Econ 100B
Intermediate Microeconomics B

Jordi Martinez-Muñoz

Note: These lecture notes will be continuously updated throughout the class.

Contents

1 Chapter 1. Technology

In this class, we will study how firms make production decisions, so it is natural to begin by understanding the production process. Given the vast number of industries worldwide, it would be impossible to study all of them. Instead, we will focus on the general concepts of a production process.

The two parts of a production process. Every production process has two elements: inputs and outputs. Inputs include all the elements necessary to produce a final good or service, such as labor (workers), capital (machines), intermediate goods, raw materials, and land. The output is the final good or service that firms sell in the market. The interaction of these inputs to produce an output is referred to as "technology" or the production process.

For example, Brenda sells cookies to her coworkers at the office. To produce these cookies, she uses ingredients (flour, sugar, butter) and mixes them in a bowl to make dough. She then rolls out the dough, cuts it using cookie cutters, places the cookies on a tray, and bakes them in the oven at 325°F for 12 minutes. The result is a batch of 15 cookies that she can sell at the office.

Studying the detailed production process of every single good produced in the market is impractical and beyond our scope; engineers are better equipped for this task, and even for them, it requires years to understand the intricacies of a particular industry. Our goal here is to simplify the study of the production process enough to derive general conclusions.

We will typically assume that firms use two inputs (labor and capital) to produce an output for the market. We will then explore the relationship between labor and capital in the production process. For example, how easily can workers be substituted with machines? What is the optimal ratio of labor to capital that maximizes profits (or minimizes costs), and how do these relationships depend on input prices and technology? Thus, we will shift our focus from the entire production process to the relationships between labor and capital as inputs of production.

1.1 The production function

We described the procedure that Brenda uses to make cookies, but for a more general approach, qualitative descriptions are not sufficient. We need to construct a mathematical framework that allows us to draw general conclusions, and it is important to understand both the insights and limitations of this quantitative framework.

Our quantitative primitive is going to be the production function. The production function, as its name suggests, is a function of the form $F(L, K)$ (sometimes denoted as $Q(L, K)$). The inputs of this function are precisely the inputs of production: labor L and capital K, and the output (denoted as q or y) is the maximum amount of final goods that can be produced given the quantities of inputs used.

For example, consider the production function $Q(L,K) = 2LK$. For this production function, we can obtain a maximum of 4 units ($q = 4$) by using two units of labor ($L = 2$) and one of capital ($K = 1$), or one unit of labor ($L = 1$) and two of capital ($K = 2$), among other possible combinations.

Technically, the firm could produce less than 4 units with the combination ($L = 2, K = 1$). However, we will assume that given the units of capital and labor available, the firm will produce as much as possible. If they wish to produce less, they should reduce some units of labor and capital. In summary, we will assume that firms do not waste resources, so the production function gives the minimum amount of inputs needed to produce a certain level of output. Do you think this is a realistic assumption for firms in real life? Why or why not?

For another production function, $Q(L,K) = 2K + L$, to produce a maximum of 4 units ($q = 4$) we would need either 2 units of capital ($K = 2$) and no labor ($L = 0$), or four units of labor and no capital ($L = 4, K = 0$), among other possible combinations.

It is important to always be aware of the units of L, K, and q. L might refer to labor hours or the number of workers. K might refer to the number of machines or the hours a machine is used. q may refer to units, hundreds of units, thousands of units, etc. We need to be careful about the units because it is not the same to hire 100 labor hours to produce 15,000 units of the final good as it is to hire 100 workers to produce 15,000 units.

Should we believe that there exists an industry where the production process is exactly $Q = 2LK$? Most likely, no. We know that other elements come into play: land, utilities, raw materials. We also know that the production process is far more complex than simply multiplying factors. Therefore, the production function does not fully represent how firms operate in the real world. Engineers study for many years to master these processes, while economists do not.

However, we can imagine industries or production processes in which the productive relation between L and K can be described by $Q = 2LK$ and differentiated from other more or less productive technologies, even if other aspects of production are left out. This simplified model is enough for economists to provide useful insights into firms' production and hiring decisions.

Now we will study important properties of production functions.

1.2 Marginal Productivity

When firms decide how many units of labor and capital to hire, it is important to understand the additional contribution of each unit of labor and capital to the production process. For example, the high additional contribution of the very first baker in a workshop with three baking machines is not the same as the additional contribution of the 10th worker, who will likely create more disruption than help given the limited number of baking machines. This leads us to the concept of marginal productivity:

Marginal Productivity of Labor (MP_L). The Marginal Productivity of Labor is the additional product generated by adding an extra unit of labor, keeping everything else (units of capital and any other inputs) constant.

Marginal Productivity of Capital (MP_K). The Marginal Productivity of Capital is the additional product generated by adding an extra unit of capital, keeping everything else (units of labor and any other inputs) constant.

A production table helps to visualize these concepts:

L	K	Output q
0	1	0
1	1	10
2	1	15
3	1	18
4	1	20
5	1	21

The first column of the table gives us the amount of labor, the second column gives the amount of capital that is fixed at one unit, and the third column gives us the output. When $L = 0$, the output is also 0. But when $L = 1$, $q = 10$, so the difference in output attributed to the first unit of labor (keeping capital fixed) is 10. Then, the marginal product of the first unit of labor is 10. For the second unit of labor, the output goes from 10 to 15, a change of +5 units. As we kept the capital fixed, we say that the marginal product of the second unit of labor is 5. Following the same reasoning, the marginal product of the third unit of labor is 3. For the fourth unit, it is 2, and for the fifth unit of labor, the marginal product is 1. So we can complete our table as follows:

L	K	Output q	Δq	Marginal product of L
0	1	0	-	-
1	1	10	10-0	10
2	1	15	15-10	5
3	1	18	18-15	3
4	1	20	20-18	2
5	1	21	21-20	1

This is the marginal product of labor when K is fixed at 1. If we had fixed K at a different number, the marginal productivity could have changed. The crucial part of the definition is to keep K fixed!

We can also calculate the marginal productivity even if we don't have one-by-one changes in the input. For example, in the following table:

L	K	Output q
0	1	0
4	1	16
10	1	28

When we go from $L = 0$ to $L = 4$ ($\Delta L = 4$), the output goes from 0 to 16 ($\Delta q = 16$). To compute the marginal product of labor (contribution of an additional unit of labor), we divide

$$\frac{\Delta q}{\Delta L} = \frac{16}{4} = 4$$

Thus, each additional unit of labor contributed four extra units of production.

Similarly, the marginal product of labor at $L = 10$ is

$$\frac{\Delta q}{\Delta L} = \frac{28 - 16}{10 - 4} = \frac{12}{6} = 2$$

How can we incorporate our notion of marginal product into the production function?

Remember that the marginal product refers to the change in production when we increase labor by one additional unit (keeping everything else fixed). But what if we want to increase labor by an infinitesimal amount? If we have a continuous production function, we can answer that question.

Imagine increasing the amount of labor by h units. To measure the change in $Q(L, K)$ when we move from L to $L+h$ units (that is $Q(L+h, K)-Q(L, K)$), to find the contribution of the next additional unit, we need to compute:

$$\frac{Q(L+h, K) - Q(L, K)}{h}$$

As h becomes infinitesimally small, the expression becomes:

$$MP_L = \lim_{h \to 0} \frac{Q(L+h, K) - Q(L, K)}{h}$$

and by the definition of a derivative, our expression is equivalent to:

$$MP_L = \frac{\partial Q(L, K)}{\partial L}$$

In other words, we define the marginal product of labor as the partial derivative of production with respect to L.

Similarly:

$$MP_K = \frac{\partial Q(L, K)}{\partial K}$$

Remember that the derivative of a function is a function itself, so the value of the marginal product of labor or capital will depend on the current units of labor and capital as explained above. Are you comfortable with this intuition?

Example:

Consider $Q(L, K) = AL^{0.5}K^{0.5}$

Here $MP_L = \frac{\partial Q(L,K)}{\partial L} = A * 0.5 * L^{0.5-1}K^{0.5} = 0.5AL^{-0.5}K^{0.5}$, which can also be represented as:

$MP_L = 0.5A\frac{K^{0.5}}{L^{0.5}}$

So here, the MP_L is a function that varies with L and K (decreases with L and increases with K). For example, when $K = 100$ and $L = 64$, $MP_L = 0.5A\frac{10}{8} = \frac{2.5A}{4} = 0.625A$. When $K = 25$ and $L = 25$, $MP_L = 0.5A$.

1.2.1 Two important properties of the marginal productivity

Since the marginal product gives us the contribution of an extra unit of labor (or capital) to the production process, it is natural to ask two questions:

For illustration, we will consider the marginal product of labor only, but the same analysis applies to the marginal product of capital.

1. Is the marginal product positive or negative?

A positive marginal product means that an additional unit will contribute to the production (the final output will increase).

A negative marginal product means that an additional unit will decrease the total output. Consider a kitchen already bustling with cooks: adding another cook might lead to supply conflicts or worsen mobility issues, hurting the total output. In general, a firm should never hire a unit of labor whose marginal product is negative.

2. Is the marginal product increasing or decreasing?

This property is more subtle; it answers the question of whether the next unit of labor contributes more or less than the previous one. Consider again the production table we saw earlier:

L	K	Output q	Δq	Marginal product of L
0	1	0	-	-
1	1	10	10-0	10
2	1	15	15-10	5
3	1	18	18-15	3
4	1	20	20-18	2
5	1	21	21-20	1

Notice that every unit of labor in this table contributed to the production, that is, each of them has a positive marginal product, but the first one contributed an additional 10 units, while the second one only an additional 5 units, the third only 3, the fourth only 2, and the last one only 1. In fact, each additional unit of labor added less additional product than the previous one.

In this case, we say that the marginal product of labor is positive but decreasing. Compare it with the following example:

L	K	Output q	Δq	Marginal product of L
0	1	0	-	-
1	1	10	10-0	10
2	1	25	25-10	15
3	1	43	43-25	18
4	1	63	63-43	20
5	1	85	85-63	22

In this case, not only is the marginal product of labor positive for each unit of labor, but it is also increasing (each additional unit of labor contributed more than the previous one to the final output).

It is important to distinguish between the sign of the marginal product (positive or negative) and the monotonicity of the marginal product (increasing or decreasing).

1.3 Average Productivity

While marginal productivity is sometimes challenging to compute in real life, a more practical measure of productivity is the average product.

The average product of labor (AP_L) is defined as the total output produced per unit of labor input. Mathematically, it is calculated as:

$$AP_L = \frac{Q(L,K)}{L}$$

The AP_L tells us how much output, on average, each unit of labor contributes to production.

Similarly, the average product of capital (AP_K) measures the total output produced per unit of capital input. It is calculated as:

$$AP_K = \frac{Q(L,K)}{K}$$

The AP_K tells us how much output, on average, each unit of capital contributes to production.

1.4 Different Ways to Produce the Same Amount of Output

Previously, we discussed that given a production function $Q(L, K)$, there may be different combinations of L and K that allow the firm to produce the same amount of output q. In fact, if $Q(L, K)$ is a continuous function, there are infinitely many combinations.

Firms are interested in knowing these combinations to pick the one that minimizes the cost of production. Imagine you committed to producing 1,000 cars. You would like to use the combination of L and K that minimizes the cost of producing those 1,000 cars given your technology.

Consider the production function $Q(L, K) = 2LK$. Let's fix the output at $\bar{q}$. That is, $\bar{q} = 2LK$. We want all the labor-capital combinations that can produce $\bar{q}$.

To do this, we isolate K from the expression $\bar{q} = 2LK$, that is:

$$K = \frac{\bar{Q}}{2L}$$

Then, for any amount of L, we know how many units of K we need to produce exactly $\bar{Q}$ units of output.

For example, if $\bar{q} = 10$, then when $L = 1$, $K = 5$; when $L = 1.5$, $K = 3.333$; when $L = 1.777$, $K = 2.814$; and so on.

We can graph this relation by putting K on the Y-axis and L on the X-axis. The resulting curve is known as the *Isoquant*, representing all possible combinations of L and K that return the same amount of output $\bar{q}$.

Isoquant. A curve that reflects all possible combinations of labor and capital that return the same amount of output.

The graph that includes isoquants for different levels of $\bar{q}$ is known as a map of isoquants. The isoquant for a higher level of production should always be drawn above the isoquant for a lower level of production. Isoquants for different levels of production can never intersect.

Another example: For the production function $Q = 2L^{0.5}K^{0.5}$, we can get the isoquant by fixing the output $\bar{q}$ and isolating K:

$$K = \left(\frac{\bar{Q}}{2L^{0.5}}\right)^{\frac{1}{0.5}} = \left(\frac{\bar{Q}}{2L^{0.5}}\right)^2$$

Note: Graphing isoquants can be challenging, as their shapes are often complex. A good sketch should reflect their monotonicity (first derivative), concavity (second derivative), and intersections with the X and Y axes.

1.5 The Marginal Rate of Technical Substitution

Imagine that a firm loses one worker but needs to produce the same output due to a previous commitment. It could rent more machines to compensate for the lost worker, but how much more capital should it rent to keep production constant? The answer is given by the *Marginal Rate of Technical Substitution* or MRTS.

If the lost labor had a marginal productivity of, say, 4 units, then production would decrease by 4. The firm wants to recover that loss. If the next unit of capital available has a marginal productivity of 2 units, the firm needs to hire $\frac{4}{2} = 2$ units of capital to recover (exactly, no more, no less) the lost production.

Alternatively, if the lost labor had a marginal productivity of 1 unit, and the next unit of capital available has a marginal productivity of 2 units, the firm needs to hire $\frac{1}{2} = 0.5$ units of capital to recover exactly the lost unit of production.

More generally, we define the marginal rate of technical substitution as follows:

Marginal Rate of Technical Substitution (MRTS). The MRTS is the amount of capital needed to replace a unit of labor while maintaining constant production levels.

