PRODUCER THEORY

DR. SANJA SAMIRANA PATTNAYAK

IIM SIRMAUR

Managerial Decisions

- Firm's Objective is to maximize profits or minimize cost
  - Total Profit = Total Revenue – Total Costs
- Firms' Actions or Decisions
  - What to produce? How to produce?
  - How much to produce? How to price?
- Firms Constraints
  - Technology drives the constraints on cost
  - Market Demand drives the constraints on revenue

Production Decisions of a Firm The Cost Side

1. Production Technology
   - Describe how inputs can be transformed into outputs
     - Inputs: land, labor, capital & raw materials
     - Outputs: cars, televisions, books, etc.
   - Firms can produce a particular level of outputs using different combinations of inputs [(2L, 3K or (3k, 2L)] etc. K refers to capital, L is labor.

Production Decisions of a Firm

2. Cost Constraints
   - Firms must consider *prices* of labor, capital and other inputs.
   - Firms want to minimize total production costs partly determined by input prices.
   - As consumers must consider budget constraints, firms must be concerned *about costs of production.*

Production Decisions of a Firm

3. Input Choices

- Given input prices and production technology, the firm must choose *how much of each input* to use in producing output.

- Given prices of different inputs, the firm may choose different combinations of inputs to minimize costs.
- [least cost combination of inputs]

  - If labor is cheap, may choose to produce with more labor and less capital.

Production Decisions of a Firm

- If a firm is a cost minimizer, we should also learn
  - How total costs of production varies with output
  - How does the firm choose the quantity to maximize its profits?
- We can represent the firm's production technology in the form of a production function.

The Technology of Production

- Production Function:
  - Indicates the highest amount of output (q) that a firm can produce for every specified combination of inputs.
  - For simplicity, we will consider only labor (L) and capital (K)
  - Shows what is technically feasible when the firm operates efficiently.

- **Core Idea:**
- Something is **technically feasible** if it is **possible to produce with the given technology and available inputs**.
- It doesn't consider costs or profits yet.
- It only asks: *"Given the production technology, can these inputs produce at least this much output?"*

- Formally, if the production function is

Formally, if the production function is

$$Q = f(L, K)$$

then an input combination (L, K) is **technically feasible** for output level Q^* if

$$f(L, K) \geq Q^*.$$

- **Example:**
- Suppose the production function is given below:

$$Q = \sqrt{L \cdot K}.$$

- If the firm wants $Q = 10$:
  - Using $L = 25, K = 4$, we get $Q = \sqrt{25 \cdot 4} = 10$. ✅ Technically feasible.
  - Using $L = 9, K = 4$, we get $Q = \sqrt{36} = 6 < 10$. ❌ Not technically feasible.

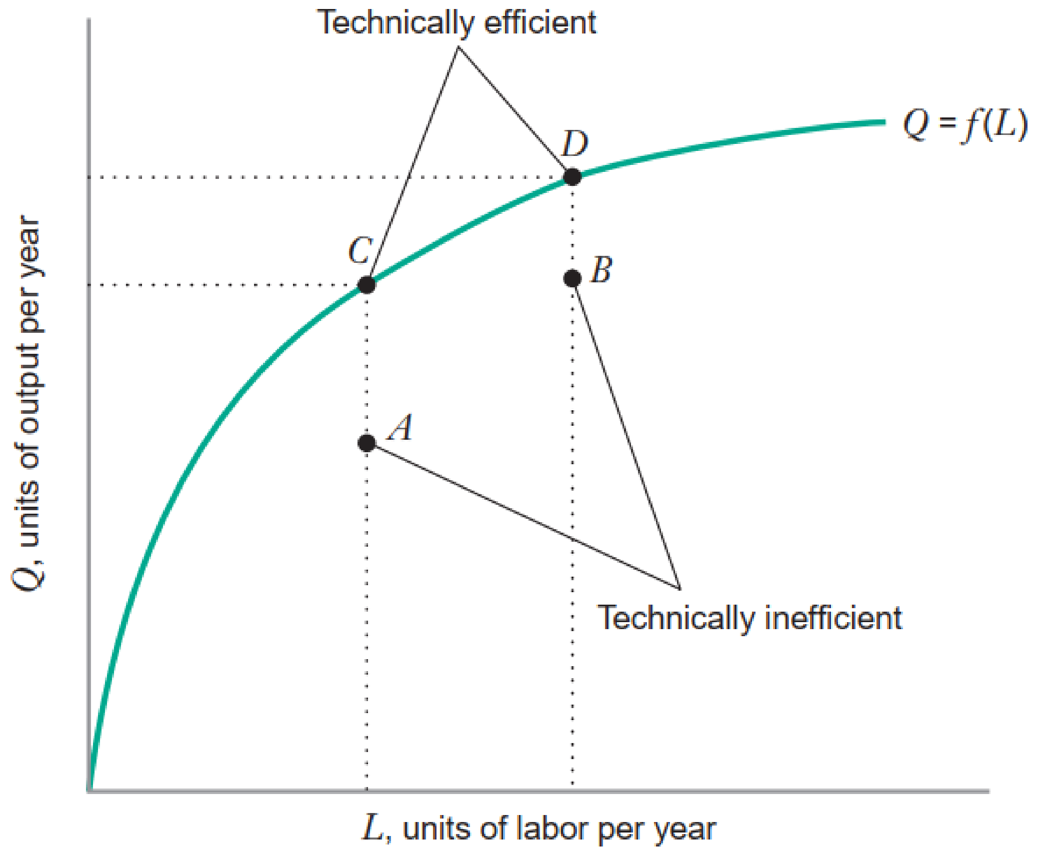
The Technology of Production

- The production function for two inputs:

$$q = F(K,L)$$

  - Output (q) is a function of capital (K) and Labor (L)
  - The production function is true for a given technology
    - If technology improved, more output can be produced for a given level of inputs

- The production function represented in equation (previous slide) is analogous to the utility function in consumer theory.
- Just as the utility function depends on exogenous consumer tastes, the production function depends on exogenous technological conditions.
- Over time, these technological conditions may change, an occurrence known as **technological progress**, and the production function may then shift.
- **Production set:** The set of technically feasible combinations of inputs and outputs.



Technically efficient
$Q = f(L)$
D
C
B
A
Technically inefficient
Q, units of output per year
L, units of labor per year

- The production function in equation tells us the *maximum* output a firm could get from a given combination of labour and capital.
- Of course, inefficient management could reduce output from what is technologically possible.
- Figure (previous slide) depicts this possibility by showing the production function for a single input, labour: $Q = f(L)$.

- Points on or below the production function make up the firm's **production set**, the set of technically *feasible* combinations of inputs and outputs.
- Points such as *A* and *B* in the production set are **technically inefficient** (i.e., at these points the firm gets less output from its labour than it could).
- Points such as *C* and *D*, on the boundary of the production set, are **technically efficient**.
- At these points, the firm produces as much output as it possibly can given the amount of labour it employs.

- If we invert the production function, we get a function $L = g(Q)$, which tells us the *minimum* amount of labour L required to produce a given amount of output Q.

- This function is the **labour requirements function**.

- If, for example, $Q = \sqrt{L}$ is the production function, then $L = Q^2$ is the labour requirements function; thus, to produce an output of 7 units, a firm will need at least $7^2 = 49$ units of labour.

The Technology of Production

- Short Run versus Long Run Decisions
  - It takes time for a firm to adjust production from one set of inputs to another
  - Firms must consider not only what inputs can be varied but over what period of time that can occur
  - We must distinguish between long run and short run

The Technology of Production

- Short Run
  - Period of time in which quantities of one or more production factors cannot be changed.
  - These inputs are called fixed inputs.

- Long-run
  - Amount of time needed to make all production inputs variable.
- Short run and long run are not time specific.

Production: One Variable Input

- We will begin by looking at the short run when only one input can be varied
- We assume capital is fixed and labor is variable
  - Output can only be increased by increasing labor
  - Must know how output changes as the amount of labor is changed (Table-1)

(Table-1)Production: One Variable Input

AMOUNT OF LABOR (L)	AMOUNT OF CAPITAL (K)	TOTAL OUTPUT (q)	AVERAGE PRODUCT (q/L)	MARGINAL PRODUCT ($\Delta q/\Delta L$)
0	10	0	—	—
1	10	15	15	15
2	10	40	20	25
3	10	69	23	29
4	10	96	24	27
5	10	120	24	24
6	10	138	23	18
7	10	147	21	9
8	10	152	19	5
9	10	153	17	1
10	10	150	15	−3
11	10	143	13	−7
12	10	133	11.08	−10

Production: One Variable Input

- Firms make decisions based on the benefits with the costs of production

- Sometimes useful to look at benefits and costs on an *incremental basis*

  - How much more can be produced when at incremental units of an input

Production: One Variable Input

- We now characterize the productivity of firm's labor input.
- Average product of Labor - Output per unit of a particular product
- Measures the productivity of a firm's labor in terms of how much, on average, each worker can produce

$$\boxed{AP = \frac{Output}{Labor\ Input} = \frac{q}{L}}$$

Production: One Variable Input

- Marginal Product of Labor – additional output produced when labor increases by one unit
- Change in output divided by the change in labor

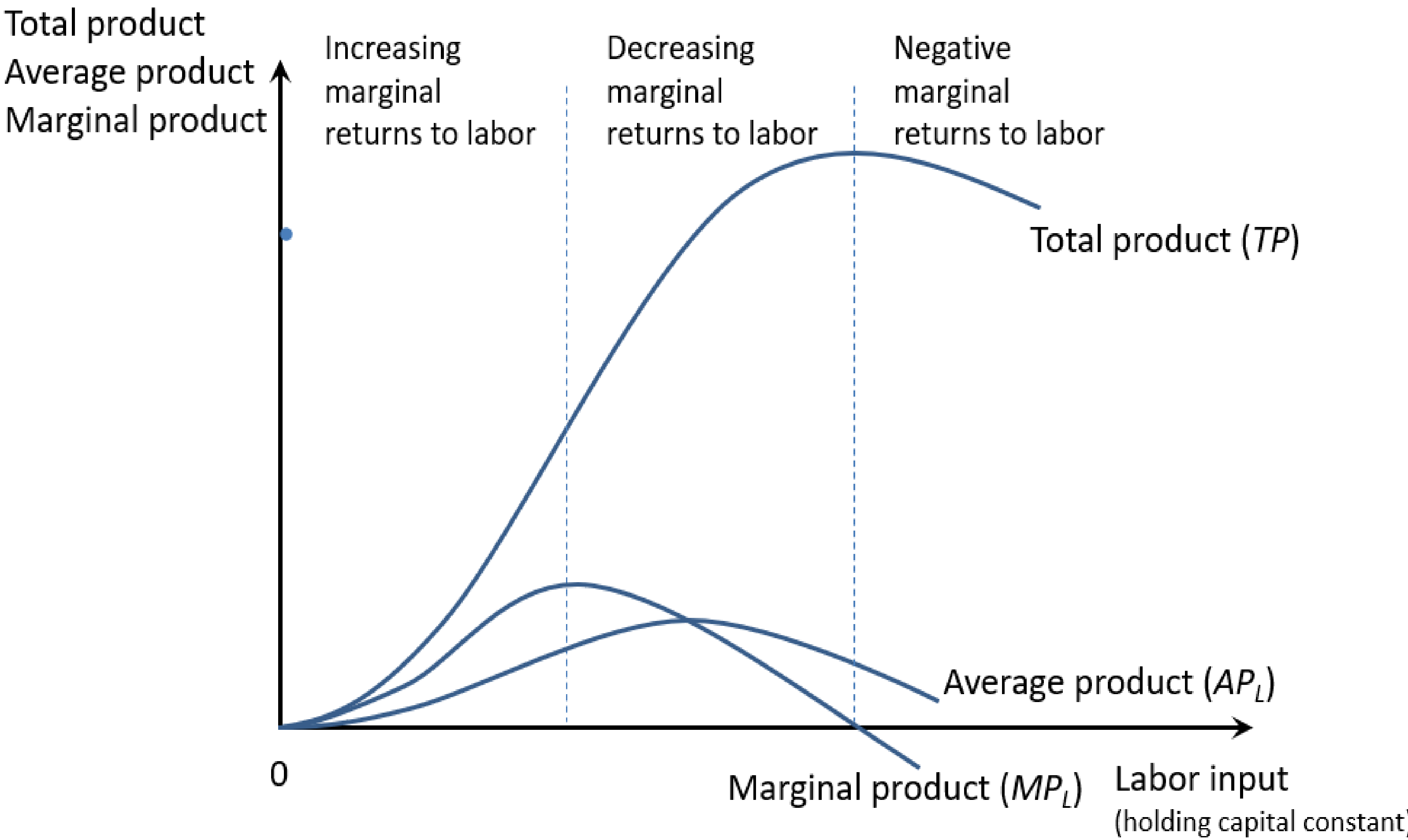
$$\boxed{MP_L = \frac{\Delta Output}{\Delta Labor\ Input} = \frac{\Delta q}{\Delta L}}$$

Average Productivity Measure

Average Productivity Measures

- Average Product of Labor
  - $AP_L = Q/L$.
  - Measures the output of an "average" worker.
  - Example: $Q = F(K,L) = \sqrt{KL}$
  - If the inputs are K = 16 and L = 16, then the average product of labor is $AP_L = \sqrt{256}/16 = 16/16 = 1$
- Average Product of Capital
  - $AP_K = Q/K$.
  - Measures the output of an "average" unit of capital.
  - Example: $Q = F(K,L) = \sqrt{KL}$
  - If the inputs are K = 16 and L = 16, then the average product of labor is $AP_L = \sqrt{256}/16 = 16/16 = 1$

Relation between Productivity Measures in Action



- When there are increasing marginal returns to labour, an increase in the quantity of labour increases total output at an increasing rate.

