Preferences and Utility

Simon Board*

This Version: October 6, 2009
First Version: October, 2008.

These lectures examine the preferences of a single agent. In Section 1 we analyse how the agent chooses among a number of competing alternatives, investigating when preferences can be represented by a utility function. In Section 2 we discuss two attractive properties of preferences: monotonicity and convexity. In Section 3 we analyse the agent's indifference curves and ask how she makes tradeoffs between different goods. Finally, in Section 4 we look at some examples of preferences, applying the insights of the earlier theory.

1 The Foundation of Utility Functions

1.1 A Basic Representation Theorem

Suppose an agent chooses from a set of goods $X = \{a, b, c, \ldots\}$. For example, one can think of these goods as different TV sets or cars.

Given two goods, x and y, the agent **weakly prefers** x over y if x is at least as good as y. To avoid us having to write "weakly prefers" repeatedly, we simply write $x \succcurlyeq y$. We now put some basic structure on the agent's preferences by adopting two axioms.[1]

Completeness Axiom: For every pair $x, y \in X$, either $x \succcurlyeq y$, $y \succcurlyeq x$, or both.

*Department of Economics, UCLA. http://www.econ.ucla.edu/sboard/. Please email suggestions and typos to sboard@econ.ucla.edu.

[1] An axiom is a foundational assumption.

Transitivity Axiom: For every triple $x, y, z \in X$, if $x \succcurlyeq y$ and $y \succcurlyeq z$ then $x \succcurlyeq z$.

An agent has complete preferences if she can compare any two objects. An agent has transitive preferences if her preferences are internally consistent. Let's consider some examples.

First, suppose that, given any two cars, the agent prefers the faster one. These preferences are complete: given any two cars x and y, then either x is faster, y is faster or they have the same speed. These preferences are also transitive: if x is faster than y and y is faster than z, then x is faster than z.

Second, suppose that, given any two cars, the agent prefers x to y if it is both faster and bigger. These preferences are transitive: if x is faster and bigger than y and y is faster and bigger than z, then x is faster and bigger than z. However, these preferences are not complete: an SUV is bigger and slower than a BMW, so it is unclear which the agent prefers. The completeness axiom says these preferences are unreasonable: after examining the SUV and BMW, the agent will have a preference between the two.

Third, suppose that the agent prefers a BMW over a Prius because it is faster, an SUV over a BMW because it is bigger, and a Prius over an SUV, because it is more environmentally friendly. In this case, the agent's preferences cycle and are therefore intransitive. The transitivity axiom says these preferences are unreasonable: if environmental concerns are so important to the agent, then she should also take them into account when choosing between the Prius and BMW, and the BMW and the SUV.

While it is natural to think about preferences, it is often more convenient to associate different numbers to different goods, and have the agent choose the good with the highest number. These numbers are called **utilities**. In turn, a **utility function** tells us the utility associated with each good $x \in X$, and is denoted by $u(x) \in \Re$. We say a utility function $u(x)$ **represents** an agent's preferences if

$$u(x) \geq u(y) \quad \text{if and only if} \quad x \succcurlyeq y \tag{1.1}$$

This means than an agent makes the same choices whether she uses her preference relation, $\succcurlyeq$, or her utility function $u(x)$.

Theorem 1 (Utility Representation Theorem)**.** *Suppose the agent's preferences, $\succcurlyeq$, are complete and transitive, and that X is finite. Then there exists a utility function $u(x) : X \to \Re$ which represents $\succcurlyeq$.*

Theorem 1 says that if an agent has complete and transitive preferences then we can associate these preferences with a utility function. Intuitively, the two axioms allow us to rank the goods under consideration. For example, if there are 10 goods, then we can say the best has a utility $u(x) = 9$, the second best has $u(x) = 8$, the third best has $u(x) = 7$ and so on. For a formal proof, see Section 1.2.

1.2 A Proof of Theorem 1[2]

The idea behind the proof is simple. For any good x, let $NBT(x) = \{y \in X | x \succcurlyeq y\}$ be the goods that are "no better than" x. The utility of x is simply given by the number of items in $NBT(x)$. That is,

$$u(x) = |NBT(x)|. \tag{1.2}$$

If there are 10 goods, then the worst has a "no better than" set which is empty, so that $u(x) = 0$. The second worst has a has a "no better than" set which has one element, so $u(x) = 1$. And so on.

We now have to verify that this utility function represents the agent's preferences. We do this in two steps: first, we show that $x \succcurlyeq y$ implies $u(x) \geq u(y)$; second, we show that $u(x) \geq u(y)$ implies $x \succcurlyeq y$.

Step 1: Suppose $x \succcurlyeq y$. Pick any $z \in NBT(y)$;[3] by the definition of $NBT(y)$, we have $y \succcurlyeq z$. Since preferences are complete, we know that z is comparable to x. Transitivity then tells us that $x \succcurlyeq z$, so $z \in NBT(x)$. We have therefore shown that every element of $NBT(y)$ is also an element of $NBT(x)$; that is, $NBT(y) \subseteq NBT(x)$. As a result,

$$u(x) = |NBT(x)| \geq |NBT(y)| = u(y)$$

as required.

Step 2: Suppose $u(x) \geq u(y)$. By completeness, we know that either $x \succcurlyeq y$ or $y \succcurlyeq x$. Using Step 1, it must then be the case that either $NBT(y) \subseteq NBT(x)$ or $NBT(x) \subseteq NBT(y)$, so the "no better than" sets cannot partially overlap or be disjoint. By the definition of utilities (1.2) we know that there are more elements in $NBT(x)$ than in $NBT(y)$, which implies that

[2] More advanced.

[3] $z \in NBT(y)$ means that z is an element of $NBT(y)$.

$NBT(y) \subseteq NBT(x)$. Completeness means that a good is weakly preferred to itself, so that $y \in NBT(y)$. Since $NBT(y) \subseteq NBT(x)$, we conclude $y \in NBT(x)$. Using the definition of the "no better than" set, this implies that $x \succcurlyeq y$, as required.

1.3 Increasing Transformations

A number system is **ordinal** if we only care about the ranking of the numbers. It is **cardinal** if we also care about the magnitude of the numbers. To illustrate, Usain Bolt and Richard Johnson came 1st and 2nd in the 2008 Olympic final of the 100m sprint. The numbers 1 and 2 are ordinal: they tell us that Bolt beat Johnson, but do not tell us that he was 1% faster or 10% faster. The actual finishing times were 9.69 for Bolt and 9.89 for Johnson. These numbers are cardinal: the ranking tells us who won, and the magnitudes tells us about the margin of the win.

Theorem 1 is ordinal: when comparing two goods, all that matters is the ranking of the utilities; the actual numbers themselves carry no significance. This is obvious from the construction: when there are 10 goods, it is clearly arbitrary that we give utility 9 to the best good, 8 to the second best, and so on. This idea can be formalised by the following result:

Theorem 2. *Suppose $u(x)$ represents the agent's preferences, $\succcurlyeq$, and $f : \Re \to \Re$ is a strictly increasing function. Then the new utility function $v(x) = f(u(x))$ also represents the agent's preferences $\succcurlyeq$.*

The proof of Theorem 2 is simply a rewriting of definitions. Suppose $u(x)$ represents the agent's preferences, so that equation (1.1) holds. If $x \succcurlyeq y$ then $u(x) \geq u(y)$ and $f(u(x)) \geq f(u(y))$, so that $v(x) \geq v(y)$. Conversely, if $v(x) \geq v(y)$ then, since $f(\cdot)$ is strictly increasing, $u(x) \geq u(y)$ and $x \succcurlyeq y$. Hence

$$v(x) \geq v(y) \quad \text{if and only if} \quad x \succcurlyeq y$$

and $v(x)$ represents $\succcurlyeq$.

Theorem 2 is important when solving problems. Suppose an agent has utility function

$$u(x) = -\frac{15}{(x_1^{1/2} + x_2^{1/2} + 10)^3}$$

Solving the agent's problem with this utility function may be be algebraically messy. Using

Theorem 2, we can rewrite the agent's utility as

$$v(x) = x_1^{1/2} + x_2^{1/2}$$

Since $u(x)$ and $v(x)$ preserve the rankings of the goods, they represent the same preferences. As a result, the agent will make the same choices with utility $u(x)$ and $v(x)$. This is useful since it is much simpler to solve the agent's choice problem using $v(x)$ than $u(x)$.

Theorem 2 is also useful for cocktail parties. For example, some people dislike the way I rank movies of a 1-10 scale. They claim that a movie is a rich artistic experience, and cannot be summarised by a number. However, Theorem 1 tells us that, if my preferences are complete and transitive, then I can represent my preferences over movies by a number. Moreover, Theorem 2 tells us that I can rescale the numbers to put them on a 1-10 scale.

Choosing from Budget Sets

Theorem 1 assumes that the consumer chooses from a finite number of goods. While this is realistic, it is more mathematically convenient to allow consumers to choose from a continuum of goods. For example, if the agent has \$10 and a hamburger costs \$2, it is easier to allow the consumer to any number between 0 and 5, rather than forcing her to choose an integer.

Suppose the choice set is given by $X \subseteq \Re^n_+$. A typical element is $x = (x_1, \ldots, x_n)$, where x_i is the number of the i^{th} good the agent consumes. In order to prove a representation theorem for this larger set of choices, we need one more (rather technical) axiom.

Continuity Axiom: Suppose $x^1, x^2, x^3, \ldots$ is a sequence of feasible choices, so that $x^i \in X$ for each i, and suppose the sequence converges to $x \in X$. If $x^i \succcurlyeq y$ for each i, then $x \succcurlyeq y$.

Theorem 3 (Representation Theorem for Budget Sets). *Suppose the agent's preferences, $\succcurlyeq$, are complete, transitive and continuous, and that $X \subseteq \Re^n_+$. Then there exists a continuous utility function $u(x) : X \to \Re$ which represents $\succcurlyeq$.*

We will not prove this result. The following example examines a case where the continuity axiom does not hold and no utility representation exists.

Suppose there are two goods and the agent has **lexicographic preferences:** when faced with two bundles the agent prefers the bundle with the most of x_1; if the two bundles have the same

x_1 then she prefers the bundle with the most of x_2. To verify that this does not satisfy the continuity axiom, consider a sequence of bundles $x^i = (1 + \frac{1}{i}, 1)$ which converges to $x = (1, 1)$ as $i \to \infty$, and let $y = (1, 2)$. For each i, x^i is preferred to y since x^i contains more of good 1. However, in the limit, the agent prefers y to x since they have the same quantity of good 1, but y has more of good 2. One can also show that there exists no utility function that represents lexicographic preferences, but this is a little tricky.