Mathematically, we define MRTS as:

$$MRTS_{L,K}(L,K) = \frac{MP_L(L,K)}{MP_K(L,K)}$$

which can also be expressed as:

$$MRTS_{L,K}(L,K) = \frac{\frac{\partial Q(L,K)}{\partial L}}{\frac{\partial Q(L,K)}{\partial K}}$$

Example. Consider the technology $Q(L,K) = L^{0.3}K^{0.7}$.

We know that $MP_L = \frac{\partial Q(L,K)}{\partial L} = 0.3L^{-0.7}K^{0.7}$.

Also, $MP_K = 0.7L^{0.3}K^{-0.3}$.

Then $MRTS_{L,K} = \frac{0.3L^{-0.7}K^{0.7}}{0.7L^{0.3}K^{-0.3}} = \frac{0.3K^{0.7}K^{0.3}}{0.7L^{0.3}L^{0.7}} = \frac{0.43K}{L}$.

The MRTS is a function itself, which means that its value depends on the current values of L and K. It also means that it can be increasing or decreasing.

For example, for $Q(L, K) = L^{0.3}K^{0.7}$ with $MRTS_{L,K} = \frac{0.43K}{L}$:

- If $K = 1$ and $L = 10$, $MRTS_{L,K} = 0.043$. This makes sense because we already have a lot of labor units (probably the last unit had a very low marginal product), and since K is scarce, the next unit of capital probably has a high marginal productivity. Thus, to keep production constant when we lose one unit of labor, it is enough to use 0.043 extra units of capital.

- If $K = 10$ and $L = 2$, $MRTS_{L,K} = 2.15$. Here, we have a lot of capital, so the next units will not contribute much (low marginal product), and we have very little labor, so its marginal productivity is likely very high. Hence, we will need several extra units of capital to substitute for the lost labor unit.

One way to determine the monotonicity of the MRTS is by taking the partial derivative of $|MRTS|$ (the absolute value of MRTS) with respect to L and K and examining the sign. If $\frac{\partial |MRTS|}{\partial L} < 0$ and $\frac{\partial |MRTS|}{\partial K} > 0$ (one of them can be zero), then the MRTS is decreasing.

The MRTS is usually decreasing. What does that mean? It means that the amount of capital required to replace a unit of labor while keeping production constant decreases as the amount of labor increases.

Similarly to the previous example, the reasoning is that the more abundant labor is, the smaller the marginal product of the last unit. Therefore, it is easier to substitute it with capital.

The MRTS is also the slope of the Isoquant curve. How so?

Every point (L, K) on an isoquant returns the same level of output, so if we remove ΔL units of labor, the new level of capital needed to keep production constant, $K + \Delta K$, should satisfy that $(L - \Delta L, K + \Delta K)$ are also points on the same isoquant. If a line crosses those two points, its slope would be $\frac{\Delta K}{\Delta L}$. For a continuous production function and considering an infinitesimal change in L, ∂L, the slope of the isoquant will be $\frac{\partial K}{\partial L}$, which is equivalent to $\frac{\frac{\partial Q(L,K)}{\partial L}}{\frac{\partial Q(L,K)}{\partial K}}$, which is the MRTS.

This means that another way to determine if the Marginal Rate of Technical Substitution (MRTS) is decreasing is by computing the isoquant, evaluating $\frac{\partial K}{\partial L}$, and examining its sign. This method may be useful when assessing $\frac{\partial |MRTS|}{\partial L}$ and $\frac{\partial |MRTS|}{\partial K}$ does not provide a clear conclusion. For example, consider the production function $F(K, L) = L^3 + LK$, where

$\frac{\partial |MRTS|}{\partial L}$ is not always ≤ 0, yet the MRTS is decreasing.

Another utility of the MRTS as the slope of the isoquant is in graphing the isoquant to determine if its slope increases, decreases, or remains constant with changes in L and K, by observing how $|MRTS|$ changes with L and K.

1.6 Returns to Scale

What happens to production if we double the amounts of every input (labor and capital)? Will it exactly double, increase less than double, or more than double? The answer to this question is given by the concept of returns to scale.

Returns to scale are defined based on the following question: If we multiply every factor of production by some number $\lambda > 1$, will production increase by exactly λ, more than λ, or less than λ?

Consider the production function $Q(L,K) = L^{0.3}K^{0.7}$. Now let's multiply both L and K by some $\lambda > 0$, that is, instead of L and K, we will use λL and λK. Thus, $Q(\lambda L, \lambda K) = (\lambda L)^{0.3} (\lambda K)^{0.7}$. *Notice that it is very important to make sure that every operation that affects L and K also affects the λs.*

Then $Q(\lambda L, \lambda K) = (\lambda L)^{0.3} (\lambda K)^{0.7} = \lambda^{0.3}L^{0.3}\lambda^{0.7}K^{0.7} = \lambda^{0.3+0.7}L^{0.3}K^{0.7} = \lambda Q(L,K)$.

For this example, when we multiply each input by λ, the result is exactly λ times the original output.

Compare with this other case:

$Q(L,K) = L^{0.7}K^{0.7}$. Now let's multiply both L and K by some $\lambda > 0$, that is, instead of L and K, we will use λL and λK.

Then $Q(\lambda L, \lambda K) = (\lambda L)^{0.7} (\lambda K)^{0.7} = \lambda^{0.7}L^{0.7}\lambda^{0.7}K^{0.7} = \lambda^{0.7+0.7}L^{0.7}K^{0.7} = \lambda^{1.4}Q(L,K)$.

Because $\lambda^{1.4} > \lambda$, we conclude that when we multiply each input by λ, the result is greater than λ times the original output.

We define returns to scale as follows:

- If $Q(\lambda L, \lambda K) < \lambda Q(L,K)$ then Q has decreasing returns to scale.
- If $Q(\lambda L, \lambda K) = \lambda Q(L,K)$ then Q has constant returns to scale.
- If $Q(\lambda L, \lambda K) > \lambda Q(L,K)$ then Q has increasing returns to scale.

Therefore, $Q(L,K) = L^{0.3}K^{0.7}$ has constant returns to scale and $Q(L,K) = L^{0.7}K^{0.7}$ has increasing returns to scale. Show that $Q(L,K) = L^{0.2}K^{0.7}$ has decreasing returns to scale.

For some functions, getting the returns to scale involves a little more work. Consider the following example:

$$Q(L,K) = L^{0.5} + 3K$$

If we multiply L and K by $\lambda > 0$, the function becomes: $Q(\lambda L, \lambda K) = \lambda^{0.5}L^{0.5} + 3\lambda K$. This is smaller than $\lambda L^{0.5} + \lambda 3K$ as $\lambda^{0.5} < \lambda$. We conclude that the function exhibits decreasing returns to scale.

1.7 Homogeneous Functions

In our previous example, we did some extra analysis to determine the returns to scale of the function because it was not possible to express $Q(\lambda L, \lambda K)$ as $\lambda^n(L^{0.5} + 3K)$ for some $n > 0$.

For functions where $Q(\lambda L, \lambda K) = \lambda^n Q(L,K)$ holds true for some $n > 0$, we classify them as *Homogeneous of degree **n***.

This property of homogeneity is important because it determines whether scaling inputs proportionally leads to an exact proportional change in output.

When $n < 1$ the function exhibits decreasing returns to scale. When $n > 1$ the function exhibits increasing returns to scale, and when $n = 1$, we talk about constant returns to scale.

1.8 Important Production Functions

Some important production functions are:

Cobb-Douglas. $Q(L,K) = AL^{\alpha}K^{\beta}$

- Isoquants are convex and do not touch the axes.
- Are homogeneous of degree $\alpha + \beta$.

Perfect Substitutes. $Q(L,K) = aL + bK$

- Linear in L and K.
- Isoquants are straight lines.
- Elasticity of substitution is infinite.

- Homogeneous of degree one (constant returns to scale).

Fixed Proportions (Leontief). $Q(L, K) = \min\{aL, bK\}$

- Isoquants are "L" shaped.
- Elasticity of substitution is zero.
- Homogeneous of degree one (constant returns to scale).

Constant Elasticity of Substitution (CES). $Q(K, L) = A\left[\alpha K^{-\rho} + (1-\alpha)L^{-\rho}\right]^{-\frac{1}{\rho}}$

1.9 Practice Problems

1. For the following production functions, calculate:

I) Average product of labor and capital.
II) Marginal product of labor and capital.
III) Are MP_L and MP_K positive? Decreasing? (justify).
IV) Marginal rate of technical substitution. Is it decreasing? (justify).

(a) $F(L, K) = 3L^{0.4}K^{0.6}$
(b) $F(L, K) = \sqrt{6}L^{0.7}K^{0.7}$
(c) $F(L, K) = \ln(3L) + 5K$
(d) $F(L, K) = \sqrt{L} + \sqrt{K}$
(e) $F(L, K) = (\sqrt{L} + \sqrt{K})^2$
(f) $F(L, K) = L + 2K$
(g) $F(L, K) = (L + 2K)^2$
(h) $F(L, K) = \sqrt{L + 2K}$
(i) $F(L, K) = L\sqrt{K}$
(j) $F(L, K) = \ln(L) + 0.5\ln(K)$
(k) $F(L, K) = 12L + K^2$
(l) $F(L, K) = L\sqrt{K} + K$
(m) $F(L, K) = 3\sqrt{K} + L$
(n) $F(L, K) = \min(L, 3K)$
(o) $F(L, K) = (L^{1/3}K^{1/3})^3$
(p) $F(L, K) = 10\sqrt{LK}$
(q) $F(L, K) = 10(L + 3K)$
(r) $F(L, K) = \min(2L, K)$
(s) $F(L, K) = 2\min(L, 5K)$
(t) $F(L, K) = 2\sqrt{L} + \sqrt{K}$

2. For each production function of problem 1, calculate the isoquant for a fixed production level $\bar{q}$. Also, provide a sketch graph. It is enough to include axes intersections, monotonicity, and concavity.

3. For each production function of problem 1, determine the returns to scale (increasing, constant, decreasing). Also, determine whether they are homogeneous and if so, the degree of homogeneity.

4. What do technologies of f, g, and h, have in common related to the substitutability between labor and capital? How are they related to each other algebraically? What conclusion can you derive from this?

5. What happens to MP_L in technologies in c, d, e, j, and p as L goes to zero? Give an economic interpretation in words. Can you find functions where a similar effect happens to MP_K when K goes to zero?

6. Consider the production function AL^aK^b, where A, a, and b are positive constants.

i) For what values of A, a, and b are there diminishing marginal returns to labor L? What about diminishing marginal returns to capital?
ii) For what values of A, a, and b does the production process have a diminishing marginal rate of technical substitution?
iii) For what values of A, a, and b are there constant/decreasing/increasing returns to scale?

7. For each function of problem 1, how does MP_L change with K, and how does MP_K change with L (suggestion: compute the cross-partial derivative)? Interpret your results.

8. Use your answers to problem 1, items (a), (c), (d), (f), (j), (l), and (o) when $L = 250$, $K = 30$ to calculate the numeric values of MP_L, MP_K, and $MRTS$. Provide an interpretation using only non-technical words for each of these numbers.

9. Consider the following production table:

L	K	Output q
0	1	0
2	1	4
4	1	14
10	1	35
20	1	45
22	1	46
23	1	46
30	1	43

I) For which values of L is MP_L increasing?
II) For which values of L is MP_L decreasing?
III) For which values of L do we know for sure that the firm should not hire any more labor?

2 Chapter 2. Cost Minimization

Imagine you are the new CEO of a computer company. You discover that the company has previously committed to producing 1,000 computers.

What should you do? Naturally, you aim to produce the computers in the most cost-effective way possible. Assuming you only need to use labor and capital to manufacture the computers, the question becomes: which combination of labor and capital allows us to produce 1,000 computers at the lowest cost? Answering this type of question is the goal of this chapter.

To determine the cost-minimizing combination of labor and capital, we need to consider the two main elements of this optimization problem: the objective function and the constraint.

We aim to minimize the cost of production, which is our objective. This cost comprises the expenses for the labor used and the capital rented. Let's focus on labor first. Suppose the market price for a unit of labor is denoted by w; hence, hiring L units of labor costs $w \times L$ dollars. Similarly, if the cost of renting a unit of capital is r and the firm rents K units of capital, then the capital cost is $r \times K$ dollars.

Therefore, the total production cost can be represented as $TC = wL + rK$. We can also denote w as p_L, r as p_K, and the total cost as $TC = p_L L + p_K K$.

An obvious way to minimize production cost might be to choose $L = K = 0$, resulting in $TC = 0$. However, this is unfeasible because the firm is committed to producing 1,000 units. This commitment serves as our constraint. Whatever quantities of L and K we choose must satisfy $Q(L, K) = 1,000$, where $Q(L, K)$ represents the firm's production technology.

Thus, our problem can be formulated as choosing L and K to minimize the production cost $wL + rK$ subject to $Q(L, K) = 1,000$.

In general, if the firm aims to produce $\bar{Q}$ units in the most cost-effective manner, the problem is to choose L and K to minimize $TC = wL + rK$ subject to $Q(L, K) = \bar{Q}$.

One crucial detail we must not overlook is the timeline for production delivery—whether it needs to be completed in a week, a month, or two years. This detail is important because it affects our flexibility in adjusting inputs. While adjusting labor hours is relatively straightforward, the same is usually not true for capital. The firm may be bound by leasing contracts for specific machines or facilities, making it impossible to terminate them prematurely. Similarly, increasing the capacity of machinery may require considerable lead time. Therefore, if the computers must be delivered within a short timeframe, the firm should focus on adjusting the labor input while utilizing existing capital. If the deadline allows for a longer period, adjustments in both labor and capital inputs can be considered.

