*, and the long-run average cost curve is *LRAC*. Economies of scale prevail over the entire length of the *LRAC* curve. At a price of 5 cents per kilowatt-hour, the quantity demanded is 4 million kilowatt-hours and one firm can produce that quantity at a cost of 5 cents per kilowatt-hour. If two firms shared the market equally, it would cost each of them 10 cents per kilowatt-hour to produce a total of 4 million kilowatt-hours.

Ownership Barrier to Entry An ownership barrier to entry occurs if one firm owns a significant portion of a key resource. An example of this type of monopoly occurred during the last century when De Beers controlled up to 90 percent of the world's supply of diamonds. (Today, its share is only 65 percent.)

FIGURE 13.1 Natural Monopoly



The market demand curve for electric power is *D*, and the long-run average cost curve is *LRAC*. Economies of scale exist over the entire *LRAC* curve. One firm can distribute 4 million kilowatt-hours at a cost of 5 cents a kilowatt-hour. This same total output costs 10 cents a kilowatt-hour with two firms. One firm can meet the market demand at a lower cost than two or more firms can. The market is a natural monopoly.

MyEconLab animation

Legal Barrier to Entry A legal barrier to entry creates a **legal monopoly:** a market in which competition and entry are restricted by the granting of a public franchise, government license, patent, or copyright.

A *public franchise* is an exclusive right granted to a firm to supply a good or service. An example is the U.S. Postal Service, which has the exclusive right to carry first-class mail. A *government license* controls entry into particular occupations, professions, and industries. Examples of this type of barrier to entry occur in medicine, law, dentistry, schoolteaching, architecture, and many other professional services. Licensing does not always create a monopoly, but it does restrict competition.

A *patent* is an exclusive right granted to the inventor of a product or service. A *copyright* is an exclusive right granted to the author or composer of a literary, musical, dramatic, or artistic work. Patents and copyrights are valid for a limited time period that varies from country to country. In the United States, a patent is valid for 20 years. Patents encourage the *invention* of new products and production methods. They also stimulate *innovation*—the use of new inventions—by encouraging inventors to publicize their discoveries and offer them for use under license. Patents have stimulated innovations in areas as diverse as soybean seeds, pharmaceuticals, memory chips, and video games.

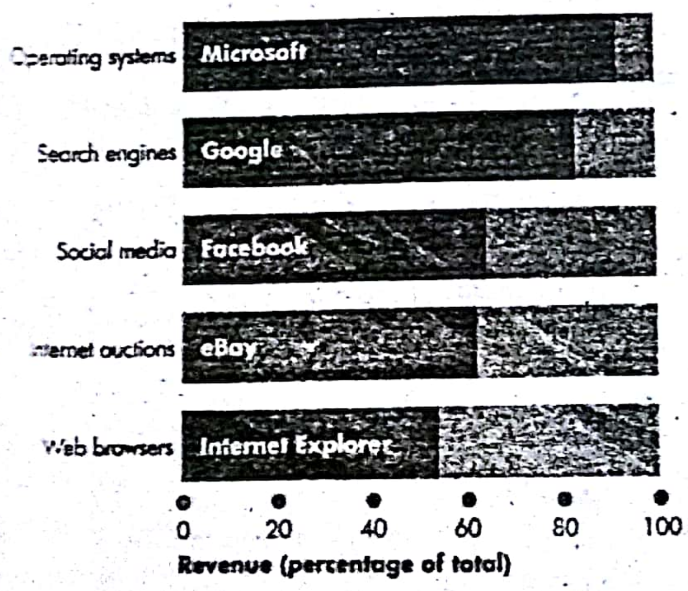
Economics in Action

Information-Age Monopolies

Information-age technologies have created four big natural monopolies—firms with large plant costs but almost zero marginal cost, so they experience economies of scale.

These firms are: Microsoft, with 92 percent of personal computers using a version of Windows and 54 percent using Internet Explorer for Web browsing; Google, which performs 83 percent of Internet searches; Facebook, with 64 percent of the social media market; and eBay, with 62 percent of the Internet auction market.

These same information-age technologies have also destroyed monopolies. FedEx, UPS, the fax machine, and e-mail have weakened the monopoly of the U.S. Postal Service; and the satellite dish has weakened cable television monopolies.