2 Properties of Preferences

In this Section we introduce two key properties of preferences: monotonicity and convexity. Throughout, we suppose $X \subseteq \Re^n_+$.

First we need a couple of definitions. If the agent weakly prefers x to y (i.e. $x \succcurlyeq y$) and weakly prefers y to x (i.e. $y \succcurlyeq x$) then she is **indifferent** between x and y and we write $x \sim y$. In terms of utilities, an agent is indifferent between x and y if and only if $u(x) = u(y)$.

If the agent weakly prefers x to y ($x \succcurlyeq y$) and is not indifferent between x and y, then she **strictly prefers** x to y and we write $x \succ y$. In terms of utilities, an agent strictly prefers x to y if and only if $u(x) > u(y)$.

2.1 Monotonicity

Preferences are **monotone** if for any two bundles $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$,

$$\left.\begin{array}{l} x_i \geq y_i \text{ for each } i \\ x_i > y_i \text{ for some } i \end{array}\right\} \text{ implies } x \succ y.$$

In words: preferences are monotone if more of any good makes the agent strictly better off.

While monotonicity is stated in terms of preferences, we can rewrite it in terms of utilities. Preferences are monotone if for any two bundles x and y,

$$\left.\begin{array}{l} x_i \geq y_i \text{ for each } i \\ x_i > y_i \text{ for some } i \end{array}\right\} \text{ implies } u(x) > u(y).$$

As we will see, the assumption of monotonicity is very useful. It implies that indifference curves are thin and downwards sloping. It implies that an agent will always spend her budget. A slightly stronger version of monotonicity also rules out inflexion points in the agent's utility function which is useful when we analyse the agent's utility maximisation problem.

2.2 Convexity

Preferences are **convex** if whenever $x \succcurlyeq y$ then

$$tx + (1-t)y \succcurlyeq y \qquad \text{for all } t \in [0,1]$$

Convexity says that the agent prefers averages to extremes: if the agent is indifferent between x and y then she prefers the average $tx + (1-t)y$ to either x or y.

We can write this assumption in terms of utilities. Preferences are convex if whenever $u(x) \geq u(y)$ then

$$u(tx + (1-t)y) \geq u(y) \qquad \text{for all } t \in [0,1] \tag{2.1}$$

Slightly confusingly, a utility function that satisfies (2.1) is called **quasi–concave**.

The assumption of convexity if important when we analyse the consumers utility maximisation problem. Along with monotonicity, it means that any solution to the agent's first–order conditions solve the agent's problem.

3 Indifference Curves

An agent's **indifference curve** is the set of bundles which yield a constant level of utility. That is,

$$\text{Indifferent Curve} = \{x \in X | u(x) = \text{const.}\}$$

An agent has a collection of indifference curves, each one corresponding to a different level of utility. By varying this level, we can trace out the agent's entire preferences.

To illustrate, suppose an agent has utility

$$u(x_1, x_2) = x_1 x_2$$

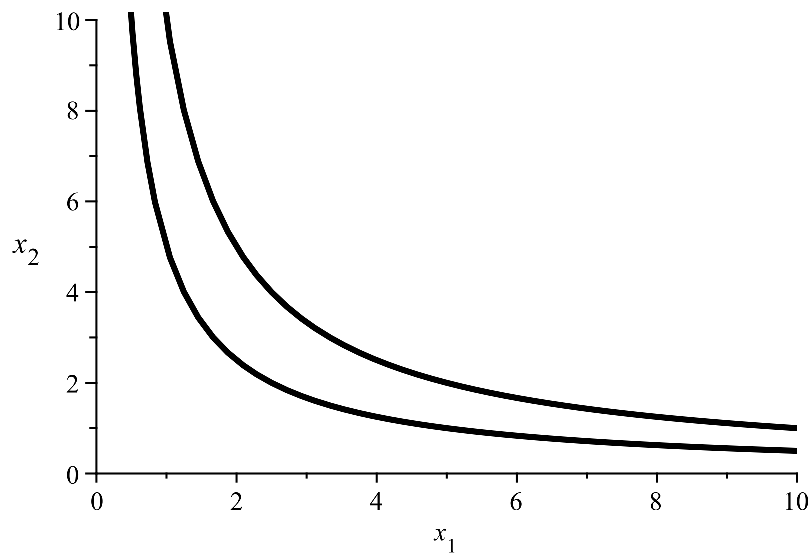

Figure 1: **Indifference Curves.** This figure shows two indifference curves. Each curve depicts the bundles that yield constant utility.

Then the indifference curve satisfies the equation $x_1x_2 = k$. Rearranging, we can solve for x_2, yielding

$$x_2 = \frac{k}{x_1} \tag{3.1}$$

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

In Section 3.1 we derive five important properties of indifference curves. In Section 3.2 we introduce the idea of the marginal rate of substitution. For simplicity, we assume there are only two goods.

3.1 Properties of Indifference Curves

We now describe five important properties of indifference curves. Throughout, we assume that preferences satisfy completeness, transitivity and continuity, so a utility function exists. We also assume monotonicity.

1. Indifference curves are thin. We say an indifference curve is thick if it contains two points x and y such that $x_i > y_i$ for all i. This is illustrated in figure 2. Monotonicity says that y must

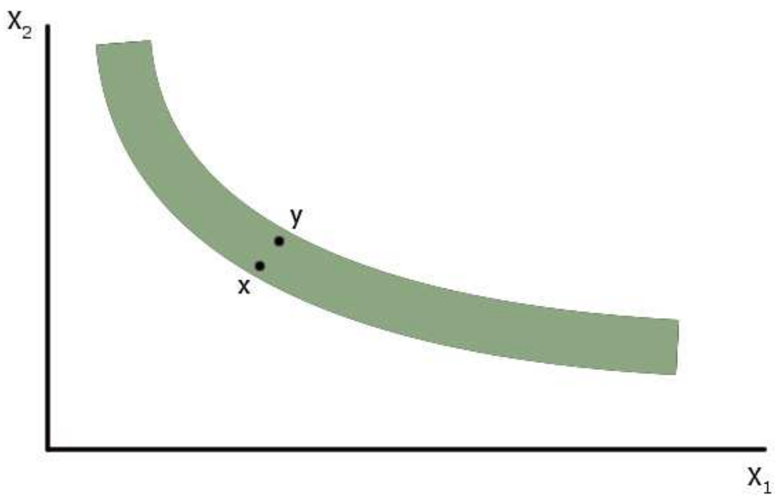

Figure 2: **A Thick Indifference Curve.** This figure shows a thick indifference curve containing points x and y.

be strictly preferred to x and therefore rules out thick indifference curves.

2. Indifference curves never cross. Suppose, by contradiction, that two indifference curves cross, as shown in figure 3. Since they lie on the same indifference curve, the agent is indifferent between points A and D, and indifferent between points B and C. In addition, by monotonicity, the agent strictly prefers A to B, and strictly prefers C to D. Putting all this together,

$$A \succ B \sim C \succ D \sim A$$

We conclude that A is strictly preferred to itself, which is false. Intuitively, two indifference curves describe the bundles that yield two different utility levels. By monotonicity, one indifference curve must always lie to the northeast of the other.

3. Indifference curves are strictly downward sloping. If an indifference curve is not strictly downward sloping, then we can find points x and y on the same indifference curve such that $y_i \geq x_i$ for all i, and $y_i > x_i$ for some i, as shown in figure 4. This contradicts monotonicity, which says the agent strictly prefers y to x.

4. Indifference curves are continuous, with no gaps. We cannot have gaps in the indifference curve, as shown in figure 5. This follows from preferences being continuous which, by Theorem 3, implies that the utility function is continuous.

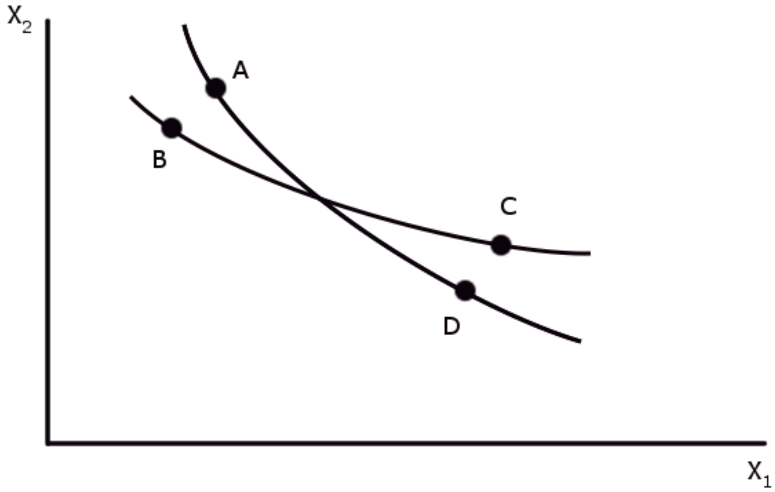

Figure 3: **Indifference Curves that Cross.**

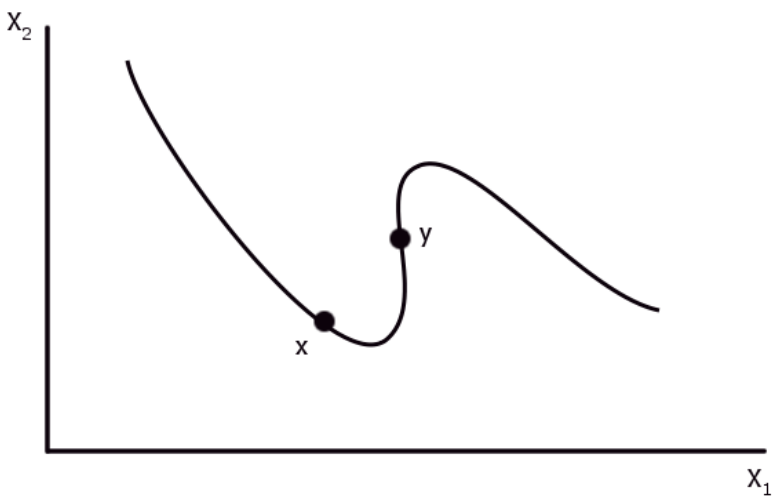

Figure 4: **An Upward Sloping Indifference Curve.**

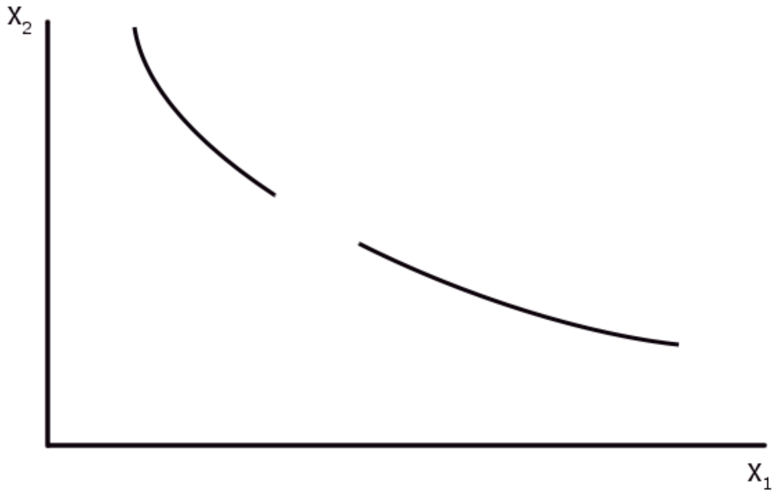

Figure 5: **An Indifference Curve with a Gap.**

5. If preferences are convex then indifference curves are convex to the origin. Suppose x and y lie on an indifference curve. By convexity, $tx + (1-t)y$ lies on a higher indifference curve, for $t \in [0, 1]$. By monotonicity, this higher indifference curve lies to the northeast of the original indifference curve. Hence the indifference curve is convex, as shown in figure 6.