- Increasing marginal returns are usually thought to occur because of the gains from specialization of labour.

- In a plant with a small work force, workers may have to perform multiple tasks.

- **For example**, a worker might be responsible for moving raw materials within the plant, operating the machines, and inspecting the finished goods (quality control) once they are produced.

- As more workers are added, workers can specialize—some will be responsible only for moving raw materials in the plant; others will be responsible only for operating the machines; still others will specialize in inspection and quality control.

- Specialization enhances the marginal productivity of workers because it allows them to concentrate on the tasks at which they are most productive.

- **Diminishing marginal returns to labour**. When there are diminishing marginal returns to labour, an increase in the quantity of labour still increases total output but at a decreasing rate.

- Diminishing marginal returns set in when the firm exhausts its ability to increase labour productivity through the specialization of workers.

- Finally, when the quantity of labour exceeds $L=5$, (see table next slide) an increase in the quantity of labour results in an increase in output but at a decreasing rate.

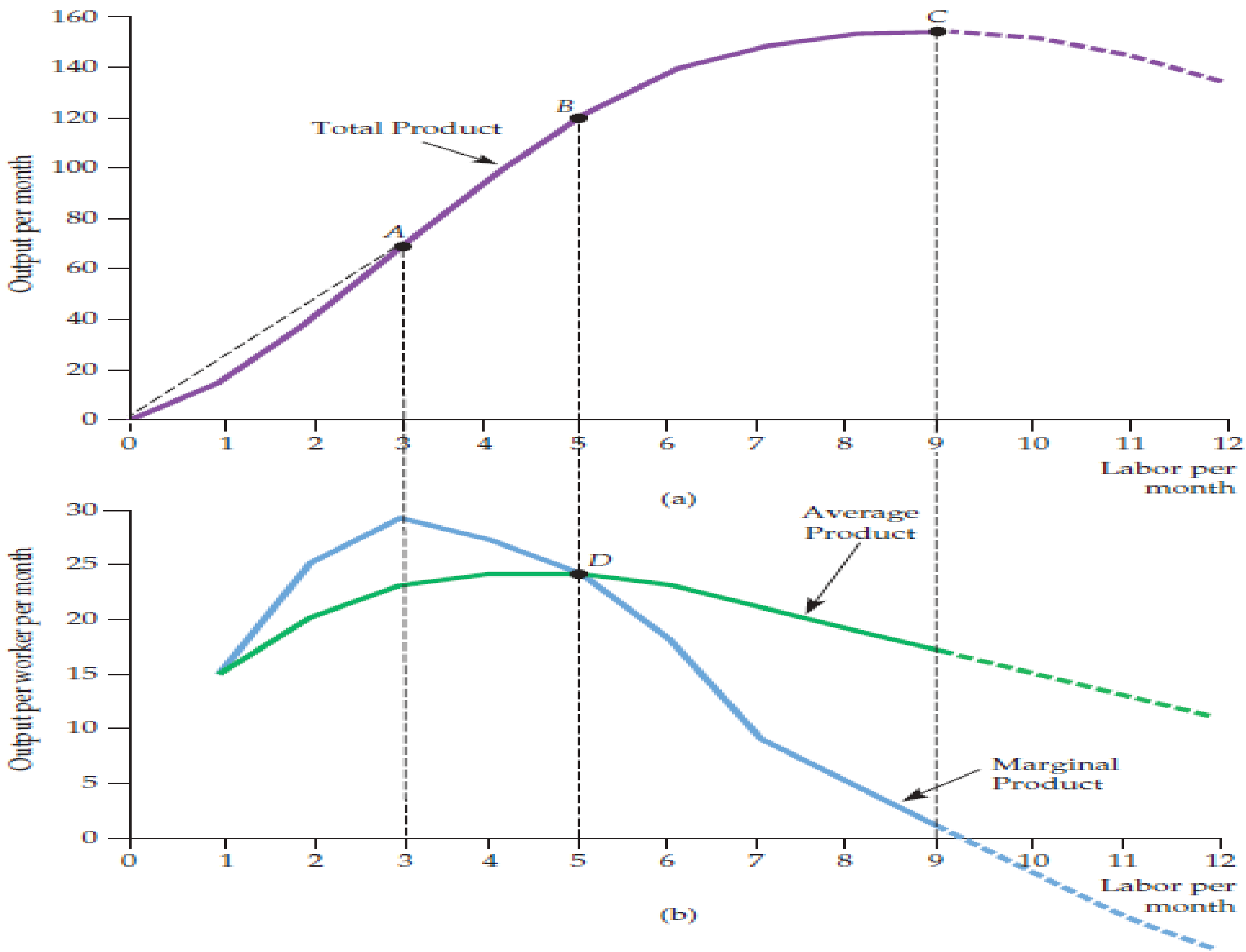
- In this region, we have **diminishing total returns to labour**.

- When there are diminishing total returns to labour, an increase in the quantity of labour decreases total output.

AMOUNT OF LABOR (L)	AMOUNT OF CAPITAL (K)	TOTAL OUTPUT (q)	AVERAGE PRODUCT (q/L)	MARGINAL PRODUCT ($\Delta q/\Delta L$)
0	10	0	—	—
1	10	15	15	15
2	10	40	20	25
3	10	69	23	29
4	10	96	24	27
5	10	120	24	24
6	10	138	23	18
7	10	147	21	9
8	10	152	19	5
9	10	153	17	1
10	10	150	15	−3
11	10	143	13	−7
12	10	133	11.08	−10

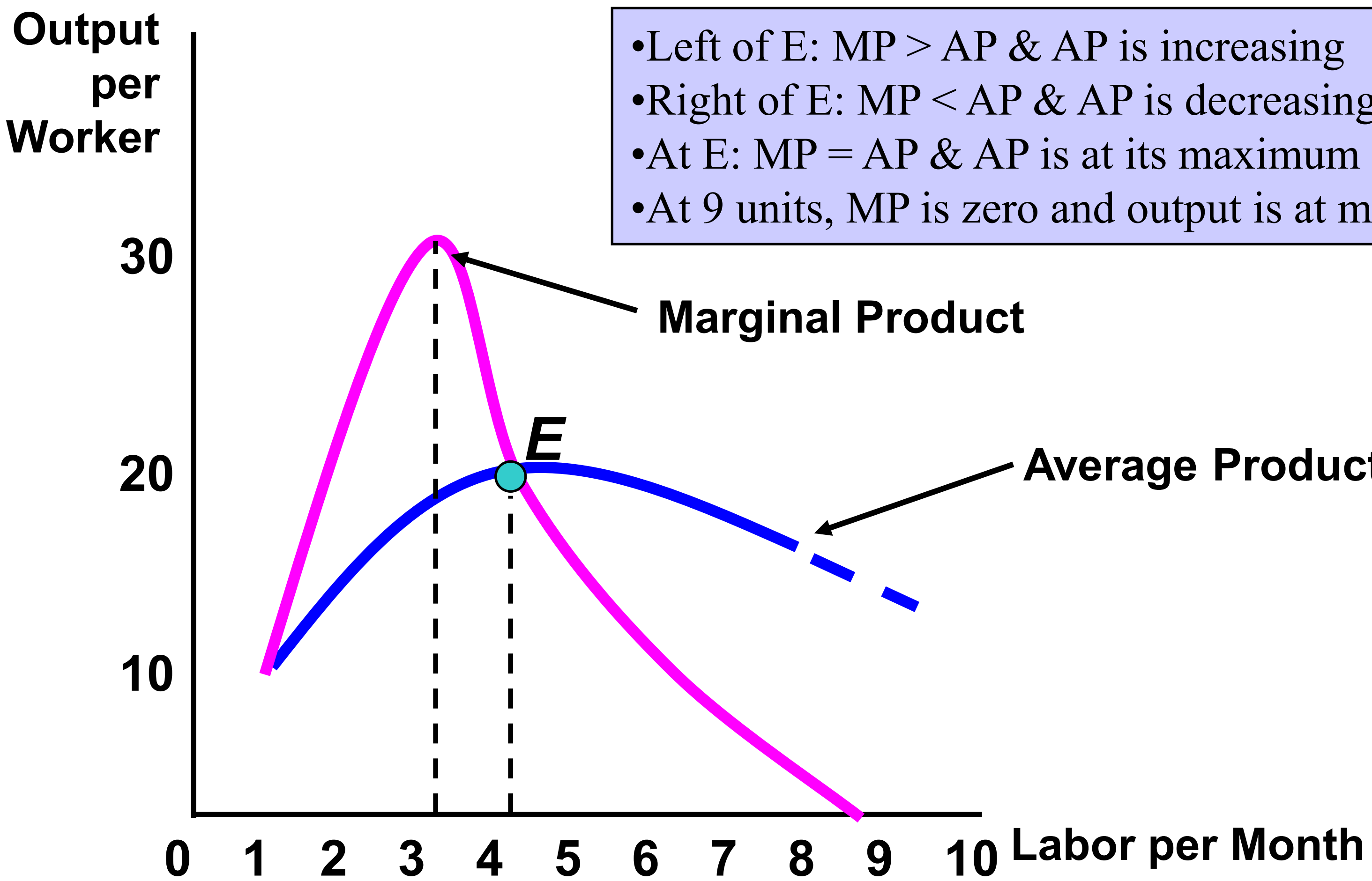
- Diminishing total returns occur because of the fixed size of the plant: if the quantity of labour used becomes too large, workers don’t have enough space to work effectively.

- Also, as the number of workers employed in the plant grows, their efforts become increasingly difficult to coordinate.


160
140
120
100
80
60
40
20
0
Output per month
Total Product
A
B
C
0 1 2 3 4 5 6 7 8 9 10 11 12
Labor per month
(a)
30
25
20
15
10
5
0
Output per worker per month
D
Average Product
Marginal Product
0 1 2 3 4 5 6 7 8 9 10 11 12
Labor per month
(b)

Production: One Variable Input

- Left of E: MP > AP & AP is increasing
- Right of E: MP < AP & AP is decreasing
- At E: MP = AP & AP is at its maximum
- At 9 units, MP is zero and output is at max



Marginal & Average Product

- When marginal product is greater than the average product, the average product is increasing
- When marginal product is less than the average product, the average product is decreasing

- When marginal product is zero, total product (output) is at its maximum

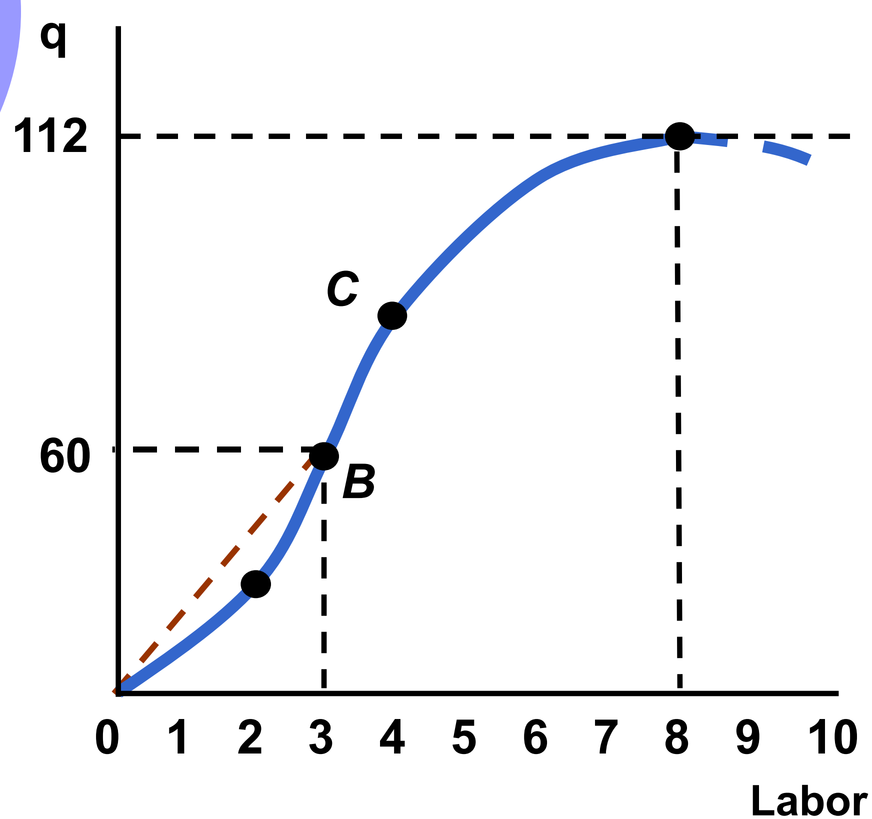
- Marginal product crosses average product at its maximum

