6 Costs

As we have mentioned before, we think of the firm as interested in **maximizing** its **profit**, that is, the difference between revenues and payments for inputs. **Revenues** from selling the firm's output depend on how much the firm decides to produce, of course, but not on **how** the firm has produced that output. That is, revenues are independent of the particular combination of inputs that the firm has chosen on the isoquant corresponding to that level of output.[5] Therefore, once the level of output is decided, the way that output is produced only affects what the firm has to pay for inputs. Maximizing profit then requires that the combination of inputs chosen to produce that output is the least expensive among all combinations in the corresponding isoquant. We may call the choice of inputs to that effect the problem of **cost minimization**.

Suppose that some level of output $q = 100$ may be produced in different ways, with more input K and less input L or vice versa, according to the following table:

L	K
20	2
15	3
10	5
5	8
1	12

That is, those five combinations are what we call **isoquant** corresponding to $q = 100$. If the firm decides to produce $q = 100$, which of these five methods should it use? Which one is the cheapest? The answer depends on the price of K, (let's represent it by r) and the price of L (let's represent it by w). If $r = 20$ and $w = 10$, then the cost of the five combinations of inputs is, respectively, \$240, \$210, \$200, \$210, and \$250. Then, the cheapest way of producing 100 units is by using 10 units of L and 5 units of K, and so the (minimum) **cost** of producing those $q = 100$ is \$200. However, if the prices of K and L were, respectively, $r = 20$ and $w = 15$, then the respective cost

[5] You may think that expenses like advertising may in fact be considered "costs" but do affect the revenues (if they allow to increase the price) even when output is kept constant. For now, we will ignore this issue.

of the five input combinations would be \$340, \$285, \$250, \$235, and \$255. The least expensive combination would be 5 units of L and 8 units of K, and so the (minimum) cost of producing those $q = 100$ would be \$235. That is, the firm would optimally **substitute** some K for the now relatively more expensive L.

Thus, minimizing the cost of producing q units amounts to finding *the combination of inputs (in this case (K, L)) that costs the least among those that satisfy $f(K, L) = q$.* At prices r and w respectively for K and L, the cost of a combination (K, L) is $w \times L + r \times K$. That is, in general, minimizing the cost of producing q units amounts to solving:

$$\begin{aligned} \min_{K,L} \quad & w \times L + r \times K \\ \text{s.t.} \quad & f(K, L) = q \end{aligned} \tag{11}$$

Take a minute to understand this expression, and why this is the representation of the problem of cost minimization for the firm. Also, make sure that you understand how to write this same problem in case there are more than two inputs.

For the case discussed above (i.e., when the production function is –partially– represented by the table –isoquant– above), and for $q = 100$, $r = 20$, and $w = 10$, we have obtained the solution to (11) as $K = 5$ and $L = 10$, and then the cost for the firm, $w \times L + r \times K = 200$. For a different value of the prices, $r = 20$, and $w = 15$, and still $q = 100$, we have obtained a different solution, $K = 8$ and $L = 5$, and then the cost for the firm, $w \times L + r \times K = 235$. If we had solved this problem for a different level of output q, of course we would have obtained (possibly) yet some different answers.

That is, the solution to (11) determines a **mapping** from (q, r, w) to combinations of inputs, K and L –and then to cost of the output–, much as the solution to problem (5) determined a mapping from prices and income to bundles of consumption.[6]

[6] However, note that this problem looks much more similar to the problem we solved when computing how much income was needed to compensate the consumer for an increase in the price of one good, when we computed the substitution effect.

6.1 Graphic analysis

As with the consumer's problem, we may first graphically illustrate the solution to problem (11). First, note that at the prices r and w, all the combinations of inputs that cost, say, \$10 satisfy

$$w \times L + r \times K = 10. \tag{12}$$

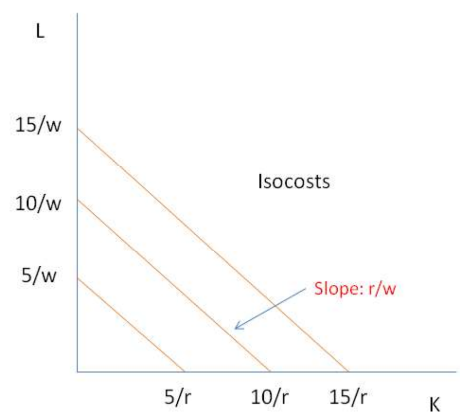


Figure 26

We can represent all those combinations and observe that they lie on a straight line. (Figure 26.) We may also call that line an **isocost** line, since indeed all its points (combinations of inputs) cost the same (\$10, in this case).