3.2 Marginal Rate of Substitution

The slope of the indifference curve measures the rate at which the agent is willing to substitute one good for another. This slope is called the **marginal rate of substitution** or **MRS**. Mathematically,

$$MRS = -\frac{dx_2}{dx_1}\Big|_{u(x_1,x_2)=\text{const.}} \tag{3.2}$$

We can rephrase this definition in words: the MRS equals the number of x_2 the agent is willing to give up in order to obtain one more x_1. This is shown in figure 7.

The MRS can be related to the agent's utility function. First, we need to introduce the idea of **marginal utility**

$$MU_i(x_1, x_2) = \frac{\partial u(x_1, x_2)}{\partial x_i}$$

which equals the gain in utility from one extra unit of good i.

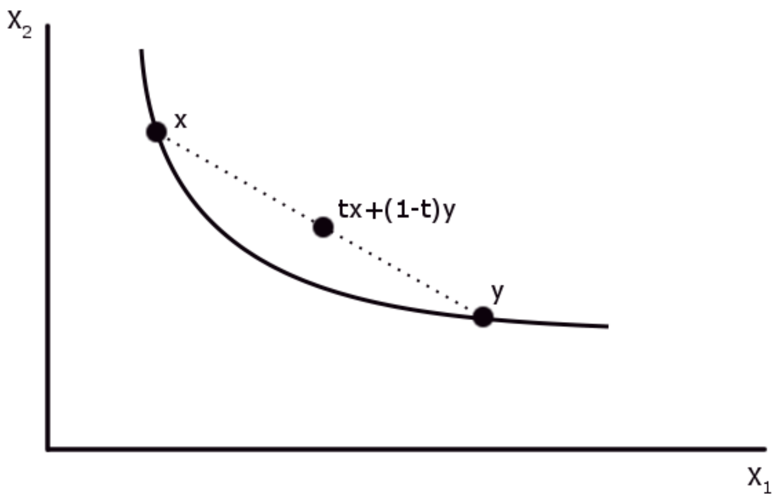

Figure 6: **Convex Indifference Curves.**

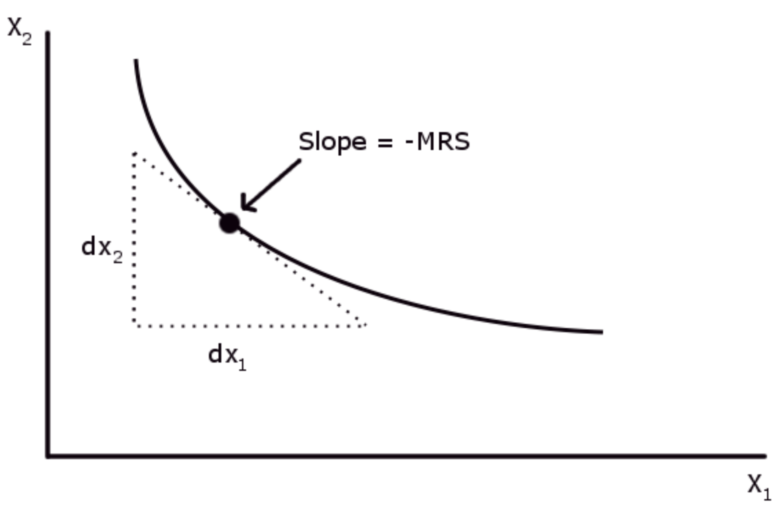

Figure 7: **Marginal Rate of Substitution.**

Let us consider the effect of a small change in the agent's bundle. Totally differentiating the utility $u(x_1, x_2)$ we obtain

$$du = \frac{\partial u(x_1, x_2)}{\partial x_1} dx_1 + \frac{\partial u(x_1, x_2)}{\partial x_2} dx_2 \tag{3.3}$$

Equation (3.3) says that the agent's utility increases by her marginal utility from good 1 times the increase in good 1 plus the marginal utility from good 1 times the increase in good 2. Along an indifference curve $du = 0$, so equation (3.3) becomes

$$\frac{\partial u(x_1, x_2)}{\partial x_1} dx_1 + \frac{\partial u(x_1, x_2)}{\partial x_2} dx_2 = 0$$

Rearranging,

$$-\frac{dx_2}{dx_1} = \frac{\partial u(x_1, x_2)/\partial x_1}{\partial u(x_1, x_2)/\partial x_2}$$

Equation (3.2) therefore implies that

$$MRS = \frac{MU_1}{MU_2} \tag{3.4}$$

The intuition behind equation (3.4) is as follows. Using the definition of MRS, one unit of x_1 is worth MRS units of x_2. That is, $MU_1 = MRS \times MU_2$. Rewriting this equation we obtain (3.4).

We can relate MRS to our earlier concepts of monotonicity and convexity. Monotonicity says that the indifference curve is downward sloping. Using equation (3.2), this means that MRS is positive.

Under the assumption of monotonicity, convexity says that the indifference curve is convex. This means that the MRS decreasing in x_1 along the indifference curve. Formally, an indifference curve defines an implicit relationship between x_1 and x_2,

$$u(x_1, x_2(x_1)) = k$$

Convexity then implies that $MRS(x_1, x_2(x_1))$ is decreasing in x_1. This is illustrated in the next section.

Finally, we can relate the MRS to the ordinal nature of the utility representation. In Theorem 2 we showed that the choices made under $u(x)$ and $v(x) = f(u(x))$ are the same, where $f : \Re \to \Re$ is strictly increasing. One way to understand this result is through the MRS. Under utility function $u(x)$ the MRS is given by equation (3.4). Under utility function $v(x)$, the MRS is given by

$$MRS^v = \frac{\partial v/\partial x_1}{\partial v/\partial x_2} = \frac{f'(u)\partial u/\partial x_1}{f'(u)\partial u/\partial x_2} = \frac{\partial u/\partial x_1}{\partial u/\partial x_2} = MRS^u$$

where the second equality uses the chain rule. This means that the agent faces the same tradeoffs under the two utility functions, has identical indifference curves, and therefore makes the same decisions.

3.3 Example: Symmetric Cobb Douglas

Suppose $u(x_1, x_2) = x_1 x_2$. We calculate the marginal rate of substitution two ways.

First, we can use equation (3.2) to derive MRS. As in equation (3.1), the equation of an indifference curve is

$$x_2 = \frac{k}{x_1} \tag{3.5}$$

Differentiating,

$$MRS = -\frac{dx_2}{dx_1} = \frac{k}{x_1^2} \tag{3.6}$$

We can now verify preferences are convex. Differentiating (3.6) with respect to x_1,

$$\frac{d}{dx_1} MRS = -2\frac{k}{x_1^3}$$

which is negative, as required.

Alternatively, we can use equation (3.4) to derive MRS. Differentiating the utility function

$$MRS = \frac{MU_1}{MU_2} = \frac{x_2}{x_1} \tag{3.7}$$

We now want to express MRS purely in terms of x_1. Using (3.5) to substitute for x_2, equation (3.7) becomes (3.6).

4 Examples of Preferences

4.1 Cobb Douglas

The Cobb–Douglas utility function is given by

$$u(x_1, x_2) = x_1^\alpha x_2^\beta \qquad \text{for } \alpha > 0,\ \beta > 0$$

A special case if the symmetric Cobb–Douglas, when $\alpha = \beta$. Using Theorem 2, we can then normalise the symmetric Cobb–Douglas to $\alpha = \beta = 1$.

The Cobb–Douglas indifference curve has equation $x_1^\alpha x_2^\beta = k$. Rearranging,

$$x_2 = k^{1/\beta} x_1^{-\alpha/\beta}$$

These indifference curves look like those in figure 1.

The marginal utilities are

$$MU_1 = \alpha x_1^{\alpha-1} x_2^\beta$$
$$MU_1 = \beta x_1^\alpha x_2^{\beta-1}$$

As a result the MRS is,

$$MRS = \frac{MU_1}{MU_2} = \frac{\alpha x_2}{\beta x_1}$$

4.2 Perfect Complements

Suppose an agent always consumes a hamburger patty with two slices of bread. If she has 5 patties and 15 slices of bread, then the last 5 slices are worthless. Similarly, if she has 7 patties and 10 slices of bread, then the last 2 patties are worthless. In this case, the agent's preferences can be represented by the utility function

$$u(x_1, x_2) = \min\{2x_1, x_2\}$$

where x_1 are patties and x_2 are slices of bread. Note the 2 goes in front of the number of patties because, intuitively speaking, each patty is twice as valuable as a piece of bread.

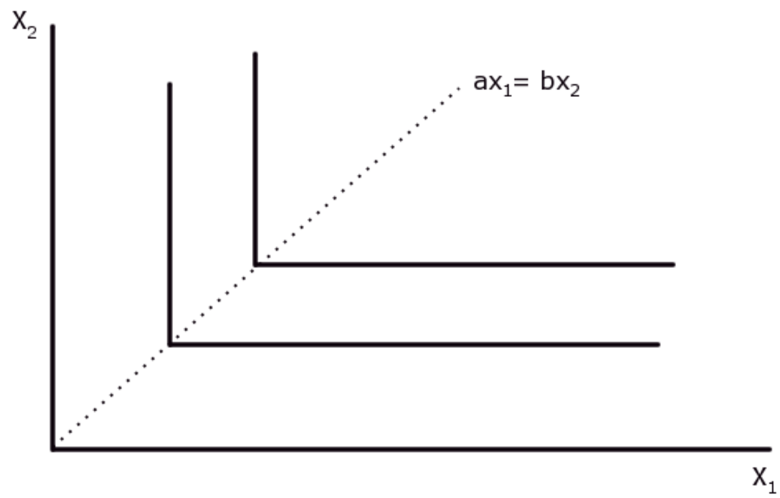

Figure 8: **Perfect Complements.** These indifference curves are L–shaped with the kink where $\alpha x_1 = \beta x_2$.

