In these notes we address the firm’s problem. We can break the firm’s problem into three questions.

1. Which combinations of inputs produce a given level of output?

2. Given input prices, what is the cheapest way to attain a certain output?

3. Given output prices, how much output should the firm produce?

We study the firm’s technology in Sections 1–2, the cost minimisation problem in Section 3 and the profit maximisation problem in Section 4.

1 Technology

1.1 Model

We model a firm as a production function that turns inputs into outputs. We assume:

1. The firm produces a single output $q \in \mathbb{R}_+$. One can generalise the model to allow for firms which make multiple products, but this is beyond this course.

*Department of Economics, UCLA. http://www.econ.ucla.edu/sboard/. Please email suggestions and typos to sboard@econ.ucla.edu.
2. The firm has $N$ possible inputs, $\{z_1, \ldots, z_N\}$, where $z_i \in \mathbb{R}_+$ for each $i$. We normally assume $N = 2$, but nothing depends on this. We can think of inputs as labour, capital or raw materials.

3. Inputs are mapped into output by a production function $q = f(z_1, z_2)$. This is normally assumed to be concave and monotone. We discuss these properties later.

To illustrate the model, we can consider a farmer’s technology. In this case, the output is the farmer’s produce (e.g. corn) while the inputs are labour and capital (i.e. machinery). There is clearly a tradeoff between these two inputs: in the developing world, farmers use little capital, doing many tasks by hand; in the developed world, farmers use large machines to plant seeds and even pick fruit.

In some examples inputs may be close substitutes. To illustrate, suppose two students are working on a homework. In this case the output equals the number of problems solved, while the inputs are the hours of the two students. The inputs are close substitutes if all that matters is the total number of hours worked (see Section 2.3).

In other cases inputs may be complements. To illustrate, suppose an MBA and a computer engineer are setting up a company. Each worker has specialised skills and neither can do the other’s job. In this case, output depends on which worker is doing the least work, and we say the inputs are perfect compliments (see Section 2.2).

The marginal product of input $z_i$ is the output from one extra unit of good $i$.

$$MP_i(z_1, z_2) = \frac{\partial f(z_1, z_2)}{\partial z_i},$$

The average product of input $i$ is

$$AP_i(z_1, z_2) = \frac{f(z_1, z_2)}{z_i}.$$

1.2 Isoquants

An isoquant describes the combinations of inputs that produce a constant level of output. That is,

$$Isoquant = \{(z_1, z_2) \in \mathbb{R}_+^2 | f(z_1, z_2) = \text{const.}\}$$
A firm has a collection of isoquants, each one corresponding to a different level of output. By varying this level, we can trace out the agent’s entire production possibilities.

To illustrate, suppose a firm has production technology

\[ f(z_1, z_2) = z_1^{1/3} z_2^{1/3} \]

Then the isoquant satisfies the equation \( z_1^{1/3} z_2^{1/3} = k \). Rearranging, we can solve for \( z_2 \), yielding

\[ z_2 = \frac{k^3}{z_1} \tag{1.1} \]

which is the equation of a hyperbola. This function is plotted in figure 1.

### 1.3 Marginal Rate of Technical Substitution

The slope of the isoquant measures the rate at which the agent is willing to substitute one good for another. This slope is called the marginal rate of technical substitution or MRTS. Mathematically,

\[ MRTS = - \frac{dz_2}{dz_1} \bigg|_{f(z_1, z_2) = \text{const.}} \tag{1.2} \]
We can rephrase this definition in words: the MRTS equals the number of $z_2$ the firm can exchange for one unit of $z_1$ in order to keep output constant.

The MRTS can be related to the firm’s production function. Let us consider the effect of a small change in the firm’s inputs. Totally differentiating the production function $f(z_1, z_2)$ we obtain

$$dq = \frac{\partial f(z_1, z_2)}{\partial z_1} dz_1 + \frac{\partial f(z_1, z_2)}{\partial z_2} dz_2$$  \hspace{1cm} (1.3)$$

Equation (1.3) says that the firm’s output increases by the marginal product of input 1 times the increase in input 1 plus the marginal product of input 2 times the increase in input 2. Along an isoquant $dq = 0$, so equation (1.3) becomes

$$\frac{\partial f(z_1, z_2)}{\partial z_1} dz_1 + \frac{\partial f(z_1, z_2)}{\partial z_2} dz_2 = 0$$

Rearranging,

$$-\frac{dz_2}{dz_1} = \frac{\partial f(z_1, z_2)/\partial z_1}{\partial u(z_1, z_2)/\partial z_2}$$

Equation (1.2) therefore implies that

$$MRTS = \frac{MP_1}{MP_2}$$  \hspace{1cm} (1.4)$$

The intuition behind equation (1.4) is as follows. Using the definition of MRTS, one unit of $z_1$ is worth MRTS units of $z_2$. That is, $MP_1 = MRTS \times MP_2$. Rewriting this equation we obtain (1.4).

### 1.4 Properties of Technology

In this section we present three properties of production functions that will prove useful.