2.1 Short Run versus Long Run

In production theory, we distinguish between two time frames: the short run and the long run.

Short Run. In this period, at least one of the inputs of production is fixed and cannot be adjusted by the firm.

We typically denote the fixed input with a "bar" notation. For example, if the amount of capital that the firm can use is fixed, we denote it as $\bar{K}$, and the production function would be written as $Q(L, \bar{K})$. This indicates that only L (labor) can vary while capital remains fixed at $\bar{K}$.

Long Run. In contrast, every input of production is flexible, and the firm can freely choose how much to employ of each input.

In the long run, all binding capital contracts are completed, and the firm has the option to either increase or decrease the size of the production plant as it sees fit.

Due to this greater flexibility in the long run, production can occur at a lower cost. In other words, the cost of producing $\bar{Q}$ units in the long run will never exceed the cost of producing it in the short run.

How to Distinguish the Long Run from the Short Run?

The easiest way is to check if any factor is fixed such that it cannot be adjusted. For instance, consider a baker who signed a six-month contract to use a production facility equipped with ovens, refrigerators, and other heavy machinery. During this period, she cannot increase or decrease the amount of capital allocated in the contract.

Clarifying Note: The baker could choose not to utilize all the capital specified in the contract, but she is still obligated to pay for its full use. Therefore, when we refer to it as "fixed" in the short run, it means that she must pay for the entirety of the specified capital amount, regardless of whether she uses all of it (or even none of it). The fixed cost in the short run is usually referred to as a sunk cost because it cannot be avoided. We will talk more about this in Chapter 3.

2.2 Cost Minimization in the Short Run

Cost minimization in the short run is straightforward because the amount of capital is fixed. To determine the optimal amount of labor needed to produce q units with technology $F(L, K)$, a fixed capital stock $\bar{K}$, and given factor prices w and r, follow these steps:

1. Substitute $\bar{K}$ into the production function $F(L, K)$. The resulting expression should have L as the only variable.

2. Solve for L from $F(L, \bar{K}) = q$ and denote this solution as L^*. Notice L^* is a function of $\bar{K}$ and q, but not of w. Why?

3. Substitute L^* into the total cost function: $TC = wL + r\bar{K}$. Here, w, r, and $\bar{K}$ are given constants. This is a function of $q, \bar{K}, w$, and r.

Example: Suppose we want to produce 200 units using the production function $F(K, L) = 3L^{0.2}K^{0.5}$. The fixed capital is $\bar{K} = 400$ units, and the input prices are $w = \$2$ and $r = \$10$. Calculate the cost-minimizing amount of labor and the total production cost.

Solution:

1. Substitute $\bar{K}$ into the production function:
$$F(400, L) = 3L^{0.2} \cdot 400^{0.5} = 3L^{0.2} \cdot 20 = 60L^{0.2}.$$

2. Set $F(L, 400) = 200$:
$$60L^{0.2} = 200.$$
Solve for L:
$$L^{0.2} = \frac{200}{60} = 3.33 \Rightarrow L = 3.33^5 = 409.47.$$
Thus, $L^* = 409.47$.

3. Calculate the total cost:
$$TC = 2 \cdot 409.47 + 10 \cdot 400 = \$40,818.94.$$

2.2.1 Short Run Cost Curves

For production decisions, costs are very important. It is also important to distinguish between different types of costs.

In the short run, the total costs (STC) can be divided into fixed costs (FC) and variable costs (VC).

The fixed cost curve refers to the fixed input costs: those that are constant and must be paid regardless of the production level.

The variable cost curve refers to the costs that depend directly on the production level q. It can be expressed as $VC(q)$, so if $q = 0$, $VC(0) = 0$ and it increases with q.

How does this relate to our cost minimization problem?

We said that our objective function to minimize (in the short run) is $STC(q, w, r, \bar{K}) = wL + r\bar{K}$. In this case, $r\bar{K}$ is a number given by r and $\bar{K}$, which are constants. Thus, $FC = r\bar{K}$. What about the wL part? We know that L is indeed a function of $\bar{K}$ and q. So the correct form of that part of the cost is $wL(\bar{K}, q)$. This clearly depends on the production level q and hence is the variable part of the cost (VC).

There are other ways to represent the STC as a function of only q. Consider, for example:

$$STC(q) = 1000 + 0.5q^2 + 20q$$

For this function, the fixed cost is $FC = 1000$ and the variable cost is $VC = 0.5q^2 + 20q$ (because it depends on q).

In fact, if we take our original problem, plug in the numbers for w, r, and $\bar{K}$, leaving q as our only variable, we can study the relationship between STC and q as:

$$STC(q) = wL(q) + FC$$

Graphs for these functions will be shown in class.

Marginal and Average Costs

Since both STC and VC are functions of q, we can compute the marginal and average costs of production.

Average Cost. It is the cost per unit. It can be obtained by dividing the cost function by q.

- Average short-run total cost, $SAC = \frac{STC}{q}$.
- Average variable cost, $AVC = \frac{VC}{q}$.
- Average fixed cost, $AFC = \frac{FC}{q}$.

Marginal Cost. It is the additional cost incurred when production is increased by one unit. Another way to say it is the cost of increasing production by one unit.

- Marginal short-run total cost, $SMC = \frac{\partial STC}{\partial q}$.
- Marginal variable cost, $MVC = \frac{\partial VC}{\partial q}$.

Notice that we don't talk about marginal fixed cost because it is always zero, as fixed cost does not depend on q.

How do the average and marginal costs relate to each other?

Imagine we produced 10 units ($q = 10$) for a total cost $STC(10) = \$500$. Then the $SAC = \frac{500}{10} = \$50$. Suppose that we increase production by one unit to $q = 11$, and this extra unit has a marginal cost of \$50. In this example, $SAC(10) = SMC(11)$. The total cost now is $STC(11) = 500 + 50 = 550$, and the new average cost is $SAC(11) = \frac{550}{11} = 50$, meaning the SAC remained constant.

What would happen if $SMC(11)$ is smaller than \$50, say \$10? In this case, $STC(11)$ would have been smaller than 550, and therefore, $SAC = \frac{STC(11)}{11}$ would have been smaller than \$50, so the SAC would have decreased from $q = 10$ to $q = 11$. In particular, for this example, $SAC(11) = \frac{500+10}{11} = 46.36$.

What if instead $SMC(11)$ is greater than \$50, say \$70? In this case, $STC(11)$ would have been greater than 550, and therefore, $SAC = \frac{STC(11)}{11}$ would have been greater than \$50, so the SAC would have increased from $q = 10$ to $q = 11$. In particular, for this example, $SAC(11) = \frac{500+70}{11} = 51.82$.

These results are not unique to the numbers in the example. In general, it can be shown that if the marginal cost is greater than the average cost, then the average cost increases as production increases. Conversely, if the marginal cost is smaller than the average cost, then the average cost decreases as production increases.

If we think about an increasing marginal cost $\frac{\partial SMC(q)}{\partial q} > 0$, then we can conclude that $SAC(q)$ takes its minimum value when $SAC(q) = SMC(q)$. Let's see why.

Suppose $SMC = SAC$ at $q = q'$. If we assume that SMC is increasing, then $SMC(q) > SMC(q')$ for every $q > q'$. In other words, $SMC(q) > SAC(q)$ for every $q > q'$, and as we said earlier, this means that SAC can only go up when $q > q'$.

In a similar fashion, for every $q < q'$, $SMC(q) < SMC(q')$ (because SMC is increasing). Then, $SMC(q) < SAC(q)$, meaning that SAC can only go down when $q < q'$. But then no such q can minimize the SAC since we can always decrease it by getting closer to q'.

Conclusion 1: When the marginal cost is increasing, the SAC is minimized at the production level q where $SAC(q) = SMC(q)$.

The same argument can be applied to the relationship between the SMC and the AVC:

Conclusion 2: When the marginal cost is increasing, the AVC is minimized at the production level q where $AVC(q) = SMC(q)$.

Finally, notice that $AVC(q) < SAC(q)$, that is, the average variable cost is always smaller than the total average cost in the short run. This makes sense because the SAC also includes the average fixed cost (AFC) which, although decreasing with q, is never negative. Hence, $AVC(q) < SAC(q)$, but they get closer and closer as q increases.

When graphing cost curves, it is important to show all these characteristics and relations.

Example: Consider the cost function given by $STC(q) = q^2 - 0.5q + 20$. Let's compute the SMC, SAC, AVC, and AFC.

$$SMC(q) = \frac{\partial STC(q)}{\partial q} = 2q - 0.5$$

$$SAC(q) = \frac{STC(q)}{q} = \frac{q^2 - 0.5q + 20}{q} = q - 0.5 + \frac{20}{q}$$

$$AVC(q) = \frac{VC(q)}{q} = q - 0.5$$

$$AFC(q) = \frac{FC}{q} = \frac{20}{q}$$

Notice how AFC decreases as q increases and SMC increases with q.

If we make $SMC = SAC$, then

$$2q - 0.5 = q - 0.5 + \frac{20}{q}$$

Then $q^2 = 20$ and $q' = \sqrt{20}$.

At $q = \sqrt{20}$, the SAC takes its minimum value.

Can we corroborate that? We know that for convex functions, we can find the minimum by taking the first derivative and making it equal to zero. SAC is convex as $\frac{\partial^2 SAC}{\partial^2 q} = 40q^{-3} > 0$ so to find the minimum we make,

$$\frac{\partial SAC}{\partial q} = 0 \Rightarrow 1 - \frac{20}{q^2} = 0 \Rightarrow q = \sqrt{20}$$

,

As we could corroborate, both methods returned the same solution.

Similarly, for AVC,

$$2q - 0.5 = q - 0.5 \Rightarrow q = 0$$

So AVC is minimized at $q = 0$.

Note: In certain cases, the short-run marginal cost (SMC) and either the short-run average cost (SAC) or the average variable cost (AVC) may intersect at multiple points (typically at zero and at a positive quantity). When this occurs, we should consider the greater intersection point as the minimizer.

Increasing Short-Run Marginal Costs and Decreasing Marginal Productivity

We still need to address whether it makes sense to have increasing short-run marginal costs. This is related to the principle of decreasing marginal productivity. As firms increase production, they need to hire additional inputs that contribute less and less to the output. This is particularly true in the short run where capital is fixed, so each additional worker will have less capital to work with, resulting in a lower marginal product of labor.

As a result, to produce each additional unit of output, the firm will need to hire increasingly more units of labor. Given that the labor price is fixed at w, the cost of producing each additional unit of output will rise. This explains why the short-run marginal cost curve is typically upward sloping. As production expands, the decreasing marginal productivity of the variable input (labor) leads to increasing marginal costs.

2.3 Cost Minimization in the Long Run

The long run is more complicated because both K and L need to be adjusted to minimize the cost.

2.3.1 The Optimality Condition

When we hire a unit of labor, we increase production by exactly the marginal product of that unit, MP_L. At the same time, we have to pay \$w for that unit. Therefore, the ratio of the extra production we get to the extra cost we have to pay is $\frac{MP_L}{w}$.

Similarly, when we hire a unit of capital, we increase production by exactly the marginal product of that unit, MP_K. At the same time, we have to pay $r for that unit. Therefore, the ratio of the extra production we get to the extra cost we have to pay is $\frac{MP_K}{r}$.

What happens if $\frac{MP_L}{w} > \frac{MP_K}{r}$? It means that the extra production per dollar we get from hiring an extra unit of labor is greater than the extra production per dollar we get from hiring an extra unit of capital. Naturally, in that situation, we should substitute capital for labor (hire more labor and less capital), as labor is more cost-efficient.

As we increase the units of labor, the marginal product of labor MP_L will decrease (assuming diminishing marginal productivity), and as we decrease the units of capital, MP_K will increase. So as we increase L and decrease K, $\frac{MP_L}{w}$ decreases and $\frac{MP_K}{r}$ increases. They will eventually be equal to each other.

What happens if $\frac{MP_L}{w} < \frac{MP_K}{r}$? It means that the extra production per dollar we get from hiring an extra unit of labor is lower than the extra production per dollar we get from hiring an extra unit of capital. Naturally, in that situation, we should substitute labor for capital (hire more capital and less labor), as labor is less cost-efficient.

As we decrease the units of labor, the marginal product of labor MP_L will increase (assuming diminishing marginal productivity), and as we increase the units of capital, MP_K will decrease. So as we decrease L and increase K, $\frac{MP_L}{w}$ increases and $\frac{MP_K}{r}$ decreases. They will eventually be equal to each other.

Then what happens if $\frac{MP_L}{w} = \frac{MP_K}{r}$? At this point, we are indifferent between hiring one more unit of capital or one more unit of labor, as both give the exact same extra product per dollar spent. This gives us the optimal ratio between labor and capital, as we are not strictly better off replacing labor with capital or capital with labor.

Optimal Capital-Labor Ratio. The optimal capital-labor ratio is given by $\frac{MP_L}{w} = \frac{MP_K}{r}$.

Notice that we can rearrange the term above to $\frac{MP_L}{MP_K} = \frac{w}{r}$, which is the same as $MRTS = \frac{w}{r}$. Hence, the optimal capital-labor ratio can be found by making the MRTS equal to the price ratio.

We would like to have an expression of the form $K = f(L)$, i.e., "K equal to some function of L".

To get the optimal capital-labor ratio in the long run, follow these steps:

1. Make $\frac{MP_L}{w} = \frac{MP_K}{r}$ or $MRTS = \frac{MP_L}{MP_K} = \frac{w}{r}$.

2. From the last expression, isolate K (solve for K in terms of L).

Example: Consider the technology $F(L, K) = 2LK$.

1. We get $MP_L = 2K$ and $MP_K = 2L$. Then our relevant expression becomes: $\frac{2K}{2L} = \frac{w}{r} \Rightarrow \frac{K}{L} = \frac{w}{r}$.