In general, preferences are perfect complements when they can be represented by a utility function of the form

$$u(x_1, x_2) = \min\{\alpha x_1, \beta x_2\}$$

The resulting indifference curves are L–shaped, as shown in figure 8, with the kink along the line $\alpha x_1 = \beta x_2$. Note that the indifference curve is not strictly decreasing along the bottom of the L. This is because these preferences do not quite obey the monotonicity condition: when the agent has 7 patties and 10 slices of bread, an extra patty does not strictly increase her utility.

The MRS in this example is a little odd. When $\alpha x_1 > \beta x_2$,

$$MRS = \frac{MU_1}{MU_2} = \frac{0}{\beta} = 0$$

When $\alpha x_1 < \beta x_2$,

$$MRS = \frac{MU_1}{MU_2} = \frac{\alpha}{0} = \infty$$

At the kink, when $\alpha x_1 = \beta x_2$, then MRS is not defined because the indifference curve is not differentiable.

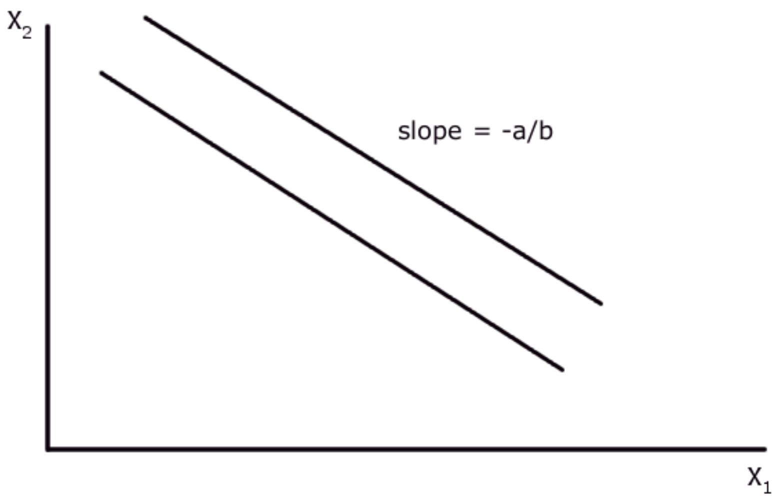

Figure 9: **Perfect Substitutes.** These indifference curves are linear with slope $-\alpha/\beta$.

4.3 Perfect Substitutes

Suppose an agent is buying food for a party. She wants enough food for her guests and considers 3 hamburgers to be equivalent to one pizza. Since each pizza is three times as valuable as a hamburger, her preferences can be represented by the utility function

$$u(x_h, x_p) = x_1 + 3x_2$$

where x_1 are hamburgers and x_2 are pizzas.

In general, preferences are perfect substitutes when they can be represented by a utility function of the form

$$u(x_1, x_2) = \alpha x_1 + \beta x_2$$

The resulting indifference curves are straight lines, as shown in figure 10. As a result, preferences are only weakly convex. The marginal rate of substitution is

$$MRS = \frac{MU_1}{MU_2} = \frac{\alpha}{\beta}$$

That is, the MRS is independent of the number of goods consumed.

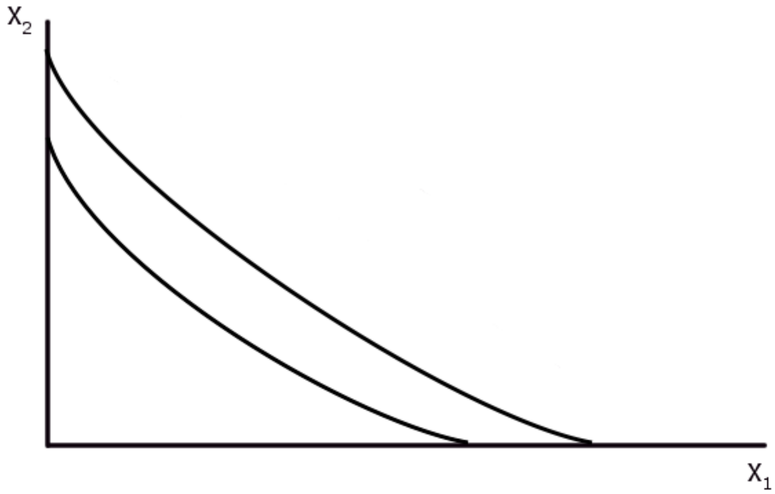

Figure 10: **CES Preferences.** In this picture $\delta > 0$ since the indifference curve intersects with the axes.

4.4 Constant Elasticity of Substitution (CES) Preferences

CES preferences have the form
$$u(x_1, x_2) = \frac{x_1^\delta}{\delta} + \frac{x_2^\delta}{\delta}$$
where $\delta \neq 0$ and $\delta < 1$.

This utility function can approximate the above examples. As $\delta \to 0$ the limit of the above utility function becomes
$$u(x_1, x_2) = \ln x_1 + \ln x_2$$
which is the same as Cobb-Douglas with equal exponents. As $\delta \to 1$, the preferences approximate perfect substitutes. As $\delta \to -\infty$, the preferences approximate perfect complements.

The MRS is
$$MRS = \frac{MU_1}{MU_2} = \frac{\delta x_1^{\delta-1}}{\delta x_2^{\delta-1}} = \frac{x_2^{1-\delta}}{x_1^{1-\delta}}.$$
The last expression is convenient since $1 - \delta > 0$. Substituting for x_2 in this equation and differentiating, one can show that MRS is decreasing in x_1, so the preferences are convex.

4.5 Additive Preferences

Additive preferences are represented by a utility function of the form

$$u(x_1, x_2) = v_1(x_1) + v_2(x_2)$$

The key property of additive preferences is that the marginal utility of x_i only depends on the amount of x_i consumed. As a result, the marginal rate of substitution is

$$MRS = \frac{MU_1}{MU_2} = \frac{v_1'(x_1)}{v_2'(x_2)}$$

For example, suppose we have

$$u(x_1, x_2) = x_1^2 + x_1^2$$

Differentiating, $MU_i = 2x_i$, so the marginal utility of each good is increasing in the amount of the good consumed. For example, one could imagine the agent becomes addicted to either good.

As shown in Figure 11, these preferences are concave. One can see this formally by showing the MRS is increasing in x_1 along an indifference curve. Differentiating,

$$MRS = \frac{MU_1}{MU_2} = \frac{2x_1}{2x_2} = \frac{x_1}{x_2} \tag{4.1}$$

The equation of an indifference curve is $x_1^2+x_2^2 = k$. Rearranging, $x_2 = (k-x_1^2)^{1/2}$. Substituting into (4.1),

$$MRS = \frac{x_1}{(k - x_1^2)^{1/2}}$$

which is increasing in x_1.

4.6 Bliss Points

Suppose preferences are represented by the utility function

$$u(x_1, x_2) = -\frac{1}{2}(x_1 - 10)^2 - \frac{1}{2}(x_2 - 10)^2$$

Figure 12 plots the resulting indifference curves which are concentric circles around the bliss point of $(x_1, x_2) = (10, 10)$. These preferences violate monotonicity; as a result the indifference

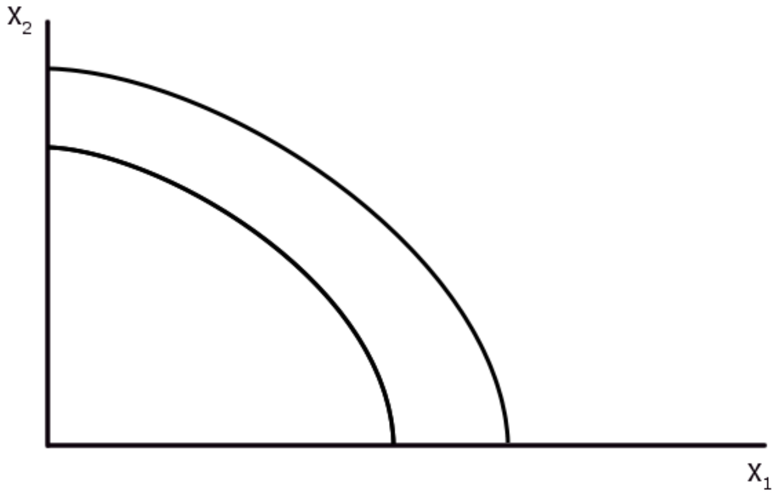

Figure 11: **Addiction Preferences.** These indifference curves are concave.

curves are sometimes upward sloping.

The marginal rate of substitution is

$$MRS = \frac{MU_1}{MU_2} = \frac{10 - x_1}{10 - x_2}$$

Hence the MRS is positive in the northeast and southwest quadrants, and is negative in the northwest and southeast quadrants. From figure 12 one can also see that preferences are convex. This is also possible to see from the MRS, but is a little tricky since monotonicity does not hold.

4.7 Quasilinear Preferences

An agent has quasilinear preferences if they can be represented by a utility function of the form

$$u(x_1, x_2) = v(x_1) + x_2$$

Quasilinear preferences are linear in x_2, so the marginal utility is constant. These preferences are often used to analyse goods which constitute a small part of an agent's income; good x_2 can then be thought of as "general consumption".

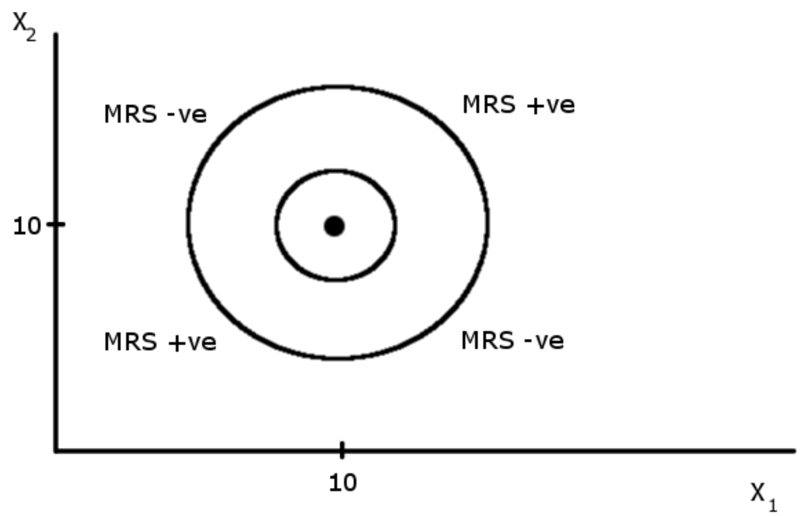

Figure 12: **Bliss Point.** Utility is maximised at $(10, 10)$. Indifference curves are circles around this bliss point.

