Production and Cost Minimization

Contents:

Theory: A firm's production depends upon (1) the amount and type of factors of production that it employs and (2) the firm's production technology.

(Throughout these notes we shall assume that a firm produces only one product.)

Assumptions often made about production: although not necessary, it is sometimes realistic to make the following assumptions concerning the behavior of a firm's production:

1) *Production increases with factor use*: Greater employment of factors of production, holding all else equal, leads to a higher level of production.

2*) Long run and short run*: We define the *short run* as a time period in which at least one of a firm's factors is used in a fixed amount (e.g. land for a farm); we define the *long run* as a time period long enough for a firm to vary its use of all of its factors.

3) *The Law of Diminishing Marginal Returns* (A short run concept): Define a factor's *marginal product* as the increase in a firm's output resulting from an incremental increase in the use of the factor (holding use of all other factors constant):

$$\text{marginal product} = \text{increase in total production} \div \text{increase in factor use}$$

The Law of Diminishing Returns states that marginal product eventually diminishes as more of a factor is employed. Diminishing marginal product means, for example, that employing the 3rd worker causes a smaller increase in production than employing the 2nd worker. (Note, however, that the 3rd worker does not reduce production; it causes total production to rise by less than the 2nd worker.)

Algebraic Representation of Production:

Q = f(factor use, technology)

Often to maintain tractability (and to be able to graph the firm's activity) we limit a firm's choice to this: the firm may employ only two types of factors; how many units of each should it employ?

<u>Production function Example 1</u>: The (long run) **Cobb-Douglas** production function represents the production possible from employing two factors, K (capital) and L (labor).

$Q = K^a L^b$

where
- K is units of capital employed
- L is units of labor employed
- Q is units of total product(ion)
- a and b are constants

<u>Example 1a</u>: The short run Cobb-Douglas production function usually has K set at a fixed amount (to represent the fixed factor), while L is allowed to vary. Below is a short run Cobb-Douglas production function with K fixed at 10 and a and b set to .5

$Q = 10^{.5} L^{.5}$

Total and marginal product: Let's use the above short run production function to construct a table illustrating how total product and marginal product vary as L changes from 0 to 5

Units of L	Units of K	Total product, Q	Marginal product (increase in Q when L rise by an additional unit)
0	10	0	--
1	10	3.162278	3.162278
2	10	4.472136	1.309858
3	10	5.477226	1.00509
4	10	6.324555	0.84733
5	10	7.071068	0.746512

How is Q calculated in the above table? Lets do an example from the third row, where L = 2 and K = 10. In this case:

$$Q = 10^{.5}2^{.5} = 4.472136$$

How is marginal product calculated in the above example? As an example, let's do the marginal product of the 3rd worker. Employing the third worker (L=3) causes total product to be 5.477226 units. With 2 workers (L=2) total product was only 4.472136 units. So employing the third worker allowed total product to rise by

$$5.477226-4.472136 = 1.00509 \text{ units}$$

Hence 1.00509 is the marginal product of the third worker.

Notice how the marginal product diminishes as more L is employed (keeping K fixed)—indicative of the law of diminishing marginal returns.

Production Function Example 2: The **Linear** production function is another (often not too realistic) representation of the production gained by employing factors K and L.

$$Q = aK + bL$$

where
- K is units of capital employed
- L is units of labor employed
- Q is units of total product(ion)
- a and b are constants

Here's a specific example of a linear production function:

$$Q = (10 K) + (3L)$$

Example 3: The **Leontief** production function is another (often not too realistic) representation of the production gained by employing factors K and L.

$$Q = \min[aK, bL]$$

where
- K is units of capital employed
- L is units of labor employed
- Q is units of total product(ion)
- a and b are constants

Here's a specific example of a Leontief production function:

Q = minimum of [1K, 6L]

Now that we've looked at production functions, let's examine the long run concept known as "returns to scale."

Returns to Scale (A Long Run Concept)
Unlike the short run, in which at least one factor is fixed, in the long run all factors can be used in varying amounts. Suppose one doubles use of all factors; will total product: *exactly* double, *more than* double, or *less than* double? The answer depends upon the specific production function representing the firm.