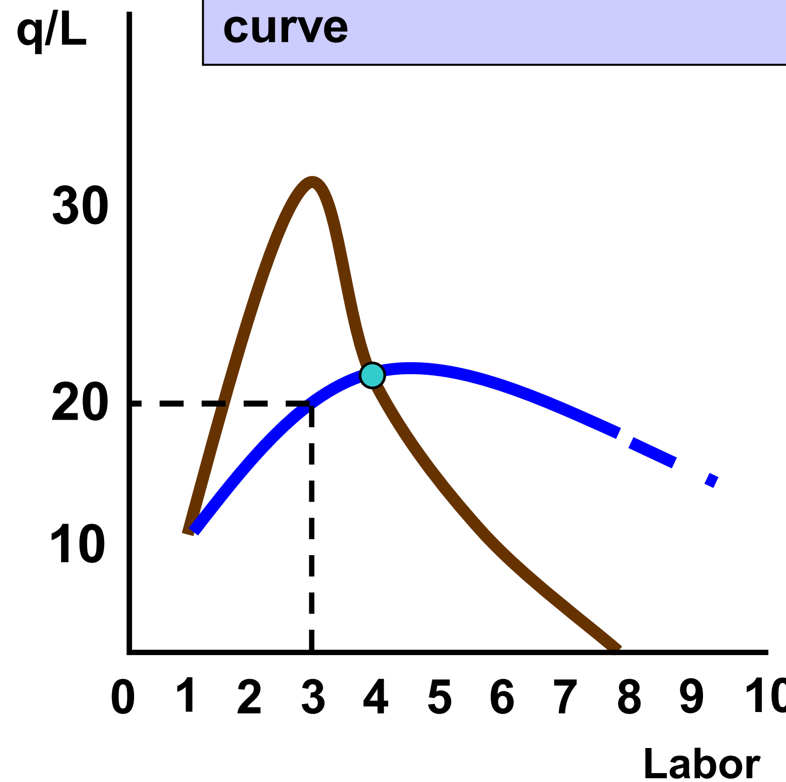
Product Curves

- A geometric relationship between the total, the average, and marginal curves
  - Slope of line from origin to any point on the total product curve is the average product
  - The marginal product is the slope of the line tangent to any corresponding point on the total product curve

Product Curves

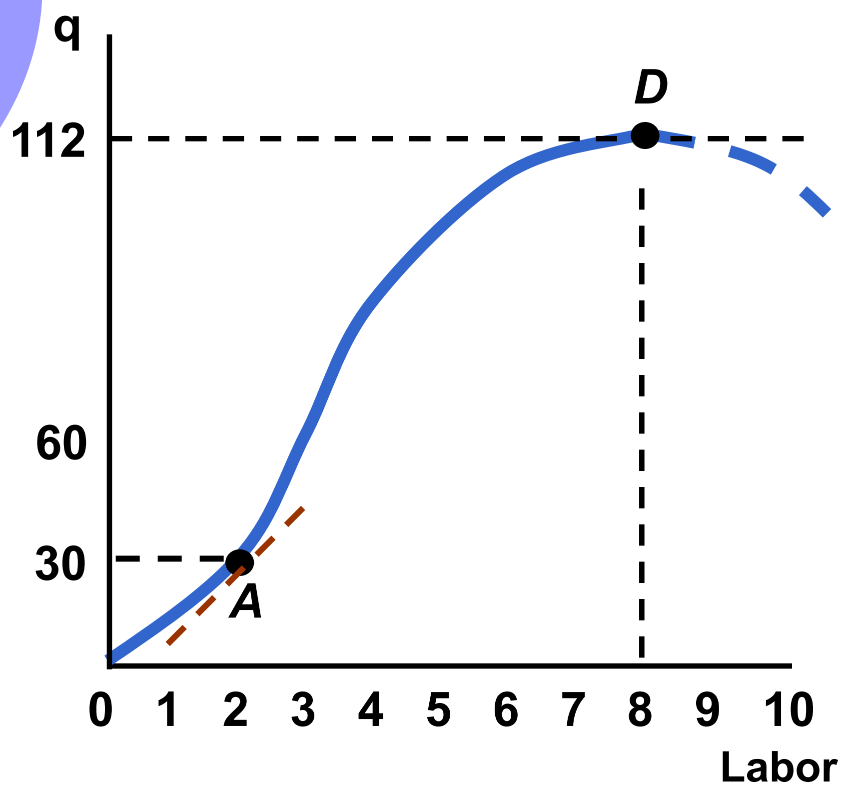
AP is slope of line from origin to point on TP curve





Product Curves

MP is slope of line tangent to corresponding point on TP curve



Law of Diminishing Marginal Returns

- Law of Diminishing Marginal Returns: As the use of an input increases with other inputs fixed, the resulting additions to output will eventually decrease.
- When the labor input is small and capital is fixed, output increases considerably.
- When the labor input is large, some workers become less efficient and MP of labor decreases.

Production: Two Variable Inputs

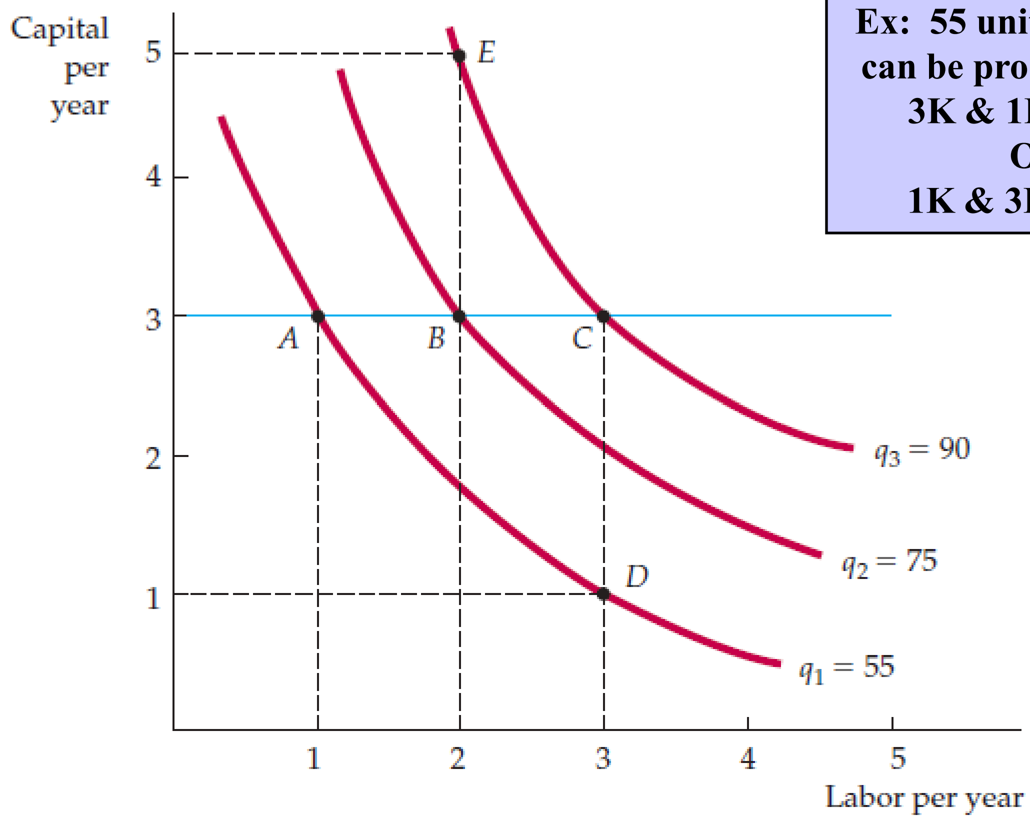
- Firm can produce output by combining different amounts of labor and capital
- In the long-run, capital and labor are both variable.
- We can look at the output we can achieve with different combinations of capital and labor – Table 2

(Table 2)Production: Two Variable Inputs

	Labor Input				
Capital Input	1	2	3	4	5
1	20	40	55	65	75
2	40	60	75	85	90
3	55	75	90	100	105
4	65	85	100	110	115
5	75	90	105	115	120

Production: Two Variable Inputs

- The information can be represented graphically using isoquants
  - Curves showing all possible combinations of inputs that yield the same output
- Curves are smooth to allow for use of fractional inputs
  - Curve 1 shows all possible combinations of labor and capital that will produce 55 units of output (figure next slide)



Ex: 55 units of output can be produced with
3K & 1L (pt. A)
OR
1K & 3L (pt. D)

Isoquant

- The combinations of inputs (K, L) that yield the producer the same level of output.

- The shape of an isoquant reflects the ease with which a producer can substitute among inputs while maintaining the same level of output.

- Convexity means that as you move down along an isoquant (using more labor and less capital), the slope becomes flatter → MRTS diminishes.

One Example

- Production Function Q = F(K,L) = $\sqrt{KL}$
- What is the algebraic form of the Isoquant for an output level of 20?

$$20 = K^{0.5}L^{0.5}$$

$$\Leftrightarrow 400 = KL$$

$$\Rightarrow K = \frac{400}{L}$$

Production: Two Variable Inputs

- Substituting Among Inputs
  - Firms must decide what combination of inputs to use to produce a certain quantity of output
  - There is a trade-off between inputs allowing them to use more of one input and less of another for the same level of output.

Production: Two Variable Inputs

- Substituting Among Inputs
  - Slope of the isoquant shows how one input can be substituted for the other and keep the level of output the same.
  - Slope is the marginal rate of technical substitution (MRTS)
    - Amount by which the quantity of one input can be reduced when one extra unit of another input is used, so that output remains constant.

- Initially, when a firm has a lot of capital and little labor, one unit of labor can replace a large amount of capital (high MRTS).

- But as labor increases and capital decreases, labor becomes less productive at replacing capital, so MRTS falls.

Production: Two Variable Inputs

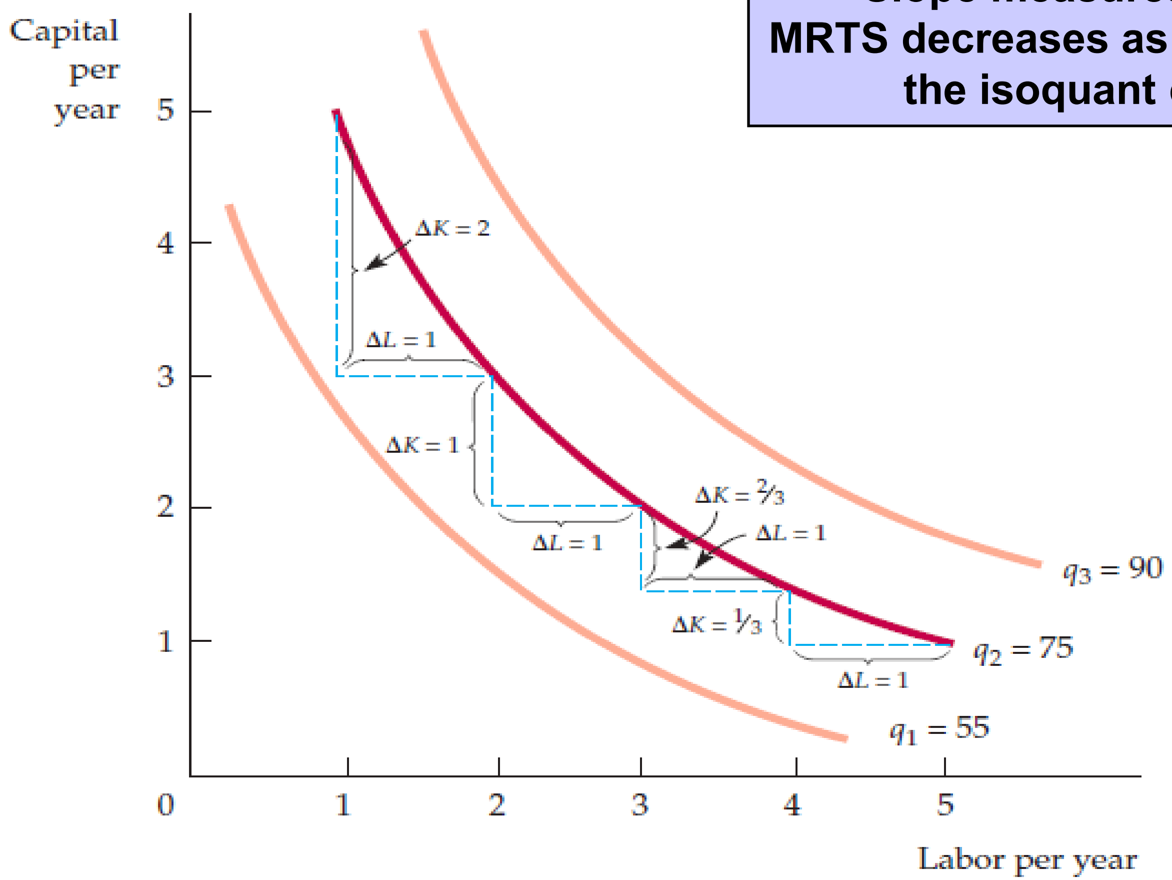
- The marginal rate of technical substitution equals:

$$MRTS = \frac{Change\,in\ Capital\,input}{Change\,in\,Labor\,input}$$

$$MRTS = -\Delta K / \Delta L \text{ (for a fixed level of } q)$$

Marginal Rate of Technical Substitution

Slope measures MRTS
MRTS decreases as move down the isoquant curve



MRTS and Isoquants

- We assume there is diminishing MRTS
  - Increasing labor in one unit increments from 1 to 5 results in a decreasing MRTS from 1 to 1/3.
- Diminishing MRTS occurs because of diminishing returns and implies isoquants are convex.
- There is a relationship between MRTS and marginal products of inputs