Market Shares

Monopoly Price-Setting Strategies

A major difference between monopoly and competition is that a monopoly sets its own price. In doing so, the monopoly faces a market constraint: To sell a larger quantity, the monopoly must set a lower price. There are two monopoly situations that create two pricing strategies:

- Single price
- Price discrimination

Single Price A **single-price monopoly** is a firm that must sell each unit of its output for the same price to all its customers. De Beers sells diamonds (of a given size and quality) for the same price to all its customers. If it tried to sell at a low price to some customers and at a higher price to others, only the low-price customers would buy from De Beers. Others would buy from De Beers' low-price customers. De Beers is a *single-price* monopoly.

Price Discrimination When a firm practices **price discrimination**, it sells different units of a good or service for different prices. Many firms price discriminate. Microsoft sells its Windows and Office software at different prices to different buyers. Computer manufacturers who install the software on new machines, students and teachers, governments, and businesses all pay different prices. Pizza producers offer a second pizza for a lower price than the first one. These are examples of *price discrimination*.

When a firm price discriminates, it looks as though it is doing its customers a favor. In fact, it is charging the highest possible price for each unit sold and making the largest possible profit.

REVIEW QUIZ

1 How does monopoly arise?
2 How does a natural monopoly differ from a legal monopoly?
3 Distinguish between a price-discriminating monopoly and a single-price monopoly.

You can work these questions in Study Plan 13.1 and get instant feedback. MyEconLab

We start with a single-price monopoly and see how it makes its decisions about the quantity to produce and the price to charge to maximize its profit.

Monopoly Regulation

Natural monopoly presents a dilemma. With economies of scale, it produces at the lowest possible cost. But with market power, it has an incentive to raise the price above the competitive price and produce too little—to operate in the self-interest of the monopolist and not in the social interest.

Regulation—rules administered by a government agency to influence prices, quantities, entry, and other aspects of economic activity in a firm or industry—is a possible solution to this dilemma.

To implement regulation, the government establishes agencies to oversee and enforce the rules. For example, the Surface Transportation Board regulates prices on interstate railroads, some trucking and bus lines, and water and oil pipelines. By the 1970s, almost a quarter of the nation's output was produced by regulated industries (far more than just natural monopolies) and a process of deregulation began.

Deregulation is the process of removing regulation of prices, quantities, entry, and other aspects of economic activity in a firm or industry. During the past 30 years, deregulation has occurred in domestic air transportation, telephone service, interstate trucking, and banking and financial services. Cable TV was deregulated in 1984, re-regulated in 1992, and deregulated again in 1996.

Regulation is a possible solution to the dilemma presented by natural monopoly but not a guaranteed solution. There are two theories about how regulation actually works: the *social interest theory* and the *capture theory*.

The **social interest theory** is that the political and regulatory process relentlessly seeks out inefficiency and introduces regulation that eliminates deadweight loss and allocates resources efficiently.

The **capture theory** is that regulation serves the self-interest of the producer, who captures the regulator and maximizes economic profit. Regulation that benefits the producer but creates a deadweight loss gets adopted because the producer's gain is large and visible while each individual consumer's loss is small and invisible. No individual consumer has an incentive to oppose the regulation but the producer has a big incentive to lobby for it.

We're going to examine efficient regulation that serves the social interest and see why it is not a simple matter to design and implement such regulation.

Efficient Regulation of a Natural Monopoly

A cable TV company is a *natural monopoly*—it can supply the entire market at a lower price than two or more competing firms can. Cox Communications, based in Atlanta, provides cable TV to households in 20 states. The firm has invested heavily in satellite receiving dishes, cables, and control equipment and so has large fixed costs. These fixed costs are part of the firm's average total cost. Its average total cost decreases as the number of households served increases because the fixed cost is spread over a larger number of households.

Unregulated, Cox produces the quantity that maximizes profit. Like all single-price monopolies, the profit-maximizing quantity is less than the efficient quantity, and underproduction results in a deadweight loss.

How can Cox be regulated to produce the efficient quantity of cable TV service? The answer is by being regulated to set its price equal to marginal cost, known as the **marginal cost pricing rule**. The quantity demanded at a price equal to marginal cost is the efficient quantity—the quantity at which marginal benefit equals marginal cost.

