Bilkent University
Econ 101 - Fall 2023
Chapter 6: Producer Theory

A. Arda Gitmez   Nuh Aygün Dalkıran

October 29, 2023

Contents

1 A Very Brief Introduction

With this chapter, we start a brand new topic – even though the concepts and analysis methods are very similar with the ones used in consumer theory. We are introducing the formal model of another economic agent: a **firm**. We consider the firm's decision to produce and sell goods and services.

Some motivation: from Chapter 1, we remember that a market is an infrastructure that facilitates **interactions** among **economic agents**. One of the most widespread economic interactions are the buying and selling of services. We spent two chapters on the decision to buy, and finally now it is time to consider selling! To this end, after spending two chapters on economic agents who buy (consumers), now we are making our transition to the economic agents who sell (firms). After we are done with the theory of the firm, we will bring the consumers and firms together and analyze a **market**.

The topics we cover here correspond to Chapters 11 and 12 of your textbook, but I am taking some liberties in the treatment (I am adding some stuff and removing them). The ideas presented here are a patch-up of several resources. Still, if you want a resource to consult, just check the section titled "Theory of the Firm" in any introductory economics textbook (it may also be called "Theory of the Producer" or "Producer Theory").

2 What is a Firm?

2.1 Some Intellectual Background for the Curious Mind

The question of what makes a firm a firm is one of the most important questions of 20th century economics. There has been at least four Economics Nobel prizes related to defining what "a firm" is (Herbert Simon, Ronald Coase, Oliver Williamson, Oliver Hart) and we still do not have a clear answer. Some of the historically popular answers are:

- "The firm is a mechanism to circumvent the **price discovery**, which is a costly activity. By setting up a firm, an entrepreneur can procure essential materials by itself without going to the market and working around the price mechanism." (early 20th century)
- "The firm is a way of ensuring **adaptation** through giving authority to a boss." (Simon)
- "The firm is a device to minimize **transaction costs** among producers who interact frequently." (Coase, Williamson)
- "The firm is a mechanism for **monitoring** performance when there is a need for joint production." (Alchian, Demsetz)
- "The firm is a way to ensure proper investments through giving **property rights** to a boss." (Hart)

In Moodle, I am posting an article that discusses some of the theories of the firm. It is somewhat advanced for Econ 101, but you can check out Section 1 to read the "informal discussion".

> Gibbons, Robert. "Four Formal(izable) Theories of the Firm?" *Journal of Economic Behavior & Organization* 58 (2005): 200–245.

But... This is Econ 101, so we are merely laying out the foundation. We will adopt a much, much simpler definition of a firm and abstract away from these concerns.

2.2 The Firm as a Technology

Here is the definition of the firm for the purposes of this class.

Definition 1. *A* **firm** *is an entity which uses inputs in order to produce outputs.*

Here, **inputs** include natural resources, labor, machinery etc. and **outputs** include goods and services (the things that consumers buy).

So, a firm is not an organization according to our definition. It is just a technology that takes some inputs and transforms it into outputs. Just like a consumer is defined by her preferences, a firm is defined by its technology. This is a tremendous simplification, or as we call it, an *abstraction of reality*. Remember that this is a model: we need abstractions!

Speaking of abstractions, here are a couple of simplifications we assume throughout this chapter:

- The firm produces a single output. In reality, Bilkent (as a firm) produces many outputs: diplomas, research papers, boarding, parking spaces... We are abstracting away from that: in this model, Bilkent only produces teaching. A pizza parlor in this model produces only a single type of pizza: medium pepperoni pizzas.

  When the firm is able to produce multiple goods at once, its decisions become more complicated to analyze but the basic insights of this simple model remain.

- The firm has already chosen which product to produce. The decision of what to produce is also extremely important, but that is not what we are focusing here. For models of product choice, you should take a class on industrial organization.

- The firm can buy as many inputs as it wants, as long as they are profitable. Effectively, we are assuming that the firm is not cash-constrained. What we have in mind is a financial market (multiple investors) willing to lend money to the firm as long as the investment is profitable.

3 The Firm as an Economic Agent

As every economic agent, a firm is also defined by its **constraints** and **preferences**. Before we define whet these are, let us introduce some notation.

A firm uses several **inputs** to produce an **output**. The output is just the good or service the firm has chosen to produce. Because there is no question about what to produce, the question is **how much to produce**. To denote the quantity of good produced by the firm, we use the letter Q. This should be familiar from consumer theory. (Some textbooks denote it with letter y.)

The next question is **how to produce** Q units of output. What inputs should the firm use? Broadly speaking, we can categorize inputs into three groups. These are also called **factors of production**:

1. **Land (or natural resources):** goods which are not created by human beings. Examples: Air, forests, minerals, soil, etc. Example: Bilkent uses the land it is on to produce teaching.

2. **(Physical) Capital:** Human-made goods used to produce other goods. Examples: machinery, computers, factories, roads, etc. Example: Bilkent uses the buildings, projectors, computers, software... for teaching.

   Note: it is common to think of *financial capital* (i.e. the money you invest and receive return on) when someone says “capital”. What we mean by “capital” is different. We mean physical capital, i.e. machines of one sort or another.

3. **Labor:** Human effort used in production. Example: Bilkent uses the effort provided by faculty members, TAs, administrative assistants... for teaching.

Typically,

- the quantity of land used in production is denoted by D,
- the quantity of capital used in production is denoted by K, and,
- the quantity of labor used in production is denoted by L.

At this point, we will forget about land/natural resources. (They are trickier to introduce to this model because they are, by definition, in limited supply. This does not bode well with the assumption we made: “The firm can buy as many inputs as it wants.”) Just assume a firm uses two inputs: machines and labor.

If you are operating a pizza parlor, you use labor (workers) and capital (pizza ovens) to produce output (medium pepperoni pizzas).

It is not crazy to think that more or less all production activities require inputs of both forms. They may be at different ratios, of course. Producing tea requires human effort and machines (tractors), but famously the human effort is a very large component. Producing Bitcoins requires a lot of capital (processors to mine Bitcoins), but also a human has to occasionally stand in front of a computer.

Note: in this model, all the inputs and outputs are measured in **flows**, i.e. their quantities are given **per unit of time**. For instance, if the unit of time is week, we will say it takes *a total of 42 labor hours per week* and *a total of 16 machine hours per week* to produce *23 hours of teaching per week*. Or we may say it takes *a total of 3 workers per day* and *3 pizza ovens per day* to produce *1100 pizzas per day*.

3.1 The Constraint

A firm faces many constraints imposed by consumers, its competitors, government, and nature. Here, we will focus on the constraints imposed by **nature**. These are some **technological constraints** that capture the following idea: there are only certain ways to produce Q units of output from labor and capital.

Given the quantities of inputs L, K, the quantity of output Q is **(technologically) feasible** if and only if

$$Q \leq f(L, K)$$

The set of feasible output quantities given (L, K) pairs is also called the **production set**. The function that specifies the maximum amount of output that can be produced using L, K, i.e., $f(L, K)$, is the **production function**.

Graphically, the production function is just the boundary of the production set. It is difficult to visually represent a function with two inputs,[1] so we will typically keep one input fixed and plot the production set by varying the other input. For instance, keep capital fixed at $\bar{K}$. Then the production set and the production function $f(L, \bar{K})$ are illustrated as in Figure 1.

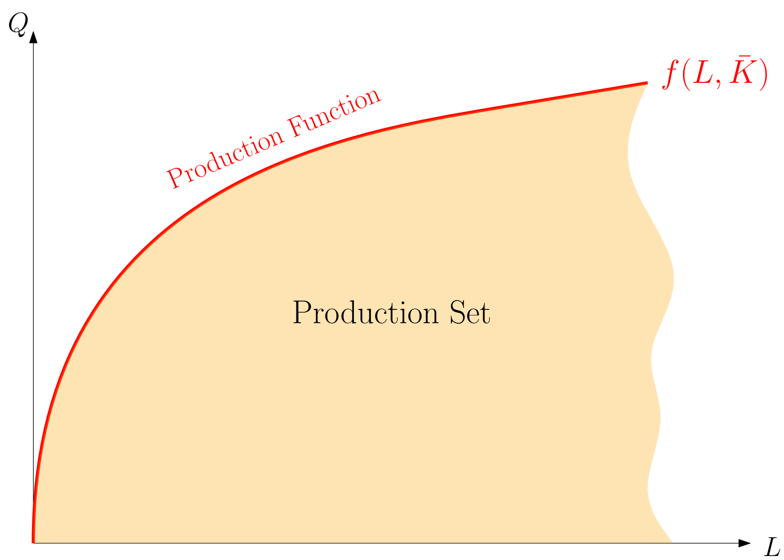

Figure 1: An example of a production set.