2. Isolating K: $K = L\frac{w}{r}$.

Hence, the optimal capital-labor ratio is given by $K = L\frac{w}{r}$.

Note: For the above procedure to work, MP_L and MP_K must be diminishing. In addition, we need $MRTS$ to be diminishing as well. If these conditions do not hold, the procedure may not result in the optimal capital-labor ratio. Also, the optimality condition may not be met in the case of corner solutions (when the optimal allocation involves either L or K equal to zero) or when the isoquant has "kinks."

2.3.2 Optimal Levels of Capital and Labor

So far, we know the optimal capital-labor ratio, but we still need to figure out the values of L^* and K^*. This is, however, simple because we know the optimal units of K for every L, so we just need to plug this into our constraint.

If we want to produce q units with input prices w and r, we follow these steps:

1. Calculate the optimal capital-labor ratio, $K = f(L)$.

2. In the constraint $F(L, K) = q$, substitute K by the capital-labor ratio. That is, write $F(L, f(L)) = q$.

3. Solve for L (isolate L). This should give you L^*.

4. Plug L^* back into the capital-labor ratio to get $K^* = f(L^*)$.

5. To get the optimal long-run total cost, just plug L^* and K^* into the total cost function: $TC(L^*, K^*) = wL^* + rK^*$.

Example: For the technology $F(L, K) = 2KL$ with input prices w and r and quantity to produce q, we get:

1. $K = L\frac{w}{r}$.

2. The constraint is $2KL = q$, so substituting K from point 1, we get:

$$2\left(L\frac{w}{r}\right)L = q \Rightarrow \frac{2w}{r}L^2 = q.$$

3. Solving for L:

$$L = \left(q\frac{r}{2w}\right)^{\frac{1}{2}}.$$

So $L^* = \left(q\frac{r}{2w}\right)^{\frac{1}{2}}$.

4. Plugging L^* back into $K = L\frac{w}{r}$, we get:

$$K^* = \left(q\frac{r}{2w}\right)^{\frac{1}{2}}\frac{w}{r} = q^{\frac{1}{2}}\left(\frac{w}{2r}\right)^{\frac{1}{2}}.$$

So $K^* = q^{\frac{1}{2}}\left(\frac{w}{2r}\right)^{\frac{1}{2}}$.

5. The optimal total cost is given by:

$$TC(L^*, K^*) = wL^* + rK^* = wq^{\frac{1}{2}}\left(\frac{r}{2w}\right)^{\frac{1}{2}} + rq^{\frac{1}{2}}\left(\frac{w}{2r}\right)^{\frac{1}{2}} = q^{\frac{1}{2}}\left(w\left(\frac{r}{2w}\right)^{\frac{1}{2}} + r\left(\frac{w}{2r}\right)^{\frac{1}{2}}\right).$$

$$= q^{\frac{1}{2}}\left(\left(\frac{wr}{2}\right)^{\frac{1}{2}} + \left(\frac{wr}{2}\right)^{\frac{1}{2}}\right) = q^{\frac{1}{2}}2\left(\frac{wr}{2}\right)^{\frac{1}{2}} = (2wr)^{\frac{1}{2}}q^{\frac{1}{2}}.$$

Hence $TC(L^*, K^*) = (2wr)^{\frac{1}{2}}q$.

Example of a corner solution. $Q(L, K) = LK + 10L$ with $w = 10$, $r = 20$, $q = 100$.

$MRTS = \frac{K+10}{L}$, so the optimality condition implies $\frac{K+10}{L} = \frac{1}{2} \Rightarrow K = \frac{L}{2} - 10 \Rightarrow \frac{L^2}{2} - 10L + 10L = 100 \Rightarrow L = \sqrt{200} \Rightarrow K = \frac{\sqrt{200}}{2} - 10 < 0$.

K cannot be negative, so we set $K = 0$, and $L = \frac{100}{10} = 10$.

Verify that the optimality condition works without issues for $q = 450$.

2.3.3 Graphical Method

Another method to compute the cost-minimizing demands for L and K is through a graphical analysis. To do this, we introduce a new concept.

Isocost line. The isocost line represents all the combinations of labor and capital that entail the same cost.

For example, the isocost line associated with cost value c_0 is formed by all the combinations of L and K that satisfy $wL + rK = c_0$.

We can express the isocost line as $K = \frac{c-wL}{r} = \frac{c}{r} - \frac{w}{r}L$. Notice that, as with the isoquant, here we also isolate K and express it as a function of L. In the case of the isocost, K is a linear function of L with slope $\frac{w}{r}$, so the graph is a straight line with that slope.

The map of isocosts consists of graphing the isocost curve for every possible cost level. The greater the cost, the higher the isocost level is. In other words, the closer the isocost is to the origin, the lower the cost it represents.

Our goal is to find the lowest isocost line that allows the firm to satisfy the restriction (production level). In other words, if the firm is producing q, then we are looking for the lowest isocost that intersects the isoquant associated with q.

If the isoquant is convex (smooth), then the lowest isocost that intersects it does so at a single point. At this point, both slopes (the isoquant's and isocost's) are equal. Because the slope of the isocost is $\frac{w}{r}$ and the slope of the isoquant is $MRTS$, the condition implies that $MRTS = \frac{w}{r}$ (same as the optimality condition).

However, the graphical method works even when the isoquants are not smooth or when the tangency condition is not met, as long as we draw our curves correctly.

2.3.4 Special Production Functions

Perfect Substitutes. In the case $Q(L, K) = aL + bK$, the MRTS is not diminishing; hence the optimality condition does not work. In this case, the rule is:

$$\begin{cases} (L = \frac{q}{a}, K = 0) & \frac{a}{w} > \frac{b}{r} \\ (L \in [\frac{q}{a}, 0], K \in [\frac{q}{b}, 0]) & \frac{a}{w} = \frac{b}{r} \\ (L = 0, K = \frac{q}{b}) & \frac{a}{w} < \frac{b}{r} \end{cases}$$

The long-run total cost function is of the form: $TC(q) = \min\{w\frac{q}{a}, r\frac{q}{b}\}$.

Fixed Proportions. In this case, we need to make sure that both inputs satisfy the production requirements. That is, if $Q(L, K) = \min\{aL, bK\}$, then $L = \frac{q}{a}$ and $K = \frac{q}{b}$, regardless of the factor prices.

The long-run total cost function is of the form: $TC(q) = w\frac{q}{a} + r\frac{q}{b}$.

2.3.5 Comparative Statics for L^* and K^*

Now that we have found the optimal levels of labor and capital given the factor prices w and r, and the technological constraint, it may be interesting to understand how they depend on these factors. For example, how does the optimal level of labor L^* depend on the price of labor w, and how does it depend on the price of capital r?

This analysis is fairly simple; all we need to do is take the partial derivative of L^* or K^* with respect to w or r.

Example: Continuing with our previous example, we know that for $F(L, K) = 2KL$ with input prices w and r and quantity to produce q, the optimal labor and capital are given by:

$L^* = \left(q\frac{r}{w}\right)^{\frac{1}{2}}$ and

$K^* = q^{\frac{1}{2}}\left(\frac{w}{r}\right)^{\frac{1}{2}}$.

How does L^* vary with w?

Let's rewrite L^* as $(qr)^{\frac{1}{2}}\left(\frac{1}{w}\right)^{\frac{1}{2}} = (qr)^{\frac{1}{2}}\, w^{-\frac{1}{2}}$.

$$\frac{\partial L^*}{\partial w} = -\frac{1}{2}w^{-\frac{3}{2}}\,(qr)^{\frac{1}{2}} < 0 \quad \text{for all values of } w, r, q > 0.$$

That is, the greater the value of the labor cost w, the less labor is demanded.

Similarly,

$$\frac{\partial L^*}{\partial r} = \frac{1}{2}r^{-\frac{1}{2}}\left(q\frac{1}{w}\right)^{\frac{1}{2}} > 0 \quad \text{for all values of } w, r, q > 0.$$

That is, the greater the value of the capital cost r, the more labor is demanded.

This tells us that labor and capital have a degree of substitutability. Labor demand falls with its own price but rises with the price of capital.

What about $\frac{\partial L^*}{\partial q}$?

$$\frac{\partial L^*}{\partial q} = \frac{1}{2} q^{-\frac{1}{2}} \left(\frac{r}{w}\right)^{\frac{1}{2}} > 0 \quad \text{for all values of } w, r, q > 0.$$

This result is very intuitive: the greater the amount of product, the more labor we will hire.

Try to replicate this analysis for the demand for capital K^*.

2.4 Long Run vs Short Run Cost Curves

2.4.1 Output Expansion Path

The output expansion path tells how the optimal capital-labor ratio changes as we increase production levels. In the short run, the output expansion path is just a horizontal line at $K = \bar{K}$, as only L increases when we increase q. In the long run, however, both inputs may be adjusted, so the output expansion path may have a positive slope.

The optimal capital-labor ratio in the short run for a given production level q_0 will lie on a higher isocost than the optimal capital-labor ratio of the same production level in the long run. This is because the optimal capital-labor combination in the short run was also feasible in the long run (the opposite is not true), so the long-run combination is at most as expensive as the short-run optimal combination.

2.4.2 Relation Between Long Run and Short Run Cost Curves

In the long run, we have the total cost (TC), average cost (AC), and marginal cost (MC).

Total and average costs. Long-run total cost is always the lowest of all possible short-run costs. TC is the lower envelope of all STC curves. See the graphs in the slides.

The long-run average cost (AC) is the lower envelope of all short-run average cost (SAC) curves. See the graphs in the slides.

Marginal costs. Consider the function $Q(L, K) = \sqrt{LK}$.

In the short run $L = \frac{q^2}{\bar{K}} \Rightarrow STC(q) = \frac{wq^2}{\bar{K}} + \bar{K}r \Rightarrow SMC = \frac{2wq}{\bar{K}}$.

In the long run $MRTS = \frac{K}{L} = \frac{w}{r} \Rightarrow K = \frac{wL}{r} \Rightarrow \sqrt{L\frac{wL}{r}} = q \Rightarrow L = q\sqrt{\frac{r}{w}}$ and $K = q\sqrt{\frac{w}{r}} \Rightarrow TC(q) = 2q\sqrt{wr} \Rightarrow MC(q) = 2\sqrt{wr}$.

With $w = r = 10$ and $\bar{K} = 5$, then $SMC = 4q$ and $MC = 20$. Therefore, when $q < 5$, $SMC < MC$ and when $q > 5$, $SMC > MC$. So even though the long-run total cost is always smaller than the short-run total cost, the same is not true for the marginal costs.

The rest of this section is better explained with graphs (see slides).

2.5 Economies of Scale

What happens to the average cost as production increases in the long run?

The answer to that question depends on how much extra input the firm needs to increase output. For example, to increase the output by a factor of $\lambda > 1$, will the firm need to increase the inputs by a factor greater or smaller than λ? Does this look familiar?

We say that the firm experiences **economies of scale** when the average cost decreases as output q goes up. That is, the more the firm produces, the cheaper the cost per unit.

In contrast, we say that the firm experiences **diseconomies of scale** when the average cost increases as output q goes up. That is, the more the firm produces, the more expensive the cost per unit.

Economies of scale are related to increasing returns to scale: to increase output, we need to increase the use of every input in a less-than-proportional amount. For example, if output doubles, but labor and capital less than double, then the cost will be less than double, and so the average cost will also be less than double.

Diseconomies of scale are related to decreasing returns to scale: to increase output, we need to increase the use of every input in a more-than-proportional amount. For example, if output doubles, but labor and capital more than double, then the cost will be more than double, and so the average cost will also be more than double.

This is why returns to scale are so important: they affect the behavior of the long-run average cost curve.

Often, we observe economies of scale, followed by constant AC, and then diseconomies of scale.

Minimum efficient scale. It is the smallest quantity at which the long-run average cost curve attains its minimum point.

2.6 Technological Progress

One of the characteristics of technological progress is that it shifts the isoquants toward the origin (fewer inputs are needed to achieve the same output).

We also said that technological progress could be neutral, labor-saving, or capital-saving. Now we will characterize these three cases:

Neutral. Technological progress is neutral when the new capital-labor optimal ratio is the same. That is, less capital and labor are needed to meet the production level q, but they are still used in the same proportion: $K_1 < K_0$ and $L_1 < L_0$, with $\frac{K_0}{L_0} = \frac{K_1}{L_1}$.

Labor-saving. Technological progress is labor-saving when the shift of the isoquants is such that the new optimal capital-labor ratio is greater than before. That is, more capital is used relative to labor after the progress ($\frac{K_0}{L_0} < \frac{K_1}{L_1}$). This happens because the reduction in labor is greater than the reduction in capital, or even because more capital may be used after technological progress (imagine increasing capital slightly in exchange for a big reduction in labor).

Capital-saving. Technological progress is capital-saving when the shift of the isoquants is such that the new optimal capital-labor ratio is lower than before. That is, less capital is used relative to labor after the progress ($\frac{K_0}{L_0} > \frac{K_1}{L_1}$).

2.7 Opportunity Costs

Is it the same if the firm owns the capital as if it rents it? We know that when the firm rents the capital, it pays rK, where r is the rental rate of capital. But what if the firm already owns the capital? Is the cost zero?

If we think only of "accounting costs" (those costs that are explicit), then when the firm owns the capital, the only cost could be wL. However, that analysis would be incomplete because explicit costs are not the only ones we should consider.

The firm could rent the capital it owns to someone else instead of using it in its own production process. It could receive a rental rate of capital r' for each unit, so $r'K$ in total. When the firm decides to use its capital instead of renting it to someone else, it forgoes a potential income of $r'K$. This is not explicit, since they do not have to pay for it, but it is implicit in the sense that they could have received it had they not used their capital. We refer to this type of cost as an **opportunity cost**.

From the perspective of the firm, the **economic costs** of operation (both accounting and opportunity costs) are $TC = wL + r^{'}K$.