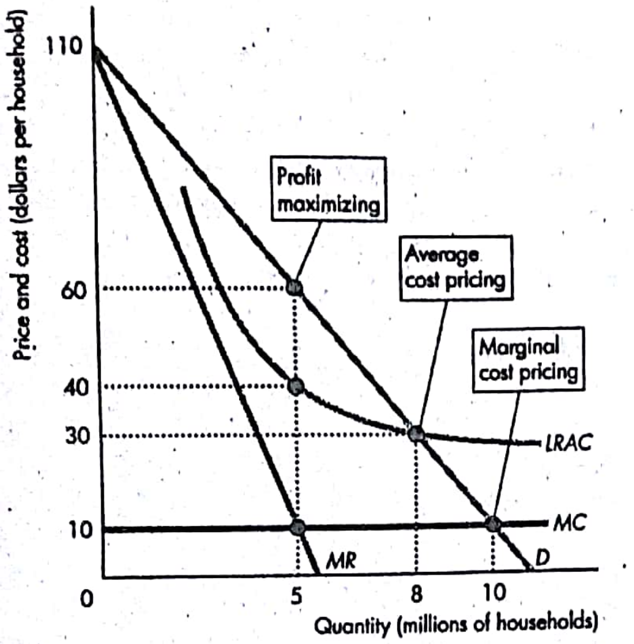
Figure 13.11 illustrates the marginal cost pricing rule. The demand curve for cable TV is D. Cox's marginal cost curve is MC. That marginal cost curve is (assumed to be) horizontal at \$10 per household per month—that is, the cost of providing each additional household with a month of cable programming is \$10. The efficient outcome occurs if the price is regulated at \$10 per household per month with 10 million households served.

But there is a problem: At the efficient output, average total cost exceeds marginal cost, so a firm that uses marginal cost pricing incurs an economic loss. A cable TV company that is required to use a marginal cost pricing rule will not stay in business for long. How can the firm cover its costs and, at the same time, obey a marginal cost pricing rule?

There are two possible ways of enabling the firm to cover its costs: price discrimination and a two-part price (called a *two-part tariff*).

For example, Verizon offers plans at a fixed monthly price that give access to the cell-phone network and unlimited free calls. The price of a call (zero) equals Verizon's marginal cost of a call. Similarly, a cable TV operator can charge a one-time connection fee that covers its fixed cost and then charge a monthly fee equal to marginal cost.

FIGURE 13.11 Regulating a Natural Monopoly



A natural monopoly cable TV supplier faces the demand curve *D*. The firm's marginal cost is constant at $10 per household per month, as shown by the curve labeled *MC*. The long-run average cost curve is *LRAC*.

Unregulated, as a profit-maximizer, the firm serves 5 million households at a price of $60 a month. An efficient marginal cost pricing rule sets the price at $10 a month. The monopoly serves 10 million households and incurs an economic loss. A second-best average cost pricing rule sets the price at $30 a month. The monopoly serves 8 million households and earns zero economic profit.

MyEconLab animation

Second-Best Regulation of a Natural Monopoly

A natural monopoly cannot always be regulated to achieve an efficient outcome. Two possible ways of enabling a regulated monopoly to avoid an economic loss are

- Average cost pricing
- Government subsidy

Average Cost Pricing The **average cost pricing rule** sets price equal to average total cost. With this rule the firm produces the quantity at which the average total cost curve cuts the demand curve. This rule results in the firm making zero economic profit—breaking even. But because for a natural monopoly average total cost exceeds marginal cost, the quantity produced is less than the efficient quantity and a deadweight loss arises.

Figure 13.11 illustrates the average cost pricing rule. The price is $30 a month and 8 million households get cable TV.

Government Subsidy A government subsidy is a direct payment to the firm equal to its economic loss. To pay a subsidy, the government must raise the revenue by taxing some other activity. You saw in Chapter 6 that taxes themselves generate deadweight loss.