At this point, it should be clear that a production set plays the role of a budget set in the consumer theory (defining what is feasible). Similarly, the production function is like the budget line: the maximum feasible

[1] But not impossible! You can illustrate them through what people call **isoquants**. See the Appendix of Chapter 11 in your textbook.

thing given the constraints. Of course, you will see that the firm's optimal choice always lies on this curve, just like the consumer's optimal bundle always lies in the curve given by the production function.

As usual, we will impose some assumptions on the production function. The following is the first property we impose.

- The production function is **monotonic**: if you increase the amount of at least one input, you can produce at least as much output as you were producing before.

  Formally, for any $L' \geq L$ and $K' \geq K$,

  $$f(L', K) \geq f(L, K) \quad \text{and} \quad f(L, K') \geq f(L, K)$$

  This makes a lot of sense: if you put more inputs you should be able to produce at least as many outputs. If you need a further defense of monotonicity, what is hidden under this property is the assumption of **free disposal of inputs**. The firm may always hire the extra worker. Even if that extra worker reduces the total output, the firm can just tell that worker to stay behind and not do anything, effectively not using that worker. In that case, the firm has disposed this particular worker and managed to keep the same amount of output as before.

  Geometrically, monotonicity imposes the production function to be **(weakly) increasing** in both L and K, just like we drawn in Figure 1.

  As you probably see, this is analogous to the monotonicity of preferences ("more is better") in the consumer theory. The only difference here is that we impose this on the constraint, not on the preferences.

3.1.1 Marginal Product of Inputs

The next property we will impose on the production function requires us to define a new object.

The **marginal product of an input** is the rate at which output increases when the "last" unit of the given input is used in production. A little less formally, this is the additional output produced as a result of employing the last "unit" of input.

Suppose K units of capital are hired, and we are considering the the marginal product of L-th unit of labor. The additional output produced as employing L units of labor rather than $L - \Delta L$ units of labor is:

$$f(L, K) - f(L - \Delta L, K)$$

Therefore, additional output *per units of labor hired* is:

$$\frac{f(L, K) - f(L - \Delta L, K)}{L - (L - \Delta L)} = \frac{f(L, K) - f(L - \Delta L, K)}{\Delta L}$$

To express the marginal product of L-th unit of labor when K units of capital hired, $MPL(L, K)$, we take the smallest increments possible. See Figure 2.

$$MPL(L, K) = \lim_{\Delta L \to 0} \frac{f(L, K) - f(L - \Delta L, K)}{L - (L - \Delta L)} = \lim_{\Delta L \to 0} \frac{f(L, K) - f(L - \Delta L, K)}{\Delta L}$$

This is just the slope of $f(L, \bar{K})$ at L. Thus, when $MPL(L, K)$ is larger, $f(K, L)$ is steeper. This means that adding an extra unit of labor increases productivity a lot.

When the units are finer, we assume that the production function is "smooth" so that its slope exists and is well-defined.

The marginal product of capital is defined analogously:

$$MPK(L, K) = \lim_{\Delta K \to 0} \frac{f(L, K) - f(L, K - \Delta K)}{\Delta K}$$

Now, we are ready to define the second property we impose on the production function.

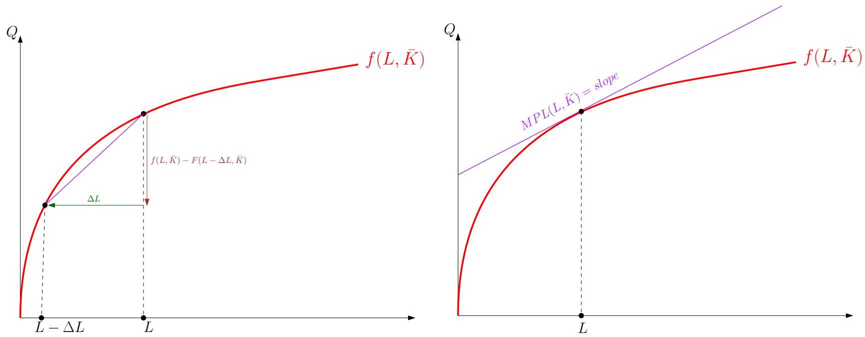

Figure 2: Marginal product of labor: Taking the limit as ΔL goes to zero in the figure on left, we obtain the figure on right.

- The production function satisfies **diminishing marginal product**: the marginal product of each input is decreasing in that input.

  This is sometimes called **the law of diminishing returns**. This is not really a law, it is just a common feature of many production functions. Why? Typically the first unit you use is very effective. But as you keep adding more and more units of the same input, their effectiveness is less prevalent. Ex: The first faculty member hired by Bilkent Economic department is very crucial: she will teach Econ 101, Econ 102, Econ 203... But the 24th faculty member will only teach some elective class if there is demand for such a class.

  Geometrically, diminishing marginal product states that the production function gets flatter as L increases. Figure 1 is not a coincidence!

  This is analogous to the diminishing MRS ("if you have more of it, the extra units you get is less valuable") in the consumer theory. Once again, the difference here is that we impose this on the constraint, not on the preferences.

  Note: Your textbook defines diminishing marginal product as the requirement that "the marginal product is diminishing *at some point*". That is, your textbook allows for MPL to be increasing for small values of L.

To wrap up: firm's constraints are given by a production set, whose boundaries is given by the production function. We assume that the production function satisfies monotonicity and diminishing marginal product of inputs.

3.2 The Preferences

We will assume that a typical firm maximizes its **profit**. This is a reasonable case for many firms in the economy. Formally,

$$profit = revenue - costs$$

The Greek letter π is typically used to denote **profit**.

If the firm produces and sells Q units of output in a given time period at a price of P per unit, its **revenue** is given by $P \cdot Q$.

If the firm uses L units of labor to produce such output at a cost of w per unit of labor per time, its **labor costs** are $w \cdot L$. Similarly, if the firm uses K units of capital to produce such output at a cost of r per unit of capital per time, its **capital costs** are $r \cdot K$. All in all, the firm's **costs** from using inputs (L, K) is: $w \cdot L + r \cdot K$.

Therefore, if the firm uses a pair of inputs (L, K) and produces Q units of outputs, its profit is:

$$\pi(Q, L, K) = P \cdot Q - (w \cdot L + r \cdot K)$$

The firm's preferences are very simple: for any two combinations (Q, L, K) and (Q', L', K'), the firm (weakly) prefers the former over the latter when $\pi(Q, L, K) \geq \pi(Q', L', K')$. That is, the firm prefers to choose combinations of inputs and outputs that yield higher profits.

3.2.1 Costs of Inputs

We have extensively talked about P when discussing consumer theory. How about w and r?

- w is somewhat easier to interpret: it is the **wage per time** paid on labor hired. If the pizza parlor hires 3 workers in a given day ($L = 3$), and if the daily wage is 98.10 TL/day ($w = 98.10$), then the labor cost is $w \cdot L = 294.30$ TL/day.
- r is a little bit more involved: it is the **rental rate of capital per time** that needs to be paid on capital used in production. The easiest way to think about this is imagining that the firm rents (does not own) the capital it uses. For instance, imagine that the pizza parlor does not own its pizza ovens, but rents it out from a wholesale pizza oven operator. If the pizza parlor rents two pizza ovens in a given day ($K = 2$) and if the **rental rate of a pizza oven per day** is 150 TL/day, then the capital cost is $r \cdot K = 300$ TL/day.

Opportunity Costs One thing that I want to emphasize is: **all the costs considered here include opportunity costs.** For a refresher on the concept, you should revisit Chapter 1. To reiterate what we had there: the wage rate includes the best alternative you give up by employing the extra unit of labor. For instance, imagine the owner of the firm is self-employed: he does not pay any salary to herself. An accountant would say that the wage paid for the labor supplied by this worker is zero. Yet, the economic logic suggests that it is not zero, because it involves the opportunity cost. Suppose, if the firm did not exist, the owner would join the labor force and find a job paying a salary of 15000 TL/month. So, by employing herself, she forgoes a salary of 15000 TL/month. Therefore, the opportunity cost is 15000 TL/month, and that should be the labor cost associated with that employee.

A similar reasoning goes for capital: if the firm owns the pizza ovens and can operate them freely, the explicit cost (i.e., the accounting cost) is zero but the opportunity cost not. This is because if the firm did not operate the ovens, it would be able to rent out these ovens and obtain a return of r. Thus the opportunity cost of using a pizza oven is still r, even though the firm does not make any explicit payments on it. Therefore, the rental rate of capital enters into the calculus even when the firm owns the capital.

Long story short: due to the costs being opportunity costs, the ownership of labor or capital (whether they are owned by the firm or merely hired/rented by the firm) does not really matter.