A production function has *constant* returns to scale if doubling use of all factors causes output to exactly double. (More technically: let f(K,L) be the firm's production function, and let z be a positive constant. The firm has constant returns to scale if zf(K,L) = f(zK,zL)).

A production function has *increasing* returns to scale if doubling use of all factors causes output to more than double. (More technically: the firm has increasing returns to scale if zf(K,L) < f(zK,zL)).

A production function has *decreasing* returns to scale if doubling use of all factors causes output to less than double. (More technically: the firm has increasing returns to scale if zf(K,L) > f(zK,zL)).

Example 1: For production function Q = 10K + 5L

Suppose the firm uses 1 unit of K and 1 unit of L. Production is:

Q = 10(1) + 5(1) = 15

Now let's double factor use to 2 units of K and 2 of L. Production is:

Q = 10(2) + 5(2) =30

Since Q exactly doubled when factor use doubles, this production function exhibits *constant* returns to scale.

Example 2: For production function $Q = KL^2$

Suppose the firm uses 1 unit of K and 1 unit of L. Production is:

$Q = 1 \times 1^2 = 1$

Now let's double factor use to 2 units of K and 2 of L. Production is:

$Q = 2 \times 2^2 = 8$

Since Q more than doubled when factor use doubles, this production function exhibits *increasing* returns to scale.

Now let's use some graphs to represent the total amount of units of outputthat a firm produces.

Total Production Graphed Using Isoquants

One can use isoquants to graph a firm's possible long run levels of total product

An *isoquant* represents all combinations of two factors K and L that result in an equal level of total product.

(A related concept is the *marginal rate of technical substitution (MRTS)*: the additional amount of K that the firm must employ, to compensate for a loss of some of factor L, in order to maintain total product at a constant level.)

Specific Example 1: Suppose a company has this linear production function:

$Q = 10K + 2L$

Let's list a few combinations of K and L that will produce 100 units of output:

K = 10 and L = 0 since 10(10) + 2(0) = 100

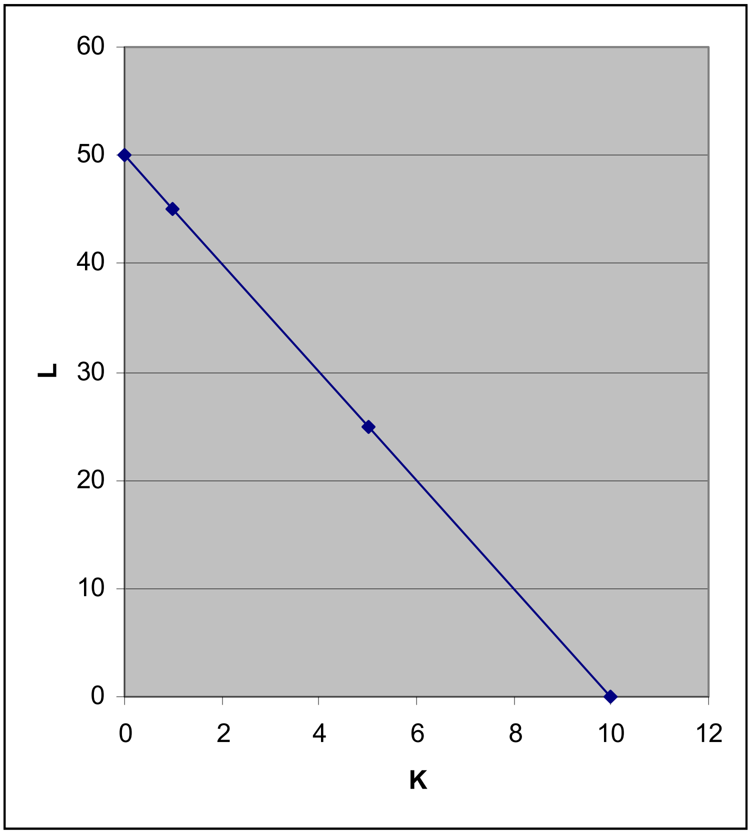
K= 0 and L = 50 since 10(0) + 2(50) = 100

K = 5 and L = 25 since 10(5) + 2(25) = 100

K = 1 and L = 45 since 10(1) + 2(45) = 100

etc.