1. **Monotonicity.** The production function is monotone if for any two input bundles $z = (z_1, z_2)$ and $z' = (z'_1, z'_2)$,

$$z_i \geq z'_i \text{ for each } i \quad \text{and} \quad z_i > z'_i \text{ for some } i \quad \implies f(z_1, z_2) > f(z'_1, z'_2).$$

In words: the production function is monotone if more of any input strictly increases the firm’s output. Monotonicity implies that isoquants are thin and downwards sloping (see the Preferences Notes). As a result, it implies that MRTS is positive.
2. **Quasi–concavity.** Let \( z = (z_1, z_2) \) and \( z' = (z'_1, z'_2) \). The production function is quasi–concave if whenever \( f(z) \geq f(z') \) then

\[
f(tz + (1 - t)z') \geq f(z') \quad \text{for all } t \in [0,1]
\]

(1.5)

Suppose \( z \) and \( z' \) are two input bundles that produce the same output, \( f(z) = f(z') \). Then (1.5) says a mixture of these bundles produces even more output. That is, mixtures of inputs are better than extremes.

Under the assumption of monotonicity, quasi–concavity says that isoquants are convex. This means that the MRTS decreasing in \( z_1 \) along the isoquant. Formally, an isoquant defines an implicit relationship between \( z_1 \) and \( z_2 \),

\[
f(z_1, z_2(z_1)) = k
\]

Convexity then implies that \( MRTS(z_1, z_2(z_1)) \) is decreasing in \( z_1 \). This is illustrated in Preferences Notes.

3. **Returns to Scale.** A production function has **decreasing returns to scale** if

\[
f(tz_1, tz_2) \leq tf(z_1, z_2) \quad \text{for } t \geq 1
\]

(1.6)

so that doubling the inputs less that doubles the output. A production function has **constant returns to scale** if

\[
f(tz_1, tz_2) = tf(z_1, z_2) \quad \text{for } t \geq 1
\]

so that doubling the inputs also doubles output. Finally, a production function has **increasing returns to scale** if

\[
f(tz_1, tz_2) \geq tf(z_1, z_2) \quad \text{for } t \geq 1
\]

so that doubling the inputs more than doubles the output.

We will sometimes use the assumption that the production function \( f(z_1, z_2) \) is **concave.** That is, for \( z = (z_1, z_2) \) and \( z' = (z'_1, z'_2) \),

\[
f(tz + (1 - t)z') \geq tf(z) + (1 - t)f(z') \quad \text{for } t \in [0,1]
\]

(1.7)

Concavity implies that the production function is quasi–concave (1.5) and hence that isoquants are convex. This follows immediately from definitions: if \( f(z) \geq f(z') \) then concavity (1.7)
implies
\[ f(tz + (1-t)z') \geq tf(z) + (1-t)f(z') \geq f(z') \]
so the production function is quasi-concave. In addition, concavity implies decreasing returns
to scale. Applying the definition of concavity (1.7) to the points \( z = sz'' \) and \( z' = 0 \) for \( s \geq 1 \),
and letting \( t = 1/s \), we obtain
\[ f \left( \frac{1}{s}(sz) + \left(1 - \frac{1}{s}\right)0 \right) \geq \frac{1}{s}f(sz) + \left(1 - \frac{1}{s}\right)f(0) \]
Using \( f(0) = 0 \) and simplifying, we obtain (1.7)

2 Examples of Production Functions

Here we present some examples of production functions. Many details are omitted since this a
repetition of the examples of utility functions.

2.1 Cobb Douglas

A Cobb–Douglas production function is given by
\[ f(z_1, z_2) = z_1^\alpha z_2^\beta \quad \text{for } \alpha \geq 0 \text{ and } \beta \geq 0 \]
Typical isoquants are shown in figure 1. The marginal products are given by
\[ MP_1 = \alpha z_1^{\alpha-1} z_2^\beta \]
\[ MP_2 = \beta z_1^\alpha z_2^{\beta-1} \]
The marginal rate of technical substitution is
\[ MRTS = \frac{MP_1}{MP_2} = \frac{\alpha z_2}{\beta z_1} \]
The returns to scale are easy to evaluate.
\[ f(tz_1, tz_2) = (tz_1)^\alpha (tz_2)^\beta t = t^{\alpha+\beta} z_1^\alpha z_2^\beta = t^{\alpha+\beta} f(z_1, z_2) \]
Figure 2: **Isoquants for Leontief Technology.** The isoquants are L–shaped, with the kink along the line \( \alpha z_1 = \beta z_2 \).

Hence there are decreasing returns if \( \alpha + \beta \leq 1 \), constant returns if \( \alpha + \beta = 1 \) and increasing returns if \( \alpha + \beta \geq 1 \).

Exercise: Assume \( \alpha + \beta \leq 1 \). Show that \( f(z_1, z_2) \) is concave.\(^1\)

### 2.2 Perfect Complements (Leontief)

A Leontief production function is given by

\[
f(z_1, z_2) = \min\{\alpha z_1, \beta z_2\}
\]

The isoquants are shown in figure 2. These are L–shaped with a kink along the line \( \alpha z_1 = \beta z_2 \). This production function exhibits constant returns to scale.

### 2.3 Perfect Substitutes

With perfect substitutes, the production function is given by

\[
f(z_1, z_2) = \alpha z_1 + \beta z_2
\]

\(^1\)For the definition of concavity with two variables, see the p. 5–6 of the math notes.
Figure 3: **Isoquants for Perfect Substitutes.** The isoquants are straight line with slope $-\alpha/\beta$.

The isoquants are shown in figure 3. These are straight lines with slope $-\alpha/\beta$. This production function exhibits constant returns to scale.