Sunk Costs What do these costs **not** include? They don't include the **sunk costs**. These are the costs that are already paid before the production decisions and cannot be recovered. Suppose, for instance, the firm paid some installation fee for the ovens: it hired some contractors to transport and install the ovens months ago. Because those costs are already paid, they are not included in the calculation of the cost of capital.

The concept of **sunk cost** is an important economic concept: it is useful to know that a rational economic agent would never take those costs into account. For instance, if you already paid for your gym membership fee, that amount you already paid should not affect your decision to go to the gym or not. Similarly, if you already bought a concert ticket, but a better alternative came up, you should not say "But I already paid for this concert ticket so I should go to the concert." Still, many economic agents are not that rational and cannot stop themselves from taking sunk costs into account when deciding. This phenomenon is aptly called the **sunk cost fallacy**. Some examples:

- A lot of people hesitate to finish their relationships because "They invested a lot in that relationship."
- Sometimes governments stick with a policy that took a lot of effort into passing, even though that policy turns out to be not very desirable.

Long story short: the costs a firm consider include opportunity costs, and exclude sunk costs.

3.2.2 Perfectly Competitive Markets

Okay, so there are prices of output and inputs in the markets the firm operate in. But what are these prices (P, w and r) determined? We will discuss this in more detail in the upcoming chapters, but for now, what you need to know is that **the firm takes these prices as given**.

The running assumption behind this statement is the following: the firm operates in a **perfectly competitive market**. We discussed this concept in Chapter 3 when we discussed exchange economies. Now we are revisiting this notion for a market with firms.

To reiterate, I know that the term "competitive" has a negative connotation: it suggests the existence of some very aggressive economic agents. Nevertheless, an economist's definition of "competitive" is much more innocuous. It merely means that there are many, many firms available in the market, and none of them are large enough to affect the prices. Formally:

Definition 2. *A* **perfectly competitive market** *for firms is a market where: (i) there are many firms selling an identical good or service to consumers, and (ii) an individual firm or consumer is not powerful enough to affect the price.*

So, following the pizza parlor example: consider a large town with many pizza parlors which produce and sell identical medium pepperoni pizzas. Our pizza parlor is a small one, so it cannot unilaterally say "I decided to sell my medium pepperoni pizzas at a different price than the other pizza parlors." Effectively, the pizza parlor is a price-taker: it cannot control the price of pizza, the only thing it can control is how many pizzas to produce.

This is the most important aspect of perfectly competitive markets: the economic agents in perfectly competitive markets are **price-takers**. We had already assumed this for consumers when we were covering consumer theory and exchange economies: we assumed that a consumer cannot affect the prices. Now we are making the same assumption for firms: a firm merely takes P, w and r as given.

Are markets perfectly competitive in real life? Most of the time, no. The closest we probably get to a competitive market is an agriculture market with many small farmers. The small apple farms in rural Ohio (i) all produce (more or less) identical products, and (ii) do not have the power to unilaterally affect the price of apple. But beyond those rare cases, most of the time the producers have some leeway in setting their prices. Yet, this is a benchmark we want to analyze first.

To be honest, most of economics deals with studying deviations from the perfectly competitive markets benchmark. The deviations are both more realistic and more interesting! But to study deviations, we need to understand the benchmark first. Hence the study of perfectly competitive markets.[2]

3.2.3 Isoprofit Curves

To summarize what we said so far: the firm takes P, w and r as given, and chooses (Q, L, K) to maximize its profit.

As you can imagine by now, there is a nice geometric illustration of the firm's preferences. These are illustrated through what we call **isoprofit line**. A small reminder: in Greek, *iso* means *equal*, so an isoprofit line is the set of points (Q, L, K) that all yield the same profit. Formally, for any two points (Q, L, K) and (Q', L', K') on an isoprofit line, we have: $\pi(Q, L, K) = \pi(Q', L', K')$.

A couple of points worth emphasizing:

- Recall that the firm wants to maximize its profit. Therefore, if two input-output combinations yield the same profit, the firm is indifferent between them. Therefore, the firm is indifferent between any two points on an isoprofit line. **Isoprofit lines are just the indifference curves for a firm.**

[2] It also deserves some emphasis that we are not making any value judgments at this stage: I am not saying "a perfectly competitive market is an ideal we want to achieve." A lot of times, it is not. We are just saying that this is a benchmark.

- Because illustrating things in three dimensional graphs are difficult, let us fix capital at a level $\bar{K}$ again. In this case, the set of points $(Q, L, \bar{K})$ that constitute an isoprofit line is the set of points where the profit $\pi(Q, L, \bar{K})$ is equal to a certain number π:

$$\pi(Q, L, \bar{K}) = \pi \iff P \cdot Q - (w \cdot L + r \cdot \bar{K}) = \pi$$
$$\iff Q = \frac{\pi}{P} + \frac{w}{P}L + \frac{r}{P}\bar{K}$$

  Therefore, in a two-dimensional graph with L in the x-axis and Q in the y-axis, and isoprofit line is a line with **intercept** $\frac{\pi}{P} + \frac{r}{P}\bar{K}$ and **slope** $\frac{w}{P}$. It is positively-sloped, because a higher L is more costly. For the profit to remain constant, the revenue must be higher as well, i.e. Q must be higher if L is higher.

  Figure 3 illustrates an isoprofit line.

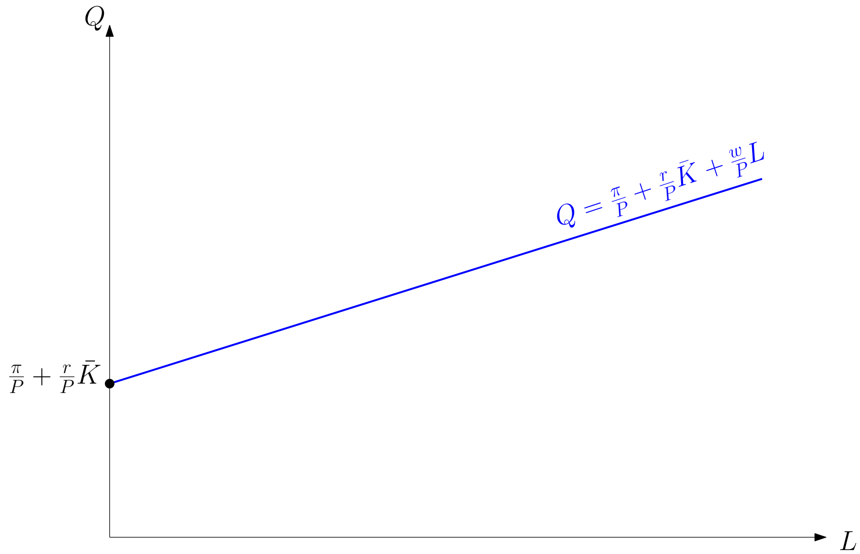

Figure 3: An isoprofit line is a line with an intercept of $\frac{\pi}{P} + \frac{r}{P}\bar{K}$ and a slope of $\frac{w}{P}$.

- Just like we plot multiple indifference curves, we can also plot multiple isoprofit lines. See Figure 4. It plots three different isoprofit lines, which correspond to different profit levels with $\pi_3 > \pi_2 > \pi_1$. Note that they are parallel to each other (after all, the slope of an isoprofit line is $\frac{w}{P}$, which the firm takes as given). They only differ in their intercepts, where a higher π corresponds to a higher intercept. Recall that the firm prefers to have higher profit, and therefore "higher" isoprofit lines correspond to input-output combinations that are more preferable for the firm. This is just like how "higher" indifference curves are more preferable to the consumer.

- To sum up, the firm just wants to find the "highest" isoprofit line it can find (subject to technological constraints), just like the consumer wanted to find the "highest" indifference curve she could afford (subject to budget constraints).

4 Profit Maximization

Okay, now we have a definition of firm's constraints and preferences. We even have a graphical representation of them! The next step is finding the input-output combinations which yield the highest profit. Before we do that, we will take a small detour.