If we graph these combinations of K and L and connect them we get the isoquant for Q = 100 for this firm:



Specific Example 2: Suppose a company has this Cobb Douglas production function:

$Q = K^{.5}L^{.5}$

Let's list a few combinations of K and L that will produce 10 units of output:

K = 10 and L = 10 since $10^{.5}10^{.5} = 10$

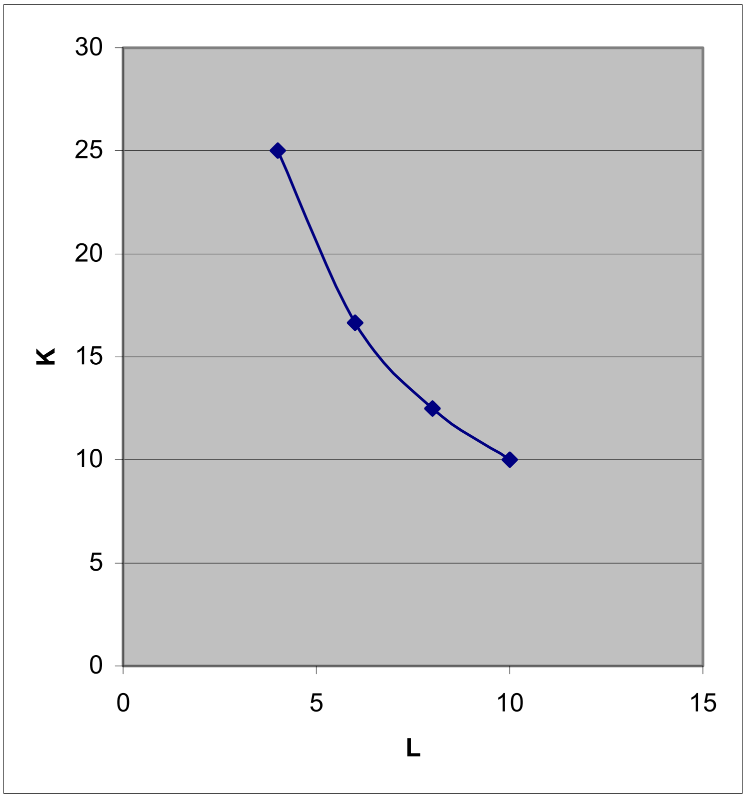
K = 12.5 and L = 8 since $12.5^{.5}8^{.5} = 10$

K = 6 and L = 16.6667 since $6^{.5}16.6667^{.5} = 10$

K = 4 and L = 25 since $4^{.5}25^{.5} = 10$

etc.

If we graph these combinations of K and L and connect then we get a portion of the isoquant for Q = 10 for this firm:



Specific Example 3: Suppose a company has this Leontief production function:

Q = min[50K, 100L]

Let's list a few combinations of K and L that will produce 1000 units of output:

K = 20 and L = 10 since min[50(20), 100(10)] = 1000

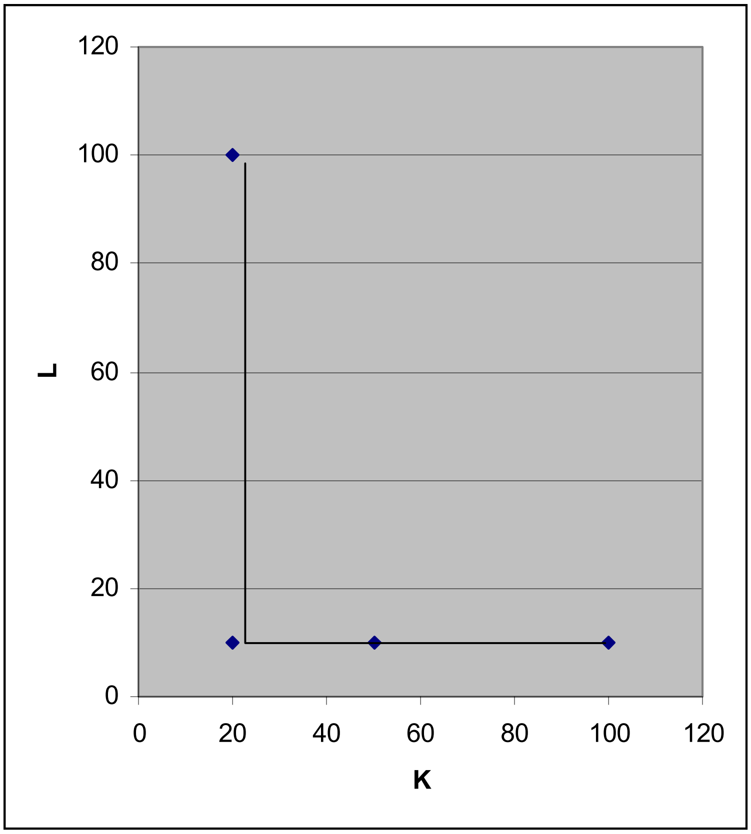
K = 20 and L = 100 since min[50(20), 100(100)] = 1000