### 3 Cost Minimisation Problem (CMP)

We make several assumptions:

1. There are $N$ inputs. For much of the analysis we assume $N = 2$ but nothing depends on this.

2. The agent takes input prices as exogenous. We assume these prices are linear and strictly positive and denote them by $\{r_1, \ldots, r_N\}$.

3. The firm has production technology $f(z_1, z_2)$. We normally assume that the production function is differentiable, which ensures that any optimal solution satisfies the Kuhn–Tucker conditions. If the production function is quasi-concave and $MP_i(z_1, z_2) > 0$ for all $(z_1, z_2)$, then any solutions to the Kuhn–Tucker conditions are optimal. See Section 4.1 of the UMP notes for more details.
3.1 Cost Minimisation Problem

The cost minimisation problem is

\[
\min_{z_1, \ldots, z_N} \sum_{i=1}^{N} r_i z_i \quad \text{subject to} \quad f(z_1, \ldots, z_N) \geq q \quad (3.1)
\]
\[
z_i \geq 0 \quad \text{for all } i
\]

The idea is that the firm is trying to find the cheapest way to attain a certain output, \( q \). The solution to this problem yields the firm’s \textbf{input demands} which are denoted by

\[z_i^*(r_1, \ldots, r_N, q)\]

The money the firm must spend in order to attain its target output is its cost. The \textbf{cost function} is therefore

\[
c(r_1, \ldots, r_N, q) = \min_{z_1, \ldots, z_N} \sum_{i=1}^{N} r_i z_i \quad \text{subject to} \quad f(z_1, \ldots, z_N) \geq q
\]
\[
z_i \geq 0 \quad \text{for all } i
\]

Equivalently, the cost function equals the amount the firm spends on her optimal inputs,

\[
c(r_1, \ldots, r_N, q) = \sum_{i=1}^{N} r_i z_i^*(r_1, \ldots, r_N, q) \quad (3.2)
\]

Note this problem is formally identical to the agent’s expenditure minimisation problem. The cost function is therefore equivalent to the agent’s expenditure function.

Given a cost function, the \textbf{average cost} is,

\[
AC(r_1, r_2, q) = \frac{c(r_1, r_2, q)}{q}
\]

The \textbf{marginal cost} equals the cost of each additional unit,

\[
MC(r_1, r_2, q) = \frac{dc(r_1, r_2, q)}{dq}
\]
3.2 Graphical Solution

The firm wishes to find the cheapest way to attain a certain output.

First, we need to understand the constraint set. The firm can choose any bundle of inputs where (a) the firm attains her target output, \( f(z_1, z_2) \geq q \); and (b) the quantities are positive, \( z_1 \geq 0 \) and \( z_2 \geq 0 \). If the firm’s production function is monotone, then the bundles that meet these conditions are the ones that lie above the isoquant with output \( q \). See figure 4.

Second, we need to understand the objective. The firm wishes to pick the bundle in the constraint set that minimises her cost. Define an isocost curve by the bundles of \( z_1 \) and \( z_2 \) that deliver constant cost:

\[
\{(z_1, z_2) : r_1 z_1 + r_2 z_2 = \text{const.}\}
\]

These isocost curves are just like budget curves and so have slope \(-r_1/r_2\). See figure 5.

Ignoring boundary problems and kinks, the solution to the CMP has the feature that the isocost curve is tangent to the target isoquant. As a result, their slopes are identical. The tangency condition can thus be written as

\[
MRTS = \frac{r_1}{r_2} \tag{3.3}
\]

This is illustrated in figure 6.
Figure 5: **Isocost.** The isocost function shows the set of inputs which cost the same amount of money.

![Isocost Diagram](image)

**Slope:** \(-r_1/r_2\)

Figure 6: **Tangency.** This figure shows that, at the optimal input combination, the isocost curve is tangent to the isoquant.

![Tangency Diagram](image)

**MRTS:** \(r_1/r_2\)
The intuition behind (3.3) is as follows. Using the fact that $MRTS = MP_1/MP_2$, equation (3.3) implies that

$$\frac{MP_1}{MP_2} = \frac{r_1}{r_1}$$

(3.4)

Rewriting (3.4) we find

$$\frac{r_1}{MP_1} = \frac{r_2}{MP_1}$$

The ratio $r_i/MP_i$ measures the cost of increasing output by one unit. At the optimum the agent equates the cost–per–unit of the two goods. Intuitively, if good 1 has a higher cost–per–unit than good 2, then the agent should spend less on good 1 and more on good 2. In doing so, she could attain the same output at a lower cost.

If the production function is monotone, then the constraint will bind,

$$f(z_1, z_2) = q.$$ 

(3.5)

The tangency equation (3.4) and constraint equation (3.5) can then be used to solve for the two input demands. In addition, one can derive the cost function using equation (3.2).

If there are $N$ inputs, the agent will equalise the cost–per–unit from each good, giving us $N - 1$ equations. Using the constraint equation (3.5), we can again solve for the firm’s input demands.

3.3 Example: Cobb Douglas

Suppose a firm has production function $f(z_1, z_2) = z_1^{1/3}z_2^{1/3}$. The MRTS is

$$MRTS = \frac{\frac{1}{3}z_1^{-2/3}z_2^{1/3}}{\frac{1}{3}z_1^{1/3}z_2^{-2/3}} = \frac{z_2}{z_1}$$

The tangency condition from the CMP is thus

$$\frac{r_1}{r_2} = \frac{z_2}{z_1}$$

Rewriting, this says $r_1z_1 = r_2z_2$, so the firm spends the same on both its inputs.
Figure 7: Cost curves. This figure shows the cost, average cost and marginal cost curves for the Cobb–Douglas example.