The marginal rate of substitution equals

$$MRS = \frac{MU_1}{MU_2} = \frac{v'(x_1)}{1} = v'(x_1)$$

Observe that MRS only depends on x_1, and not x_2. This means that the indifference curves are vertical parallel shifts of each other, as shown in figure 13. As a consequence, preferences are convex if and only if $v(x_1)$ is a concave function, so the marginal utility of x_1 decreases in x_1.

As we will see later, quasilinear preferences have the attractive property that the consumption of x_1 is independent of the agent's income (ignoring boundary constraints). This makes the consumer's problem simple to analyse and provides an easy way to calculate consumer surplus.

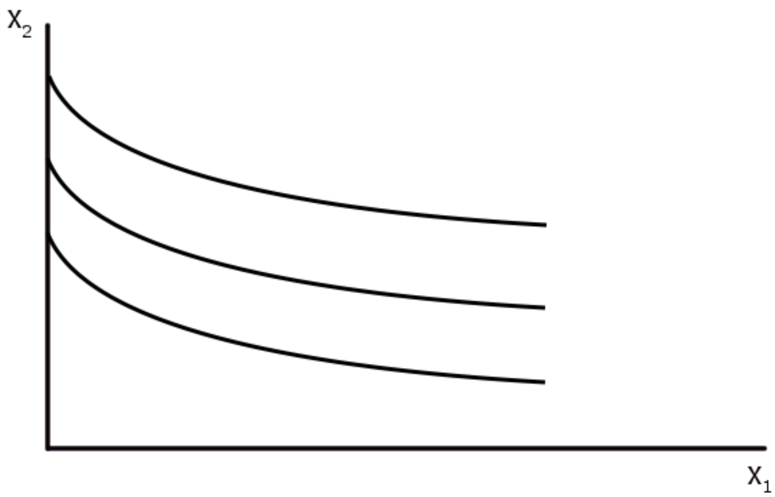

Figure 13: **Quasilinear Preferences.** These indifference curves are parallel shifts of each other.

Utility Maximisation Problem

Simon Board*

This Version: September 20, 2009
First Version: October, 2008.

The utility maximisation problem (UMP) considers an agent with income m who wishes to maximise her utility. Among others, we are interested in the following questions:

- How do we determine an agent's optimal bundle of goods?
- How do we derive an agent's demand curve for a particular good?
- What is the effect of an increase in income on an agent's consumption?

1 Model

We make several assumptions:

1. There are N goods. For much of the analysis we assume $N = 2$, but nothing depends on this.
2. The agent takes prices as exogenous. We normally assume prices are linear and denote them by $\{p_1, \ldots, p_N\}$.
3. Preferences satisfy completeness, transitivity and continuity. As a result, a utility function exists. We normally assume preferences also satisfy monotonicity (so indifference curves are well behaved) and convexity (so the optima can be characterised by tangency conditions).

*Department of Economics, UCLA. http://www.econ.ucla.edu/sboard/. Please email suggestions and typos to sboard@econ.ucla.edu.

4. The consumer is endowed with income m.

The utility maximisation problem is:

$$\max_{x_1,\dots,x_N} u(x_1,\dots,x_N) \qquad \text{subject to} \qquad \begin{array}{l} \sum_{i=1}^{N} p_i x_i \leq m \\ x_i \geq 0 \quad \text{for all } i \end{array} \tag{1.1}$$

The idea is that the agent is trying to spend her income in order to maximise her utility. The solution to this problem is called the **Marshallian demand** or uncompensated demand. It is denoted by

$$x_i^*(p_1,\dots,p_N,m)$$

The most utility the agent can attain is given by her **indirect utility function**. It is defined by

$$v(p_1,\dots,p_N,m) = \max_{x_1,\dots,x_N} u(x_1,\dots,x_N) \qquad \text{subject to} \qquad \begin{array}{l} \sum_{i=1}^{N} p_i x_i \leq m \\ x_i \geq 0 \quad \text{for all } i \end{array} \tag{1.2}$$

Equivalently, the indirect utility function equals the utility the agent gains from her optimal bundle,

$$v(p_1,\dots,p_N,m) = u(x_1^*,\dots,x_N^*).$$

1.1 Example: One Good

To illustrate the problem, suppose $N = 1$. For example, the agent has income m and is choosing how many cookies to consume. The agent's utilities are given by table 1.

In general, we solve the problem in two steps. First, we determine which bundles of goods are affordable. The collection of these bundles is called the **budget set**. Second, we find which bundle in the budget set the agent most prefers. That is, which bundle gives the agent most utility.

Suppose the price of the good is $p_1 = 1$ and the agent has income $m = 4$. Then the agent can afford up to 4 units of x_1. Given this budget set, the agent's utility is maximised by choosing $x_1^* = 4$, yielding utility $v = 28$.

Units of x_1	Utility
1	10
2	18
3	24
4	28
5	30
6	29
7	26
8	21

Table 1: **Utilities from different bundles.** Observe that this agent is satiated at 5 units.

Next, suppose the price of the good is $p_1 = 1$ and the agent has income $m = 8$. Then the agent can afford up to 8 units of x_1. Given this budget set, the agent's utility is maximised by choosing $x_1^* = 5$, yielding utility $v = 30$. In this example, the consumer can afford 8 units but chooses to consume 5. If the agent's preferences are monotone, then she will always spend her entire budget.

Finally, suppose the price of the good is $p_1 = 2$ and the agent has income $m = 8$. Then the agent can afford up to 4 units of x_1, as in the original case. This illustrates that the budget set is determined jointly by the prices and income: doubling both does not change the agent's budget set. When maximising her utility, the agent once again chooses $x_1^* = 4$.

2 Budget Sets

As in Section 1.1, we will solve the agent's problem in two steps. First, we determine which bundles of goods are affordable. Second, we find which of these bundles yields the agent the highest utility. In this section we look at the first step.

2.1 Standard Budget Sets

In the standard model, we assume there are unit prices $\{p_1, p_2\}$ for the 2 goods. The **budget set** is the collection of bundles (x_1, x_2) such that (a) the quantities are positive; and (b) the bundle is affordable. Mathematically, the budget set is

$$\{(x_1, x_2) \in \Re^2_+ : p_1 x_1 + p_2 x_2 \leq m\}$$

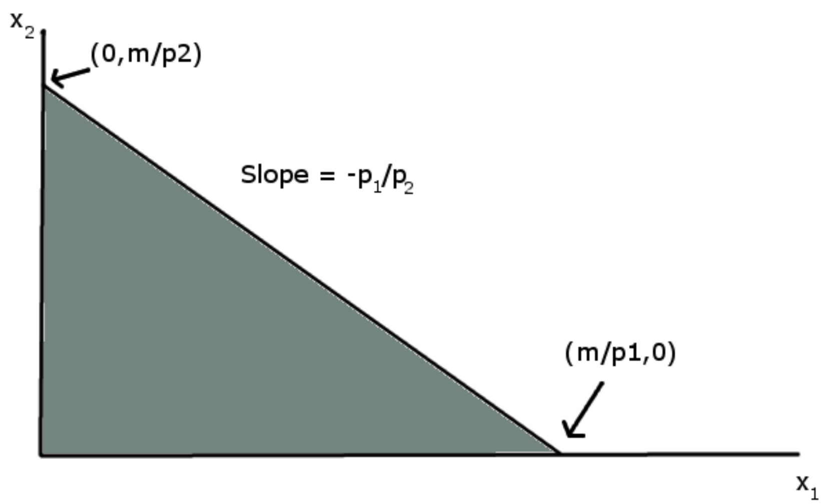

Figure 1: **Budget Set.**

where $\Re_+$ is the positive part of the real line, and $\Re^2_+$ is the positive orthont in $\Re^2$.

Figure 1 illustrates such a budget set. The equation where the budget binds is given by

$$p_1x_1 + p_2x_2 = m \tag{2.1}$$

We can rearrange this to be in the form of a standard linear equation

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}x_1 \tag{2.2}$$

Hence the budget line is linear with intercept m/p_2 and slope $-p_1/p_2$. Crucially, the slope only depends on the relative prices.

The two endpoints are easy to calculate. If the agent spends all her money on x_1 she can afford

$$x_1 = \frac{m}{p_1} \qquad \text{and} \qquad x_2 = 0$$

If the agent spends all her money on x_2 she can afford

$$x_1 = 0 \qquad \text{and} \qquad x_2 = \frac{m}{p_2}$$

Figure 2 shows that an increase in the agent's income leads the budget line to make a parallel

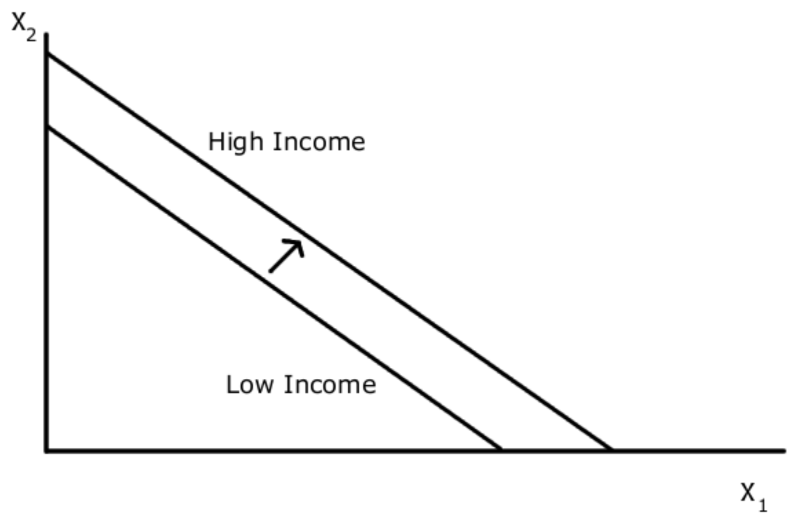

Figure 2: **An Increase in Income.**

shift outwards. Mathematically, this can be seen from equation (2.2). Intuitively, if the agent's budget doubles then she can double her consumption of both goods. Since relative prices do not change, the new budget line is parallel to the old one.

Figure 3 shows that an increase in p_1 leads the budget curve to pivot around it's left endpoint. Mathematically, this can be seen from equation (2.2). Intuitively, if the agent only buys x_2, then her purchasing power is unaffected by the increase in p_1. As a result, the left endpoint does not move. If the agent only buys x_1, then the increase in p_1 reduces the amount she can buy, forcing the right endpoint to shift in. As a result, the budget line become steeper, reflecting the change in the relative prices.