K = 50 and L = 10 since min[50(50), 100(10)] = 1000

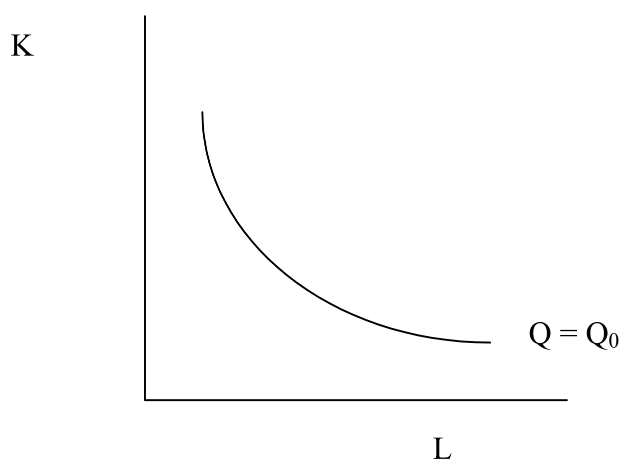
K = 100 and L = 10 since min[50(100), 100(10)] = 1000

etc.

If we graph these combinations of K and L and connect then we get a portion of the isoquant for Q = 10 for this firm:



General Example 1: In general, if one draws an isoquant for a Cobb-Douglas production function then it will look as this:

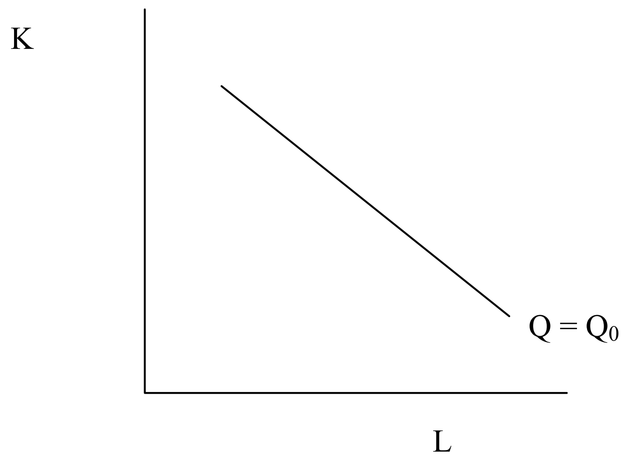


Note in the above graph:

--the negative slope: This indicates that the if one takes some L from the firm, then it must be given more K in order to maintain its total product at its initial level.

--the convex curve: This indicates diminishing MRTS

General Example 2: : In general, if one draws an isoquant for a linear production function then it will look as this:

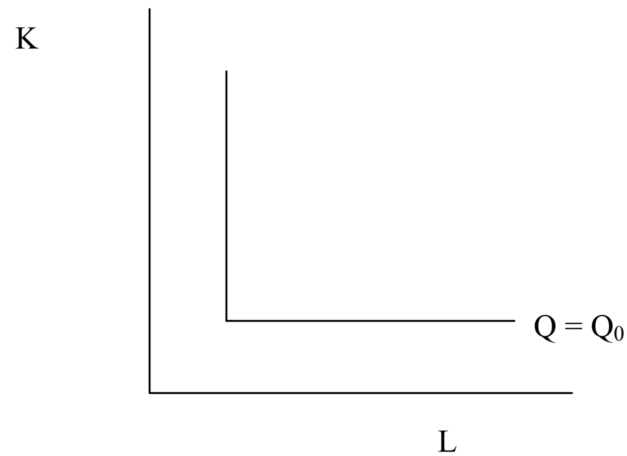


Note:

--the negative slope: This indicates that the if one takes some L from the firm, then it must be given more K in order to maintain its total product at its initial level.