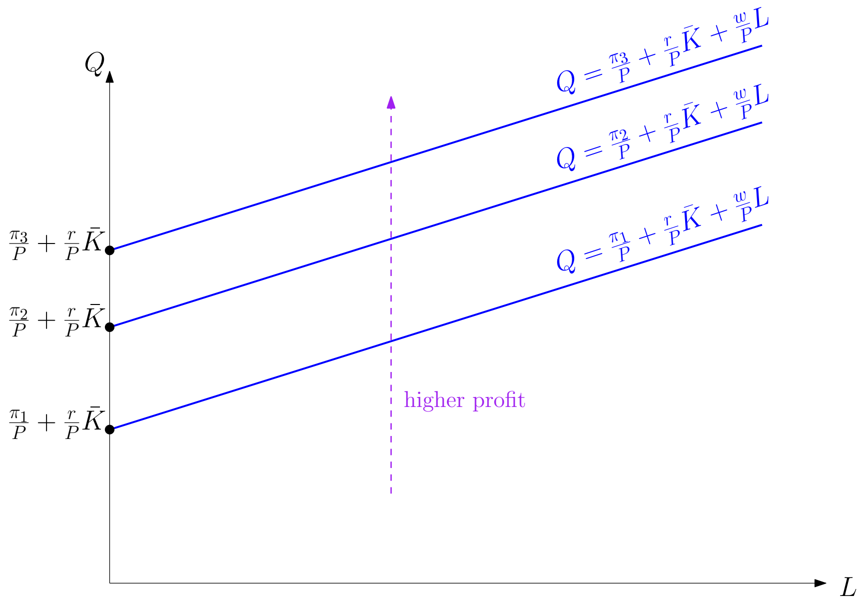

Figure 4: Multiple isoprofit lines corresponding to multiple profit levels, with $\pi_3 > \pi_2 > \pi_1$. The "higher" isoprofit lines correspond to higher profits, and therefore are more preferable to the firm. (Just like how "higher" indifference curves are more preferable to the consumer.)

4.1 Short-Run versus Long-Run

So far in our discussions, we usually kept capital fixed at $\bar{K}$ and allowed L to vary. One reason behind this choice is the easiness of illustration. But there is a deeper reason: there is indeed a substantial difference between capital and labor. It is the following: **the firm can easily change the quantity of labor it uses, but adjusting the quantity of capital takes time**. This is because hiring/firing labor, asking the workers to work extra shifts etc. is much easier to do than buying/selling/building/destroying machines and buildings. Therefore, the time frame used in the analysis matters.

- **Short run:** The time frame where the quantity of capital used in production cannot change.
- **Long run:** The time frame where all the inputs used in production can change.

The factors of production that cannot be adjusted in the short run are sometimes referred to as **fixed factors**. The factors of production that can be adjusted in the short run are referred to as **variable factors**. The terminology of *fixed versus variable* will be important in a few pages.

4.2 Short-Run Profit Maximization

In the short-run, the firm keeps capital fixed at $\bar{K}$ and only chooses (Q, L).

Given the output price P, input prices w, r, and a capital level $\bar{K}$, an input-output combination (Q^*, L^*) *profit-maximizing in the short-run* if and only if

- $Q \leq f(L^*, \bar{K})$ (i.e., (Q^*, L^*) is feasible), and
- for any input-output combination (Q, L), if $Q \leq f(L, K)$ (i.e., if (Q, L) is feasible), then $\pi(Q^*, L^*, \bar{K}) \geq \pi(Q, L, \bar{K})$ (i.e., (Q^*, L^*) yields at least as high a profit as any feasible input-output combination).

There is actually a mathematical derivation of the profit-maximizing input-output combination (just like the derivation of the optimal bundle in consumer theory), but we will skip it because we will cover a similar argument in a few pages. Instead, we will only give a geometrical argument.

Geometrically, the firm is trying to find the highest isoprofit line it can find, but has to adhere by the production set (i.e., (Q^*, L^*) has to satisfy $Q \leq f(L^*, \bar{K})$). **First** of all, a visual inspection suffices to argue that $Q = f(L^*, \bar{K})$, i.e.,the firm must be choosing an input-output combination in the boundary of the production set.

Next, imagine the (Q, L) combination on the production function that lies on the highest isoprofit line. You will see that at that point, the isoprofit line barely touches the production function, i.e., it is tangent to the production function. See Figure 5 below.

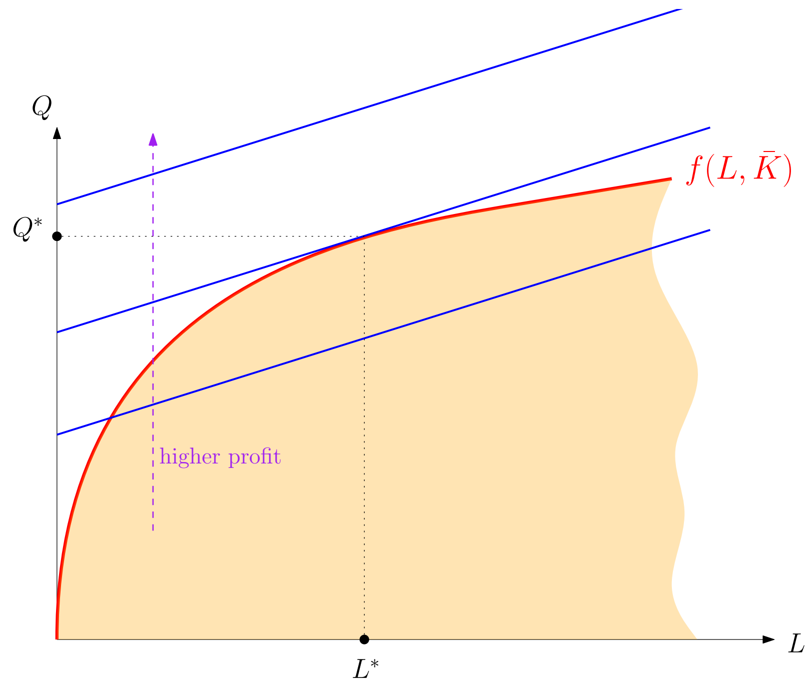

Figure 5: The profit-maximizing input-output combination (Q^*, L^*) is on the production function, and the isoprofit line passing through it is tangent to the production function.

Here is the interesting thing about this figure: at (Q^*, L^*), the isoprofit line is tangent to the production function. Therefore, slope of production function = slope of the isoprofit line.

- The slope of the production function at L^* is: $MPL(L^*, \bar{K})$.
- The slope of isoprofit line is: $\frac{w}{P}$.

Therefore, the profit-maximizing input-output combination satisfies:

$$MPL(L^*, \bar{K}) = \frac{w}{P}$$

Rearrange this to get:

$$MPL(L^*, \bar{K}) \cdot P = w \tag{1}$$

The left-hand side of this equation is "the revenue brought by the last unit of labor hired" (the output produced by the last unit times the output price). It is sometimes referred to as the **marginal revenue product of labor** or **value of the marginal product of labor (VMPL)**.

The right-hand side is simply the cost of the last unit of labor hired.

Therefore, the profit-maximizing input-output combination has the following feature: **the revenue brought by the last unit of labor hired is equal to the cost of the last unit of labor hired**. Intuitively: the firm starts by hiring some small amount of labor. As long as the next unit of labor hired brings higher revenue than its cost, the firm keeps hiring. But remember: the marginal product of labor decreases as more labor is hired. Therefore, the revenue brought by the last unit of labor keeps decreasing, and the firm stops at the point where it is no longer profitable to hire the next unit of labor.

What is nice about this graphical interpretation is that: you can also change the parameters of the model to see how (Q^*, L^*) changes!

- Suppose, for instance, w is higher. You can easily see that this means: isoprofit lines are steeper. Just play around with Figure 5 or Equation (1) so see that: L^* is lower when w is higher! This makes perfect sense: labor is more costly, so the firm hires less labor (or, stops earlier when hiring labor).
- Suppose, for instance, P is higher. You can easily see that this means: isoprofit lines are flatter. Just play around with Figure 5 or Equation (1) so see that: L^* is higher when P is higher! Since $Q^* = f(L^*, \bar{K})$, Q^* is also higher when P is higher. This also makes perfect sense: output is more expensive, so the firm hires produces more (or, stops later when hiring labor) and sells more.

  Indeed, following this argument, one can construct a **firm supply schedule** which gives a list of profit-maximizing P, Q^* pairs. One can then draw a **firm supply curve**. This is eventually what we will do, but let's not get ahead of ourselves.

4.3 Long-Run Profit Maximization

Profit maximization in the long-run is very similar to the profit maximization in the short-run, except that the firm also chooses K in the long-run. But the basic idea remains the same.

The profit-maximizing input-output combination (Q^*, L^*, K^*) now satisfies two equalities:

$$MPL(L^*, K^*) \cdot P = w$$
$$MPK(L^*, K^*) \cdot P = r$$

Thus, the marginal revenue brought by each input must be equal to the cost of the respective input. Hopefully, this convinced you a little bit more towards how valuable *thinking at the margin* is. It is a powerful tool of analysis!

It is worth noting that the firm chooses Q^*, L^* and K^* simultaneously: after all, this is an optimization problem with three variables. Yet, these equations carry a sense of *the firm choosing* (L^*, K^*) *first and then producing the maximum output* Q^* *that can be produced with* (L^*, K^*)*, i.e.,* $Q^* = f(L^*, K^*)$. Let me reiterate: this is just an illusion, this is not what these equations mean, but this is typically how people think about these optimality conditions.