2.2 Nonlinear Budget Sets

While we focus on linear budget constraints, agents often face nonlinear prices. Here we present some examples.

Figure 4 shows an example of quantity discounts. In this example, the agent has income $m = 30$. Good 1 has per–unit price $p_1 = 2$ for $x_1 < 10$, and per–unit price $p_1 = 1$ for $x_1 \geq 10$. Good 2 has a constant price, $p_2 = 2$. Lets consider 2 cases. First, when the agent buys $x_1 < 10$, the price of good 1 is $p_1 = 2$ and the equation of the budget line is therefore $2x_1 + 2x_2 = 30$ or

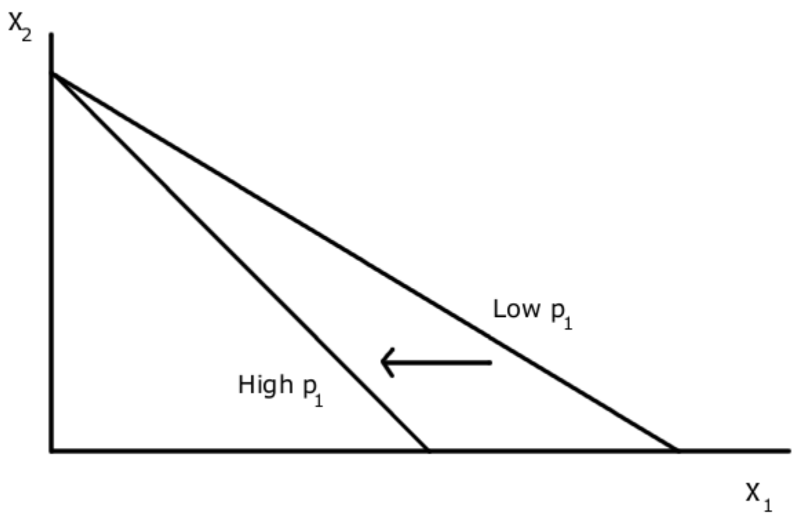

Figure 3: **An Increase in the Price of Good 1.**

$x_2 = 15 - x_1$. For example, when the agent spends all her money on good 2, she can afford $x_2 = 15$. Second, when $x_1 \geq 10$ the agent spends \$20 on the first 10 units of x_1 and \$1 per unit thereafter. Hence her budget constraint is

$$20 + (x_1 - 10) + 2x_2 = 30$$

Figure 5 shows an example of rationing. In this example, the agent has income $m = 30$. Good 1 has per–unit price $p_1 = 2$ for $x_1 \leq 10$, but she is only allowed to purchase 10 units. Good 2 has a constant price, $p_2 = 2$. When the agent buys $x_1 \leq 10$, the price of good 1 is $p_1 = 2$ and the budget line is $2x_1 + 2x_2 = 30$. For example, when the agent spends all her money on good 2, she can afford $x_2 = 15$. The agent is unable to buy more than 10 units of x_1, so the budget set is cut off at $x_1 = 10$.

Exercise: Excess tax on x_1. Suppose $m = 30$, $p_2 = 2$, and $p_1 = 2$ for the first 10 units and $p_1 = 3$ for each additional unit. Draw the agent's budget set.

Exercise: Food stamps. Suppose $m = 30$, $p_2 = 2$, and $p_1 = 0$ for the first 10 units and $p_1 = 2$ for each additional unit. Draw the agent's budget set.

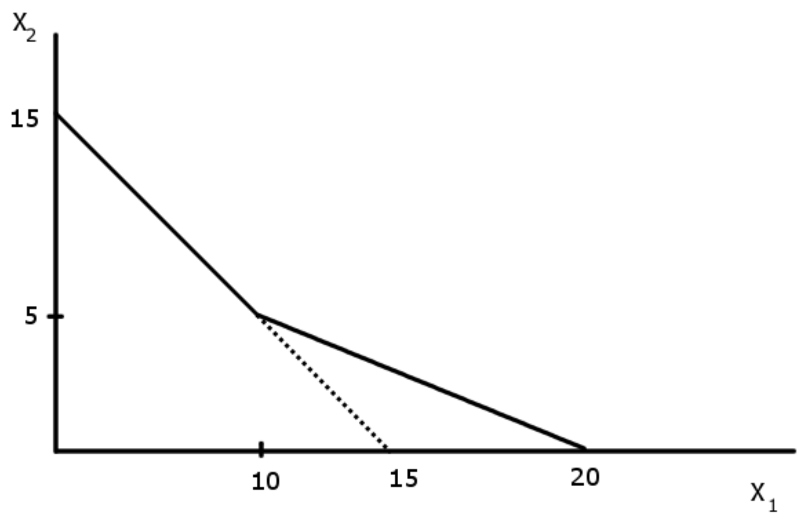

Figure 4: **Quantity Discounts.** The dark line shows the agent's budget set with $p_1 = 1$ for $x_1 \geq 10$. The dotted line shows her budget set if $p_1 = 2$ for all units.

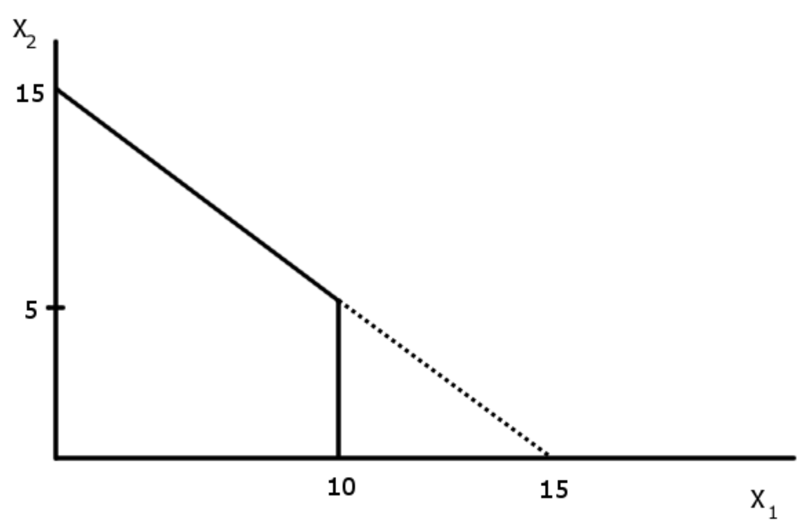

Figure 5: **Rationing.** The dark line shows the agent's actual budget set given she can only buy 10 units of x_1. The dotted line shows her budget set without rationing.

3 Solving the Utility Maximisation Problem

In this section we solve the agent's utility maximisation problem. We make a number of simplifying assumptions which we explore in Section 4. In particular, we assume:

- The agents utility function is differentiable. As a result there are no kinks in the indifference curve.
- The agent's preferences are monotone. As a result, she spends her entire budget.[1]
- The agent's preferences are convex. As a result, any solution to the tangency conditions constitute a maximum.
- There is an interior solution to the agent's maximisation problem.

3.1 Solution Method 1: Graphical Approach

The agent wishes to choose a point in her budget set to maximise her utility. That is, the agent wishes to choose a point in her budget set that lies on the highest indifference curve.[2]

Figure 6 characterises the agent's optimal choice. Graphically, one can imagine the indifference curve flying in from the top right corner (where utility is highest) and stopping when it touches the budget set.

To understand this further, consider figure 7. There are 3 indifference curves. I_1 yields the highest utility, but never intersects with the budget set. I_2 is corresponds to the agent's optimal choice (point A). I_3 yields a lower level of utility which is attainable but not desirable.

At the optimal point, the budget line is tangential to the indifference curve. As a result the budget line and the indifference curve have the same slope. This **tangency condition** means that

$$\text{MRS}(x_1^*, x_2^*) = \frac{p_1}{p_2} \tag{3.1}$$

[1] We actually assume $\partial u(x_1, x_2)/\partial x_i > 0$ for each i. See Section 4.3.

[2] Recall monotonicity and convexity implies that indifference curves are thin, downward sloping and convex.

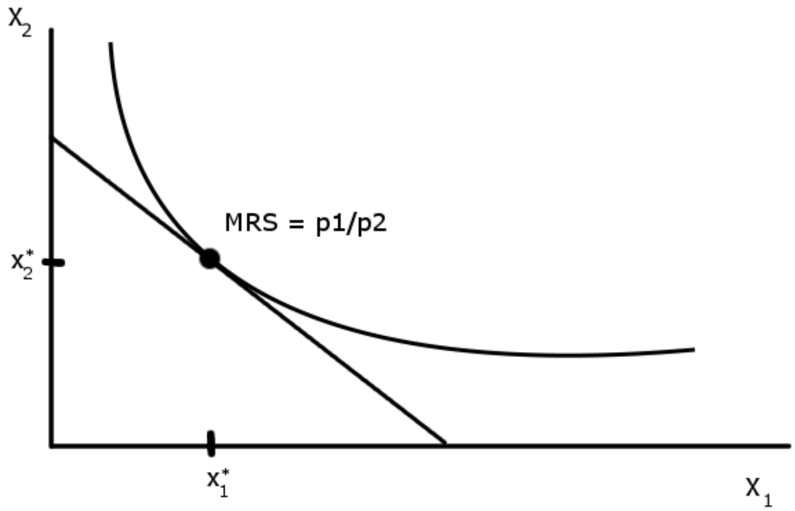

Figure 6: **The Agent's Optimal Demand.** This figure shows how the agent's optimal demand is characterised by the tangency condition.