--the lack of a curve: K and L are perfect substitutes in production.

General Example 3: In general, if one draws an isoquant for a Leontief production function then it will look as this:



Note:

--the goofy "L" shape of the isoquant: This indicates that the if one gives some extra L to the firm without giving it any more K, then its total product does not rise! Hence K and L are perfect complements.

Cost minimization and isoquants: For any production function, there are actually an infinite number of isoquants—one for each possible production level, from Q = 0 to Q = infinity. If a firm desires to produce a certain level of total product in the cheapest possible manner, then we say that the firm is using a combination of K and L that minimizes costs for a given isoquant.

The isoquant and the MRTS: Take any point on an isoquant. The MRTS is the (absolute value of) the slope of a line tangent to the isoquant at that point.

The MRTS and marginal product (or MP): The ratio of the marginal products of L and K equals the marginal rate of technical substitution:

$$MRTS = MP_L/MP_K$$

(Reading the above equation: "the marginal rate of technical substitution equals the marginal product of labor divided by the marginal product of capital.")

This makes sense, since the extra number of units of L that one must employ, after losing some K, in order to maintain a constant production level, depends upon the relative production gained/lost from an exchange of K for L

Example:

Suppose for a hypothetical firm the MP_L = 10 units of output and the MP_K = 20 units of output. This means that each unit of capital is currently twice as productive as each unit of labor.

Now take away a unit of K from the firm. How much extra L do they need to maintain their production? 2 units of L, since K is twice as productive as L. That is: when you takeaway 1 units of K, the firm loses 20 units of output; to replace that output requires 2 units of L, since each unit of L adds 10 units of output.

So the MRTS = 1K / 2L = .5. This is the same ratio as the MP_L/MP_K.

(Did you really follow the last bit of stuff? In you did not, then don't just skip to the next section! Go back over the last section again and again until you get it! If you don't get it after that then give me a call or come see me in my office!)

Constraints facing the firm

No firm can attain an infinitely high level of production. This is because each firm faces some barriers, or constraints, that limit its feasible choices. Generally the faces *time* constraints and *budget* constraints.

We shall often ignore (explicitly, anyway) the time constraint and consider only the budget constraint.

Budget constraint: example

Consider our firm that can employ two factors, L and K, at price w per unit of L and r per unit of K.[1] Suppose that this firm has dollar amount C to pay for K and L.

This firm's budget constraint can be written

$$rK + wL = C$$

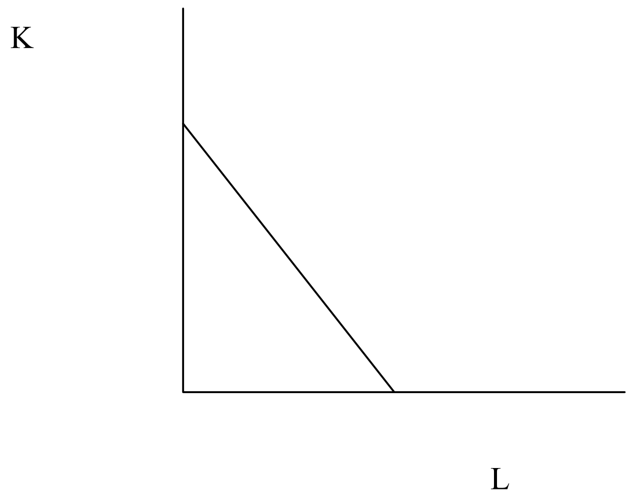
For ease of graphing, we can solve this equation for K:

$$K = C/r - (w/r)L$$

Hence a budget constraint is a negatively-sloped line with vertical intercept C/r and slope –(w/r):

(When graphed, the budget constraint is sometimes called an isocost line)

[1] "w" is short for "wage," and "r" is short for "rental cost of capital."