The constraint equation is $q = z_1^{1/3} z_2^{1/3}$. This means that

$$q^3 = z_1 z_2 = z_1 \frac{r_1 z_1}{r_2}$$

where the second equality uses the tangency condition. Rearranging, we find the optimal input demands are

$$z_1^* = \left( \frac{r_1}{r_2} \right)^{1/2} q^{3/2} \quad \text{and} \quad z_2^* = \left( \frac{r_2}{r_1} \right)^{1/2} q^{3/2}$$

The cost function is

$$c(r_1, r_2, q) = r_1 z_1^* + r_2 z_2^* = 2(r_1 r_2)^{1/2} q^{3/2}$$

The average and marginal costs are

$$AC(r_1, r_2, q) = 2(r_1 r_2)^{1/2} q^{1/2} \quad \text{and} \quad MC(r_1, r_2, q) = 3(r_1 r_2)^{1/2} q^{1/2}$$

These are illustrated in figure 7.

3.4 Lagrangian Solution

Using a Lagrangian, we can encode the tangency conditions into one formula. As before, let us ignore boundary problems. The CMP can be expressed as minimizing the Lagrangian

$$\mathcal{L} = r_1 z_1 + r_2 z_2 + \lambda[q - f(z_1, z_2)]$$
As usual, the term in brackets can be thought as the penalty for violating the constraint. That is, the firm is punished for falling short of the target output.

The FOCs with respect to $z_1$ and $z_2$ are

$$
\frac{\partial L}{\partial z_1} = r_1 - \lambda \frac{\partial u}{\partial z_1} = 0
$$

(3.6)

$$
\frac{\partial L}{\partial z_2} = r_2 - \lambda \frac{\partial u}{\partial z_2} = 0
$$

(3.7)

If the production function is monotone then the constraint will bind,

$$
f(z_1, z_2) = q
$$

(3.8)

These three equations can then be used to solve for the three unknowns: $z_1$, $z_2$ and $\lambda$.

Several remarks are in order. First, this approach is identical to the graphical approach. Dividing (3.6) by (3.7) yields

$$
\frac{\partial u/\partial z_1}{\partial u/\partial z_2} = \frac{r_1}{r_2}
$$

which is the same as (3.4). Moreover, the Lagrange multiplier is exactly the cost–per–unit,

$$
\lambda = \frac{r_1}{MP_1} = \frac{r_2}{MP_2}.
$$

Second, if preferences are not monotone, the constraint (3.5) may not bind. If it does not bind, the Lagrange multiplier in the FOCs will be zero.

Third, the approach is easy to extend to $N$ inputs. In this case, one obtains $N$ first–order conditions and the constraint equation (3.5).

### 3.5 Properties of Cost Functions

We now develop six properties of the cost function. The first four are identical to the properties of the expenditure function: see the EMP Notes for more details.

1. **The cost function is homogenous of degree one in prices.** That is,

$$
c(r_1, r_2, q) = c(tp_1, tp_2, q) \quad \text{for } t > 0
$$


This figure shows how the cost function lies under the pseudo–cost function. Intuitively, if the prices of \( r_1 \) and \( r_2 \) double, then the cheapest way to attain the target output does not change. However, the cost of attaining this output doubles.

2. The cost function is increasing in \((r_1, r_2, q)\). If we increase the target output then the constraint becomes harder to satisfy and the cost of attaining the target increases. If we increase \( r_1 \) then it costs more to buy any bundle of inputs and it costs more to attain the target output.

3. The cost function is concave in input prices \((r_1, r_2)\). Fix the target utility \( q \) and prices \((r_1, r_2) = (r'_1, r'_2)\). Solving the CMP we obtain input demands \( z'_1 = z^*_1(r'_1, r'_2, q) \) and \( z'_2 = z^*_2(r'_1, r'_2, q) \). Now suppose we fix demands and change \( r_1 \), the price of input 1. This gives us a pseudo–cost function

\[
c_{z'_1, z'_2}(p_1) = r'_1 z'_1 + r'_2 z'_2
\]

which is linear in \( r_1 \). Of course, as \( r_1 \) rises the firm can reduce her costs by rebalancing her input demand towards the input that is cheaper. This means that real cost function lies below the pseudo–cost function and is therefore concave. See figure 8.

4. Sheppard’s Lemma: The derivative of the cost function equals the input demand. That is,

\[
\frac{\partial}{\partial r_1} c(r_1, r_2, q) = z^*_1(r_1, r_2, q)
\]  

(3.9)

The idea behind this result can be seen from figure 8. At \( r_1 = r'_1 \) the cost function is tangential.
to the pseudo–cost function. The pseudo–cost is linear in \( r_1 \) with slope \( z_1^*(r'_1, r'_2, q) \), so the expenditure function also has slope \( z_1^*(r'_1, r'_2, q) \).

The intuition behind Sheppard’s Lemma is as follows. When \( r_1 \) increases by \( \Delta r_1 \) there are two effects. First, holding input demand constant, the firm’s cost rises by \( z_1^*(r_1, r_2, q) \times \Delta r_1 \). Second, the firm rebalances its demands, buying less of input 1 and more of input 2. However, this has a small effect on the firm’s costs since it is close to indifferent buying the optimal quantity and nearby quantities.