Production Function VS Utility Function

Production Function	Utility Function
Isoquant(Defn: all possible combinations of inputs that just suffice to produce a given amount of output)	Indifference Curve
Marginal Rate of Technical Substitution	Marginal Rate of Substitution

MRTS and Marginal Products

- If we increase labor and decrease capital to keep output constant, we can see how much the increase in output is due to the increased labor
  - Amount of labor increased times the marginal productivity of labor

$$= (MP_L)(\Delta L)$$

MRTS and Marginal Products

- Similarly, the decrease in output from the decrease in capital can be calculated
  - Decrease in output from reduction of capital times the marginal productivity of capital

$$= (MP_K)(\Delta K)$$

MRTS and Marginal Products

- If we are holding output constant, the net effect of increasing labor and decreasing capital must be zero
- Using changes in output from capital and labor we can see

$$\boxed{(MP_L)(\Delta L) + (MP_K)(\Delta K) = 0}$$

MRTS and Marginal Products

- Rearranging equation, we can see the relationship between MRTS and MPs

$$(MP_L)(\Delta L) + (MP_K)(\Delta K) = 0$$

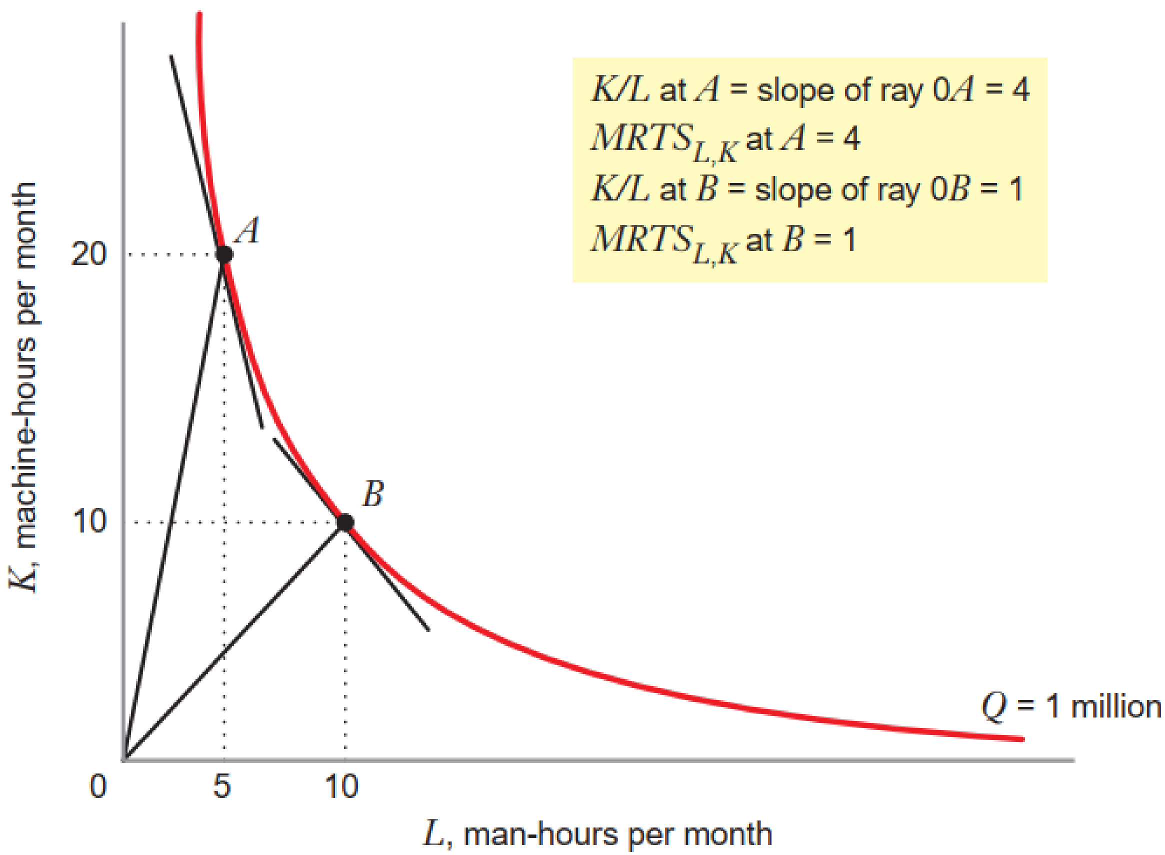
$$(MP_L)(\Delta L) = -(MP_K)(\Delta K)$$

$$\frac{(MP_L)}{(MP_K)} = -\frac{\Delta K}{\Delta L} = MRTS$$

Elasticity of Substitution

- It is a measure of how easy it is for a firm to substitute labour for capital.
- Specifically, the elasticity of substitution measures how quickly the marginal rate of technical substitution of labour for capital changes as we move along an isoquant. [see figure next slide].
- As labour is substituted for capital, the ratio of the quantity of capital to the quantity of labour, known as the **capital–labour ratio**, K/L, must fall.

Figure-A


K/L at A = slope of ray 0A = 4
$MRTS_{L,K}$ at A = 4
K/L at B = slope of ray 0B = 1
$MRTS_{L,K}$ at B = 1
A
B
20
10
0
5
10
Q = 1 million
K, machine-hours per month
L, man-hours per month

- The marginal rate of substitution of capital for labour, $MRTS_{LK}$, also falls, as we saw in the previous section.

- The elasticity of substitution, often denoted by , measures the percentage change in the capital–labour ratio for each 1 percent change $MRTS_{LK}$ in as we move along an isoquant:

$$\sigma = \frac{\textit{percentage change} \text{ in capital–labor ratio}}{\textit{percentage change} \text{ in } MRTS_{L,K}}$$

$$= \frac{\%\,\Delta\left(\frac{K}{L}\right)}{\%\,\Delta MRTS_{L,K}}$$

- Figure A illustrates the elasticity of substitution.
- Suppose a firm moves from the input combination at point A (L =5 man-hours per month, K 20 machine hours per month) to the combination at point B (L=10, K =10). The capital–lab ratio K/L at A is equal to the slope of a ray from the origin to A (slope of ray $0A$ =4);

- The $MRTS_{LK}$ at A is equal to the negative of the slope of the isoquant at A (slope of isoquant = -4, thus, $MRTS_{LK}$=4)
- At B, the capital–labour ratio equals the slope of ray $0B$, or 1;

- The $MRTS_{LK}$ equals the negative of the slope of the isoquant at B, also 1.

- The percent change in the capital–labour ratio from *A* to *B* is -75 percent (from 4 down to 1), as is the percent change in the $MRTS_{LK}$ between those points.

- Thus, the elasticity of substitution over this interval is 1 (-75%/-75% 1).

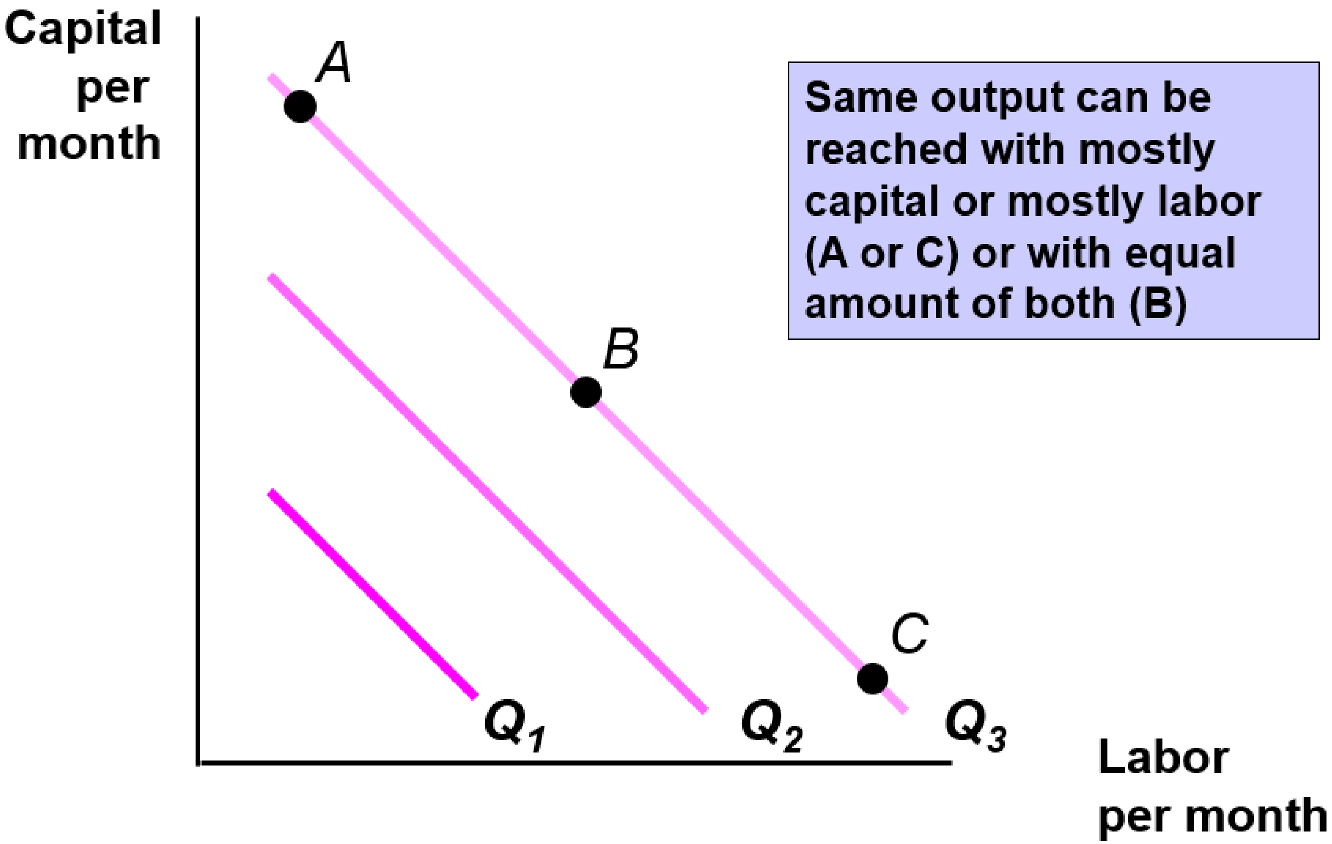
- Example: to be discussed in the class.

Isoquants: Special Cases

- Two extreme cases show the possible range of input substitution in production.

1. Perfect substitutes
   - MRTS is constant at all points on isoquant
   - Sometimes a firm may find one type of input may be perfectly substituted for another type
   - Example: manufacturing process may require energy in the form of natural gas or fuel oil, and a given amount of natural gas can always be substituted for each liter of fuel oil.

Perfect Substitutes



The production function:

$$Q = 20H + 10L$$

can be explained as follows:

A. Perfect Substitutes (e.g., $Q = 20H + 10L$)

- **Isoquants** are **straight lines** with negative slope.
- A firm can use **all H**, all **L**, or **any mix** as long as total output is maintained.

Example: Isoquant for $Q = 100$

$$20H + 10L = 100 \quad \Rightarrow \quad L = 10 - 2H$$

- In **perfect substitutes**, the firm has **flexibility**: if one input becomes expensive, it can switch to the other.

This is a **linear production function**, where output Q is a **linear combination** of two inputs:

- H: Units of skilled labor (e.g., high-skilled workers)
- L: Units of unskilled labor (e.g., low-skilled workers)

- The coefficient **20** means each unit of **skilled labor (H)** contributes **20 units of output**.
- The coefficient **10** means each unit of **unskilled labor (L)** contributes **10 units of output**.

So, **skilled labor is twice as productive** as unskilled labor in this function.