Opportunity cost. For any resources dedicated to an activity, the opportunity cost is the value of the next best activity those resources could have been used for.

Opportunity cost does not apply only to firms. The reader of these lecture notes is also paying an opportunity cost. What is the best thing you could be doing right now instead of reading these notes? How much do you value that alternative activity? That is your opportunity cost of reading these notes.

So, does it matter if the firm owns or rents the capital? We are going to assume that $r = r^{'}$, i.e., the price of renting capital from someone else is the same price you would get if you owned the capital and rented it out. Under this assumption, the costs faced by the firm are $TC = wL + rK$ regardless of whether it owns or rents the capital.

Example. A college student rents an office ($850 per month) and a computer ($250 per month) to provide high school tutoring. When thinking about the cost of her business, she estimates it at $1,100.

Is that estimate complete?

No. If she were not tutoring, she could be employed at a job and receive a salary instead. Suppose the best salary she could get is $900 per month. Then she faces an opportunity cost of $900 for dedicating her time to tutoring instead of working for someone else. The total costs she faces are therefore $2,000 ($1,100 accounting cost + $900 opportunity cost).

2.8 Practice Problems.

1. For each of the following production functions, assuming the cost of labor is w and the cost of capital is r:

I) Compute the short-run cost-minimizing demand for labor when K is fixed at $\bar{K}$ and output is equal to q. Hint: your answer should be a function of $\bar{K}$ and q.
II) Sketch an inverse graph L^* as a function of $\bar{K}$.
III) Compute the short-run total, average, and marginal cost functions.
IV) Compute the variable cost, average variable cost, and average fixed cost functions.
V) Sketch a graph with the (short-run total cost) STC, (short-run marginal cost) SMC, (short-run average cost) SAC, and (average variable cost) AVC.

VI) Sketch a graph with the STC, fixed cost, and variable cost.
VII) If $\bar{K} = 100$, $w = 50$, and $r = 10$, draw the graphs of parts V) and VI), labeling the intersections of the different cost curves.
VIII) Now assume $\bar{K} = 100$ and $q = 4000$. Which is the cheapest technology to produce in the short run, and which is the most expensive?

2. For the same production functions of problem 1, now assuming the firm is in the long run, with input prices w and r:

I) Compute the demand for labor and capital as functions of w, r, and q.
II) Compute the total, average, and marginal cost functions.
III) Sketch the total, average, and marginal cost functions. Label the intersections.
IV) For $w = 50$, $r = 10$, draw graphs comparing STC vs TC, SAC vs AC, and SMC vs MC.
V) Now assume $q = 4000$. Which is the cheapest technology to produce in the long run, and which is the most expensive? How did your answer change with respect to the short-run exercise?

3. For the following short-run cost functions, find and graph: AFC, AVC, SAC, and SMC. Find the minimum of the SAC.

(a) $STC(q) = q^2 + 2q + 9$
(b) $STC(q) = 50q + 9$
(c) $STC(q) = 5q^2 - 3q + 11$

4. For each function of problem 1, check if it exhibits economies or diseconomies of scale by (I) looking at its returns to scale and (II) looking at how AC changes when you move from producing q to λq for some $\lambda > 1$.

5. For the function $Q(L, K) = L^{0.4}K^{0.6}$, with $q = 100$ and $w = r = 10$, check if the following technologies represent technological progress and of which type:

- $Q_1(L, K) = (2L)^{0.4}K^{0.6}$
- $Q_2(L, K) = (0.5L)^{0.4}K^{0.6}$
- $Q_3(L, K) = L^{0.4}(3K)^{0.6}$
- $Q_4(L, K) = 1.5L^{0.4}K^{0.6}$

6. A man inherited a piece of land. He plans to build a house and estimated its cost as the sum of construction permit fees plus the cost of building the house (workers, materials, machines). Is this analysis complete? Why or why not?

3 Chapter 3, Profit Maximization in Perfect Competition

3.1 The profit maximization problem

In the last chapter, we assumed that the firm wanted to produce q units of output and our goal was to find the combination of L and K that minimized the cost of producing q.

In this chapter, we will go further and consider how firms decide the level of q. For this, we need to define the goal of the firm. We assume that the goal of the firm is to maximize its profits ($\pi(q)$), defined as revenue ($Rev(q)$)—what the firm gets from selling its products in the market—minus costs $TC(q)$—what the firm pays to produce. Is this a reasonable assumption? There may be reasons why a firm may behave as if it had goals other than the maximization of profits, but here we will assume profit maximization is the only goal of the firm.

Then, we can define the firm's optimization problem as:

$$\max_q \pi(q) = Rev(q) - TC(q)$$

This is an unconstrained problem (there are no constraints), and it is well-defined as long as $\pi(q)$ is concave in q (so a maximum exists).

If $\pi(q)$ is concave, we know from calculus that we can solve the maximization problem by setting $\pi^{'}(q) = 0$ (set the derivative equal to zero).

Also, $\frac{d\pi(q)}{dq} = \frac{dRev(q)}{dq} - \frac{dTC(q)}{dq}$ because the derivative of a sum is equal to the sum of the derivatives.

We are already familiar with $\frac{dTC(q)}{dq}$ from the previous chapter. It is the marginal cost!

We interpret $\frac{dRev(q)}{dq}$ as the change in total revenue when the production q increases by one unit, or in other words, the marginal revenue.

3.1.1 Revenue

Revenue is what the firm gets from selling its products in the market. If the firm produces q units of output and sells them at a price $p(q)$, then the revenue is defined as $Rev(q) = q*p(q)$.

In a previous econ class, you may have studied that the price at which a good is sold depends on the quantity of that good. We want to capture that idea by allowing p to depend on q, as $p(q)$.

Then the marginal revenue can be defined mathematically as $\frac{dRev(q)}{dq} = p(q) + q * \frac{dp(q)}{dq}$. Notice that we applied the product rule of the derivative because $Rev(q)$ is the product of two functions of q: q itself and $p(q)$.

Then the profit maximization problem of the firm can be solved by the first-order condition:

$$p(q) + q * \frac{dp(q)}{dq} - \frac{dTC(q)}{dq} = 0$$

and solving for q from there.

The second-order condition (checks if the problem is well-defined) is that $\pi(q)$ is concave. That is, $\frac{d^2\pi(q)}{dq^2} < 0$ or, equivalently, $\frac{d^2Rev(q)}{dq^2} - \frac{d^2TC(q)}{dq^2} < 0$.

3.2 Perfectly Competitive Markets

We focus on a particular structure of markets: perfect competition.

In a perfectly competitive market, the following four conditions hold:

1. All firms sell the same standardized product, and consumers perceive the product as identical (no differentiation).
2. The market has many buyers and sellers, each very small relative to the total (no player is too big to make a difference).
3. Technology and inputs are equally accessible by all firms (no technological advantages).
4. Buyers and sellers are perfectly informed about quality, prices, etc. (no information advantages).

Notice that all of these characteristics must hold for a market to be perfectly competitive. Is this realistic?

Probably none (or very few) markets would satisfy all four properties. However, it is important to study these structures first as a theoretical benchmark, as it is in perfectly competitive markets where social welfare is maximized (more on this in later chapters).

These characteristics have the following three implications:

- **Sellers and buyers act as price takers.**

Because each buyer and seller is too small, none of them can individually influence the price—by producing too few or too many units. The individual production decision does not affect the market price.

This means mathematically that

$$\frac{dp(q)}{dq} = 0$$

- **Law of one price.**

Sellers can't influence the price by differentiating from other buyers (products are identical) or by implementing more efficient production procedures (technology is the same for all sellers). The transactions between buyers and sellers occur at a single market price.

For practical purposes, it means we can use just p without worries.

- **There is free entry.**

If it is profitable for new firms to enter the industry, they will eventually do so. In the long-run equilibrium, profits will be zero (or very close to zero) because if there are positive profits, new firms will enter and drive the price down, reducing the profits to zero.

$$\pi(q) = Rev(q) - TC(q) = 0 \Rightarrow Rev(q) = TC(q)$$

3.2.1 The long-run zero profit condition

It probably sounds very strange that firms earn zero profit in perfect competition in the long run. Who would want to sell in such a market?

It is important to remember from the last chapter that when we talk about total cost $TC(q)$, we include both explicit and opportunity costs. If an entrepreneur opens a business and invests their own money or their own work, the opportunity cost of such investment is included in the total cost (either the market value of investing the money in the best alternative or the salary they would earn by working elsewhere).

When we talk about positive profits $Rev(q) - TC(q) > 0$, we mean that even after the firm pays its employees, the rent of the capital, and the opportunity cost of the firm owners, there is still extra money. We refer to that extra as economic profit. In perfect competition, that extra is zero in the long run.

3.2.2 Profit maximization in perfect competition

Because perfect competition implies that

$$\frac{dp(q)}{dq} = 0$$

the profit maximization first-order condition:

$$p(q) + q * \frac{dp(q)}{dq} - \frac{dTC(q)}{dq} = 0$$

becomes:

$$p - \frac{dTC(q)}{dq} = 0 \Rightarrow p = MC(q)$$

.

The second-order condition becomes:

$$-\frac{d^2TC(q)}{dq^2} < 0 \Rightarrow \frac{d^2TC(q)}{dq^2} > 0$$

That is, in perfectly competitive markets, to find the profit-maximizing quantity, all we need to do is set the price equal to marginal cost and solve for q.

Also, to check if the problem is well-defined, all we need to check is that the total cost function is convex.

Steps for profit maximization in perfectly competitive markets

1. Compute the marginal cost $MC(q) = \frac{dTC(q)}{dq}$.
2. Set $p = MC(q)$.
3. Solve for q.

To check if a maximum exists (the problem is well-defined):

- Check if $\frac{d^2TC(q)}{dq^2} > 0$.

Example 1. Given $TC(q) = 3q^2 + 10q$, find the supply curve of the firm in perfect competition.

First, we find the q that maximizes profits, given some price p:

- $MC(q) = TC'(q) = 6q + 10$

- $MC(q) = p \Rightarrow 6q + 10 = p \Rightarrow q = \frac{p-10}{6}$

Since q can't be negative, the firm supply is given by:

$$\begin{cases} \frac{p-10}{6} & p > 10 \\ 0 & p \leq 10 \end{cases}$$

Try to graph this firm supply curve (q on the x-axis and price on the y-axis).

Then we know how much will be produced for each possible price. For instance:

- If $p = \$4$, $q = 0$. $TC(q) = 0$, $Rev(q) = 0$.
- If $p = \$34$, $q = 4$. $TC(4) = 3 * 16 + 10 * 4 = 88$. $Rev(4) = 34 * 4 = 136$. $\pi(4) = 136 - 88 = \$48 > 0$.

Example 2. Given $TC(q) = 4q^2 + 10q + 100$, find the supply curve of the firm in perfect competition.

First, we find the q that maximizes profits, given some price p:

- $MC(q) = TC'(q) = 8q + 10$
- $MC(q) = p \Rightarrow 8q + 10 = p \Rightarrow q = \frac{p-10}{8}$

Since q can't be negative, the firm supply is given by:

$$\begin{cases} \frac{p-10}{8} & p > 10 \\ 0 & p \leq 10 \end{cases}$$

Now let's see what happens for different values of p:

- If $p = \$30$, $q = 2.5$. $TC(2.5) = 4 * (2.5)^2 + 10 * (2.5) + 100 = 150$. $Rev(2.5) = 30 * 2.5 = 75$. $\pi(2.5) = 75 - 150 = -\$75 < 0$.
- If $p = \$80$, $q = 8.75$. $TC(8.75) = 4 * (8.75)^2 + 10 * (8.75) + 100 \sim 493.7$. $Rev(8.75) = 700$. $\pi(8.75) \sim 700 - 493.7 = \$206.3 > 0$.

Since the maximum profits at $p = 30$ are negative (-75), should the firm shut down (exit the market at that price)?

We will answer that question in the next section.

3.3 Shutdown Rule

3.3.1 Sunk versus Avoidable Fixed Cost

In the previous chapter, we said that the fixed cost was only present in the short run. Here, we are going to allow the fixed cost to be also present in the long run, but we will distinguish between these two situations.

We mentioned in the previous chapter that the fixed cost was "sunk," that is, it has to be paid regardless of the production decision. Even if $q = 0$, the sunk cost must still be paid because it is an already incurred cost. We will refer to the fixed cost in the short run as the **fixed sunk cost**.

The fixed cost that a firm may have in the long run is different: it is fixed in the sense that it does not depend on q as long as $q > 0$, but if $q = 0$, then $FC(0) = 0$. That is, the fixed cost is paid only when the firm is producing. If the firm does not produce, the fixed cost is completely avoided. For this reason, we refer to it as an **fixed avoidable cost**.

For example, a baker signs a year lease for a facility and machinery. She committed to paying the rent for this place every month, and this holds even if she closes her business and produces nothing. To be able to sell her bakery goods, the city government demands a fixed selling permit of \$200 per month. The permit is only paid if she sells her product. If she shuts down, no payment is needed.

Here, both the rent of the facility and the selling permit are fixed costs, as neither depends on q. However, the selling permit only applies when $q > 0$ (can be avoided when no production happens), but the rent has to be paid regardless (it is a sunk cost).

Note: The sunk cost, more generally, refers to any cost that has already been expended and cannot be recovered regardless of the current choices. Sunk costs should not determine our production/consumption choices. For example, if you started a series on Netflix, and after three episodes you realize you don't like it, should you stop watching it? Some people may say: "Well, since I already spent three hours watching it, I will keep doing it." That analysis is fallacious since those three hours are already gone. You will not have them back if you keep watching the show. Instead, you will be spending more hours on something you don't like for no good reason.

The sunk cost fallacy (making decisions based on a sunk cost) applies to many areas of decision-making. Can you think of an occasion in which you fell for the sunk cost fallacy?