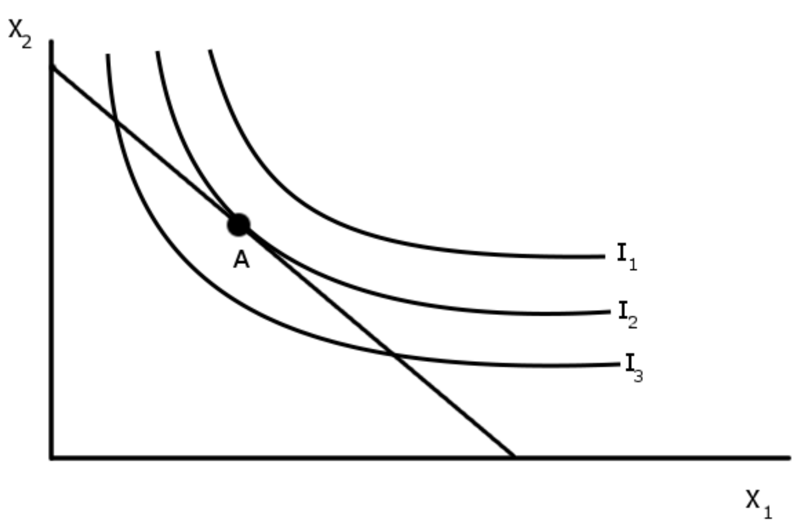

Figure 7: **Understanding the Tangency Condition.** In this figure I_2 is the highest attainable indifference curve. I_3 is higher but unaffordable; I_1 is affordable but not optimal.

where the marginal rate of substitution is evaluated at the optimal choice, (x_1^*, x_2^*). Since $MRS = MU_1/MU_2$,[3] we can write the tangency condition as

$$\frac{MU_1(x_1^*, x_2^*)}{MU_2(x_1^*, x_2^*)} = \frac{p_1}{p_2} \tag{3.2}$$

The intuition for this result is as follows. The MRS equals the number of x_2 the agent is willing to give up to get one more unit of x_1. The price ratio equals the number of x_2 the agent has to give up in order to get one more unit of x_1 if she wishes to stay within her budget. If $MRS > p_1/p_2$, then the agent is more willing to give up x_2 than the market requires, so she can increase her utility by consuming less x_2 and more x_1. If $MRS < p_1/p_2$, then the agent is less willing to give up x_2 than the market requires, so she can increase her utility by consuming more x_2 and less x_1.

One can rewrite the tangency condition as:

$$\frac{MU_1(x_1^*, x_2^*)}{p_1} = \frac{MU_2(x_1^*, x_2^*)}{p_2} \tag{3.3}$$

Equation (3.3) says that the agent equalises the marginal utility per dollar, or the **bang–per–buck** of the two goods. If the bang–per–buck from good 1 is higher than that from good 2, then the agent buys more of good 1. If the bang–per–buck from good 1 is lower than that from good 2, then the agent buys less of good 1. At the optimal choice, the bang–per–buck of the two goods is equal.

3.2 Example: Symmetric Cobb–Douglas.

Suppose $u(x_1, x_2) = x_1 x_2$. The MRS is

$$\frac{MU_1}{MU_2} = \frac{x_2}{x_1}$$

The tangency condition (3.1) therefore becomes

$$\frac{x_2}{x_1} = \frac{p_1}{p_2}$$

[3] Recall $MU_i(x_1, x_2) = \partial u(x_1, x_2)/\partial x_i$.

Rewriting, we see that

$$p_1 x_1 = p_2 x_2 \tag{3.4}$$

which means that the agent's spends the same money on each good. This is a special property of the symmetric Cobb–Douglas model and helps make it so tractable. The budget constraint states that

$$p_1 x_1 + p_2 x_2 = m$$

Substituting equation (3.4) into the budget constraint implies that $p_1 x_1 = p_2 x_2 = m/2$, so the agent spends half her income on each good. As a result, the Marshallian demands are

$$x_1^*(p_1, p_2, m) = \frac{m}{2p_1} \qquad \text{and} \qquad x_2^*(p_1, p_2, m) = \frac{m}{2p_2} \tag{3.5}$$

We can also calculate the agent's indirect utility, her utility from the optimal bundle. Using the demands in equation (3.5), we have

$$v(p_1, p_2, m) = u(x_1^*, x_2^*) = x_1^* x_2^* = \left(\frac{m}{2p_1}\right)\left(\frac{m}{2p_2}\right) = \frac{m^2}{4p_1 p_2}$$

3.3 Solution Method 2: Lagrangian Approach

A second way to solve the agent's utility maximisation problem is to use a Lagrangian. This approach is equivalent to the tangency approach but can be more convenient, especially with complex problems.

At the optimal solution, equation (3.3) tells us that the agent equalises the bang–per–buck from each good. Let λ equal the marginal utility the agent derives from \$1 at her optimal bundle. We then have

$$\lambda = \frac{MU_1(x_1^*, x_2^*)}{p_1} \tag{3.6}$$

$$\lambda = \frac{MU_2(x_1^*, x_2^*)}{p_2} \tag{3.7}$$

We can encode equations (3.6) and (3.7) as the first–order conditions of one single equation.

Consider maximising the **Lagrangian**,

$$\mathcal{L} = u(x_1, x_2) + \lambda[m - p_1x_1 - p_2x_2] \tag{3.8}$$

In this equation λ is called a **Lagrange multiplier**. Intuitively, we are maximising the agent's utility plus a penalty term which punishes the agent for exceeding her budget (if $m - p_1x_1 - p_2x_2 < 0$). If the penalty λ is too low the agent will spend more than her income; if the penalty λ is too high the agent will spend less than her income. We must therefore choose the penalty λ so the agent exactly spends her budget.

Mechanically, we solve the problem as follows. First, we derive the first–order–conditions of the Lagrangian (3.8). This yields:

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial u(x_1, x_2)}{\partial x_1} - \lambda p_1 = 0 \tag{3.9}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial u(x_1, x_2)}{\partial x_2} - \lambda p_2 = 0 \tag{3.10}$$

We also know the agent's budget constraint binds:

$$p_1x_1 + p_2x_2 = m \tag{3.11}$$

We now have three unknowns (x_1, x_2, λ) and three equations: (3.9), (3.10) and (3.11). We can therefore solve for the agent's optimal demands.

We can relate these results to those in Section 3.1. Rearranging, equations (3.9) and (3.10) yield equations (3.6) and (3.7), which means we can interpret the optimal λ as the "bang–per–buck" of the optimal bundle. In addition, dividing (3.9) by (3.10) yields

$$\frac{\partial u(x_1, x_2)/\partial x_1}{\partial u(x_1, x_2)/\partial x_2} = \frac{p_1}{p_2}$$

which is the tangency condition (3.2) from Section 3.1.

3.4 Solution Method 3: Substitution

The agent's problem is to maximise her utility subject to her budget constraint. When there are two goods one can solve for the agent's optimal bundle by substituting the budget constraint directly into the objective function.

Recalling equation (2.2), the budget constraint can be written as

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}x_1$$

Substituting this into the agent's utility function, she chooses x_1 to maximise

$$u\left(x_1, \frac{m}{p_2} - \frac{p_1}{p_2}x_1\right)$$

The first–order condition is

$$\frac{\partial u(x_1,x_2)}{\partial x_1} + \frac{\partial u(x_1,x_2)}{\partial x_1}\left(-\frac{p_1}{p_2}\right) = 0$$

Rearranging,

$$\frac{\partial u(x_1,x_2)/\partial x_1}{\partial u(x_1,x_2)/\partial x_2} = \frac{p_1}{p_2}$$

which is the tangency condition (3.2) from Section 3.1. Using the budget constraint, we can therefore solve the for agent's optimal bundle (x_1^*, x_2^*).

3.5 Many goods

The tangency condition (3.3) extends to many goods. In this case, the agent equates the bang–per–buck from each of the N goods. That is,

$$\frac{MU_1}{p_1} = \frac{MU_2}{p_2} = \ldots = \frac{MU_N}{p_N}$$

One can then derive the agent's optimal demand using these $(N-1)$ equations and the budget constraint.

Equivalently, one can set up the Lagrangian

$$\mathcal{L} = u(x_1, x_2, \ldots, x_N) + \lambda[m - p_1x_1 - p_2x_2 - \ldots - p_Nx_N]$$

One can then solve for the agent's optimal demand and the Lagrange multiplier using the N first–order–conditions and the budget constraint.

4 Problems with the UMP

In this Section we investigate the tangency conditions in more depth. In Section 4.1 we state a formal version of the theorem we were implicitly using in Section 3. We then look at the four assumptions we made at the start of the last section: kinks, the monotonicity of preferences, the convexity of preferences, and boundary problems.

4.1 Kuhn–Tucker Conditions

Suppose the agent is solving the problem:

$$\max_{x_1,x_2} u(x_1,x_2) \qquad \text{subject to} \qquad \begin{aligned} &p_1x_1+p_2x_2 \leq m \\ &x_1 \geq 0 \\ &x_2 \geq 0 \end{aligned} \tag{4.1}$$

Form the Lagrangian

$$\mathcal{L} = u(x_1,x_2) + \lambda[m - p_1x_1 - p_2x_2] + \mu_1x_1 + \mu_2x_2 \tag{4.2}$$

where the last two terms are the penalties associated with constraints $x_1 \geq 0$ and $x_2 \geq 0$.

Suppose $u(x_1,x_2)$ is continuously differentiable and (x_1^*, x_2^*) solves (4.1). Then four **Kuhn–Tucker conditions** hold:

(a) The first order conditions hold: $\frac{\partial \mathcal{L}}{\partial x_1} = 0$ and $\frac{\partial \mathcal{L}}{\partial x_2} = 0$.

(b) The Lagrange multipliers are positive: $\lambda, \mu_1, \mu_2 \geq 0$.

(c) The constraints hold: $m - p_1x_1 - p_2x_2 \leq 0$, $x_1 \geq 0$ and $x_2 \geq 0$.

(d) Complimentary slackness holds: $\lambda[m - p_1x_1 - p_2x_2] = 0$, $\mu_1x_1 = 0$ and $\mu_2x_2 = 0$.

The idea behind these conditions is exactly the same as in Section 3.3. Part (a) say that the agent is choosing (x_1, x_2) to maximise her utility plus the penalty functions. Part (b) says that the penalties are positive. Part (c) says that the agent's choice must be feasible. Part (d) says that we cannot have a constraint slack and have the associated Lagrange multiplier being positive. For example, consider the budget constraint and recall that the Lagrange multiplier can be interpreted as the bang–per–buck. Complimentary slackness says that if the budget constraint at the optimal solution is slack, then the bang–per–buck equals zero and hence $\lambda = 0$.

We now present the formal result.

Theorem 1. *Suppose the utility function is continuously differentiable.*
(i) Any optimal solution satisfies the conditions (a)–(d).
(ii) If preferences are convex and $MU_i(x_1, x_2) > 0$ for each i, then the solutions to (a)–(d) are optimal.

In Section 4.2 we look at the issue of kinks, where the optimal solution may not satisfy the Kuhn–Tucker conditions. In Sections 4.3–4.4 we investigate what happens if monotonicity and convexity do not hold. In this case the there may be multiple solutions to the Kuhn–Tucker conditions, some of which are not optimal. Finally, in Section 4.5 we look at the issue of boundary constraints, where the nonnegativity constraints may bind.

4.2 Kinks

When deriving the tangency condition, we assumed that the utility function (and hence the indifference curve) is differentiable.