5. If \( f(z_1, z_2) \) is concave then \( c(r_1, r_2, q) \) is convex in \( q \). Intuitively, concavity of the production function, implies that the marginal product of an input is decreasing in the amount of the input used:

\[
\frac{d}{dz_i} MP_i(z_1, z_2) = \frac{d^2}{dz_i^2} f(z_1, z_2) \leq 0
\]

Therefore, as the firm expands, it needs more inputs to produce each additional unit of output. As a result, the cost of producing this unit increases, and the total cost is convex. When there is only one input this is easy to see formally: if \( f(z) \) is concave, then \( c(q) = rf^{-1}(q) \) is convex.

6. \( AC(q) \) is increasing when \( MC(q) \geq AC(q) \), is flat when \( MC(q) = AC(q) \) and is decreasing when \( MC(q) \leq AC(q) \). Suppose the firm currently produces \( n \) units of output, and that the marginal cost of the \((n+1)\)st unit is higher than the average cost of the first \( n \) units. Then the average cost of producing \( n + 1 \) units is higher that producing \( n \) units since the costs is being dragged up by the final unit. To prove this result formally, we can differentiate the \( AC \) curve,

\[
\frac{d}{dq} AC(q) = \frac{d}{dq} \frac{c(q)}{q} = \frac{c'(q)q - c(q)}{q^2}
\]

Hence \( AC(q) \) is increasing if and only if \( c'(q)q \geq c(q) \). Rearranging, this condition is just \( MC(q) \geq AC(q) \), as required.

3.6 Pictures of Cost Functions

Figure 7 shows the cost curves associated with a concave production function. One can see that the cost function is convex and, as a result, the marginal cost is increasing and exceeds the average cost.

Figure 9 shows the cost curves associated with a production function which is concave for
positive quantities but requires a fixed cost needed to initiate production. The marginal cost of the first unit is infinite and is therefore not shown in the picture; the marginal cost of each subsequent unit is increasing. The average cost is U–shaped: it starts at infinity, is minimised at $q'$ and then rises as the higher marginal cost drags up the average cost. Note that the marginal cost intersects the average cost at its lowest point: this follows from property 6 from Section 3.5.

Figure 10 shows the cost curves associated with a second nonconcave production function. The cost curve is S–shaped. As a result, the marginal cost and average cost functions are U–shaped. For the first unit, the marginal cost and average cost coincide; for low levels of output, the marginal cost is decreasing and lies below the average cost; for high levels of output, the marginal cost is increasing, exceeding the average cost for $q \geq q'$.

### 3.7 Long Run vs. Short Run Costs

The cost of a firm depend on which factors of production are flexible. We differentiate between four cases, and then illustrate them with an example.

---

2 For example, try $f(z) = (z - 1)^{1/2}$.
3 For example, try $f(z) = 100z - 16z^2 + z^3$.
4 While the idea of short and long run is standard, different authors mean different things by the “short run” and “long run”.

---
1. In the **very short run** all the factors of production are fixed, and output is fixed.

2. In the **short run** some factors are flexible, while others are fixed. For example, the firm may be able hire some more workers, but may not be able to order new capital equipment. Any fixed costs are also sunk, so that they cannot be avoided even if the firm ceases production.

3. In the **medium run** all factors are flexible, but fixed costs are sunk.

4. In the **long run** all factors are flexible and fixed costs are not sunk. Hence the firm can costlessly exit.

In practice, the meaning of short and long run depend on the application. For example, consider a farmer who wishes to increase her output. It may take her a few days to hire an extra worker, a few weeks to lease an extra tractor and a few months for a new farmer to buy land and enter the business (or for an old one to exit).

To illustrate, suppose a firm has production function\(^5\)

\[ f(z_1, z_2) = (z_1 - 1)^{1/3}(z_2 - 1)^{1/3} \]

This firm has Cobb–Douglas production, except that the first unit of both inputs is useless, inducing a fixed cost.

\(^5\)Since negative outputs are impossible, we should say that \( q = 0 \) if either \( z_1 < 1 \) or \( z_2 < 1 \).
First, let us solve for the long–run cost function. The firm’s Lagrangian is

\[ \mathcal{L} = r_1 z_1 + r_1 z_2 + \lambda [q - (z_1 - 1)^{1/3}(z_2 - 1)^{1/3}] \]

Differentiating, this induces the tangency condition \( r_1 (z_1 - 1) = r_2 (z_2 - 1) \). Using the constraint, \( q = (z_1 - 1)^{1/3}(z_2 - 1)^{1/3} - 1 \), we obtain

\[ z_1^* = \left( \frac{r_1}{r_2} \right)^{1/2} (q)^{3/2} + 1 \quad \text{and} \quad z_2^* = \left( \frac{r_2}{r_1} \right)^{1/2} (q)^{3/2} + 1 \]

The cost function is

\[ c(r_1, r_2, q) = r_1 z_1^* + r_2 z_2^* = 2(r_1 r_2)^{1/2} (q)^{3/2} + (r_1 + r_2) \]

In addition, the firm can shutdown and produce zero at cost \( c(r_1, r_2, 0) = 0 \). Observe that this cost function is the same as that in Section 3.3 with a startup cost of \( r_1 + r_2 \).