Next, we will turn this around and imagine that the firm chooses Q^*. We will essentially be writing down the same optimization problem, but there will be a single decision variable Q. To this end, we will define the cost of producing Q. This cost will include the amount of (L, K) necessary to produce Q, and the costs of those inputs (w, r). But we will sideline this at this stage (see the Appendix of Chapter 11 in your book if you wonder about it), and just take the cost function as given.

5 The "Cost Approach" to Production

The key object of analysis in this section will be the cost of producing Q. The cost is implicitly derived through production function. Formally:

Definition 3. *The* **cost function** *is the relation between the quantity produced* Q *and the cost of the optimal combination of inputs needed to produce the given quantity.*

A brief reminder that the costs that the firm considers are *opportunity costs*: they include the best alternative use of inputs (working elsewhere, renting away capital etc.) Moreover, they don't include *sunk costs*.

5.1 Fixed and Variable Costs

The total cost of producing Q units of output with the optimal combination of inputs is the **total cost**, denoted by $TC(Q)$.

We now decompose $TC(Q)$ into two components:

1. **Fixed Cost:** This is the component of total cost that does not change with the quantity produced. The fixed cost is due to the **fixed factors of production** (i.e., capital in the short run).

   Because the fixed cost, by definition, does not depend on Q, we will denote it with FC.

2. **Variable Cost:** This is the component of the total costs that changes as the quantity produced changes. The variable cost is due to the **variable factors of production** (i.e., labor in the short run).

   We will denote the variable cost with $VC(Q)$.

All in all, we have:

$$\text{Total Cost} = \text{Fixed Cost} + \text{Variable Cost}$$

or, to use the notation,

$$TC(Q) = FC + VC(Q)$$

A brief remark: As we discussed before, in the long run, all factors of production are variable. Thus, in the long run, all costs are variable costs, i.e., $FC = 0$ in the long run.

5.2 Marginal Cost

As I have hopefully convinced you by now, it is useful to think at the margin! To incorporate marginal thinking into firm's problem, we define **marginal cost**.

Formally, **Marginal Cost** of producing the Q-th unit is the rate at which cost changes when output increases by a "small" amount so that the final output is Q units. We denote the marginal cost function with $MC(Q)$.

The increase in total cost when Q units of output rather than $Q - \Delta Q$ units is produced is:

$$TC(Q) - TC(Q - \Delta Q)$$

Therefore, the increase in total cost *per unit* is:

$$\frac{TC(Q) - TC(Q - \Delta Q)}{Q - (Q - \Delta Q)} = \frac{TC(Q) - TC(Q - \Delta Q)}{\Delta Q}$$

To express the marginal cost of producing Q-th unit, we take the smallest increments possible.

$$MC(Q) = \lim_{\Delta Q \to 0} \frac{TC(Q) - TC(Q - \Delta Q)}{\Delta Q} = \frac{dTC(Q)}{dQ}$$

This is literally the derivative of $TC(Q)$. Because FC is a constant, we also have:

$$MC(Q) = \frac{dTC(Q)}{dQ} = \frac{d(FC + VC(Q))}{dQ} = \frac{dFC}{dQ} + \frac{dVC(Q)}{dQ} = \frac{dVC(Q)}{dQ}$$

So, $MC(Q)$ is also the derivative of $VC(Q)$.

Below is a very crucial observation about the marginal cost:

- If the production function satisfies diminishing marginal product, then $MC(Q)$ is increasing. So it looks like the one in Figure 6.

  Intuitively, this is because each additional input is less productive. Therefore, it takes more inputs to produce the marginal output. Thus, the cost of producing the last unit of output is higher.

  Note: Your textbook assumes "diminishing marginal product **at some point**" (this is a natural implication of assuming "diminishing marginal product *at some point.* So, according your textbook's definition, that $MC(Q)$ is increasing **after some value of** Q, i.e., it may be decreasing at first. In these lecture notes, I will adopt a simpler requirement and assume that $MC(Q)$ is always increasing. But the overall conclusions would not change.

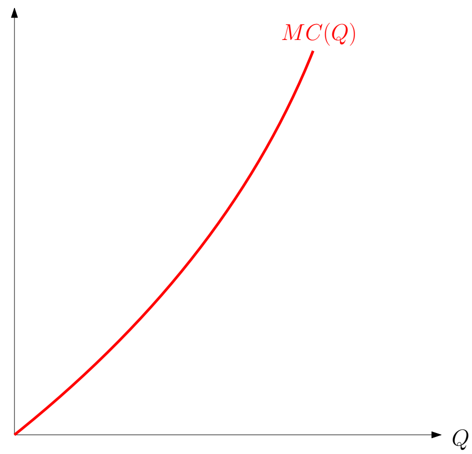

Figure 6: An increasing marginal cost function.

5.3 Average Costs

Beyond marginal cost, it turns out it is also useful to think about "averages costs", i.e. costs per unit produced. Please note that average is different than marginal: this is **not** the cost of **last** unit produced, but the average of **all** units produced!

To calculate the average costs, we take the equation:

$$TC(Q) = FC + VC(Q)$$

and divide it by Q to calculate average costs:

$$\frac{TC(Q)}{Q} = \frac{FC}{Q} + \frac{VC(Q)}{Q}$$

We now define:

- **Average Total Cost** of producing Q units: total cost of producing Q units divided by Q.

$$ATC(Q) = \frac{TC(Q)}{Q}$$

- **Average Fixed Cost** of producing Q units: the fixed cost divided by Q.

$$AFC(Q) = \frac{FC}{Q}$$

- **Average Variable Cost** of producing Q units:variable cost of producing Q units dived by Q.

$$AVC(Q) = \frac{VC(Q)}{Q} \quad .$$

Therefore, we have:

$$ATC(Q) = AFC(Q) + AVC(Q)$$

A couple of notes about average costs are in order.

- By construction, $AFC(Q) = \frac{FC}{Q}$ is decreasing in Q.
- $AVC(Q)$ is the average of $MC(Q')$ of all $Q' \leq Q$.
  - To see this, first realize that: $VC(Q)$ is the sum of marginal costs of first Q units.
  - Then, $AVC(Q)$ is the average of the marginal costs of first Q units.

  Since $MC(Q)$ is increasing, $AVC(Q)$ is increasing in Q. (When calculating $AVC(Q)$, for larger values of Q, the firm is taking average of more items, where the items added later are larger quantities.)
- $AC(Q)$ is the sum of $AFC(Q)$ and $AVC(Q)$. Thus, it is first decreasing and then increasing in Q.

In light of these discussions, we end up with a figure like Figure 7.

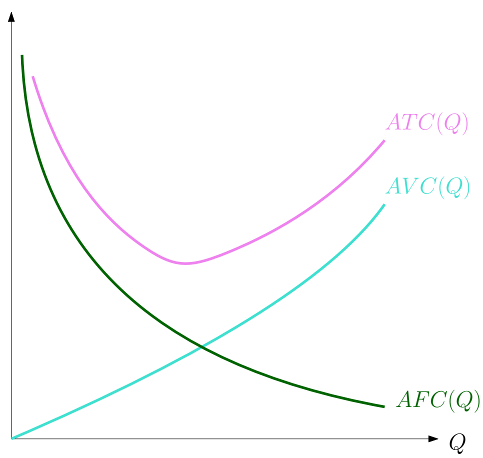

Figure 7: Average cost functions when $MC(Q)$ is increasing.

5.4 Putting Marginal and Averages Together

To sum up, here is what he have:

$$\begin{aligned}
MC(Q) &= \frac{d}{dQ}TC(Q) = \frac{d}{dQ}\left(FC + VC(Q)\right) = \frac{d}{dQ}VC(Q) \\
AFC(Q) &= \frac{FC}{Q} \\
AVC(Q) &= \frac{VC(Q)}{Q} \\
ATC(Q) &= \frac{TC(Q)}{Q} = \frac{FC + VC(Q)}{Q} = \frac{FC}{Q} + \frac{VC(Q)}{Q} = AFC(Q) + AVC(Q)
\end{aligned}$$

where $MC(Q)$ and $AVC(Q)$ are increasing, $AFC(Q)$ is decreasing, and $ATC(Q)$ is first decreasing and then increasing.

Here is another set of critical observations:

- $AVC(Q) \leq MC(Q)$ for every Q.
  - This is because, when calculating $AVC(Q)$, the firm is taking average of $MC(Q')$ for all $Q' \leq Q$.

– Since marginal cost is increasing, $MC(Q') \leq MC(Q)$. Therefore, the firm is taking average of a bunch of items, each less than $MC(Q)$.

- $MC(Q)$ intersects $ATC(Q)$ at the point where $ATC(Q)$ is minimized.