- Since the function is linear, **H and L are perfect substitutes**.
- The **marginal rate of technical substitution (MRTS)** is constant:

$$\mathrm{MRTS}_{HL} = \frac{\partial Q/\partial H}{\partial Q/\partial L} = \frac{20}{10} = 2$$

This means: the firm is always willing to trade **2 units of unskilled labor for 1 unit of skilled labor**, to keep output constant.

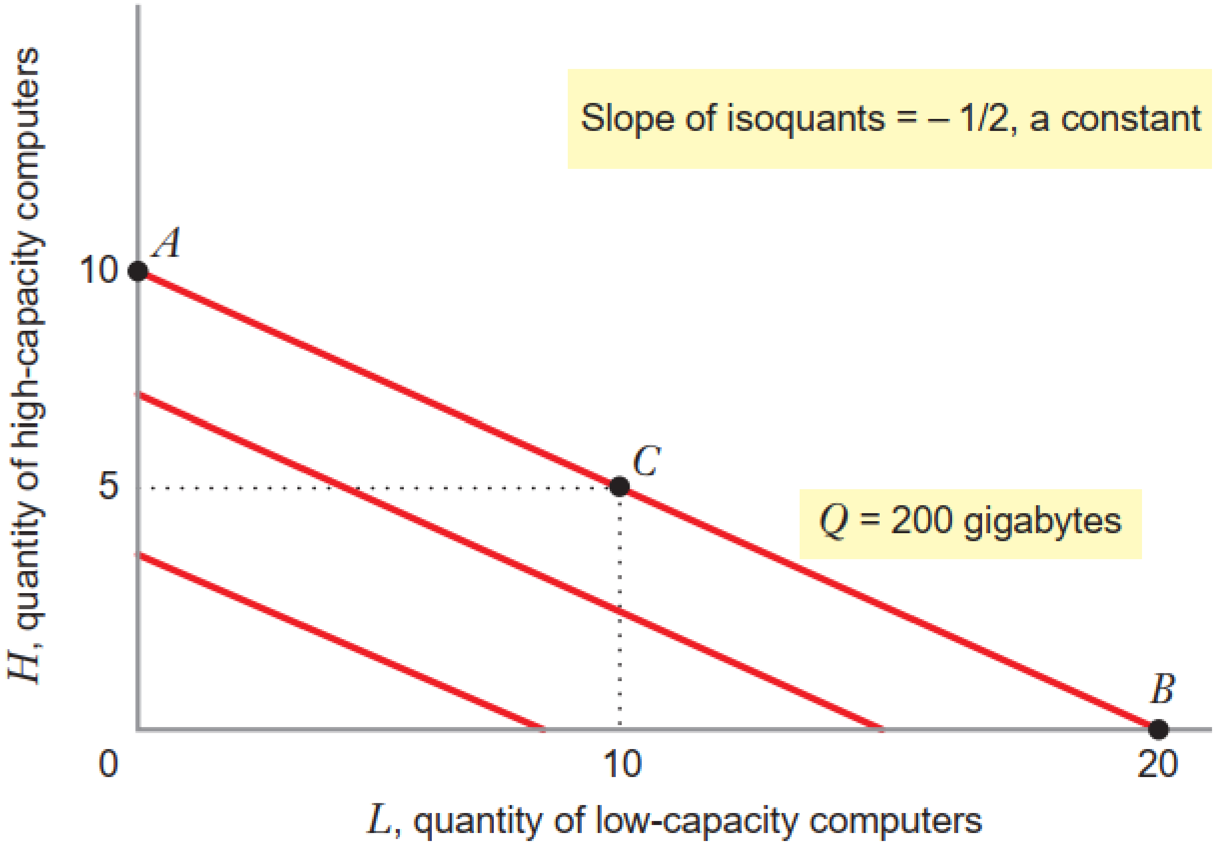
Suppose a firm hires:

- $H = 3$ units of skilled labor
- $L = 4$ units of unskilled labor

Then:

$$Q = 20(3) + 10(4) = 60 + 40 = 100$$

Figure: Linear Isoquants



Elasticity of Substitution

- Because $MRTS_{LH}$ does not change as we move along an isoquant, $\Delta MRTS_{LH}$=0.
- Using the equation below:

$$\sigma = \frac{\textit{percentage change} \text{ in capital–labor ratio}}{\textit{percentage change} \text{ in } MRTS_{L,K}}$$

$$= \frac{\%\,\Delta\left(\frac{K}{L}\right)}{\%\,\Delta MRTS_{L,K}}$$

The elasticity of substitution for a linear production $(\sigma = \infty)$.

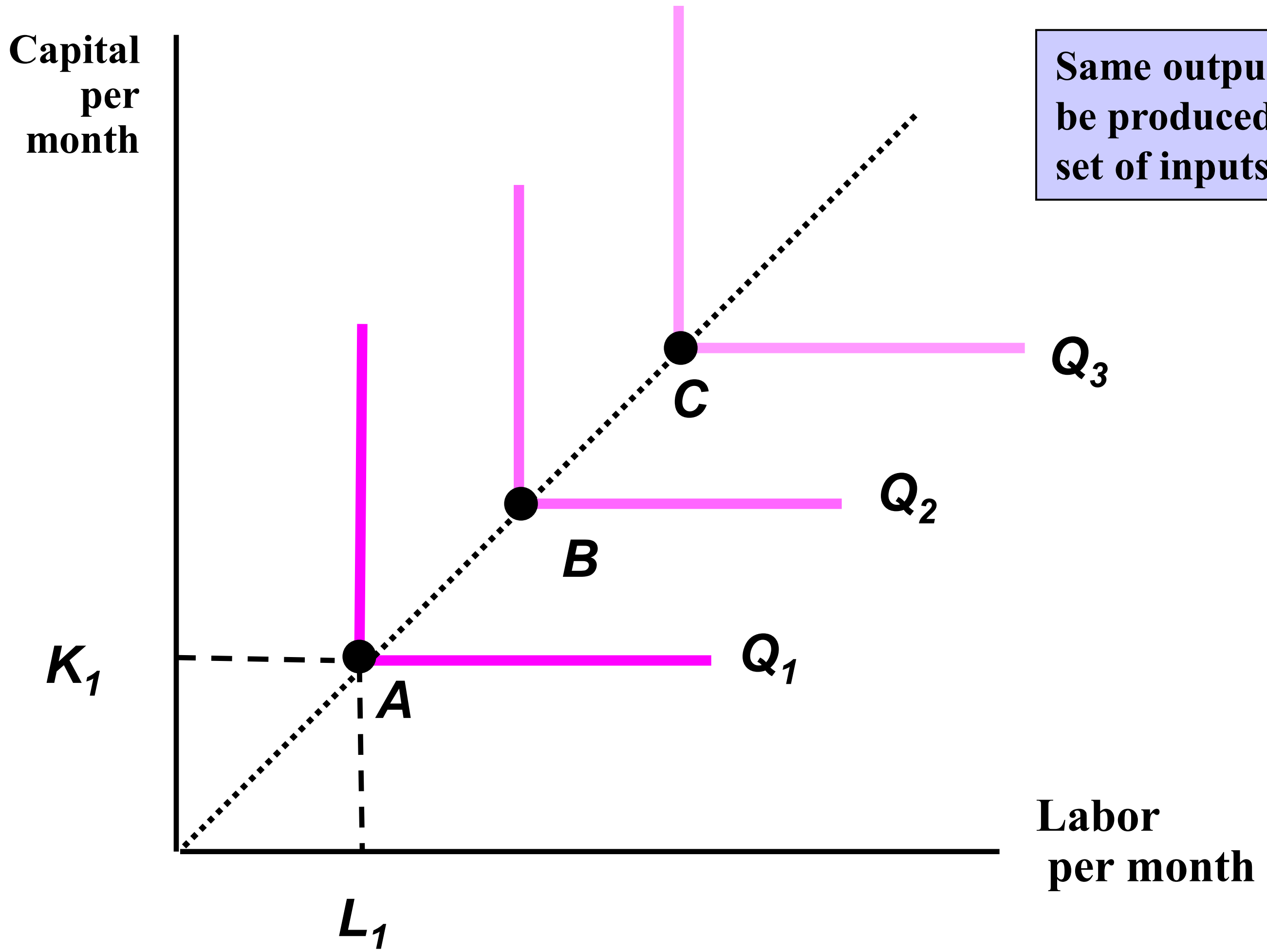
- In other words, the inputs in a linear production function are infinitely (perfectly) substitutable for each other.
- When we have a linear production function, we say that the inputs are **perfect substitutes**.
- In our computer example, the fact that low-capacity and high-capacity computers are perfect substitutes means that in terms of data storage capabilities, two low-capacity computers are just as good as one high-capacity computer.
- Or, put another way, the firm can perfectly replicate the productivity of one high-capacity computer by employing two low-capacity computers.

Isoquants: Special Cases

- Extreme cases (cont.)

2. Perfect Complements
   - Fixed proportions production function
   - There is no substitution available between inputs
   - The output can be made with only a specific proportion of capital and labor
   - Cannot increase output unless increase both capital and labor in that specific proportion

Fixed-Proportions Production Function



Same output can only be produced with one set of inputs.

A **Fixed-Proportions Production Function** (also known as a **Leontief production function**) assumes that inputs must be used in **fixed, rigid ratios** to produce output — like ingredients in a recipe. This reflects a **perfect complements** relationship between inputs.

General Form:

$$Q = \min\left(\frac{H}{a}, \frac{L}{b}\right)$$

Where:

- H: Units of skilled labor
- L: Units of unskilled labor
- a, b: Fixed input proportions
- Q: Quantity of output

1. **No Substitutability**:

   You **cannot substitute** one input for another. More of one input alone does not increase output.

2. **Fixed Input Ratios**:

   The firm requires inputs in **a strict proportion** (e.g., 1 unit of H for every 2 units of L). For example:

$$Q = \min\left(H, \frac{L}{2}\right)$$

   requires **1 skilled worker and 2 unskilled workers** to produce 1 unit of output.

3. **L-shaped Isoquants**:

   Isoquants are **right-angled (L-shaped)**, reflecting fixed ratios. There's **no marginal rate of technical substitution (MRTS)** along the kink — it's either zero or infinite.

4. **Inefficiency of Excess Input**:

   Having extra of one input doesn't help. For example, if you have 5 skilled workers but only enough unskilled labor for 3 units of output, the extra skilled workers go unused.

Let's take:

$$Q = \min\left(\frac{H}{2}, \frac{L}{3}\right)$$

This means:

- It takes **2 units of H** and **3 units of L** to produce **1 unit of Q**.
- If you have $H = 10$ and $L = 9$, then:

$$Q = \min\left(\frac{10}{2}, \frac{9}{3}\right) = \min(5, 3) = 3$$

You are constrained by the amount of unskilled labor (L), so you can only produce 3 units of output.

- Manufacturing Example: Car Assembly Line

To assemble **1 car**, you need:

- **4 tires**
- **1 engine**
- **2 workers**

This is a **Leontief production process** because you cannot substitute tires for workers, or engines for tires — you need them in **fixed proportion**.

- Write a production function:

$$Q = \min\left(\frac{T}{4}, E, \frac{L}{2}\right)$$

Where:

- Q: number of cars produced
- T: number of tires
- E: number of engines
- L: number of laborers

If you have:

- 20 tires
- 5 engines
- 10 workers

Then:

$$Q = \min\left(\frac{20}{4}, 5, \frac{10}{2}\right) = \min(5, 5, 5) = 5$$

You can produce **5 cars**. All inputs are used in their fixed proportion — no wastage.

🎓 Why Leontief is Useful:

- Accurately models **rigid production processes** where **engineering or chemical constraints** fix input ratios.
- Helps in **input-output analysis** (originated by Wassily Leontief, Nobel laureate).
- Essential in industries like **automotive, construction, electricity, and pharmaceuticals**, where inputs can't be easily substituted.

MRTS for Perfect Complements

The **Marginal Rate of Technical Substitution (MRTS)** measures how much of one input (e.g., labor) can be reduced when **one additional unit** of another input (e.g., capital) is used, **holding output constant**.

Mathematically:

$$\text{MRTS}_{KL} = -\frac{MP_K}{MP_L}$$

Where:

- MP_K = Marginal Product of capital
- MP_L = Marginal Product of labor

MRTS is undefined or Discontinuous

- At the **kink point** (efficient input combination):

  The isoquant is **L-shaped**, and MRTS is **not defined** — there's **no slope** because the isoquant changes direction.