Another budget constraint example:

Suppose the firm has $100 to spend on K and L. L costs $2 each and K costs $4 each. The budget constraint is

$$100 = 4K + 2L$$

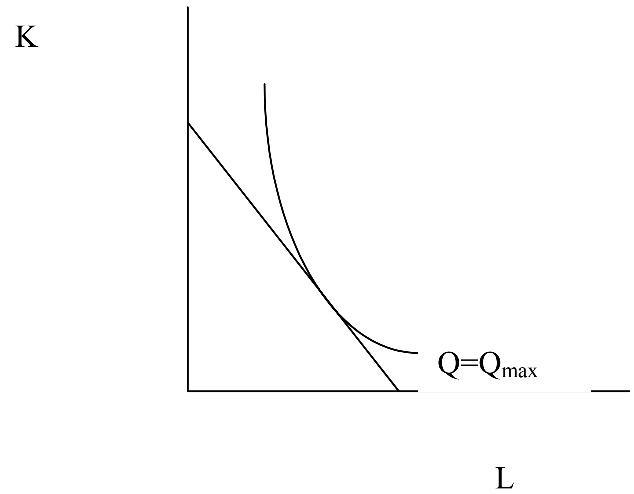
This is the equation of a line that intersects the K axis at K=25 and the L axis at L = 50. So the slope = 25/50 = 1/2. Note that this also equals the price ratio of labor to capital, 2/4 = ½.

Interpretation of the budget constraint. With its limited budget, the firm can only afford to employ combinations of K and L *on or inside* the isocost line. (By "inside the isocost line," I mean between it and the origin.) This is sometimes called the *feasible area*.

Cost Minimization Graphed

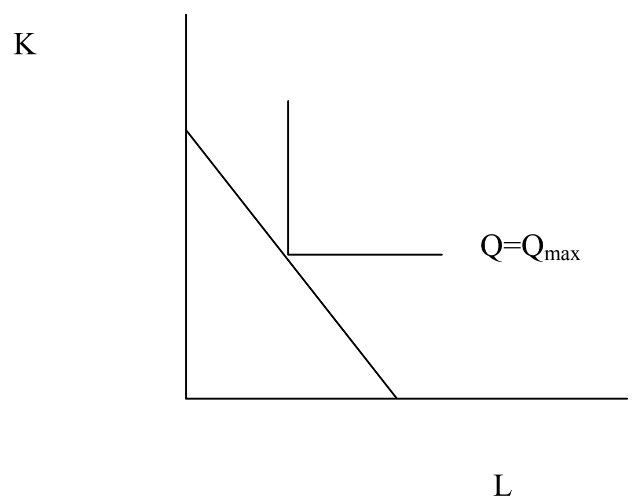
The highest isoquant that the firm can "get on" must have at least one of its points in the feasible area. Usually, in fact, it only has 1 point in the feasible area.

Example 1 (the most realistic): Cobb-Douglas

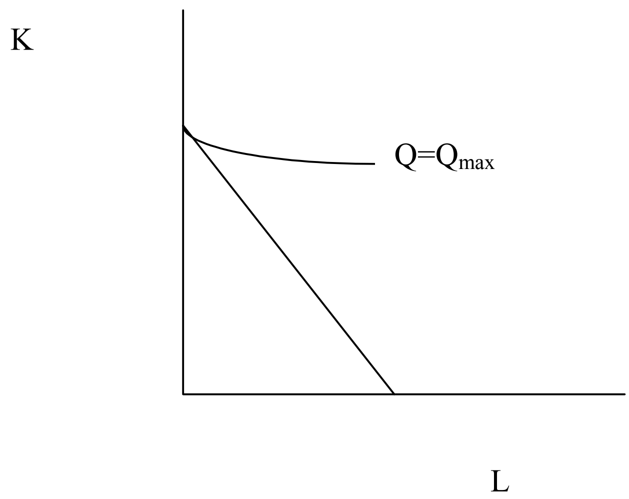


Note: The point at which the isoquant representing quantity Qmax is tangent to the budget constraint represents the least costly way to produce Qmax. This is the firm's cost-minimizing point.