Note: althought they may look similar, this is not a budget constraint.

We can represent other isocosts. (In Figure 26, we have represented the isocosts \$5 and \$15, as well.) As you can see, isocosts closer to the origin represent lower spending in inputs. Computing the slope of all these isocost curves is straightforward: by solving for L the condition (12) that defines them:

$$L = \frac{10}{w} - \frac{r}{w}K,$$

we get that the slope (for any iscost curve) is the ratio of input prices, $\frac{r}{w}$.

Thus, solving problem (11) means finding the point (combination of inputs) that lies on the isocost closest to the origin among all that allow the firm to produce q units. That is, among all the combinations of inputs on isoquant q.

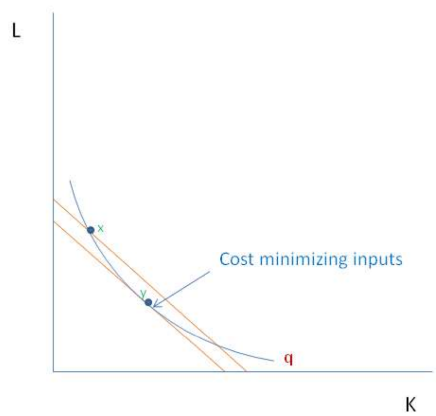


Figure 27

Combination x in Figure 27 lies on the q isoquant, and so allows the firm to produce q units, but there are cheaper combinations on the isoquant. The fact that the isocost where x lies *cuts* the q isoquant means that moving along the isoquant it is possible to reach below that isocost, that is, to input combinations on isocosts closer to the origin. In fact, combination y, where the isocost does not cut but is tangent to the isoquant, is the cheapest combination in that isoquant. That sounds familiar, doesn't it?

Indeed, in a fashion reminiscent of the consumer problem, **if** the solution to (11) contains positive amounts of both inputs (interior solution), then the slope of the constraint (the isoquant) equals the slope of the level curves of the objective function (the isocosts). That is, the MRTS equals the ratio of input prices.

Again, note that the isoquant (the "curvy" curve) is now the constraint (the restriction is that we need to produce q units) and the isocosts (the straight lines) represent the levels of spending (which correspond to the iso-*labels*, in the consumer's problem).

6.2 Analytical solution

Let us now analyze the solution to (11) by using the Lagrange method and looking at the first order conditions that define that solution. Similarly as in the discussion of the consumer problem, we may write the Lagrangian of that problem as

$$\mathcal{L}(K, L, \lambda) = w \times L + r \times K - \lambda \left(f(K, L) - q \right),$$

and so the first order conditions that characterize the (interior) cost minimizing combination of inputs for producing a level q of output are

$$\begin{aligned} r - \lambda \frac{\partial f(K,L)}{\partial K} &= 0 \\ w - \lambda \frac{\partial f(K,L)}{\partial L} &= 0 \\ f(K,L) - q &= 0. \end{aligned}$$

If we divide the first of these equations by the second as we often did in the past, we obtain

$$\frac{r}{w} = \frac{\frac{\partial f(K,L)}{\partial K}}{\frac{\partial f(K,L)}{\partial L}}. \tag{13}$$

The right hand side is what we called the **MRTS** (marginal product of L divided by the marginal product of K: how many units of K the firm may substitute per unit of L without reducing output). That is, the **slope of the isoquant** (at the corresponding combination of inputs). So, the cost minimizing combination of inputs (when the solution is interior) indeed satisfies that the *MRTS equals the ratio of input prices.*

The interpretation of this condition is similar to the interpretation of MRS=ratio of prices, in consumption. The ratio of input prices is the rate at which a firm *can substitute one input for the other in the market* for inputs: if the firm buys one less unit of L it saves w dollars that it can then use to acquire w/r more units of K. The MRTS is the rate at which the firm is able to substitute one input for the other and still produce output q: if the firm reduces the use of labor it will reduce output (per unit) by $\frac{\partial f(K,L)}{\partial L}$ and so it will need to increase the use of K (which per unit increases output by $\frac{\partial f(K,L)}{\partial K}$) by $\frac{\partial f(K,L)}{\partial L}/\frac{\partial f(K,L)}{\partial K}$ units.