In the medium run, the fixed cost \( r_1 + r_2 \) is sunk. The medium run cost curve is therefore

\[ c(r_1, r_2, q) = 2(r_1 r_2)^{1/2} (q)^{3/2} + (r_1 + r_2) \]

where \( c(r_1, r_2, 0) = r_1 + r_2 \).

In the short run, \( z_1 \) is flexible but \( z_2 \) is fixed at \( z_2' \). The fixed cost is also sunk. The constraint in the CMP becomes

\[ q = (z_1 - 1)^{1/3}(z_2' - 1)^{1/3} \]

Rearranging,

\[ z_1^* = \frac{q^3}{z_2' - 1} + 1 \]

The cost function is therefore given by

\[ c(r_1, r_2; q; z_2') = r_1 z_1^* + r_2 z_2' = r_1 \frac{q^3}{z_2' - 1} + r_1 + r_2 z_2' \]

Figure 11 illustrates the short run cost curves for three different levels of \( z_2 \). Observe that the long run cost curve is given by the lower envelope of the short run cost curves. To see why this is the case, fix an output level \( q' \) and calculate the optimal input demands when both factors are flexible, denoted by \( z_1' \) and \( z_2' \). Now suppose we fix \( z_2 \) at \( z_2' \) and consider the cost of attaining different output levels. If \( q = q' \) then the firm is using the optimal amount of input 2 and the short–run cost will coincide with the long–run cost. If \( q > q' \) then the firm is using too little of
Figure 11: **Long Run and Short Run Costs.** This figure shows the long run cost curve and the short run cost curves corresponding to three levels of the second input.

$z_2$ and too much of $z_1$, raising the short–run cost over the long–run cost. If $q < q'$ then the firm is using too much of $z_2$ and too little of $z_1$, again raising the short–run cost over the long–run cost.

In the very short run, inputs are fixed at $z_1 = z'_1$ and $z_2 = z'_2$. Hence the firm can produce $q' = (z'_1 - 1)^{1/3}(z'_2 - 1)^{1/3}$ at cost $r_1 z'_1 + r_2 z'_2$, but is unable to produce anything else.

## 4 Profit Maximisation Problem (PMP)

**Assumptions:**

1. There is one output good, with linear price $p$. This means that the firm is a price–taker in the output market.

2. There are two input goods with linear prices $r_1$ and $r_2$. The firm is therefore a price–taker in the input market.

3. The firm has production technology $f(z_1, z_2)$. We normally assume that the production function is differentiable, which ensures that any optimal solution satisfies the first–order conditions.
The firm’s profit equals its revenue from selling the output minus its cost:

$$\pi = pf(z_1, z_2) - r_1 z_1 - r_2 z_2$$

We now explore two ways of solving this problem.

### 4.1 One–Step Solution

The firm’s **profit maximisation problem** is

$$\max_{z_1, z_2} pf(z_1, z_2) - r_1 z_1 - r_2 z_2 \quad \text{subject to } z_i \geq 0 \text{ for all } i$$  \hspace{1cm} (4.1)

The first–order conditions are

$$\frac{d\pi}{dz_1} = p \frac{\partial f(z_1, z_2)}{\partial z_1} - r_1 = 0$$  \hspace{1cm} (4.2)

$$\frac{d\pi}{dz_2} = p \frac{\partial f(z_1, z_2)}{\partial z_2} - r_2 = 0$$  \hspace{1cm} (4.3)

Together (4.2) and (4.3) define the optimal input demands of the firm, $z_1^*(p, r_1, r_2)$ and $z_2^*(p, r_1, r_2)$. We can then derive the optimal output:

$$q^*(p, r_1, r_2) = f(z_1^*, z_2^*)$$

which is called the **supply function**. We can also derive the firm’s optimal profit,

$$\pi^*(p, r_1, r_2) = pq^* - r_1 z_1^* - r_2 z_2^*$$

which is called the **profit function**.

Observe that solving (4.1) is much easier than solving the utility maximisation problem. With the UMP, the consumer maximises her utility subject to spending no more than her income. With the PMP, the firm’s expenses directly enter the firm’s objective function, so we only have to solve an unconstrained optimisation problem.

In order for the FOCs (4.2) and (4.3) to characterise a maximum, the second–order conditions
must hold. That is, \(f(z_1, z_2)\) must be locally concave, which implies
\[
\frac{\partial^2}{\partial z_1^2} f(z_1, z_2) = \frac{\partial}{\partial z_1} MP_1(z_1, z_2) \leq 0
\]
\[
\frac{\partial^2}{\partial z_2^2} f(z_1, z_2) = \frac{\partial}{\partial z_2} MP_2(z_1, z_2) \leq 0
\]

If \(f(z_1, z_2)\) is globally concave, then any solution to the FOCs is a maximum.