Example 2: Leontief



Example 3: A "corner solution"



An algebraic cost-minimizing rule:

If one has a nicely-behaved production function, such as a Cobb-Douglas, then at the cost-minimizing point there will probably be a tangency between the isoquant and the budget constraint. In other words, cost-minimization requires

MRTS = slope of budget constraint

or

$MP_L/MP_K = w/r$

rewritten

$MP_L/w = MP_K/r$

In this last form, this can be interpreted as: the incremental production *per dollar* spent on labor equals the incremental production *per dollar* spent on capital.

The calculus of cost minimization: an example

Consider a firm with Cobb-Douglas production function

$Q = K^aL^{(1-a)}$ (Note that in this special case of a Cobb-Douglas production function, the exponents a + b = 1.)

and budget constraint

$C = rK + wL$

We can set up a constrained maximization problem to calculate how many units of X and Y that this person will buy.

maximize $Q = K^aL^{(1-a)}$

subject to: $C = rK + wL$

Solution of this problem, using the Lagrangean method, is undertaken in the textbook, so I will avoid reprinting the details here. (You do not need to be able to actually do this type of maximization problem, but you should understand and remember the results of it.) Suffice it to say that the results are:

Cost-minimizing Factor Demand equations:

$K = aC/r$ $L = (1-a)C/w$

(Note also a useful feature of this special Cobb-Douglas production function: the cost-minimizing firm spends fraction "a" of its budget on K, and fraction "1-a" of its budget on L.)

Specific example:

Suppose a firm has production function

$Q = K^{.8}L^{.2}$

K costs $5 each and L costs $6 each.

This firm's factor demand equations are:

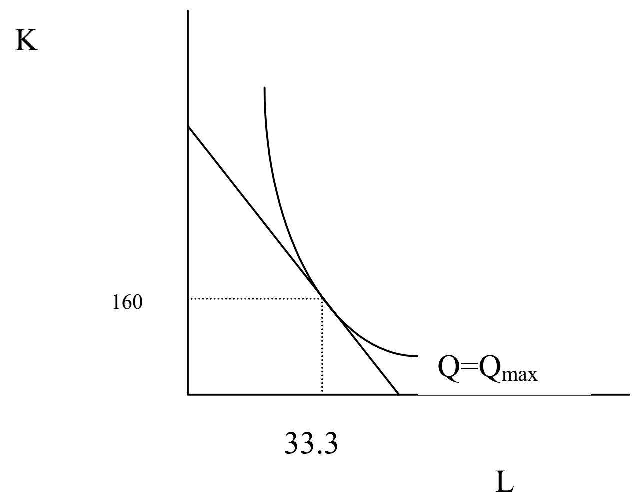
$K = .8C/5$ $L = .2C/6$

For example, if the firm has $1000 to spend on K and L, it would minimize costs by hiring

K = .8(1000)/5 = 160 units of K

L = .2(1000)/6 = 33 1/3 units of L

This would look like so on the graph (not drawn to precise scale):



Cost minimization vs. profit maximization: they ain't the same thing
The analysis of cost-minimization above gives a firm a rule to follow to minimize the cost of producing any level of output that it chooses to produce. But the rule cannot tell it the precise level of output that it should produce!!! This choice depends on factors external to the firm's production—specifically, it depends upon the demand for its product. In later class notes, we shall develop a profit-maximizing rule which will tell the firm the best level of output to produce.

Costs and scale economies
Suppose a cost-minimizing firm decides to double its production. Will its costs: exactly double, more than double, or less than double?

A firm has *economies of scale* if it can double its output with its costs less than doubling.

A firm has *diseconomies of scale* if when it doubles its output, its costs more than double.

A firm has *neither economies nor diseconomies of scale* if when it doubles its output, its costs exactly doubles.

Average costs and scale economies
Define *average* costs, a.k.a. cost per unit of output, as follows:

Average cost = total cost / Q

(Example: If a lemonade stand produces 300 glasses of lemonade at a cost of \$900, then its average cost is \$900/300 = \$3 per glass. And that had better be some danged good lemonade!)

If a firm has economies of scale, its average costs *fall* as its production rises.
If a firm has diseconomies of scale, its average costs *rise* as its production rises.
If a firm has neither economies nor diseconomies of scale, its average costs *do not change* as its production rises.