If these rates are not equal, the firm can produce q with a less expensive input combination. For instance, if $\frac{w}{r} < \frac{\frac{\partial f(K,L)}{\partial L}}{\frac{\partial f(K,L)}{\partial K}}$, then $\frac{r}{w} > \frac{\frac{\partial f(K,L)}{\partial K}}{\frac{\partial f(K,L)}{\partial L}}$ and so the firm can obtain in the markets $\frac{r}{w}$ units of labor for each unit that it reduces the purchase of capital (without changing the spending in inputs), and that is more than what is needed to keep the output unchaged ($\frac{\frac{\partial f(K,L)}{\partial K}}{\frac{\partial f(K,L)}{\partial L}}$). So the firm does not even have to buy so much extra labor. That is, it may save money and still produce q. (Try to argue that there is a way of saving money also if $\frac{w}{r} > \frac{\frac{\partial f(K,L)}{\partial L}}{\frac{\partial f(K,L)}{\partial K}}$). Therefore, unless (13) holds, the firm is NOT minimizing costs when using both inputs.

6.3 Conditional input demands and the cost function

As in the consumer problem, and as we have mentioned above, if the circumstances for the firm change, its choice will also change. The three "circumstances" in problem (11) are the prices w and r, and the desired output q. The output is **not** a given for the firm: it is the firm that decides output. However, remember, we are only analyzing what combination of inputs is the least expensive **if the firm wants** to produce q units. Thus, for the problem of cost minimization, q is indeed a given. (You can imagine that central headquarters decide how much to produce, and then the firm's purchases and production divisions decide how to produce this output, which for them is a given.)

Then, and also as in the consumer problem, the solution to problem (11) defines a mapping from circumstances (w, r, and q) to input purchases, L and K. These mappings, i.e., the solution of the problem above for each triple (q, r, w), which we can write $L(q, r, w)$ and $K(q, r, w)$, are called **conditional input demands**.

Remember: conditional input demands represent how much the firm will buy of each of the inputs in order to produce q units in the least expensive way when the prices of the inputs are r and w. Thus, the lowest spending in inputs necessary to produce q units of output (when the prices of the inputs are r and w) is

$$C(q, r, w) = w \times L(q, r, w) + r \times K(q, r, w),$$

the dollars that the firm has to pay to acquire the combinations of inputs $L(q, r, w)$ and $K(q, r, w)$. This new mapping is what we call the **cost function**, the central concept in this discussion.

In the following pages we will be analyzing this function, and discussing several ways of obtaining and presenting information contained in it.

6.4 Marginal and average cost

From now on, we will focus on how the costs change with the level of output. (Remember, that choice, the choice of output, is one that we still need to –and will later– analyze.) Therefore, we will ignore the input prices, which we will keep unchanged, and, for compactness, we will write the (total) **cost function** as $C(q)$.

Remember the information that this function offers us: **if** the firm decides to produce q units, the least expensive way to do so requires spending $C(q)$ dollars in inputs. That is, $C(q)$ is the (economic, opportunity) cost of producing q units (given the technology as represented by $f(K, L)$ and the input prices r and w).

We may present this information in "per-unit" terms. **If** the firm decides to produce q units, the least expensive way of doing so requires a spending in inputs of $C(q)/q$ dollars per unit of output. Note that the two sentences that begin with a (bold) if are equivalent. They give exactly the same information: if we know $C(q)$ we get the per-unit cost by dividing this amount by q, and if we know $C(q)/q$ we obtain the (total) cost by multiplying this amount by q. We use the name **average cost** function, and write it $AC(q)$, for this per-unit way of presenting the information. That is, $AC(q) = C(q)/q$.

Sometimes we will be also interested in representing how the cost changes with the level of output. That is, how much (extra) it costs to produce an extra unit. (Or, more formally, what is the per-unit cost of increasing output slightly.)