And the Second-Best Is ... Which is the better option, average cost pricing or marginal cost pricing with a government subsidy? The answer depends on the relative magnitudes of the two deadweight losses. Average cost pricing generates a deadweight loss in the market served by the natural monopoly. A subsidy generates deadweight losses in the markets for the items that are taxed to pay for the subsidy. The smaller deadweight loss is the second-best solution to regulating a natural monopoly. Making this calculation in practice is too difficult and average cost pricing is generally preferred to a subsidy.

Implementing average cost pricing presents the regulator with a challenge because it is not possible to be sure what a firm's costs are. So regulators use one of two practical rules:

- Rate of return regulation
- Price cap regulation

Rate of Return Regulation Under **rate of return regulation**, a firm must justify its price by showing that its return on capital doesn't exceed a specified target rate. This type of regulation can end up serving the self-interest of the firm rather than the social interest. The firm's managers have an incentive to inflate costs by spending on items such as private jets, free baseball tickets (disguised as public relations expenses), and lavish entertainment. Managers also have an incentive to use more capital than the efficient amount. The rate of return on capital is regulated but not the total return on capital, and the greater the amount of capital, the greater is the total return.

What Is Monopolistic Competition?

You have studied perfect competition, in which a large number of firms produce at the lowest possible cost, make zero economic profit, and are efficient. You've also studied monopoly, in which a single firm restricts output, produces at a higher cost and price than in perfect competition, and is inefficient.

Most real-world markets are competitive but not perfectly competitive, because firms in these markets have some power to set their prices, as monopolies do. We call this type of market *monopolistic competition*.

Monopolistic competition is a market structure in which

- A large number of firms compete.
- Each firm produces a differentiated product.
- Firms compete on product quality, price, and marketing.
- Firms are free to enter and exit the industry.

Large Number of Firms

In monopolistic competition, as in perfect competition, the industry consists of a large number of firms. The presence of a large number of firms has three implications for the firms in the industry.

Small Market Share In monopolistic competition, each firm supplies a small part of the total industry output. Consequently, each firm has only limited power to influence the price of its product. Each firm's price can deviate from the average price of other firms by only a relatively small amount.

Ignore Other Firms A firm in monopolistic competition must be sensitive to the average market price of the product, but the firm does not pay attention to any one individual competitor. Because all the firms are relatively small, no one firm can dictate market conditions, and so no one firm's actions directly affect the actions of the other firms.

Collusion Impossible Firms in monopolistic competition would like to be able to conspire to fix a higher price—called *collusion*. But because the number of firms in monopolistic competition is large, coordination is difficult and collusion is not possible.

Product Differentiation

A firm practices **product differentiation** if it makes a product that is slightly different from the products of competing firms. A differentiated product is one that is a close substitute but not a perfect substitute for the products of the other firms. Some people are willing to pay more for one variety of the product, so when its price rises, the quantity demanded of that variety decreases, but it does not (necessarily) decrease to zero. For example, Adidas, Asics, Diadora, Etonic, Fila, New Balance, Nike, Puma, and Reebok all make differentiated running shoes. If the price of Adidas running shoes rises and the prices of the other shoes remain constant, Adidas sells fewer shoes and the other producers sell more. But Adidas shoes don't disappear unless the price rises by a large enough amount.

Competing on Quality, Price, and Marketing

Product differentiation enables a firm to compete with other firms in three areas: product quality, price, and marketing.

Quality The quality of a product is the physical attributes that make it different from the products of other firms. Quality includes design, reliability, the service provided to the buyer, and the buyer's ease of access to the product. Quality lies on a spectrum that runs from high to low. Some firms—such as Dell Computer Corp.—offer high-quality products. They are well designed and reliable, and the customer receives quick and efficient service. Other firms offer a lower-quality product that is poorly designed, that might not work perfectly, and that is not supported by effective customer service.

Price Because of product differentiation, a firm in monopolistic competition faces a downward-sloping demand curve. So, like a monopoly, the firm can set both its price and its output. But there is a tradeoff between the product's quality and price. A firm that makes a high-quality product can charge a higher price than a firm that makes a low-quality product.

Marketing Because of product differentiation, a firm in monopolistic competition must market its product. Marketing takes two main forms: advertising and packaging. A firm that produces a high-quality demand (almost infinite) is a soft drink from two campus machines located side by side. If the two machines offer the same soft drinks for the same price, some people buy from one machine and some from the other. But if one machine's price is higher than the other's, by even a small amount, no one buys from the machine with the higher price. Drinks from the two machines are perfect substitutes. The demand for a good that has a perfect substitute is perfectly elastic.