  We can prove this by taking derivatives and using simple calculus. Let Q^* be the point where $ATC(Q)$ reaches its minimum value. If the cost functions are "smooth",

  $$\frac{d}{dQ}ATC(Q^*) = 0$$

  By definition of $ATC(Q)$, this is equivalent to:

  $$\frac{d}{dQ}\left(\frac{TC(Q^*)}{Q^*}\right) = 0$$

  which is equivalent to:

  $$\frac{\frac{d}{dQ}TC(Q^*)Q^* - TC(Q^*)}{(Q^*)^2} = 0$$

  which implies:

  $$\frac{d}{dQ}TC(Q^*)Q^* - TC(Q^*) = 0$$

  By rearranging,

  $$\frac{d}{dQ}TC(Q^*) = \frac{TC(Q^*)}{Q^*}$$

  But recall that $\frac{d}{dQ}TC(Q^*) = MC(Q^*)$ and $\frac{TC(Q^*)}{Q^*} = ATC(Q^*)$. Therefore, at the quantity Q^* where $ATC(Q)$ attains its minimum value,

  $$MC(Q^*) = ATC(Q^*)$$

  Pretty cool, huh? Not really surprising, though: for $Q \leq Q^*$, $MC(Q)$ is small so it is pulling $ATC(Q)$ downwards. For $Q \geq Q^*$, $MC(Q)$ becomes large enough so that it pulls $ATC(Q)$ upwards. At $Q = Q^*$, $MC(Q)$ pulls $ATC(Q)$ neither downwards nor upwards.

All in all, the cost curves look like the ones in Figure 8.

6 Profit Maximization Using Cost Funcstions

Okay, now let's put whet we have to work. Recall that the firm chooses its quantity Q to maximize its profit, where:

$$\text{Profit} = \text{Revenue} - \text{Cost} \ .$$

and

$$\text{Revenue} = \text{Price} \times \text{Quantity} \ .$$

We know the "cost" and "quantity", but what about the price? If the firm would like to sell Q units, it must set its price according to *demand*. Recall that the demand curve gives the relationship between quantity demanded (Q) and prices (P).

Consider the demand curve for a firm (which may be different from the demand curve for an entire market). This curve gives a the price of the good as a function of the quantity produced. Let $P(Q)$ denote the price level at which Q units of the good will be demanded.

Now,

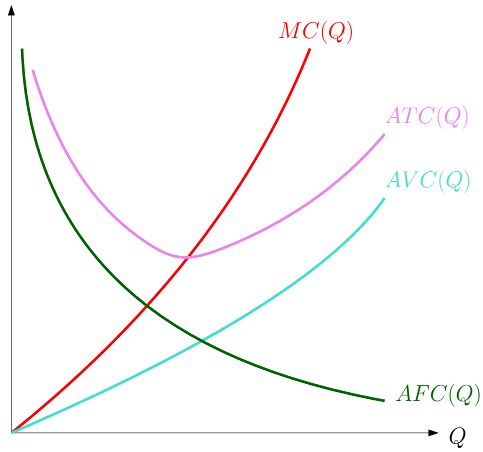

Figure 8: Cost functions when $MC(Q)$ is increasing.

- The **profit** from producing and selling Q units of output is: $\pi(Q)$
- The **total revenue** from selling Q units of outputs is: $TR(Q) = P(Q) \cdot Q$.
- The **total cost** of producing Q units of output is: $TC(Q)$.

Then,

$$\pi(Q) = TR(Q) - TC(Q)$$

6.1 Marginal Revenue

Should I remind you about the value of thinking at the margin?

So profit has two components: revenue and cost. We know what marginal cost is. How about we define **marginal revenue**?

Formally, **Marginal Revenue** brought by the Q-th unit is the rate at which revenue changes when output increases by a "small" amount so that the final output is Q units. We denote marginal revenue function with $MR(Q)$.

The increase in total revenue when Q units of output rather than $Q - \Delta Q$ units is produced is:

$$TR(Q) - TR(Q - \Delta Q)$$

Therefore, the increase in total revenue *per unit* is:

$$\frac{TR(Q) - TR(Q - \Delta Q)}{Q - (Q - \Delta Q)} = \frac{TR(Q) - TR(Q - \Delta Q)}{\Delta Q}$$

To express the marginal revenue of producing Q-th unit, we take the smallest increments possible.

$$MR(Q) = \lim_{\Delta Q \to 0} \frac{TR(Q) - TR(Q - \Delta Q)}{\Delta Q} = \frac{dTR(Q)}{dQ}$$

6.2 Profit-Maximizing Quantity

The firm is choosing Q to maximize:

$$\pi(Q) = TR(Q) - TC(Q)$$

How do we go around solving this problem? The classical method is taking the derivative and identifying the first-order condition. I will skip the steps of the argument here, but you should realize that if $Q^* > 0$ is a profit-maximizing quantity, then it should satisfy:

$$\begin{aligned}\frac{d\pi(Q^*)}{dQ} = 0 &\implies \frac{d(TR(Q^*) - TC(Q^*))}{dQ} = 0 \\ &\implies \frac{dTR(Q^*)}{dQ} - \frac{TC(Q^*)}{dQ} = 0 \\ &\implies MR(Q^*) - MC(Q^*) = 0 \\ &\implies MR(Q^*) = MC(Q^*)\end{aligned}$$

Let's express this as a theorem. (I relegate the proof to the Appendix in case you are interested.)

Theorem 1. *If Q^* is a profit-maximizing quantity and $Q^* > 0$, then*

$$MR(Q^*) = MC(Q^*) \tag{2}$$

Equation (2) is a very famous equality. It is colloquially called the **golden rule of profit maximization**!

The intuitive way to understand Theorem 1 is to visualize the following process. The firm keeps producing and selling output as long as the revenue brought by the last unit is higher than its cost. The firm stops at the point where the marginal revenue is not higher any more. Under reasonable conditions (such as increasing marginal cost and decreasing marginal revenue), this process stops at some point.

7 Firms in a Perfectly Competitive Market

Suppose the firm operates in a perfectly competitive market. This means that the firm is a **price-taker**: it takes P as given and cannot change it. More formally, the firm's demand curve is a perfectly elastic curve, which is a horizontal line at P.

In this case, since $P(Q)$ does not depend on Q, we have:

$$TR(Q) = P \cdot Q$$

and

$$MR(Q) = P$$

Therefore, by Theorem 1 the profit-maximizing quantity $Q^* > 0$ satisfies:

$$P = MC(Q^*) \tag{3}$$

What does Equation (3) mean? Basically it says: "In a competitive market, if the firm is producing anything at all, it produces the quantity such that the price equals marginal cost. (We will take care of the "if the firm is producing anything at all..." part in the next subsection.)

In a competitive market, the firm keeps producing and selling output as long as the price is higher than the marginal cost. The firm stops at the point where the price equals the marginal cost. If marginal cost of increasing, this process stops at some point.

7.1 Shut Down Criterion

Let's go back to the annoying clause we had above: "if the firm is producing anything at all..."

Note that the firm always has the option of shutting down and not producing anything ($Q = 0$ is always an option for the firm). Clearly, this would keep $VC(0) = 0$: if the firm does not produce anything, it does not need to hire any variable factors of production. Still, in the short run, FC is still positive: even when the firm does not produce anything it needs to pay for the fixed factors of production. Thus, $TC(0) = VC(0) + FC = FC$. Therefore, shutting down yields a profit of $\pi(0) = P \cdot 0 - TC(0) = P \cdot 0 - FC = -FC$ in the short run.

When does the firm produce anything at all? Recall that the maximum profit the firm receives from producing anything at all is attained when the firm produces Q^*. This gives a profit of:

$$\pi(Q^*) = P \cdot Q^* - TC(Q^*) = P \cdot Q^* - (VC(Q^*) - FC)$$

The firm shuts down in the short run if not producing anything yields a higher profit than producing Q^*, i.e., if $\pi(0) > \pi(Q^*)$. This is the case when:

$$-FC > P \cdot Q^* - (VC(Q^*) - FC)$$

which is equivalent to:

$$0 > P \cdot Q^* - VC(Q^*)$$

which is equivalent to:

$$P < \frac{VC(Q^*)}{Q^*} = AVC(Q^*)$$

Therefore, **the firm chooses to shut down in the short run if the price is so low that it doesn't even cover the average variable costs**. The firm is getting a revenue of P per unit produced, and paying a variable cost of $AVC(Q^*)$ per unit! If $P < AVC(Q^*)$, there is no reason for the firm to produce anything: the revenue does not even cover the cost of variable inputs!