- **Along the vertical segment** (extra capital, insufficient labor):

$$\text{MRTS} = 0 \quad \text{(no substitution possible)}$$

- **Along the horizontal segment** (extra labor, insufficient capital):

$$\text{MRTS} = \infty \quad \text{(no substitution possible)}$$

- Intuition:

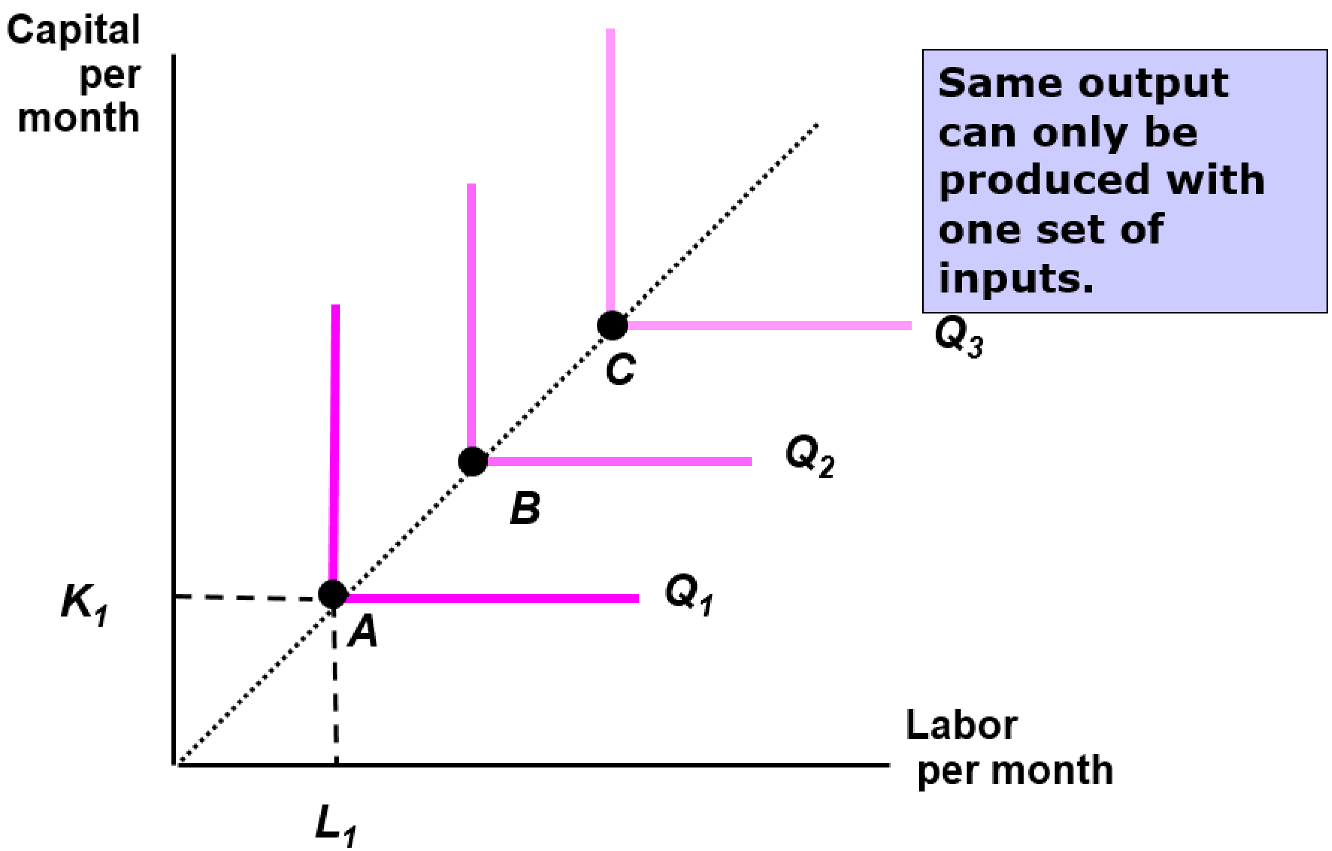
In a Leontief function, **substitution is impossible.** So the MRTS, which depends on substitutability, becomes:

- **Zero** when one input is in excess
- **Infinite** when the other input is in excess
- **Undefined** at the efficient combination (kink)

🎓 **Example:**

For $Q = \min(K, L)$

- At $K = L$: Efficient point (kink), MRTS is **undefined**
- At $K > L$: Output limited by labor; MRTS = 0 (extra capital doesn't help)
- At $L > K$: Output limited by capital; MRTS = ∞ (extra labor doesn't help)



Capital per month
Same output can only be produced with one set of inputs.
Q_3
C
Q_2
B
K_1
Q_1
A
Labor per month
L_1

Cobb-Douglas Production Function

- **Cobb–Douglas production function** is intermediate between a linear production function and a fixed-proportions production function.

- The Cobb–Douglas production function is given by the formula $Q = AL^{\propto}K^{\beta}$
- where A, α, and β are positive constants. In the figure (next slide) their values are 100, 0.4, and 0.6, respectively).

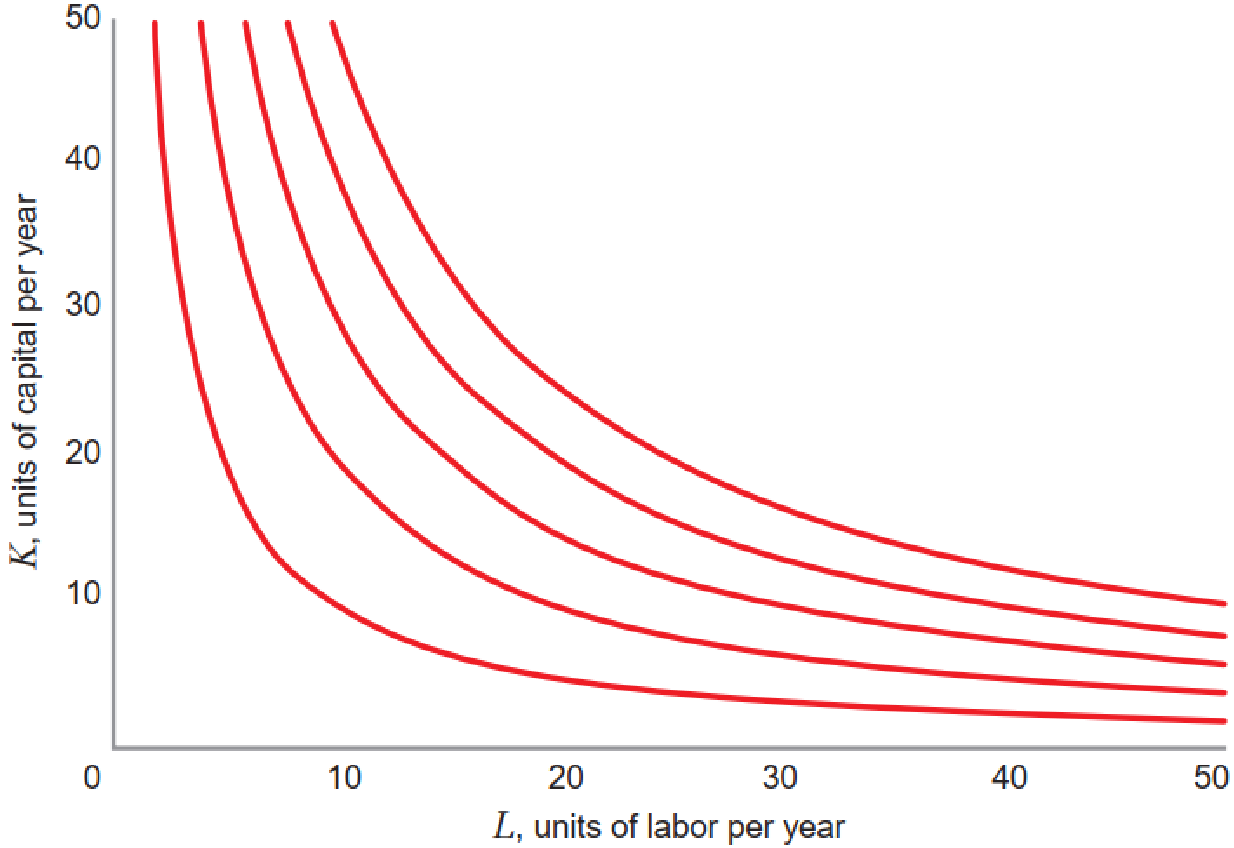
- With the Cobb–Douglas production function, capital and labour can be substituted for each other.

- Unlike a fixed-proportions production function, capital and labour can be used in variable proportions.

- Unlike a linear production function, though, the rate at which labour can be substituted for capital is not constant as you move along an isoquant.

- This suggests that the elasticity of substitution for a Cobb–Douglas production function falls somewhere between 0 and infinity.
- In fact, it turns out that the elasticity of substitution along a Cobb–Douglas production function is always equal to 1.

Figure: Cobb-Douglas Production Function



Real World Example

- Example: Textile Manufacturing Firm in India

- A textile factory uses **capital** (e.g., weaving machines, looms, power supply) and **labour** (e.g., workers operating machines, quality control staff) to produce garments.
- Let ’s assume the production function is: $Q = 2 \cdot L^{0.6} \cdot K^{0.4}$

- Interpretation:
- For every combination of labour and capital, output is scaled by a factor of 2.

- The output is more sensitive to labour (0.6) than to capital (0.4), which might be typical in labour-intensive industries like textiles in India.

- Since0.6+0.4=1, the production function has **constant returns to scale**. Doubling both labour and capital will double the output.

Measuring Productivity

- Multiple research find that the Cobb–Douglas production function is thought to be a plausible/suitable way of characterizing many real-world production processes, economists often use it to study issues related to input productivity.

- For example, Nicholas Bloom, Raffaella Sadun, and John Van Reenen estimated Cobb–Douglas production functions to study the ability of U.S. and European companies to exploit information technology (IT) to raise productivity.

- Reference: Nicholas Bloom, Raffaella Sadun, and John Van Reenen, "Americans Do I.T. Better: U.S. Multinationals and the Productivity Miracle," NBER Working Paper W13085 (May 2007), available at SSRN, http://ssrn.com/abstract986935.

Returns to Scale

- In addition to discussing the tradeoff between inputs to keep production at the same level, we must know how increases in all input quantities affect the quantity of output the firm can produce.
  - Can change the scale of production by increasing all inputs in proportion
  - If we double inputs, output will most likely increase but by how much?

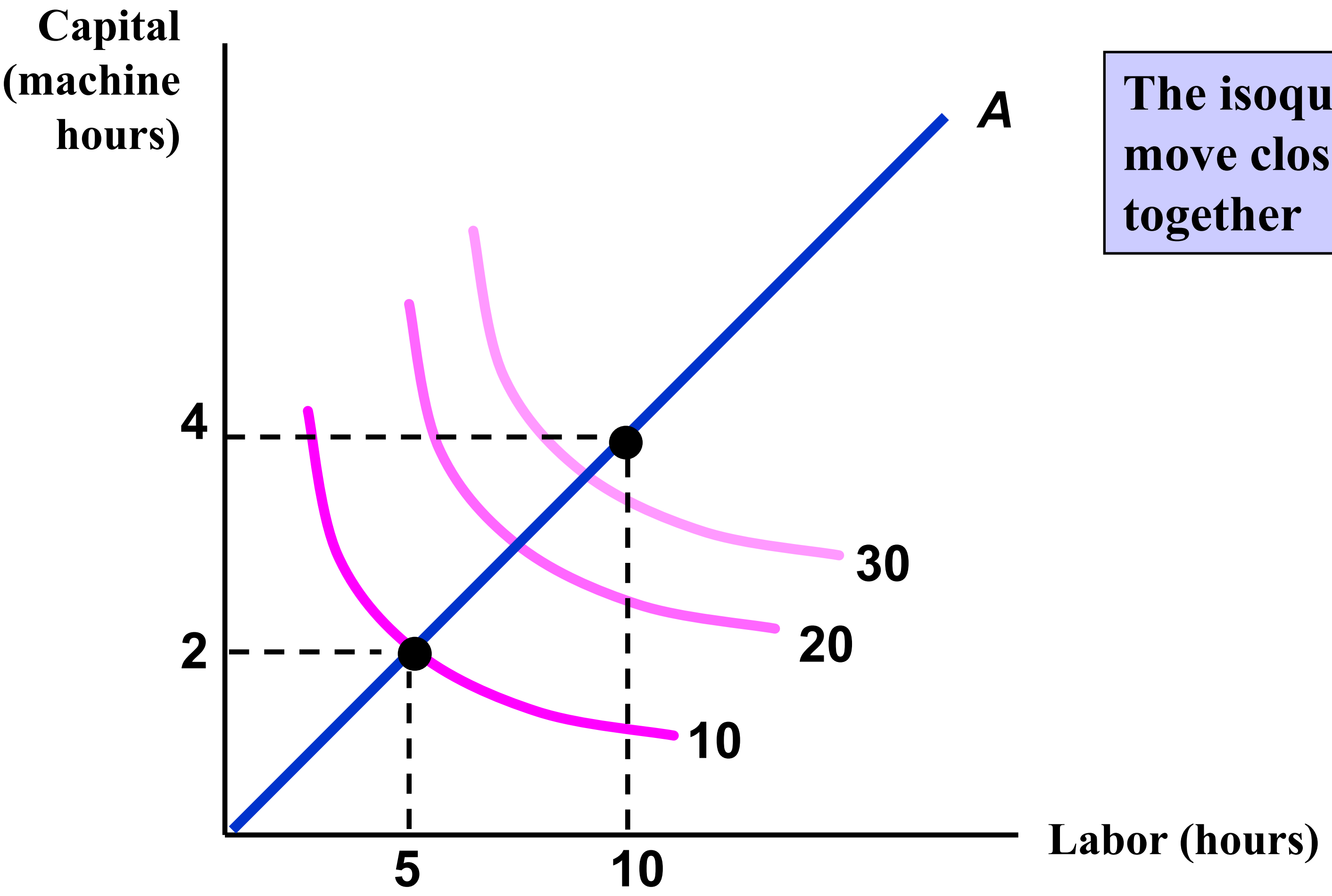
Returns to Scale

- Rate at which output increases as inputs are increased proportionately can be translated into the following:
  - Increasing returns to scale
  - Constant returns to scale
  - Decreasing returns to scale