Between unit elastic demand and perfectly elastic demand is another general case in which *the percentage change in the quantity demanded exceeds the percentage change in price.* In this case, the price elasticity of demand is greater than 1 and the good is said to have an **elastic demand.** Automobiles and furniture are examples of goods that have elastic demand.

Figure 4.2 shows three demand curves that cover the entire range of possible elasticities of demand that you've just reviewed. In Fig. 4.2(a), the quantity demanded is constant regardless of the price, so this demand is perfectly inelastic. In Fig. 4.2(b), the percentage change in the quantity demanded equals the percentage change in price, so this demand is unit elastic. In Fig. 4.2(c), the price is constant regardless of the quantity demanded, so this figure illustrates a perfectly elastic demand.

You now know the distinction between elastic and inelastic demand. But what determines whether the demand for a good is elastic or inelastic?

The Factors that Influence the Elasticity of Demand

The elasticity of demand for a good depends on

- The closeness of substitutes
- The proportion of income spent on the good
- The time elapsed since the price change

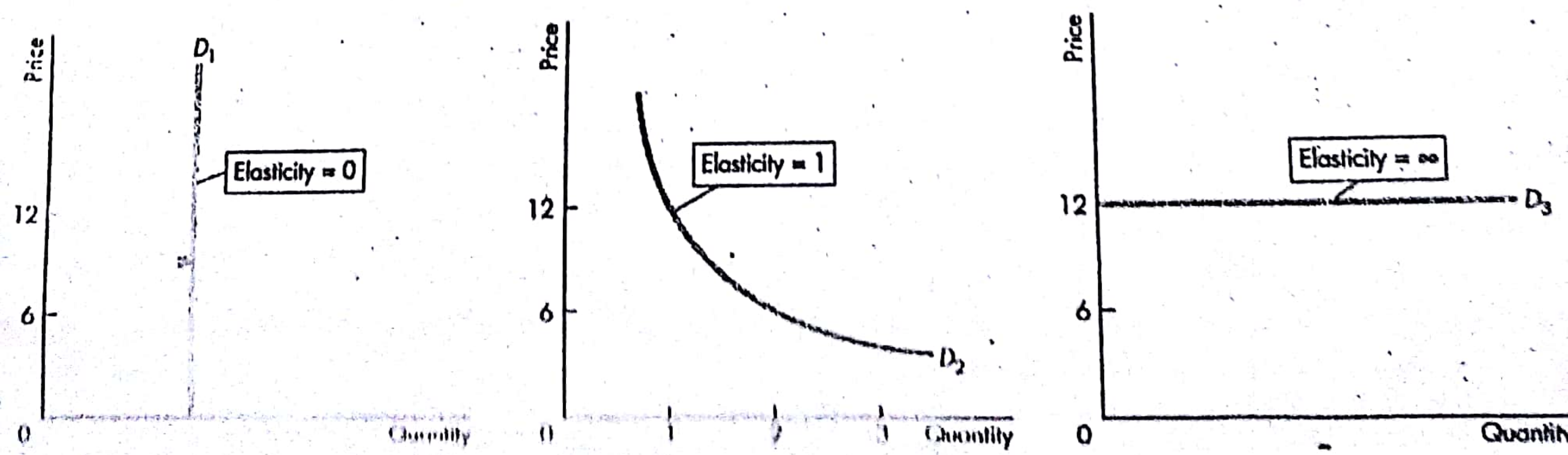
Closeness of Substitutes The closer the substitutes for a good, the more elastic is the demand for it. Oil as fuel or raw material for chemicals has no close substitutes so the demand for oil is inelastic. Plastics are close substitutes for metals, so the demand for metals is elastic.

The degree of substitutability depends on how narrowly (or broadly) we define a good. For example, a personal computer has no close substitutes, but a Dell PC is a close substitute for a Hewlett-Packard PC. So the elasticity of demand for personal computers is lower than the elasticity of demand for a Dell or a Hewlett-Packard.