3.3.2 Shutdown rule: Short run

We may be tempted to think that if a firm's profits are negative, it should shut down. After all, the revenue is not enough to cover the costs of production.

We must consider, however, that in the short run, even if $q = 0$ (shutdown), the firm still must pay the fixed cost (because it is sunk). So $\pi(0) = -FC < 0$.

Can the firm do better by producing a positive amount? That depends on whether the firm can cover the totality of its variable cost.

Consider a firm that produces $q = 100$, gets revenues of \$1000, its variable cost is \$700 and its fixed cost is \$500 (for a total cost of \$1200). Clearly, profits are negative ($\pi(100) = -200$), but the full variable cost of \$700 is covered by the revenue. Not only that, because there is a surplus of \$300, they could cover part of the fixed cost.

If the firm had produced $q = 0$, then they would get 0 revenue and 0 variable cost, but $\pi(0) = -FC = -500$ which is worse than -200.

Hence, if the revenue covers at least the variable cost, the firm should operate.

What is the minimum price for which a firm should operate in the short run? The key to this question is in the averages!

We just defined the following rule:

Operate if: $Rev(q) - VC(q) > 0 \Rightarrow Rev(q) > VC(q) \Rightarrow pq > VC(q)$.

By the definition of the average variable cost, $AVC(q) = \frac{VC(q)}{q}$, which rearranging means that $VC(q) = AVC(q) * q$.

Then we can re-express our rule as:

Operate if: $pq > AVC(q) * q$.

Dividing both sides of the inequality by q, we get:

Operate if: $p > AVC(q)$.

So for any given p, if the firm can find a production level q for which $p > AVC(q)$, then the firm should operate, and shut down otherwise.

Now let's think about $minAVC(q)$, which is the minimum value that the AVC can have and its associated production q_{min}.

If $p > minAVC(q)$, then the firm could produce $q = q_{min}$ to get $AVC(q) = minAVC(q)$.

That would satisfy our rule. Is q_{min} the optimal production level? We don't know that. We would need to find the optimal q^* by following the steps at the beginning of this chapter, but we know that for whatever the optimal q is, the firm will be as good as with q_{min} and so $p > AVC(q^*)$.

If $p < minAVC(q)$, then for no level of q is $p > AVC(q)$ true. Since it is not true for the lowest value of $AVC(q)$, it can't be true for any other value.

If $p = minAVC(q)$, then the rule $p > AVC(q)$ is only satisfied for $q = q_{min}$.

Hence we can write the shutdown rule as follows:

Shutdown rule in the short run. *In the short run, the firm will operate as long as $p \geq minAVC(q)$, and shut down otherwise.*

It's also worth noting that, even though the short-run shutdown rule depends on $Rev(q) > VC(q)$ instead of $\pi(q) > 0$, we can still determine if the profits are negative or positive by looking at an average, this time the $SAC(q)$, the short-run total average cost. Because $SAC(q) = \frac{STC(q)}{q}$, then $SAC(q) = STC(q)/q$. Then $\pi(q) = Rev(q) - STC(q) = p * q - STC(q)$. Therefore, $\pi(q) > 0 \Rightarrow p * q > STC(q) \Rightarrow p > STC(q)$.

Example. For the previous example, $TC(q) = 4q^2 + 10q + 100$, with sunk $FC = 100$, $AVC = 4q + 10$, which is minimized at $q = 0$. As $AVC(0) = 0$, the firm will produce at any price $p \geq 0$. As $SAC = 4q + 10 + \frac{100}{q}$ with minimum at $q = 5$ and $SAC(5) = 50$, we get that $\pi > 0$ only for $p \geq 50$.

3.3.3 Shutdown rule: Long run

In the long run, all the costs are avoidable, so there is no reason to keep producing if $\pi(q) < 0$.

For this reason, the rule for the long run is:

Operate if: $\pi(q) > 0 \Rightarrow Rev(q) > TC(q)$.

By doing the same algebraic rearrangements we did in the short-run case, we can express the rule as:

Operate if: $p > AC(q)$.

Following the same type of arguments as in the short-run section, the firm will produce if $p \geq minAC(q)$ and shut down otherwise.

Shutdown rule in the long run. *In the long run, the firm will operate as long as $p \geq minAC(q)$, and shut down otherwise.*

Example. Assume the same cost function as before, $TC(q) = 4q^2 + 10q + 100$, but now $TC = 100$ is avoidable. Then we need $minAC$ with $AC(q) = 4q + 10 + \frac{100}{q}$. One way to find the minimum is to set $AC(q) = MC(q)$ so $4q + 10 + \frac{100}{q} = 8q + 10 \Rightarrow q_{min} = 5$. Plugging it back into AC, $AC(5) = 50$. In the long run, produce if $p \geq 50$.

It is important to notice two details:

- The shutdown rule in the short run depends on the $minAVC$ (minimum of the average variable cost), while in the long run, it depends on the $minAC$ (minimum of the average total cost).
- To get the shutdown rule threshold p, we will need to compute either $minAVC$ or $minAC$. We learned how to do that in the previous chapter. Here is a reminder:
    1. Set $AVC = MC$ if short run, or set $AC = MC$ if long run, and solve for q.
    2. Insert the q back into AVC or AC to get the minimum value.
    3. Alternatively, find q by setting $\frac{dAVC}{dq} = 0$ or $\frac{dAC}{dq} = 0$.

3.4 Producer Surplus versus Profit

Producer Surplus (PS). We define the producer surplus as the difference between the revenue and the avoidable cost (variable cost in the short run or variable+fixed cost in the long run).

The $PS(q)$ gives us the net gain from producing q units.

In contrast, profits $\pi(q)$ are the difference between the revenue and the total cost. They give us the net gain from being in the market.

Notice that $PS(q) = \pi(q)$ in the long run, as all costs are avoidable. In the short run, $\pi(q) = PS(q) - \text{Sunk Cost}$.

PS should be non-negative, $PS \geq 0$, for every firm in the market, but π may be negative in the short run.

Another way to express the shutdown rule is:

A firm should operate in the market as long as $PS(q^*) \geq 0$.

Note: If you took an Intro to Econ class in the past, you probably know PS as the area between the market price and the supply curve. That definition is equivalent to the one presented in this chapter.

3.5 Market Supply and Market Equilibrium in the Short Run

When several firms produce in a market, we can determine the aggregated supply of the market as the summation of the individual firms' supply. Let's express the supply of an individual firm i as $q_i(p)$. Then we can define the market supply $Q(p) = \sum_i q_i(p)$.

Because in perfect competition all the firms' supply functions are identical (they all have the same technology), if there are N firms in the market, the aggregate supply is $Q^S(p) = Nq(p)$.

We already have a market supply. Now we must bring in the consumers (market demand) to compute the market equilibrium price and quantity. If the market demand is given by $Q^D(p)$, then we can find the market equilibrium price by making $Nq(p) = Q^D(p)$.

To compute the aggregated supply in the short run, we need to follow these steps:

Given the number of firms N, the market demand function $Q^D(p)$, and cost function $TC(q)$,

1. Compute the firm supply curve $q(p)$ by making $MC(q) = p$. Remember that q can't be negative.
2. Calculate p_{exit}, that is the shutdown price by finding $minAVC$.
3. Compute the aggregated supply $Q^S(p) = Nq(p)$.
4. Find the market price equilibrium by setting $Q^S = Q^D$ and solving for p_{eq}.
5. Is $p_{eq} > p_{exit}$?
    - If $p_{eq} > p_{exit}$, then plug p_{eq} into $q(p)$ to get the individual firm supply. Verify that $Nq(p_{eq}) = Q^D(p_{eq})$.
    - If $p_{eq} < p_{exit}$, then $q = Nq = Q^D = 0$.
    - If $p_{eq} = p_{exit}$, then q will be either $q(p_{eq})$ or 0, and the market supply will be any number between $\{0, q(p_{eq}), 2q(p_{eq}), \ldots, Nq(p_{eq})\}$.

Step 5 is important. Once we know the market equilibrium price, we must check if it is greater than or equal to the shutdown price. If this is not the case, no firm will want to produce in that market, so the equilibrium quantity will be zero (the entire market shuts down).

If the market price is at least as great as the shutdown price, then we can find the individual firm's supply by plugging the market price into the firm's supply function.

From p_{eq} and $q(p_{eq})$, we can compute the profits of the firm, which in the short run can be either positive, negative, or zero.

Example. Let's come back to the cost function $TC(q) = 4q^2 + 10q + 100$. Assume 160 firms and $Q^D = 800 - 30p$.

Let's find the market equilibrium.

1. $MC = p \Rightarrow 8q + 10 = p \Rightarrow$

   - $$\begin{cases} \frac{p-10}{8} & p \geq 10 \\ 0 & p \leq 10 \end{cases}$$

2. $AVC = 4q + 10$, which is minimized at $q = 0$ and $AVC(0) = 10$, so $p_{exit} = 10$

3. $Q^S = 160q(p)$

   - $$\begin{cases} 20p - 200 & p \geq 10 \\ 0 & p \leq 10 \end{cases}$$

4. Set $Q^D = Q^S$. $20p - 200 = 800 - 30p$. Then $p_{eq} = 20$.

5. As $20 > 10$, we get a positive firm supply of $q(20) = 1.25$.

6. The market quantity equilibrium is $Q = 160 * 1.25 = 200$.

At this price, the profits of the firm are $\pi(1.25) = 1.25 * 20 - 118.75 < 0$, which we could have figured out by noticing that $p_{eq} < minSAC = 50$. In this case, each firm gets negative profits.

Example. In the previous example, what if $Q^D = 2800 - 30p$?

In this case, we set the equilibrium $Q^D = Q^S$. $20p - 200 = 2800 - 30p$. Then $p_{eq} = 60$.

Because $p_{eq} > minSAC$, we expect the firms to have positive profits. Let's check. At $p = 60$, $q(60) = 6.25$ and revenue is 375. $TC(6.25) = 4(6.25)^2 + 10(6.25) + 100 = 318.75$, so $\pi(6.25) > 0$.

In this example, profits are positive. Remember that this is the short run. We also mentioned previously that, in the long run, perfect competition implies zero profits.

So this equilibrium with positive profits is clearly not sustainable over time. We study the long run in the next section.

3.6 Market Supply and Market Equilibrium in the Long Run

In the last example, the firms have positive profits in equilibrium. Remember that in perfect competition all firms have free access to the production technology, meaning that there are no barriers to entry to these markets in the long run. Everyone who wants to produce in this market will eventually be able to do so, and everyone who does not want to produce in this market can eventually exit. How realistic do you think this assumption is?

The free entry and exit condition of perfect competition means that when $\pi(q) > 0$, more firms will enter the market motivated by positive profits, so the number of firms will increase. By our formula of the market supply, $Q^S(p) = Nq(p)$, if N goes up, the whole market supply will expand, which can be represented graphically by a shift to the right of the supply curve.

If supply increases, the market will meet at a new equilibrium with greater quantity and lower price. Think if this makes sense to you.

A lower price will decrease the individual firm's supply and, more importantly, its profits. By how much should the profits go down? As long as $\pi > 0$, new firms will keep entering the market and the equilibrium will continue changing. This process stops when $\pi = 0$. Hence, the optimal number of firms in the market in the long run is the one that leads to $\pi = 0$ (no profits for the individual firm).

What is this magical number of firms? Remember from the shutdown rule that profits are zero when $p = minAC$ (that's why $p = minAC$ is the shutdown price in the long run). So we have our desired equilibrium price in the long run $p = minAC$. We can plug the price into the individual firm's supply to get the quantity produced by each firm. We can also plug the price into the demand function to get the equilibrium quantity. As $Q^D = Q^S(p) = Nq(p)$, we can solve for N.

To compute the number of firms in the long-run equilibrium we need to follow these steps:

1. Compute the long-run shutdown price $p_{exit} = minAC$. This is going to be our equilibrium price, so $p^*_{LR} = minAC$.
2. Plug p^*_{LR} into the demand curve $Q^D(p)$ to get the market equilibrium quantity.

3. Remember that in equilibrium $Q^D(p^*_{LR}) = Q^S(p^*_{LR})$, so we now have a number for $Q^S(p^*_{LR})$.

4. Plug p^*_{LR} into the individual firm's supply $q(p)$ to get the amount that each firm will produce.

5. Because we now have the quantity produced by each firm and the equilibrium quantity, we can use the formula $Q^S(p) = Nq(p)$ and solve for $N_{LR} = \frac{Q^S(p^*_{LR})}{q(p^*_{LR})}$.

6. (Optional) Verify that $\pi(q(p^*_{LR})) = 0$.

Example. Continuing with the example of $TC(q) = 4q^2+10q+100$ and $Q^D = 2800-30p$, we have:

1. Compute the long-run shutdown price $p_{exit} = minAC$. $AC(q) = 4q + 10 + \frac{100}{q}$. Set $AC(q) = MC(q) = 8q + 10$, so $q = 5$ and $AC(5) = 50$.

2. $Q^D(50) = 2800 - 30 * 50 = 1300 = Q^S(50)$.

3. $q(50) = \frac{50-10}{8} = 5$.

4. $1300 = N * 5$ and solve for $N_{LR} = \frac{1300}{5} = 260$.

What happens if the profits in the short run are negative?

The above steps work well for that case too. What changes is the dynamics: instead of new firms entering the market and prices going low, what will happen is that some firms will exit the market and the price will go up until it reaches $minAC$.

In the last chapter, we defined the minimum efficient scale as the smallest quantity at which the long-run average cost curve attains its minimum point. Another way to say this is the smallest q (in case there is more than one) that results in $minAC$. This concept is relevant because it tells us the efficient "size" of the market (number of firms) in the long run.

3.7 Demand Shocks

How does the market adjust to demand shocks?