Increasing Returns to Scale

- Increasing returns to scale: output more than doubles when all inputs are doubled
  - The isoquants get closer together

Increasing Returns to Scale

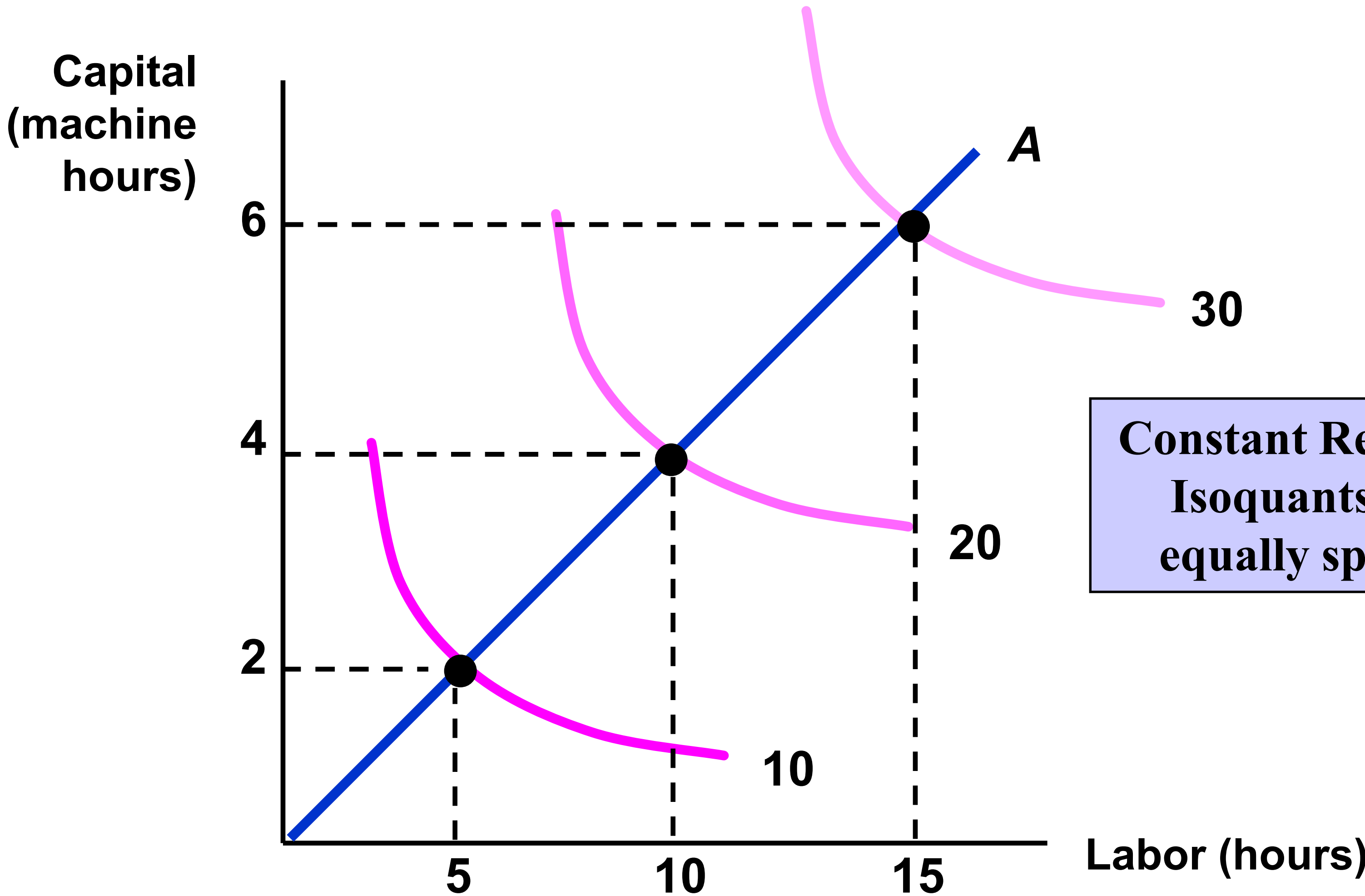


The isoquants move closer together

Constant Returns to Scale

- Constant returns to scale: output doubles when all inputs are doubled
  - Size does not affect productivity: one plant using a particular production process can easily be replicated, two plants produce twice as much output. (example: travel agent.)
  - Isoquants are equidistant apart

Constant Returns to Scale

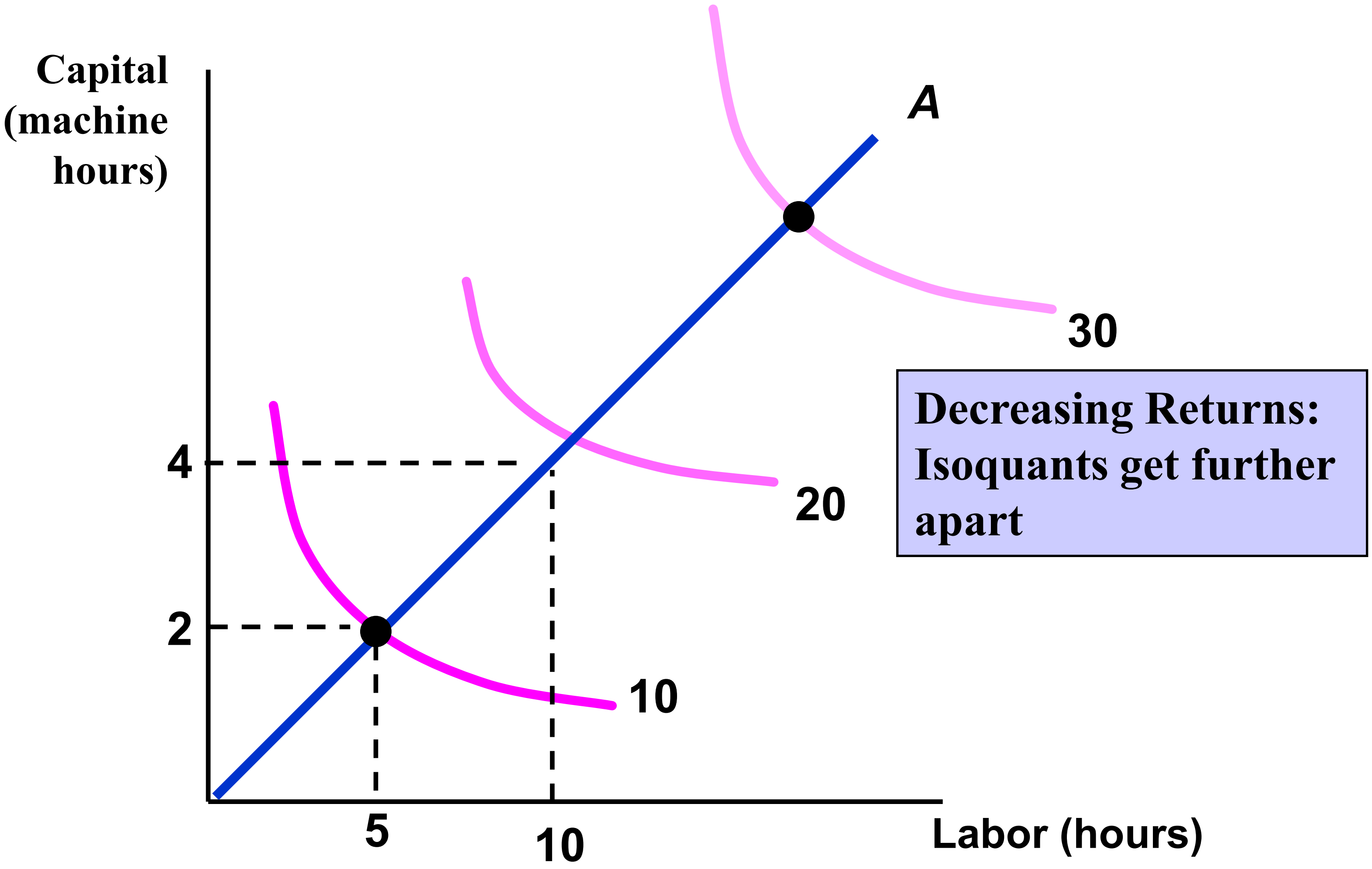


Constant Returns: Isoquants are equally spaced

Decreasing Returns to Scale

- Decreasing returns to scale: output less than doubles when all inputs are doubled
  - Decreasing efficiency with large size: difficulties in organizing and running a large-scale operation may lead to decreased productivity to both labor and capital.
  - Isoquants become farther apart

Decreasing Returns to Scale



Technological Progress

- The notion of **technological progress** captures the idea that production functions can shift over time.

- As knowledge in the economy evolves and as firms acquire technical know-how through experience and investment in research and development, a firm's production function will change.

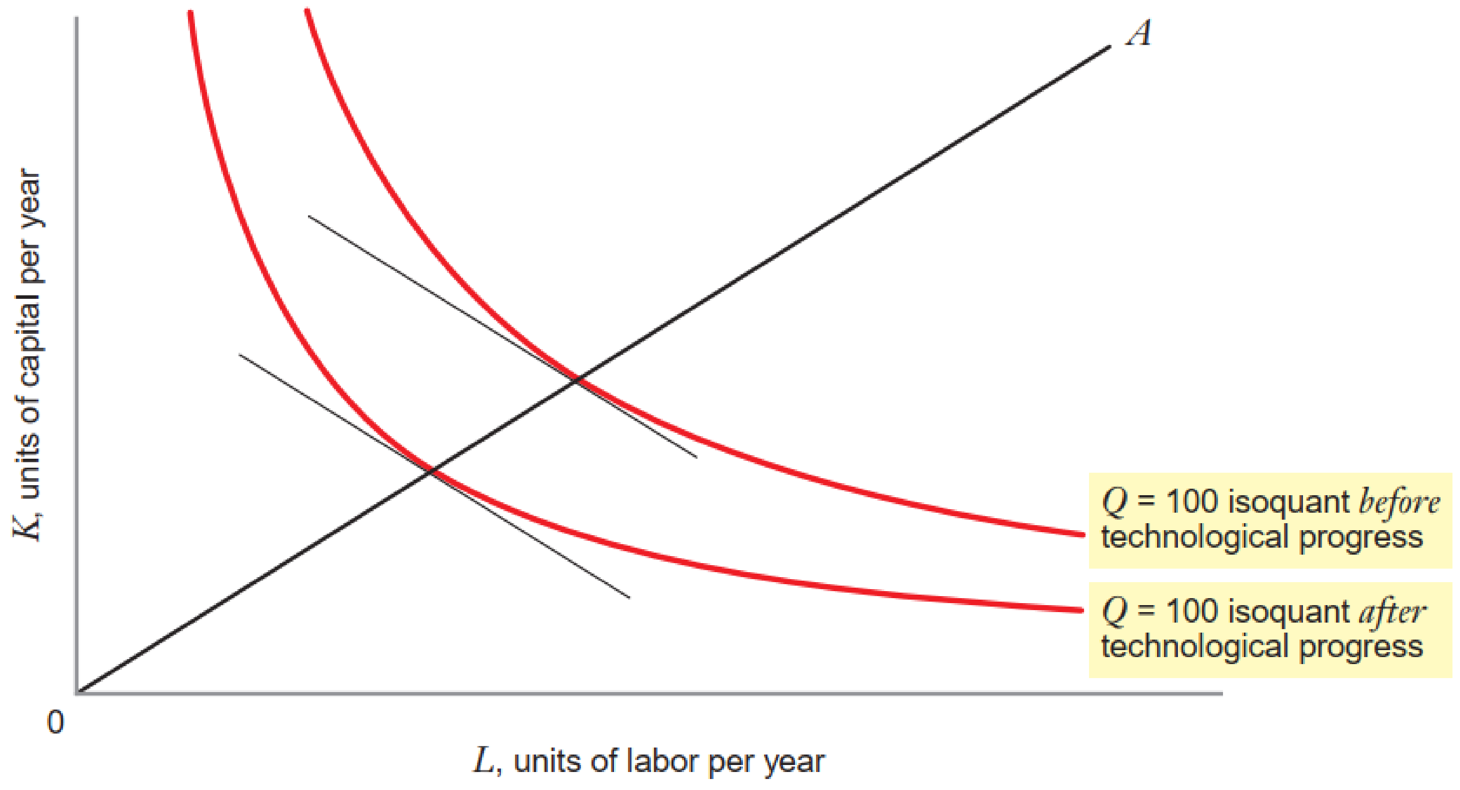
- In particular, technological progress refers to a situation in which a firm can achieve more output from a given combination of inputs, or equivalently, the same amount of output from lesser quantities of inputs.