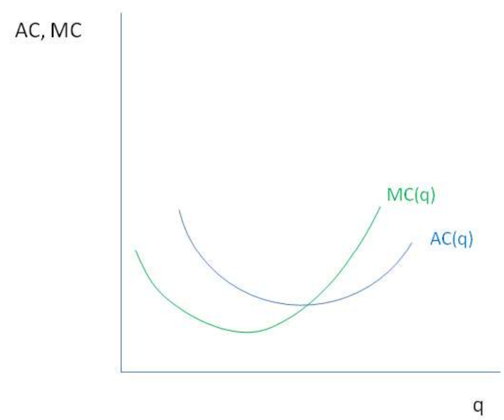
For any function, that is exactly what the derivative measures. So, that information is given by the derivative of the cost function with respect to output: $C'(q)$. We call that derivative the **marginal cost function**, $MC(q) \equiv C'(q)$.

Of course, both the AC and the MC are obtained from the cost function, $C(q)$, and so are simply obtained from the solution to our problem (11).

There is a very straightforward relationship between the marginal and the average cost of any firm. It is simply a consequence of the behavior of any average as we add elements. Consider what happens when we obtain, say, the average height of people sitting in a class and a new person enters the classroom. If the height of that newcomer is above the average, the average increases after the person joins. If the newcomer´s height is below the average of the class, then the latter decreases. Likewise, if the MC (the cost of the new unit) is above the AC at some output level q, then the AC

decreases at that output level, and vice versa. That is why you usually see the AC and the MC curves represented as in Figure 28.

28



$_2 9.pdf$

Figure 28

The only general point there is what we mentioned above: at q's such that $MC(q)$ is below $AC(q)$, the AC is decreasing, and at q's such that $MC(q) > AC(q)$, the AC is increasing. Also, and as a consequence, if the AC has a local minimum (a valley) as in the figure, that is, if there is a q such that the AC is decreasing for lower output levels and increasing for higher output levels, then the MC must switch sides exactly at q. That is, the MC must cross the AC at that minimum point. (What if the AC has a maximum?)

6.5 Returns to scale revisited

We have already defined what returns to scale are. Now, suppose that we have a production function, $f(K, L)$ characterized by increasing returns to scale. That, remember, implies that if we increase the amount of all inputs by some (the same) factor $\lambda > 1$, then the output increases by a factor of more than λ. Another way to put it is that in order to increase the output by a factor λ we could simply increase the amount of all inputs by a factor of **less than** λ!

That is, suppose that producing output 100 can be done (in the cheapest way) using 10 units of K and 20 units of L. That is the conditional demand of inputs (at the input prices and) for $q = 100$, are $L(100, w, r) = 20$ and $K(100, w, r) = 10$, and so $C(100, w, r) = 10 \times r + 20 \times w$. If the production function is characterized by increasing returns to scale, producing $q = \lambda \times 100$ can be done using **less** than $K = \lambda \times 10$ and $L = \lambda \times 20$. That is, **a** way to produce $q = \lambda \times 100$ costs less than

$$\lambda \times 10 \times r + \lambda \times 20 \times w = \lambda \times (10 \times r + 20 \times w) = \lambda \times C(100, w, r),$$

and so (the least expensive way cannot cost more than that and so) $C(\lambda \times 100, w, r) < \lambda \times C(100, w, r)$. That is, if we talk about the per unit cost (AC) and ignore the prices of inputs for compactness (as before), we have that

$$AC(\lambda \times 100) = \frac{C(\lambda \times 100)}{\lambda \times 100} < \frac{\lambda \times C(100)}{\lambda \times 100} = AC(100).$$

Of course, this has nothing to do with the particular q that we have assumed, or the particular conditional demands for that q. This is general. That is, if the production function is characterized by **increasing** returns to scale, then $AC(q)$ is **decreasing** (since $\lambda \times q > q$ for any $\lambda > 1$), and vice versa.

For the same token, if the production function is characterized by **decreasing** returns to scale, then $AC(q)$ is **increasing**, and if it is characterized by constant returns to scale, then $AC(q)$ is constant.

(You can observe that in Figure 28 we have represented the cost curves corresponding to a production function that is characterized by increasing returns to scale up to some output level –the level at which the AC attains a minimum– and by decreasing returns to scale for larger levels of output.)