Suppose that the market is at its long-run equilibrium ($p = minAC$ and let's denote the equilibrium quantity as Q^*_0) when there is an increase in demand (the curve shifts to the right).

This demand shock puts us in a new "short run". The increase in demand will increase the price, so in the new short run, where the number of firms is fixed, profits will become greater than zero.

Because of the free entry property of competitive markets, eventually more firms will enter, increasing the market supply and lowering the price again. How low? Again, until no more firms keep entering, that is $p = minAC$. At that moment, the new long-run equilibrium will be $p = minAC$ and Q_1^*, where $Q_1^* > Q_0^*$ because both the market demand and supply increased.

If you represent this adjustment in a graph, you will notice that from the original long-run equilibrium point to the new one, the only element that changed is the quantity; the price is the same ($p = minAC$). Indeed, the market supply curve in the long run is a horizontal line at $p = minAC$.

- The individual firm's supply curve is upward-sloping.
- The short-run market supply curve is upward-sloping.
- The long-run market supply curve is horizontal.

What if there are barriers to entry?

If there are barriers to entry, even if profits are positive, no new firms will enter the market in the long run. So, will the short- and long-run market supply look identical?

Not necessarily. Remember that in the long run, each firm has flexibility over its capital, so it could, for example, increase its plant size. Motivated by positive profits, the firms will increase their production capacity in the long run (remember that $p = SMC$ - short run - will not be the same supply as $p = MC$ - long run -).

With each firm increasing production, in the long run, the market supply will shift to the right. But this shift may not be enough to completely offset the positive profits; they will go down, but not necessarily to zero. In this case, the long-run equilibrium will support positive profits unless new firms enter the market.

3.8 Profit-Maximizing Demand for Inputs

In Chapter 2 we used the following notation:

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

$$TC = wL + rK$$

If we bring this notation to the profit maximization problem, then

$$\pi(q) = pq - TC(q)$$

becomes

$$\pi(L, K) = pF(L, K) - wL - rK$$

And the profit maximization problem can be written as:

Choose L and K to maximize $\pi(L, K) = pF(L, K) - wL - rK$,

If the maximization problem is well defined, that is, the firm has decreasing returns to scale (and thus also diminishing marginal products),

Then the first-order conditions of this problem are:

- $\frac{\partial \pi}{\partial L} = 0$
- $\frac{\partial \pi}{\partial K} = 0$

That is:

- $p\frac{\partial F(L,K)}{\partial L} - w = 0$
- $p\frac{\partial F(L,K)}{\partial K} - r = 0$

Which can also be expressed as:

- $pMP_L = w$
- $pMP_K = r$

We call pMP_L the marginal revenue product of labor (MRP_L), which tells us the change in revenue associated with hiring an additional unit of labor.

Similarly, pMP_K is the marginal revenue product of capital (MRP_K).

The first-order condition then tells us that the extra revenue we get from hiring an additional unit of labor should be equal to the cost of that unit, w, and similarly for capital. Does this make sense? What happens if $MRP_L > w$? What if $MRP_L < w$?

Notice also that $MRP_L = p * MP_L$ and $MRP_K = p * MP_K$ because we are in a perfectly competitive market, so marginal revenue is equal to p. More generally, we should define $MRP_L = \frac{\partial Rev}{\partial L}$ and $MRP_K = \frac{\partial Rev}{\partial K}$.

The solutions of this problem, $L^*(p, w, r)$ and $K^*(p, w, r)$, give us the profit-maximizing demand for labor and capital. Notice that they depend on p, w, and r.

How are these demands related to the cost-minimizing demands $L^*(q, w, r)$ and $K^*(q, w, r)$?

We need to notice that the profit-maximizing demands depend on p, but the cost-minimizing demands depend on q, so they depend on different parameters and hence, they are different objects.

Also, they are the solutions to different optimization problems, so we should not think they are, in general, the same objects.

However, they are equivalent at a very special point. If we compute the profit-maximizing quantity $q^*(p)$ and calculate the cost-minimizing demands of labor and capital, $L^*(q^*(p), w, r)$ and $K^*(q^*(p), w, r)$, the following equivalences hold: $L^*(q^*(p), w, r) = L^*(p, w, r)$ and $K^*(q^*(p), w, r) = K^*(p, w, r)$. In other words, at the profit-maximizing quantity, the profit-maximizing and cost-minimizing demands for inputs are equivalent.

Example: $F(L, K) = L^{1/3}K^{1/3}$.

The profit function can be written as:

$$\pi(L, K) = pL^{1/3}K^{1/3} - wL - rK$$

To maximize $\pi(L, K)$, we set the partial derivatives with respect to L and K to zero.

The first-order conditions (FOCs) for maximization are:

$$\frac{\partial \pi}{\partial L} = \frac{p}{3}L^{-2/3}K^{1/3} - w = 0$$

$$\frac{\partial \pi}{\partial K} = \frac{p}{3}L^{1/3}K^{-2/3} - r = 0$$

From $\frac{\partial \pi}{\partial L} = 0$:

$$\frac{p}{3}L^{-2/3}K^{1/3} = w \implies K^{1/3} = \frac{3w}{p}L^{2/3} \implies K = \left(\frac{3w}{p}\right)^3 L^2$$

From $\frac{\partial \pi}{\partial K} = 0$:

$$\frac{p}{3}L^{1/3}K^{-2/3} = r \implies L^{1/3} = \frac{3r}{p}K^{2/3} \implies L = \left(\frac{3r}{p}\right)^3 K^2$$

Substitute $K = \left(\frac{3w}{p}\right)^3 L^2$ into $L = \left(\frac{3r}{p}\right)^3 K^2$:

$$L = \left(\frac{3r}{p}\right)^3 \left(\left(\frac{3w}{p}\right)^3 L^2\right)^2$$

$$L = \left(\frac{3r}{p}\right)^3 \left(\frac{3w}{p}\right)^6 L^4$$

$$1 = \left(\frac{3r}{p}\right)^3 \left(\frac{3w}{p}\right)^6 L^3$$

$$L^3 = \frac{1}{\left(\frac{3r}{p}\right)^3 \left(\frac{3w}{p}\right)^6}$$

$$L = \frac{p^3}{27rw^2}$$

Substitute L back into $K = \left(\frac{3w}{p}\right)^3 L^2$:

$$K = \left(\frac{3w}{p}\right)^3 \left(\frac{p^3}{27rw^2}\right)^2$$

$$K = \frac{p^3}{27r^2w}$$

The optimal values of L and K that maximize profits are:

$$L = \frac{p^3}{27rw^2}, K = \frac{p^3}{27r^2w}$$

Notice how both L and K depend negatively on both w and r. The idea is that if the price of any of the inputs increases (increase of cost), the firm will maximize profits by producing less, and therefore, demanding less of both inputs. As we would expect, the reaction of labor is stronger to changes in w (square factor) and the reaction of capital is stronger to changes in r.

Now let's solve the cost-minimization problem with the same production function and some level of production q.

Minimize the cost function:

$$C = wL + rK$$

subject to the constraint:

$$L^{1/3}K^{1/3} = q$$

From the optimal capital-labor ratio:

$$K = \frac{w}{r} L$$

Substitute $K = \frac{w}{r} L$ into the production constraint:

$$L^{1/3} \left(\frac{w}{r} L \right)^{1/3} = q$$

$$L = q^{3/2} \left(\frac{r}{w} \right)^{1/2}$$

Substitute L back into $K = \frac{w}{r} L$:

$$K = q^{3/2} \left(\frac{w}{r} \right)^{1/2}$$

The cost-minimizing values of L and K are:

$$L = q^{3/2} \left(\frac{r}{w} \right)^{1/2}, K = q^{3/2} \left(\frac{w}{r} \right)^{1/2}$$

Notice that in this case, labor has a positive relation with r, as r increases, capital gets substituted by labor, and similarly, K has a positive relation with w. This substitution effect is so evident because q is fixed, so the only thing the firm can do is substitute one input for the other. However, when the firm can choose the output level like in the profit-maximizing case, an increase in any of the input's prices will decrease q and hence decrease the demand for both inputs (demands are more elastic to input prices when we use the profit-maximizing level of analysis).

The above demands of inputs give us the cost function:

$$TC(q) = 2q^{3/2} \sqrt{wr}$$

Notice that we can use this total cost function to compute the marginal cost and get the profit-maximizing supply. That is:

$MC(q) = 3q^{\frac{1}{2}} (wr)^{\frac{1}{2}}$,

Hence $MC(q) = p \Rightarrow p = 3q^{\frac{1}{2}} (wr)^{\frac{1}{2}}$ and $q^*(p) = \frac{p^2}{9wr}$.

Now we can plug that particular q into our cost-minimizing demands of labor and capital:

$$L = q^{3/2} \left(\frac{r}{w} \right)^{1/2}, K = q^{3/2} \left(\frac{w}{r} \right)^{1/2}$$

to get:

$$L = \left(\frac{p^2}{9wr}\right)^{3/2} \left(\frac{r}{w}\right)^{1/2} = \frac{p^3}{27rw^2}$$

$$K = \left(\frac{p^2}{9wr}\right)^{3/2} \left(\frac{w}{r}\right)^{1/2} = \frac{p^3}{27r^2w}$$

They are identical to the profit-maximizing demands of labor and capital.

Graphical Analysis of input price changes – Different levels of analysis.

As we observed in the previous example, the demands from inputs coming from cost-minimizing and profit-maximizing approaches react differently to changes in input prices. Which analysis we use will give us different answers. This means we need to be careful about which level of analysis is more appropriate for the situation we are studying.

Considering some extra factors makes the analysis even more complex. If w or r increases, then each (identical) firm will maximize profits by producing less. As a result, in the short run, the whole supply will shift to the left and the price will increase. Whether the equilibrium quantity reduces a lot or a little will depend on the price elasticity of the demand. In the long run, the price should come back to $minAC$; however, a permanent increase in r or w will increase the entire $AC(q)$ curve, so the new $minAC$ will occur at a greater point, and the long-run equilibrium price will increase.

We may even think that if w rises, the workers will have more income, and if the workers are also consumers of that good, and the good is normal (demand increases with income), then as a result of an increase in w, the demand curve will shift to the right. This, by itself, will increase the equilibrium price and quantity, but now it also interacts with a decrease in quantity due to the contraction of the supply (as w increased).

This situation is known as a double shock: both the supply and the demand are independently impacted by the increase in w. Whether the equilibrium quantity will increase or decrease with respect to the initial situation will depend on which effect is greater. Check the graphs in class.

3.9 Practice Problems

1. Consider the following total cost functions:

   (a) $TC(q) = q^2 + 2q + 9$

   (b) $TC(q) = 3q^2 + 10q + 300$

(c) $TC(q) = 8q^2 + 3q + 80$

Assume we are in the short run (and all fixed costs are sunk), so $TC(q)$ is indeed $STC(q)$.

(I) Find and graph the supply of the firm.

(II) Find the profits if $p = 8$. Draw a graph representing revenue, variable cost, fixed costs, and profits with rectangles as we did in class. Should the firm shut down?

(III) Find the producer surplus when $p = 1$ and represent it in a graph with rectangles.

(IV) Repeat (II) and (III) when $p = 11$.

(V) Find the shutdown (or exit) price for the firm.

2. For each cost function of Problem 1, now assume all the fixed costs are avoidable.

(I) Should the firm produce when $p = 8$? What about $p = 11$? Hint: Use your answers from Problem 1.

(II) Find the producer surplus when $p = 1$ and represent it in a graph with rectangles.

(III) Find the shutdown (or exit) price for the firm.

3. Now repeat Problem 1 but assume that only half of the fixed cost is sunk (4.5 for (a), 150 for (b), and 40 for (c)).

4. Why is $p_{shutdown}$ in the short run always less than or equal to $p_{shutdown}$ in the long run?

5. For each of the cost functions in Problem 1, imagine there is only one firm in the market (so the supply of that firm is the supply of the market). Assume the demand is given by $Q^D = 100 - 2p$. Find the market price and determine if the firm will produce at that price or not.

6. Consider the following total cost functions:

(a) $TC(q) = q^2 + 5q + 100$

(b) $TC(q) = 3q^2 + 10q + 300$

(c) $TC(q) = 8q^2 + 3q + 80$

Demand in this market is given by $Q^D = 330 - 12p$. Consider 12 firms operating in the short run. For each of them, do the following:

(i) Find the firm- and industry-level supply functions, assuming that the fixed cost is sunk (short-run scenario).

(ii) Solve for the market equilibrium, and for consumer surplus and producer surplus. How much does each firm produce, and what are its producer surplus and profit?

(iii) If the total cost function stays the same in the long run (but none of the cost is sunk anymore), what will happen in the market, and what will be the profit per firm?

(iv) Write (carefully) the long-run supply function at the firm level.

(v) Now repeat steps (ii) and (iii) after an expansion of demand to $Q^D = 3300 - 12p$.

(vi) Now repeat steps (ii) and (iii) after a contraction of demand to $Q^D = 110 - 4p$.

(vii) Explain and illustrate with graphs the short- and long-run equilibrium outcomes of scenarios (v) and (vi) with market entry barriers.

7. For the following production functions:

(a) $F(L, K) = \sqrt{L} + \sqrt{K}$

(b) $F(L, K) = L^{0.25}K^{0.25}$

(c) $F(L, K) = L^{0.6}K^{0.5}$

(i) For a fixed level $\bar{K}$, find the short-run profit-maximizing demand for labor $L^*(p, w, r, \bar{K})$.

(ii) Find the profit-maximizing input demands for labor and capital in the long run, as functions of the prices of the output and inputs: $L^*(p, w, r), K^*(p, w, r)$.

(iii) For a given level of q, find the cost-minimizing demands for labor and capital in the long run: $L^*(q, w, r), K^*(q, w, r)$.

(iv) Use your results in (iii) to find the total cost function. Simplify as much as possible.