- We can classify technological progress into three categories:
  - Neutral technological progress,
  - Labour-saving technological progress, and
  - Capital-saving technological progress.

Neutral Technological Progress

- In this case, an isoquant corresponding to a given level of output (100 units in the figure: next slide) shifts inward (indicating that lesser amounts of labour and capital are needed to produce a given output),
- But the shift leaves $MRTS_{LK}$, the marginal rate of technical substitution of labour for capital, unchanged along any ray OA from the origin.
- Under neutral technological progress, each isoquant corresponds to a higher level of output than before, but the isoquants themselves retain the same shape.

Figure A



A
K, units of capital per year
0
L, units of labor per year
Q = 100 isoquant *before* technological progress
Q = 100 isoquant *after* technological progress

- **Why “Neutral?**
- The term *neutral* means the technological change does not distort the **relative contribution** of capital and labor—it shifts the production possibility without altering factor shares in output (at constant factor prices).
- **Real-world example:**
- Introduction of better management software that improves overall efficiency of a factory, boosting both machine utilization (capital) and worker productivity equally.
- A more advanced machine interface that allows each worker to operate twice as many machines at the same time, effectively making labor more efficient.

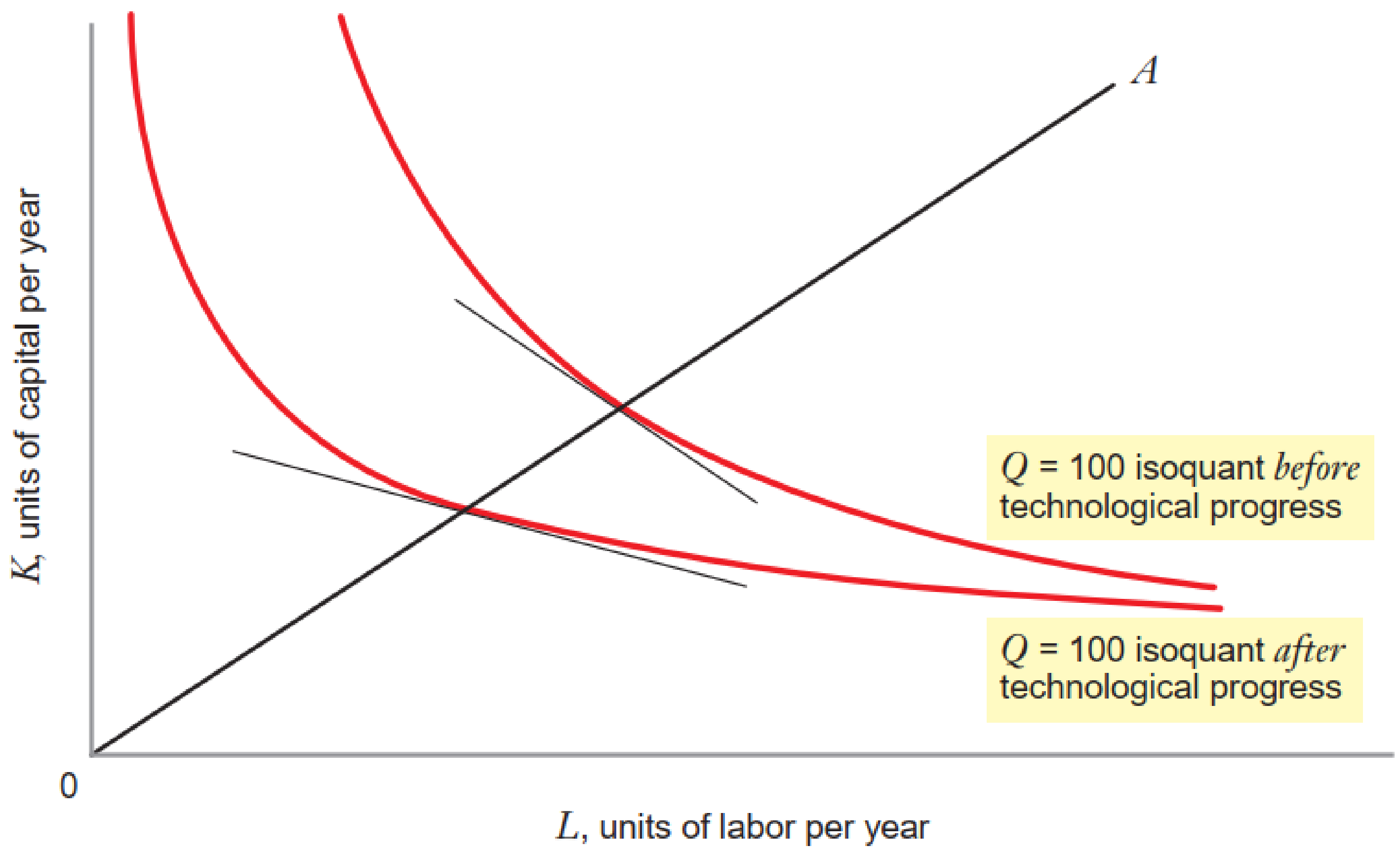
Labour Saving Technological Progress

- In this case, too, the isoquant corresponding to a given level of output shifts inward, but now along any ray from the origin, the isoquant becomes flatter, indicating that the $MRTS_{LK}$ is less than it was before.

- You should recall that $MRTS_{LK} = \frac{MP_L}{MP_K}$, so the fact that the $MRTS_{LK}$ decreases implies that under this form of technological progress the marginal product of capital increases more rapidly than the marginal product of labour.

- This form of technological progress arises when technical advances in capital equipment, robotics, or computers increase the marginal productivity of capital relative to the marginal productivity of labour.
- Figure B (next slide) represents labour saving technological progress.

Figure B



A
Q = 100 isoquant before technological progress
Q = 100 isoquant after technological progress
K, units of capital per year
0
L, units of labor per year

- **Why This Is Labor-Saving?**
- **Technical shift:** The new technology increases the **marginal productivity of capital** (robots) much more than that of labor.
- **Effect on production:** At the same level of output, **less labor is needed**; alternatively, for the same amount of labor, much more output can be produced.
- **Factor substitution:** Firms substitute capital for labor in response to the technology—this changes the capital–labor ratio in favor of capital.

- **Other common labor-saving technological progress examples:**
- **ATM machines in banking** – Reduced the need for bank tellers for cash withdrawal and deposit functions.
- **Self-checkout kiosks in supermarkets** – Fewer cashiers are required as customers process their own purchases.
- **Automated port container handling** – Cranes and automated guided vehicles move containers with minimal human intervention.

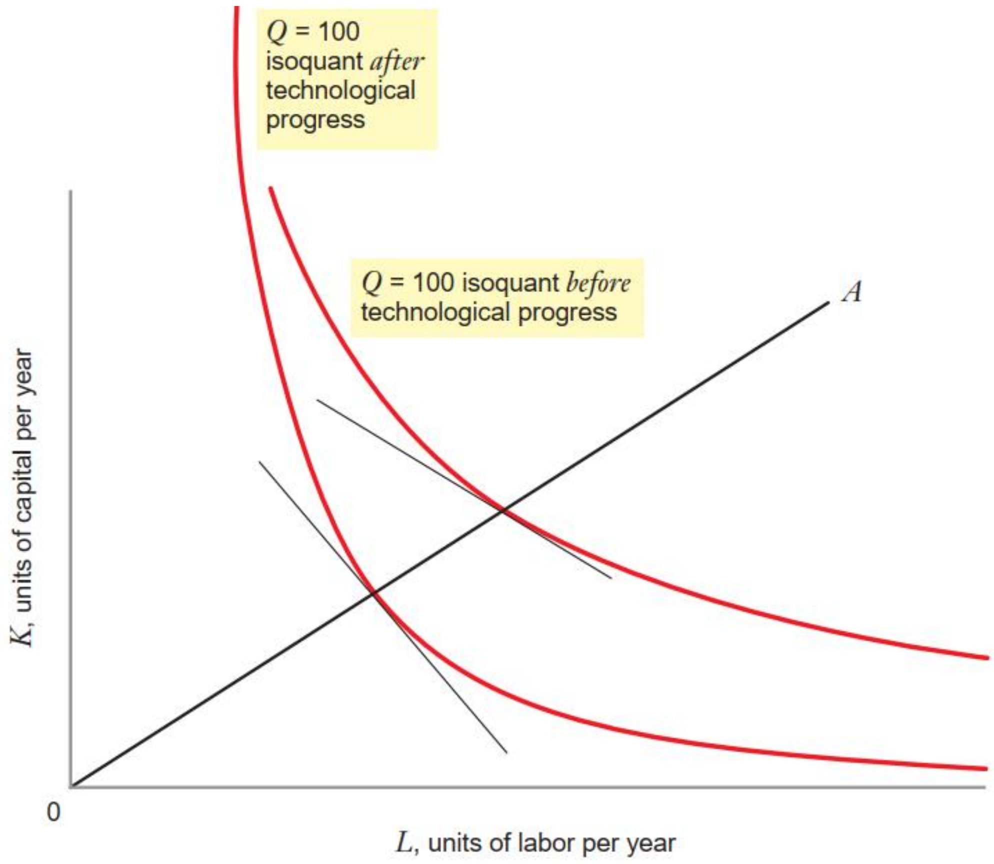
Capital saving technological Progress

- Technological progress that causes the marginal product of labour to increase relative to the marginal product of capital.

- Here, as an isoquant shifts inward, $MRTS_{LK}$ increases, indicating that the marginal product of labour increases more rapidly than the marginal product of capital.

- This form of technological progress arises if, for example, the educational or skill level of the firm's actual (and potential) work force rises, increasing the marginal productivity of labour relative to the marginal product of capital.

Figure C



Q = 100 isoquant after technological progress
Q = 100 isoquant before technological progress
A
K, units of capital per year
0
L, units of labor per year

- **Why This Is Capital-Saving?**

- The innovation increases the **marginal productivity of labor** relative to capital.

- The same output can be achieved with **less capital equipment** than before.

- The capital–labor ratio shifts in favor of **labor**, because technology substitutes labor for expensive capital machinery.

- **Real-World Examples**
- **Renewable energy (solar panels in rural areas)**: Earlier, electrification required heavy capital investments in grids and power plants. Now, solar panels (less capital-intensive) combined with local labor for installation/maintenance represent capital-saving progress.
- **Service-sector digital platforms**: Platforms like Uber, Zomato, or fintech reduce the need for large fleets of physical assets (capital) while relying more on organizational labor, and human interaction.

Returns to Scale in Electric Power Generation

- Returns to scale have been thoroughly studied in electric power generation, where the pioneering work was done by economist Marc Nerlove.
- (Ref: Marc Nerlove, "Returns to Scale in Electricity Supply," Chapter 7 in Carl F. Christ, ed., Measurement in Economics: Studies in Honour of Yebuda Grunfeld (Stanford, CA: Stanford University Press, 1963): 167-198.

- Summary: using the data from 145 electric utilities in the US during year 1955, Nerlove estimated the exponents of a Cobb-Douglas production Function and found that their sum was greater than 1. This implies that electricity generation is subject to increasing returns to scale.

Learning by Doing

- Returns to scale for a Cobb-Douglas Production function.
- Problem:

Does a Cobb-Douglas production function, $Q = AL^{\alpha}K^{\beta}$, exhibit increasing, decreasing, or constant returns to scale?

Solution: Let L_1 and K_1 denote the initial quantities of labor and capital and Q_1 denote the initial output, so $Q_1 = AL_1^{\alpha}K_1^{\beta}$. Now let's increase all input quantities by the same proportional amount λ, where $\lambda > 1$, and let Q_2 denote the resulting volume of output:

Learning by Doing

$$Q_2 = A(\lambda L_1)^{\alpha}(\lambda K_1)^{\beta} = \lambda^{\alpha+\beta} A L_1^{\alpha} K_1^{\beta} = \lambda^{\alpha+\beta} Q_1$$

From this, we can say that if:

$\alpha + \beta > 1$, then $\lambda^{\alpha+\beta} > \lambda$ and so $Q_2 > \lambda Q_1$ (increasing returns to scale)

$\alpha + \beta = 1$, then $\lambda^{\alpha+\beta} = \lambda$, and so $Q_2 = \lambda Q_1$ (constant returns to scale)

$\alpha + \beta < 1$, then $\lambda^{\alpha+\beta} < \lambda$, and so $Q_2 < \lambda Q_1$ (decreasing returns to scale)

This shows that the sum of the exponents $\alpha + \beta$ in the Cobb-Douglas production function determines whether returns to scale are increasing, constant or decreasing.