In everyday language we call goods such as food and shelter *necessities* and goods such as exotic vacations *luxuries*. A necessity has poor substitutes, so it generally has an inelastic demand. A luxury usually has many substitutes, one of which is not buying it. So a luxury generally has an elastic demand.

Proportion of Income Spent on the Good Other things remaining the same, the greater the proportion of income spent on a good, the more elastic (or less inelastic) is the demand for it.

FIGURE 4.2 Inelastic and Elastic Demand



Think about your own elasticity of demand for chewing gum and housing. If the price of gum rises, you consume almost as much as before. Your demand for gum is inelastic. If apartment rents rise, you look for someone to share with. Your demand for housing is not as inelastic as your demand for gum. Why the difference? Housing takes a big chunk of your budget, and gum takes little. You barely notice the higher price of gum, while the higher rent puts your budget under severe strain.

Time Elapsed Since Price Change The longer the time that has elapsed since a price change, the more elastic is demand. When the price of oil increased by 400 percent during the 1970s, people barely changed the quantity of oil and gasoline they bought. But gradually, as more efficient auto and airplane engines were developed, the quantity bought decreased. The demand for oil became more elastic as more time elapsed following the huge price hike.

Elasticity Along a Linear Demand Curve

Elasticity of demand is not the same as slope. And a good way to see this fact is by studying a demand curve that has a constant slope but a varying elasticity.

The demand curve in Fig. 4.3 is linear, which means that it has a constant slope. Along this demand curve, a $5 rise in the price brings a decrease of 10 pizzas an hour.

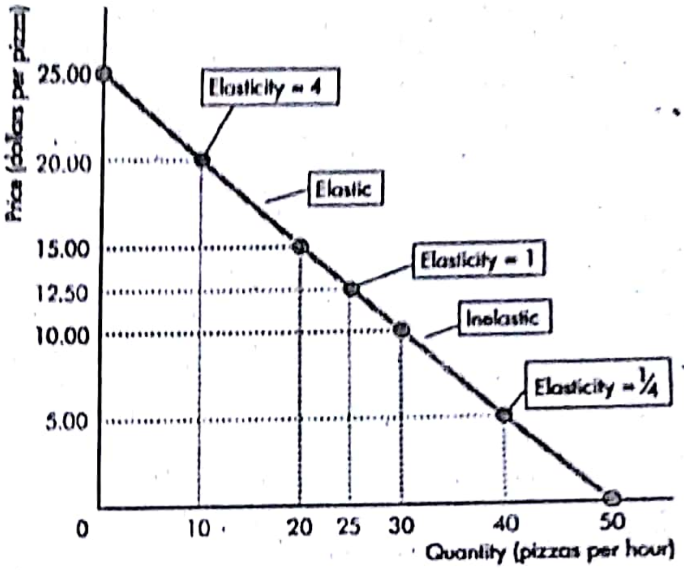
But the price elasticity of demand is not constant along this demand curve. To see why, let's calculate some elasticities.

At the midpoint of the demand curve, the price is $12.50 and the quantity is 25 pizzas per hour. If the price rises from $10 to $15 a pizza the quantity demanded decreases from 30 to 20 pizzas an hour and the average price and average quantity are at the midpoint of the demand curve. So

$$\text{Price elasticity of demand} = \frac{10/25}{5/12.50} = 1.$$

That is, at the midpoint of a linear demand curve

FIGURE 4.3 Elasticity Along a Linear Demand Curve



On a linear demand curve, demand is unit elastic at the midpoint (elasticity is 1), elastic above the midpoint, and inelastic below the midpoint.

MyEconLab animation

the average quantity is 10 pizzas. Putting these numbers into the elasticity formula,

$$\text{Price elasticity of demand} = \frac{\Delta Q/Q_{ave}}{\Delta P/P_{ave}} = \frac{20/10}{10/20} = 4.$$

That is, the price elasticity of demand at an average price of $20 a pizza is 4.

At prices *below* the midpoint, the price elasticity of demand is less than 1: Demand is inelastic. For example, if the price rises from zero to $10 a pizza, the quantity demanded decreases from 50 to 30 pizzas an hour. The average price is now $5 and the average quantity is 40 pizzas an hour. So