### 4.2 Example: Cobb Douglas

Suppose a firm has production function \(f = \frac{z_1}{z_2}^{1/3}\). Profit is given by
\[
\pi = pz_1^{1/3}z_2^{1/3} - r_1z_1 - r_2z_2
\]

The FOCs are
\[
\frac{1}{3}pz_1^{-2/3}z_2^{1/3} = r_1
\]
\[
\frac{1}{3}z_1^{1/3}z_2^{-2/3} = r_2
\]

Solving these two equations yields input demands:
\[
z_1^*(p, r_1, r_2) = \frac{1}{27} \frac{p^3}{r_1^2r_2^2} \quad \text{and} \quad z_2^*(p, r_1, r_2) = \frac{1}{27} \frac{p^3}{r_1r_2^3}
\]

The optimal supply is
\[
q^*(p, r_1, r_2) = (z_1^*)^{1/3}(z_2^*)^{1/3} = \frac{1}{9} \frac{p^2}{r_1r_2}
\]

The profit function is
\[
\pi^*(p, r_1, r_2) = pq^* - r_1z_1^* - r_2z_2^* = \frac{1}{27} \frac{p^3}{r_1r_2}
\]

### 4.3 Two–Step Solution

**Step 1. Find the cheapest way to attain output \(q\).** Recall the cost function is given by
\[
c(q, r_1, r_2) = \min_{z_1, z_2} r_1z_1 + r_2z_2 \quad \text{subject to} \quad f(z_1, z_2) \geq q
\]
\[
z_i \geq 0 \quad \text{for all } i
\]
Step 2. Find the profit–maximising output. Given a cost function, the firm’s problem is

$$\max_q \pi = pq - c(q, r_1, r_2) \quad \text{subject to } q \geq 0$$

The first–order condition for this problem is

$$\frac{d\pi}{dq} = p - \frac{d}{dq} c(q, r_1, r_2) = 0$$

That is,

$$p = MC(q, r_1, r_2) \quad (4.4)$$

The idea behind this result is shown in the left panel of figure 12, which shows the firm’s revenue and costs as a function of output, $q$. The firm wishes to maximise the vertical distance between the two lines so, at the optimum, they are parallel. The slope of the revenue line is $p$ while the slope of the cost function is $MC$, which yields (4.4).

One can also look at this result with the right panel of figure 12. The difference $p - MC$ equals the profit the firm makes on the last unit. The FOC (4.4) says that the firm will keep producing while the profit–per–unit is positive and will stop when it falls to zero. Note that, in this picture, one can measure profits two ways. First, profit equals the price obtained per unit minus the average cost of a unit multiplied by the number of units sold:

$$\pi(q) = pq - c(q) = pq - AC(q)q = [p - AC(q)]q$$

In the picture, this equals the areas given by A+B+C.

Second, the profit of a marginal unit is $p - MC(q)$. Hence the total profit of the firm, ignoring fixed costs, is the area below the price and above the MR curve. That is,

$$\pi(q) = pq - c(q) = \int_0^q pd\tilde{q} - \int_0^q MC(\tilde{q})d\tilde{q} - F = \int_0^q [p - MC(\tilde{q})] d\tilde{q} - F$$

where $F$ is the fixed cost. Hence the firm’s profit is A+B+D+E minus the fixed cost, $F$.

In order for the FOC (4.4) to constitute an optimum, the second–order condition should hold:

$$\frac{d^2\pi}{dq^2} = -\frac{d^2}{dq^2} c(q, r_1, r_2) = -\frac{d}{dq} MC(q, r_1, r_2) \leq 0$$

So the marginal cost needs to be locally increasing. Conversely, if the cost function is convex,
Figure 12: **Profit maximisation.** The left panel shows that profit is maximised when the revenue line is parallel to the cost line. The vertical gap, is then equal to the firm’s profit. The right panel shows that profit is maximised when the price equals to marginal cost. Profit then equals A+B+C.

which is guaranteed by the concavity of $f(z_1, z_2)$, then any solution to the FOC (4.4) is an optimum.

### 4.4 Example: Cobb Douglas

We now return to the example in Section 4.2, deriving the same results using the two-step approach.

Suppose $f(z_1, z_2) = z_1^{1/3} z_2^{1/3}$. Using the results in Section 3.3, the cost function is

$$c(q, r_1, r_2) = 2(r_1 r_2)^{1/2} q^{3/2}$$

The first-order condition (4.4) yields

$$p = 3(r_1 r_2 q)^{1/2}$$

Rearranging, the supply curve is given by

$$q^*(p, r_1, r_2) = \frac{p^2}{9 r_1 r_2}$$
The profit function is then
\[
\pi^*(p, r_1, r_2) = pq^* - r_1 z_1^* - r_2 z_2^* = \frac{1}{27} \frac{p^3}{r_1 r_2}
\]
as in Section 4.2.

4.5 Examples of Supply Functions

Figure 13 shows the supply function that results from a convex cost function with no fixed cost.\(^6\) The marginal cost is increasing and is always above the average cost. For any given price, the firm chooses quantity such that \(p = MC(p)\). Hence the supply curve coincides with the \(MC\) curve.

Figure 14 shows the supply function that results from a convex cost function with a fixed cost.\(^7\) The marginal cost function is increasing so, if the firm produces, its supply curve coincides with \(MC(q)\). However, when the price lies below the average cost, the firm makes negative profits. Hence the firm’s supply curve coincides with the \(MC(q)\) curve above the \(AC(q)\) curve and is zero elsewhere.

Figure 15 shows the supply function that results from a U–shaped marginal cost function without a fixed cost.\(^8\) For prices below \(p'\) the marginal cost is below the average cost, so the firm cannot make a profit and it chooses to produce \(q^*(p) = 0\). At \(p = p'\) the firm is indifferent between producing 0 and \(q'\). For price above \(p'\) the firm produces on the increasing part of the marginal cost function.