The flip side of this is: **the firm operates in the short run if** $P \geq AVC(Q^*)$.[3]

The condition of $P < AVC(Q^*)$ is sometimes referred as the **shut down criterion**: this is the condition for the firm to shut down in the short run. Note that the fixed costs are not included in this calculation. This is because they are like **sunk costs** in the short run: the firm has to pay them no matter what it produces, and therefore they should not be taken into account in the decision to produce or not in the short run.

Reassuringly, when $MC(Q)$ is increasing, the shut down criterion is never satisfied. This is because Q^* satisfies $P = MC(Q^*)$. But recall that $AVC(Q) \leq MC(Q)$ for all Q. Therefore,

$$AVC(Q^*) \leq MC(Q^*) = P$$

So, if $MC(Q)$ is increasing for all Q, we do not need to worry about the firm shutting down in the short run.

The shut down criterion becomes a larger concern if we have "eventually increasing marginal cost" as your textbook assumes. Then, $MC(Q)$ may not always be increasing, and we may have $AVC(Q) > MC(Q)$ for some values of Q. I am adding a discussion of eventually increasing marginal cost to the Appendix, if you are interested.

[3] When $P = AVC(Q^*)$, the firm is indifferent between producing or not. I am breaking the tie in favor of producing. Because this is a knife-edge case, it doesn't really matter.

Appendix

A An Example with Discrete Increases: Pizza Parlor

This section contains an example that may help with your understanding of production function, marginal product of inputs, and the cost functions. The good thing about it is: this is a concrete example. The bad thing is: because the inputs and outputs in this example are only in discrete amounts, we cannot get the "finest" increments.

I took this example from your textbook, but changed the numbers slightly.

Consider a pizza parlor that uses two inputs: workers (L, measured in the number of workers per day) and pizza ovens (K, measured in the number of pizza ovens used per day). It produces medium sized pepperoni pizzas (Q, measured in the number of pizzas per day). Since the units are discrete, the production function $f(L,K)$ can be represented through a table:

No. of Workers \ No. of Pizza Ovens	0	1	2	3
0	0	0	0	0
1	0	225	450	600
2	0	250	550	900
3	0	260	600	1100
4	0	265	625	1200
5	0	265	640	1250

As you can check, the production function satisfies monotonicity: employing more inputs generate (weakly) more outputs. Now, let us fix $\bar{K} = 2$: the pizza parlor has two pizza ovens it has to operate. then, we focus on the column with $\bar{K} = 2$.

No. of Workers \ No. of Pizza Ovens	0	1	2	3
0	0	0	0	0
1	0	225	450	600
2	0	250	550	900
3	0	260	600	1100
4	0	265	625	1200
5	0	265	640	1250

Graphically, this looks like Figure 9.

When the smallest increment we can have is one unit, the marginal product takes a simple form.

$$MPL(L,K) = \frac{f(L,K) - f(L-1,K)}{L-(L-1)} = f(L,K) - f(L-1,K)$$
$$MPK(L,K) = \frac{f(L,K) - f(L,K-1)}{K-(K-1)} = f(L,K) - f(L,K-1)$$

For instance, the Marginal Product of Labor (MPL) of the 3rd worker when using 2 ovens is:

$$MPL(3,2) = \frac{600-550}{3-2} = 50 \text{ pizzas per day per worker.}$$

Marginal product of the 2nd oven (when there are 3 workers) is

$$MPK(3,2) = \frac{600-260}{2-1} = 340 \text{ pizzas per day per oven.}$$

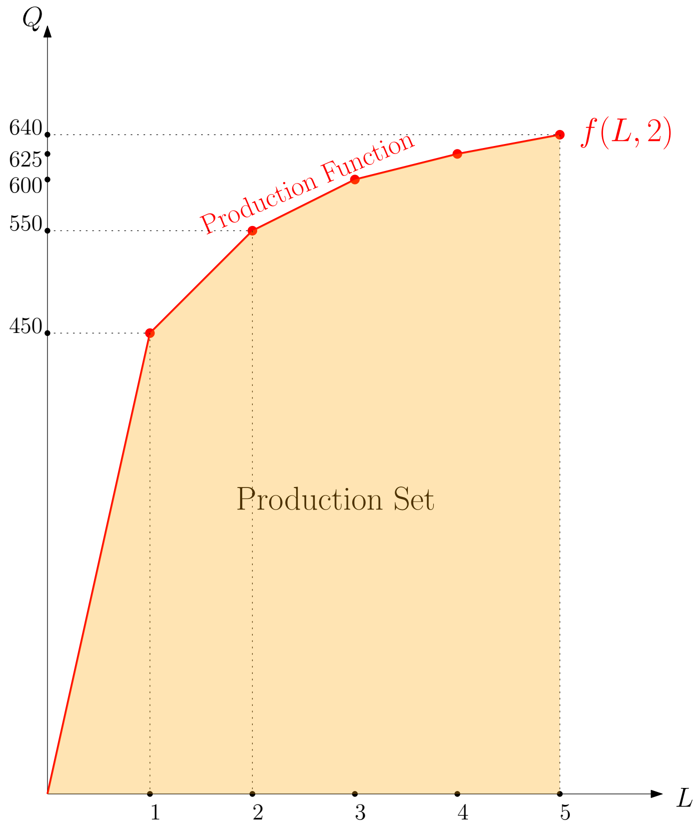

Figure 9: Production set in the pizza parlor example.

Here is a table that contains $MPL(L,2)$ for different values of L:

L (# of Workers)	Q (# of Pizzas)	$MPL(L,2)$
0	0	–
1	450	$\dfrac{450-0}{1-0}=450$
2	550	$\dfrac{550-450}{2-1}=100$
3	600	$\dfrac{600-550}{3-2}=50$
4	625	$\dfrac{625-600}{4-3}=25$
5	640	$\dfrac{640-625}{5-4}=15$

Therefore, the production function satisfies diminishing marginal product of labor.

Finally, let's tie this discussion into the cost functions. Suppose that the cost of operating an oven is 100 TL per day and the wage for labor is 20 TL per day. The pizza parlor owns 2 ovens and can not buy any more in a one week period (short-run analysis). Then, the fixed cost of the firm is $2 \times 100 = 200$.

The cost functions in this example look like:

# of Pizzas (Q)	FC	$VC(Q)$	$TC(Q)$
0	200	$0 \times 20 = 0$	$200 + 0 = 200$
450	200	$1 \times 20 = 20$	$200 + 20 = 220$
550	200	$2 \times 20 = 40$	$200 + 40 = 240$
600	200	$3 \times 20 = 60$	$200 + 60 = 260$
625	200	$4 \times 20 = 80$	$200 + 80 = 280$
640	200	$5 \times 20 = 100$	$200 + 100 = 300$

Based on this, we can calculate $ATC(Q)$, $AVC(Q)$ and $AFC(Q)$. They will have the same formula: just divide the respective costs with Q.

B Proof of Theorem 1

This proof assumes that total revenue and cost are "smooth" functions, and thus their derivatives exist.

Begin by noting:

$$\begin{aligned}\pi(Q) - \pi(Q - \Delta Q) &= TR(Q) - TC(Q) - (TR(Q - \Delta Q) - TC(Q - \Delta Q)) \\ &= TR(Q) - TR(Q - \Delta Q) - (TC(Q) - TC(Q - \Delta Q))\end{aligned}$$

Then, dividing both sides by ΔQ,

$$\frac{\pi(Q) - \pi(Q - \Delta Q)}{\Delta Q} = \frac{TR(Q) - TR(Q - \Delta Q)}{\Delta Q} - \frac{TC(Q) - TC(Q - \Delta Q)}{\Delta Q} \tag{4}$$

Suppose Q^* is a profit-maximizing quantity and $Q^* > 0$.