(v) Using your total cost function from (iv), find the profit-maximizing supply $q^*(p, w, r)$.

Evaluate your input demands from (iii) and verify that they are identical to your input demands from (ii).

(vi) How do $L^*(p, w, r), K^*(p, w, r)$ and $L^*(q, w, r), K^*(q, w, r)$ depend on w? Is the effect the same? If they differ, what is the intuition behind it?

8. Illustrate with graphs the different effects of an increase in wages on the demands for L and K under the following scenarios:

(a) Fixed q, short run

(b) Fixed q, long run

(c) Flexible q, short run

(d) Flexible q, long run

Compare the different effects.

4 Chapter 4: Market Interventions

These notes focus more on formulas and calculations than on conceptual lecture notes, unlike the previous chapters.

4.1 Market Equilibrium

The market is at equilibrium at the pair (p^*, q^*), such that at price p^*, $Q^D(p^*) = Q^S(p^*) = q^*$.

At the competitive market equilibrium:

- Every consumer willing to buy at the price p^* gets the chance to do so.
- Every seller willing to sell at the price p^* gets the chance to do so.
- $MC(q^*) = MB(q^*) = p^*$. *Marginal benefit = Marginal cost = Market price.*

Steps to Find Market Equilibrium:

1. Set the supply function Q_S equal to the demand function Q_D:

$$Q_S = Q_D$$

2. Solve the equation for the equilibrium price P^*.

3. Substitute P^* back into either the supply or demand function to find the equilibrium quantity Q^*.

Consumer Surplus (CS) and Producer Surplus (PS)

Consumer surplus is the difference between what consumers are willing to pay and what they actually pay. Producer surplus is the difference between what producers receive and the minimum amount they are willing to accept.

Consumer Surplus:

$$CS = \frac{1}{2} \times (P_{\max} - p^*) \times q^*$$

where $P_{\max}$ is the maximum price consumers are willing to pay, p^* is the equilibrium price, and q^* is the equilibrium quantity.

Producer Surplus:

$$PS = \frac{1}{2} \times (p^* - P_{\min}) \times q^*$$

where $P_{\min}$ is the minimum price at which producers are willing to supply the good.

Minimum and Maximum Prices:

- The minimum price ($P_{\min}$) is the lowest price at which producers are willing to supply the good. - The maximum price ($P_{\max}$) is the highest price consumers are willing to pay.

Finding $P_{\min}$:

$$Q_S = 0 \Rightarrow P - 2.5 = 0 \Rightarrow P_{\min} = 2.5$$

Finding $P_{\max}$:

$$Q_D = 0 \Rightarrow 20 - 2P = 0 \Rightarrow P_{\max} = 10$$

Example:

Consider the following supply and demand functions:

$$Q_S = P - 3, \quad Q_D = 18 - 2P$$

For easier calculations, we use the inverse supply and demand functions:

Inverse Supply and Demand Functions:

$$P_S = Q_S + 3, \quad P_D = 9 - \frac{Q_D}{2}$$

Step-by-Step Solution Using Inverse Functions:

1. Set the inverse supply function equal to the inverse demand function to find P^*:

$$Q + 3 = 9 - \frac{Q}{2} \Rightarrow \frac{3Q}{2} = 6 \Rightarrow Q^* = 4$$

2. Substitute Q^* into either inverse function to find P^*:

$$P^* = 4 + 3 = 7$$

3. Calculate consumer surplus:

$$CS = \frac{1}{2} \times (P_{\max} - P^*) \times Q^* = \frac{1}{2} \times (9 - 7) \times 4 = 4$$

4. Calculate producer surplus:

$$PS = \frac{1}{2} \times (P^* - P_{\min}) \times Q^* = \frac{1}{2} \times (7 - 3) \times 4 = 8$$

5. Equilibrium price and quantity are $P^* = 7$ and $Q^* = 4$.

4.2 Price Controls

Price Ceiling

A price ceiling is a maximum price set below the equilibrium price. It aims to make goods more affordable but can create shortages.

General Formulas for a Price Ceiling:

Define:

- p_c = price ceiling
- q_c = quantity supplied at p_c
- z = demand at quantity q_c

- New Consumer Surplus:

$$CS = \frac{1}{2} \times (P_{\max} - z) \times q_c + (z - p_c) \times q_c$$

- New Producer Surplus:

$$PS = \frac{1}{2} \times (p_c - P_{\min}) \times q_c$$

- Deadweight Loss:

$$DWL = \frac{1}{2} \times (z - p_c) \times (q^* - q_c)$$

Example:

Consider a price ceiling at $p_c = 6$.
1. Find q_c and z:

$$q_c : 6 = Q_S + 3 \Rightarrow q_c = 3$$
$$z = 9 - \frac{3}{2} = 7.5$$

2. Calculate the new consumer surplus:

$$CS = 0.5 \times (9 - 7.5) \times 3 + (7.5 - 6) \times 3$$

3. Calculate the new producer surplus:

$$PS = 0.5 \times (6 - 3) \times 3$$

4. Calculate the deadweight loss:

$$DWL = 0.5 \times (7.5 - 6) \times (4 - 3)$$

Price Floor

A price floor is a minimum price set above the equilibrium price. It is intended to ensure that producers receive a minimum income but can lead to an excess of supply.

General Formulas for Price Floor

Let's define:

- p_f, the price floor.
- q_f, the value of demand at p_f.
- z, the value obtained by plugging q_f into the supply function.

- New Consumer Surplus:

$$CS = \frac{1}{2} \times (P_{\max} - p_f) \times q_f$$

- New Producer Surplus:

$$PS = \frac{1}{2} \times (z - P_{\min}) \times q_f + (p_f - z) \times q_f$$

- Deadweight Loss:

$$DWL = \frac{1}{2} \times (p_f - z) \times (q^* - q_f)$$

Example:

Following the previous example, consider a price floor at $P_f = 8$.

1. Find p_f, q_f, and z:

$$p_f = 8$$
$$q_f = 18 - 2 \cdot 8 = 2$$
$$z = 2 + 3 = 5$$

2. $CS = 0.5 \times (9 - 8) \times 2$

3. $PS = 0.5 \times (5 - 3) \times 2 + (8 - 5) \times 2$

4. $DWL = 0.5 \times (8 - 5) \times (4 - 2)$

4.3 Per-Unit Tax Analysis

Consider a market with the following inverse demand and supply functions:

$$p = f_d(Q_d) \quad \text{(Inverse Demand)}$$
$$p = f_s(Q_s) \quad \text{(Inverse Supply)}$$

When a per-unit tax t is imposed on producers, the supply curve shifts upward by the amount of the tax:

$$p = f_s(Q_s) + t$$

Note: Alternatively, the tax could be levied on consumers, which is modeled as:

$$p = f_d(Q_d) - t$$

Equilibrium without Tax

To find the equilibrium price and quantity without the tax, set the inverse demand equal to the inverse supply:

$$f_d(Q) = f_s(Q)$$

Solve for q^* and substitute back to find the equilibrium price p^*.

Equilibrium with Tax

To find the new equilibrium price and quantity with the tax, set the inverse demand equal to the new inverse supply:

$$f_d(Q) = f_s(Q) + t$$

Solve for q_t and substitute back to find the price paid by consumers p_{cons} and the price received by producers p_{prod}. Remember that $p_{cons} - p_{prod} = t$.

Consumer Surplus (CS)

$$CS = \frac{1}{2} \times q_t \times (P_{\max} - p_{cons})$$

Producer Surplus (PS)

$$PS = \frac{1}{2} \times q_t \times (p_{prod} - P_{\min})$$

Government Revenue

Government revenue is the tax amount multiplied by the equilibrium quantity:

$$\text{Revenue} = t \times q_t$$

Deadweight Loss (DWL)

$$DWL = \frac{1}{2} \times t \times (q^* - q_t)$$

Tax Incidence

The incidence of the tax on consumers and producers can be calculated using elasticities:

$$\text{Consumer Tax Incidence} = \frac{E_s}{E_d + E_s}$$

$$\text{Producer Tax Incidence} = \frac{E_d}{E_d + E_s}$$

Example

Consider the following demand and supply functions:

$$Q_d = 300 - 3p$$
$$Q_s = 4p - 100$$

The inverse functions are:

$$p = 100 - \frac{Q_d}{3}$$
$$p = 25 + \frac{Q_s}{4}$$

Equilibrium without Tax

Set the inverse demand equal to the inverse supply:

$$100 - \frac{Q}{3} = 25 + \frac{Q}{4}$$

Solve for Q:

$$75 = \frac{4Q + 3Q}{12} \implies q^* = 128.57$$

Substitute Q back to find p:

$$p^* \approx 57.14$$

Equilibrium with \$5 Tax

Set the inverse demand equal to the new inverse supply:

$$100 - \frac{Q}{3} = 30 + \frac{Q}{4}$$

Solve for q_t:

$$70 = \frac{7Q}{12} \implies q_t = 120$$

Find the prices:

$$p_c = 100 - \frac{120}{3} = 60$$
$$p_p = 60 - 5 = 55$$

Consumer Surplus (CS)

$$CS = \frac{1}{2} \times 120 \times (100 - 60) = 2400$$

Producer Surplus (PS)

$$PS = \frac{1}{2} \times 120 \times (55 - 25) = 1800$$

Government Revenue

$$\text{Revenue} = 5 \times 120 = 600$$

Deadweight Loss (DWL)

$$DWL = \frac{1}{2} \times 5 \times (128.57 - 120) \approx 21.43$$

Tax Incidence

Calculate elasticities at equilibrium:

$$E_d = \left| -3 \times \frac{57.14}{128.57} \right| \approx 1.33$$
$$E_s = \left| 4 \times \frac{57.14}{128.57} \right| \approx 1.78$$

Incidence on Consumers:

$$\frac{E_s}{E_d + E_s} = \frac{1.78}{1.33 + 1.78} \approx 0.57$$

Incidence on Producers:

$$\frac{E_d}{E_d + E_s} = \frac{1.33}{1.33 + 1.78} \approx 0.43$$

4.4 Per-Unit Subsidy Analysis

Consider a market with the following inverse demand and supply functions:

$$p = f_d(Q_d) \quad \text{(Inverse Demand)}$$
$$p = f_s(Q_s) \quad \text{(Inverse Supply)}$$

When a per-unit subsidy s is given to producers, the supply curve shifts downward by the amount of the subsidy:

$$p = f_s(Q_s) - s$$

Note: Alternatively, the subsidy could go to consumers, modeled as:

$$p = f_d(Q_d) + s$$

Equilibrium without Subsidy

Set the inverse demand equal to the inverse supply:

$$f_d(Q) = f_s(Q)$$

Solve for q^* and substitute back to find p^*.

Equilibrium with Subsidy

Set the inverse demand equal to the new inverse supply:

$$f_d(Q) = f_s(Q) - s$$

Solve for q_s and substitute back to find the price paid by consumers p_{cons} and the price received by producers p_{prod}.

Consumer Surplus (CS)

$$CS = \frac{1}{2} \times q_s \times (p_{\max} - p_{cons})$$

Producer Surplus (PS)

$$PS = \frac{1}{2} \times q_s \times (p_{prod} - P_{\min})$$

Government Revenue (GR)

$$Revenue = -s \times q_s$$

Deadweight Loss (DWL)

$$DWL = \frac{1}{2} \times s \times (q_s - q^*)$$

Example

Consider the following demand and supply functions:

$$Q_d = 300 - 3p$$
$$Q_s = 4p - 100$$

The inverse functions are:

$$p = 100 - \frac{Q_d}{3}$$
$$p = 25 + \frac{Q_s}{4}$$

Consider a $5 subsidy to producers.

Equilibrium without Subsidy

$$100 - \frac{Q}{3} = 25 + \frac{Q}{4} \implies q^* = 128.57, \quad p^* \approx 57.14$$

Equilibrium with $5 Subsidy

$$100 - \frac{Q}{3} = 20 + \frac{Q}{4} \implies q_s = 137.14$$

Find the prices:

$$p_{cons} = 100 - \frac{137.14}{3} \approx 54.29$$
$$p_{prod} = 54.29 + 5 \approx 59.29$$

Consumer Surplus (CS)

$$CS \approx 3137.63$$

Producer Surplus (PS)

$$PS \approx 2345.01$$

Government Revenue

$$Revenue = -5 \times 137.14$$

Deadweight Loss (DWL)

$$DWL \approx 21.43$$

4.5 Practice Problems

Problem 1.Consider a market where supply is $Q_S = P - 2.5$ and demand is $Q_D = 20 - 2P$.

(a) Find the market equilibrium, producer surplus, and consumer surplus.

(b) If a tax of $t = 3$ per unit is imposed on producers, find the new equilibrium quantity, the price faced by the demand side, and the price faced by the supply side. Find producer and consumer surplus, deadweight loss, and the incidence of the tax on consumers and producers.

(c) Now analyze the scenario of the same tax $t = 3$ to consumption instead of production. Corroborate that the welfare effects are identical to part (b).

(d) If we wanted to achieve an equilibrium with the same market quantity as (b), but by using a price ceiling instead of a tax, what maximum price $P_{ceiling}$ would we set? Find PS, CS, and DWL.

(e) Can we achieve the same quantity with a price floor? If so, find P_{floor}, PS, CS, and DWL.

(f) Now suppose a subsidy of $3 per unit was imposed in this market instead: for every unit sold, producers receive $3. What would be the new quantity produced and the deadweight loss? Does the sum of producer and consumer surplus (PS+CS) increase or decrease? By how much?

Problem 2. Consider again the same market ($Q_S = P - 2.5$ and $Q_D = 20 - 2P$) and the following cases:

(a) What happens if there is a price ceiling of 4.5? What about $P_{ceiling} = 8$?

(b) What happens if there is a price floor of 4.5? What about $P_{floor} = 8$?