Figure 16 shows the supply function that results from a nonconvex cost curve.\(^9\) For low prices the supply curve coincides with the first part of the \(MC\) curve. At a price \(p'\) the supply jumps to the right. Intuitively, if the firm is going to pay to produce the expensive units in region A then it should also produce the cheap units in region B. At the optimum, the area of A equals the area of B, so the profit lost by producing the expensive units is exactly offset by the profit gained by producing the cheap units.

One can also use these figures to understand the difference between the short–run and long–run supply curves. In the very short run, supply is fixed and the supply curve is vertical. In

\(^6\)For example, try \(c(q) = q + q^2\).
\(^7\)For example, try \(c(q) = 1 + q + q^2\).
\(^8\)For example, try \(c(q) = 15q - 12q^2 + q^3\).
\(^9\)For example, try \(c(q) = 20q^2 - 8q^3 + q^4\).
Figure 13: **Supply Curve with Convex Costs.** This figure shows how the supply curve coincides with the marginal cost curve.

![Supply Curve with Convex Costs](image)

Figure 14: **Supply Curve with Nonconvex Costs I: Fixed Costs.** This figure shows how the supply curve coincides with the marginal cost curve when it lies above the average cost.

![Supply Curve with Nonconvex Costs I: Fixed Costs](image)
Figure 15: **Supply Curve with Nonconvex Costs II: U–Shaped Marginal Cost.** This figure shows how the supply curve coincides with the marginal cost curve when it lies above the average cost.

Figure 16: **Supply Curve with Nonconvex Costs III.** This figure shows how the supply curve coincides with the marginal cost curve when it lies above the average cost.
the short–run, some of the inputs are fixed and the supply curve coincides with the short–run marginal cost. In the medium–run, the firm can change all its inputs, but cannot close down. Hence the supply curve coincides with the marginal cost curve above the average variable cost. In the long–run the firm can shut down, so the supply curve coincides with the marginal cost above the average cost.

4.6 Properties of the Profit Function

The profit function $\pi^*(p, r_1, r_2)$ has four key properties:

1. $\pi^*(p, r_1, r_2)$ is homogenous of degree one in $(p, r_1, r_2)$. If all prices double then the optimal production choices remain unchanged and profit also doubles. Intuitively, if currency is denominated in a different currency this should not affect the firm’s choices.

2. $\pi^*(p, r_1, r_2)$ is increasing in $p$ and decreasing in $(r_1, r_2)$. An increase in $p$ increases profits for any output $q$, and therefore increases profit for the optimal output choice. An increase in $r_1$ increases costs and decreases profits for any output $q$, and therefore decreases profit for the optimal output choice.

3. $\pi^*(p, r_1, r_2)$ is convex in $(p, r_1, r_2)$. Let us first consider changes in $p$, and ignore the input prices. Fix $p = p'$ and solve for the optimal output $q' = q^*(p')$. Now suppose we fix the output and change $p$, yielding a pseudo–profit function $pq' - c(q')$ which is linear in $p$. Of course, as $p$ rises the firm can increase her output, so the real cost function lies above this straight line and is therefore convex. See figure 16. Second, the profit function is convex in $(r_1, r_2)$ because profit is equal $\pi = pq - c(q, r_1, r_2)$ and $c(q, r_1, r_2)$ is concave in $(r_1, r_2)$.

4. Hotelling’s Lemma: The derivative of the profit function with respect to the output price equals the optimal output. That is,

$$\frac{\partial}{\partial p}\pi^*(p, r_1, r_2) = q^*(p, r_1, r_2) \quad (4.5)$$

The idea behind this result can be seen from figure 16. At $p = p'$ the profit function is tangential to the pseudo–profit function. The pseudo–profit is linear in $p$ with slope $q'(p')$. Hence the expenditure function also has slope $q^*(p)$.

The intuition behind Hotelling’s Lemma can be seen in figure 17. We start at $p = p'$, with
Figure 17: **Convexity of Profit Functions** This figure shows how the profit function equals the upper envelope of the pseudo-profit functions, \( pq - c(q) \).

profit equal to area A.\(^{10}\) When the price increases to \( p'' \) there are two effects. First, holding output constant, the firm’s profit rises by \( q^*(p) \times (p'' - p') \), illustrated by area B. Second, the firm increases its output, yielding extra profit C. However, for small price changes this second effect is small, which yields Hotelling’s Lemma. One can also see from this picture that profit is convex in price: output is higher when the price is higher, so the change in profit induced by a 1¢ increase in the price is higher when the price is higher.

### 4.7 Properties of Supply Functions

There are two important properties of the supply function.

1. **Supply** \( q^*(p, r_1, r_2) \) is homogenous of degree zero in \( (p, r_1, r_2) \). If prices are denominated in a different currency this will not affect the firm’s optimal output.

2. **Law of Supply**: \( q^*(p, r_1, r_2) \) is increasing in \( p \). The supply curve is always upward sloping. Intuitively, an increase in the price increases the benefits to producing and so increases the optimal output. Formally, Hotelling’s Lemma implies that

\[
\frac{d}{dp} q^*(p, r_1, r_2) = \frac{d^2}{dp^2} \pi^*(p, r_1, r_2) \geq 0
\]

\(^{10}\)Note there are no fixed costs in this picture
Figure 18: Convexity of Profit Functions This figure shows how the profit function is convex in the price and that the derivative equals the current supply.

where the inequality come from the convexity of the profit function.