- Consider any $\Delta Q > 0$ such that $Q^* - \Delta Q > 0$. Q^* being the profit-maximizing quantity implies: $\pi(Q^*) \geq \pi(Q^* - \Delta Q)$. Rearranging:

  $$\pi(Q^*) - \pi(Q^* - \Delta Q) \geq 0$$

  Dividing both sides by $\Delta Q > 0$ gives:

  $$\frac{\pi(Q^*) - \pi(Q^* - \Delta Q)}{\Delta Q} \geq 0$$

  Substituting (4),

  $$\frac{TR(Q^*) - TR(Q^* - \Delta Q)}{\Delta Q} - \frac{TC(Q^*) - TC(Q^* - \Delta Q)}{\Delta Q} \geq 0$$

  Since the above inequality is true for any positive $\Delta Q < Q^*$, it is also true for "small" ΔQ. But this implies

  $$MR(Q^*) - MC(Q^*) \geq 0$$

  or

  $$MR(Q^*) \geq MC(Q^*) \tag{5}$$

- Now, consider $\Delta Q < 0$. Once again, Q^* being the profit-maximizing quantity implies: $\pi(Q^*) \geq \pi(Q^* - \Delta Q)$. Rearranging:

  $$\pi(Q^*) - \pi(Q^* - \Delta Q) \geq 0$$

  Dividing both sides by $\Delta Q < 0$ gives:

  $$\frac{\pi(Q^*) - \pi(Q^* - \Delta Q)}{\Delta Q} \leq 0$$

Substituting (4),

$$\frac{TR(Q^*) - TR(Q^* - \Delta Q)}{\Delta Q} - \frac{TC(Q^*) - TC(Q^* - \Delta Q)}{\Delta Q} \leq 0$$

Since the above inequality is true for any $\Delta Q < 0$, it is also true for "small" ΔQ. But this implies

$$MR(Q^*) - MC(Q^*) \leq 0$$

or

$$MR(Q^*) \leq MC(Q^*) \tag{6}$$

Combining (5) and (6) gives the following statement. If Q^* is a profit-maximizing quantity of output and $Q^* > 0$, then

$$MR(Q^*) \leq MC(Q^*) \text{ and } MR(Q^*) \geq MC(Q^*) \implies MR(Q^*) = MC(Q^*)$$

C "Eventually" Increasing Marginal Cost

A reminder that your textbook (and many textbooks) merely imposes "diminishing marginal product **at some point**" on the production function. This translates into "increasing marginal cost **at some point**", i.e., $MC(Q)$ only has to be increasing after some value of Q. That is, $MC(Q)$ only needs to be **eventually** increasing. This is a weaker requirement than increasing $MC(Q)$: it is sufficient if $MC(Q)$ is *eventually* increasing. (Indeed, "increasing marginal cost" is just a special case of "eventually increasing marginal cost" – it amounts to saying that "eventually" hits very early.)

From a broad point of view, not much changes by adopting the requirement of "eventually increasing $MC(Q)$".

- $AVC(Q)$ is *eventually* increasing in Q.
- Moreover, $AVC(Q) \geq MC(Q)$ for small values of Q and $AVC(Q) \leq MC(Q)$ for larger values of Q. In other words, $AVC(Q) \leq MC(Q)$ for sufficiently large values of Q.

All in all, when $MC(Q)$ is eventually increasing, the cost curves look like the ones in Figure 10.

Another observation: $MC(Q) = AVC(Q)$ at the point where $AVC(Q)$ is minimized, just like $MC(Q) = ATC(Q)$ at the point where $ATC(Q)$ is minimized. Once again, not a coincidence! We can indeed also prove this by taking derivatives and using simple calculus. I'm leaving it as an exercise.

C.1 Shut Down Criterion and "Eventually" Increasing Marginal Cost

When $MC(Q)$ is eventually increasing, we actually need to worry about the shut down criterion. Therefore, for P small enough so that $P < AVC(Q^*)$, the firm actually shuts down: the firm chooses $Q = 0$ instead of Q^*. This means, for small enough P, the firm produces nothing. Once P exceeds $AVC(Q^*)$, the firm chooses Q^* such that $P = MC(Q^*)$.

But when is $P < AVC(Q^*)$? Recall that $MC(Q^*) = P$, so that the shut down criterion is: $MC(Q^*) < AVC(Q^*)$. Now, recall that $MC(Q) = AVC(Q)$ at the point where $AVC(Q)$ is minimized. Therefore, $MC(Q) < AVC(Q)$ at the quantities smaller than the quantity that minimizes $AVC(Q)$. We end up with the following observation:

> In the short run, the firm shuts down if P is less than the minimum value of $AVC(Q)$.

See Figure 11 below, which should illustrate.

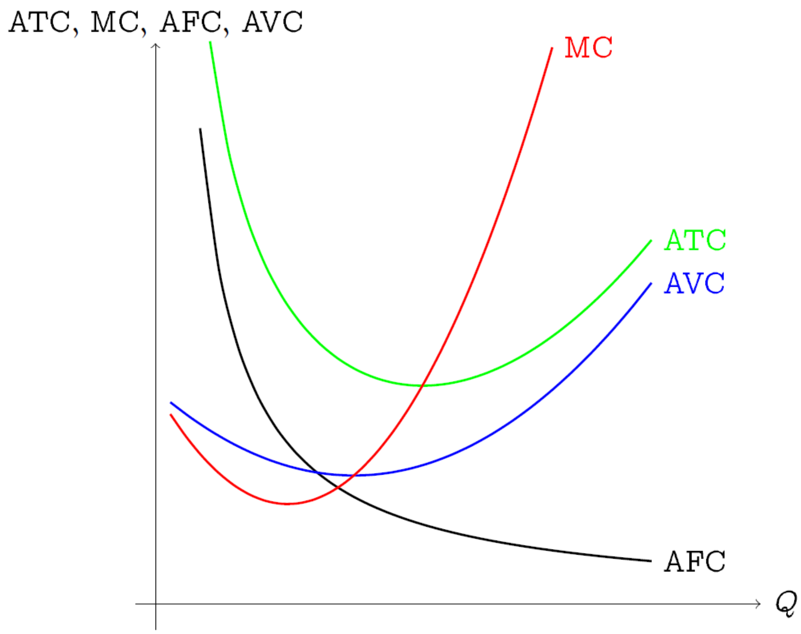

Figure 10: Cost functions when $MC(Q)$ is eventually increasing.

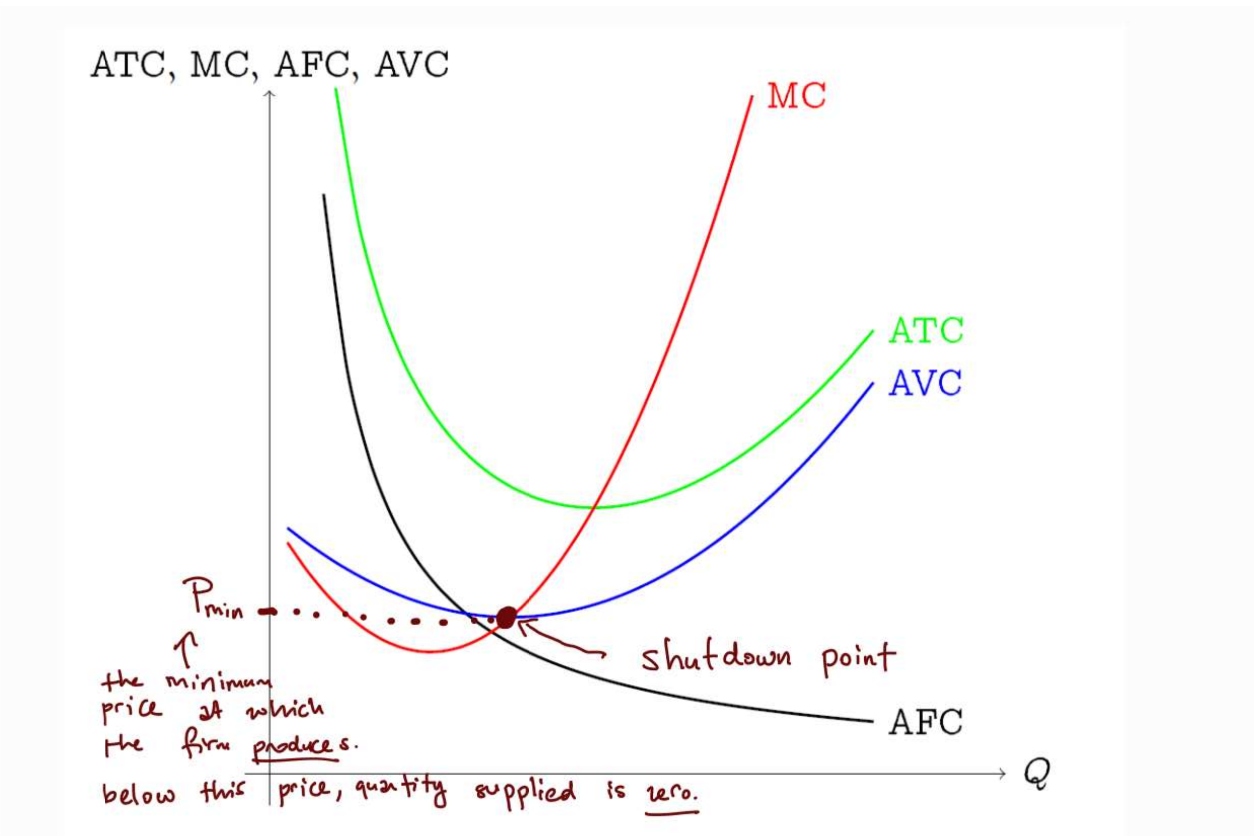

Figure 11: Shut Down Criterion for Eventually Increasing marginal cost.