Suppose there is a kink in the indifference curve along the as shown in figure 8. At the kink, the MRS to the left and right are different. Let the MRS to the left be denoted MRS^L and than to the right be denoted MRS^R. Then the solution is at the kink if

$$MRS^L(x_1^*, x_2^*) \geq \frac{p_1}{p_2} \geq MRS^R(x_1^*, x_2^*) \tag{4.3}$$

Noting that $MRS = MU_1/MU_2$, equation (4.3) says that to the left of (x_1^*, x_2^*) we have

$$\frac{MU_1}{p_1} \geq \frac{MU_2}{p_2}$$

so the agent wishes to increase x_1. While to the right of (x_1^*, x_2^*) we have

$$\frac{MU_1}{p_1} \leq \frac{MU_2}{p_2}$$

so the agent wishes to decrease x_1. Putting this together, when $x_1 < x_1^*$, the agent wishes to

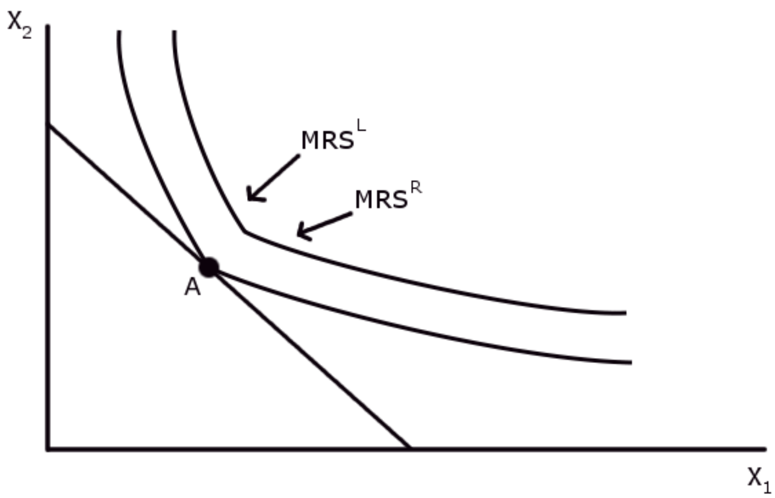

Figure 8: **Indifference Curve with Kinks.** The agent chooses point A. At the optimum, $MRS^L \geq p_1/p_2 \geq MRS^R$.

buy more x_1; when $x_1 > x_1^*$, the agent wishes to buy less x_1. Hence the optimal solution is at $x_1 = x_1^*$.

One way to approach this is to think of the MRS at the kink as being an entire set of numbers $[MRS^R, MRS^L]$. The tangency condition then says that the solution is at the kink if p_1/p_2 falls in the set.

4.2.1 Example: Perfect Complements

Suppose $u(x_1, x_2) = \min\{\alpha x_1, \beta x_2\}$. In this case utility is not differentiable along the line $\alpha x_1 = \beta x_2$, as shown in figure 9. The agent's optimal bundle clearly has the property that $\alpha x_1 = \beta x_2$. One can then use the agent's budget constraint to solve for the optimal bundle.

To illustrate, suppose an agent's utility is

$$u(x_1, x_2) = \min\{2x_1, x_2\}$$

The agent's income is $m = 30$ and prices are $p_1 = 1$ and $p_2 = 1$. At the optimal bundle, $x_1 = 2x_2$. Using the budget constraint we thus have $x_1^* = 10$ and $x_2^* = 10$.

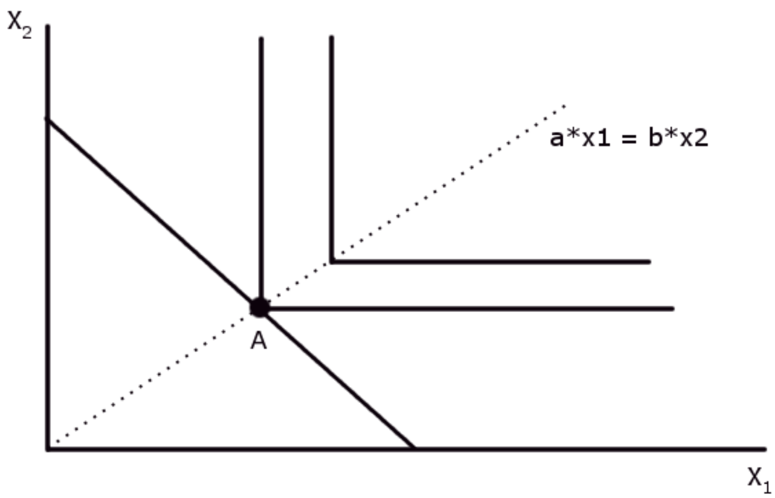

Figure 9: **Perfect Complements.** With perfect complements there is a kink in the indifference curve where $\alpha x_1 = \beta x_2$. The agent maximises her utility at the kink, at point A.

4.3 Monotonicity

When deriving the tangency conditions, we assumed that each good has a strictly positive marginal utility. This is an attractive property for two reasons.

First, if preferences are monotone then indifference curves are thin and downward sloping. From the perspective of the utility maximisation problem, monotonicity also ensures the agent spends her entire budget. In Section 4.3.1 we look at an example where monotonicity fails and the agent does not always wish to spend her budget.

Second, the fact that marginal utilities are strictly positive implies that the Kuhn–Tucker conditions pick out the optimal solution. In Section 4.3.2 we look at an example where there are two solutions to the Kuhn–Tucker conditions, only one of which is optimal.

4.3.1 Example: Bliss Point

Suppose $u(x_1, x_2) = -\frac{1}{2}(x_1 - 10)^2 - \frac{1}{2}(x_2 - 10)^2$. Figure 10 plots the corresponding indifference curves which are concentric circles around the bliss point of $(x_1, x_2) = (10, 10)$.

Suppose prices are $p_1 = 1$ and $p_2 = 1$ and the agent has income $m = 10$. At this point, the

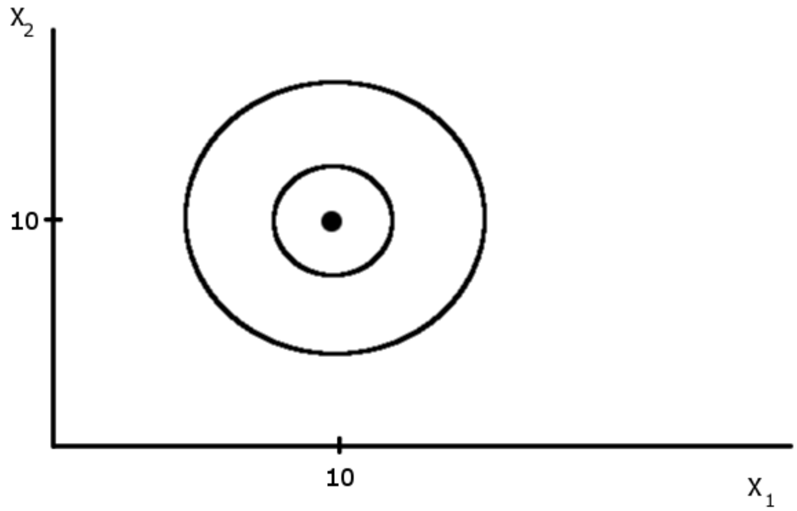

Figure 10: **Bliss Point.** Utility is maximised at $(10, 10)$. Indifference curves are circles around this bliss point.

agent cannot afford her bliss point (which costs \$20), so her budget constraint will bind. The price ratio is $p_1/p_2 = 1$, so the tangency condition implies

$$\frac{MU_1}{MU_2} = \frac{-(x_1 - 10)}{-(x_2 - 10)} = 1$$

Rearranging, we see that $x_1 = x_2$. The budget constraint states that $x_1 + x_2 = 10$. Hence we have $x_1^* = 5$ and $x_2^* = 5$.

Next, suppose prices are $p_1 = 1$ and $p_2 = 1$ and the agent has income $m = 30$. In this case the agent can afford her bliss point and will buy $x_1^* = 10$ and $x_2^* = 10$.

One can derive the same results from the Kuhn–Tucker conditions. First, suppose that the budget constraint binds. Ignoring boundary constraints (which are not an issue here), the FOCs of the Lagrangian are

$$-(x_1 - 10) - \lambda = 0 \tag{4.4}$$

$$-(x_2 - 10) - \lambda = 0 \tag{4.5}$$

As above, the FOCs imply $x_1 = x_2$. If $m = 10$, then the budget constraint implies that $(x_1^*, x_2^*) = (5, 5)$, and we are done. If $m = 30$, then the budget constraint implies that $(x_1^*, x_2^*) = (15, 15)$, which we know to be wrong. Substituting back into (4.4) or (4.5), we find $\lambda = -5$,

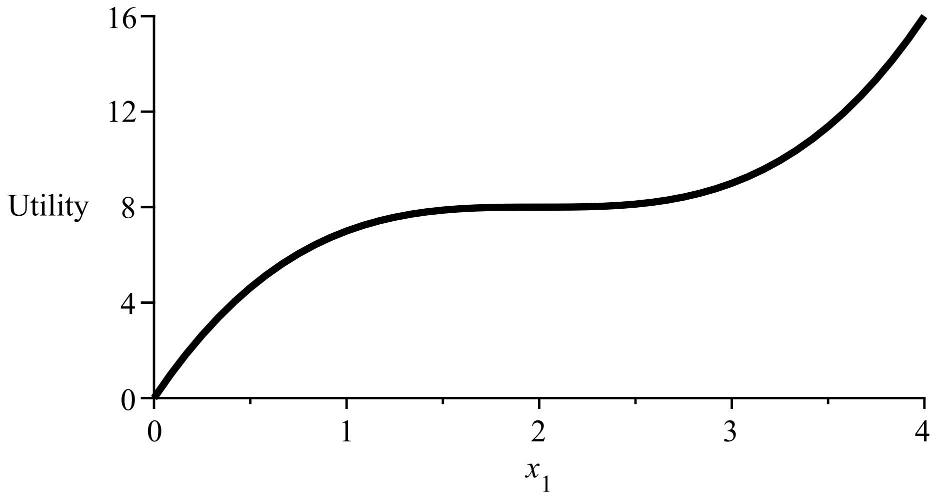

Figure 11: **Inflexion.** The utility function is strictly increasing but has an inflexion at $x_1 = 2$. As a result, there are two points that satisfy the Kuhn–Tucker conditions, only one of which is optimal.

which breaks part (c) of the Kuhn–Tucker conditions. Economically, this result means that the bang–per–buck is negative, so the solution is clearly not optimal. We thus know that the budget constraint does not bind, so complimentary slackness implies $\lambda = 0$. The agent thus maximises

$$\mathcal{L} = -(x_1 - 10)^2 - (x_2 - 10)^2$$

which yields $(x_1^*, x_2^*) = (10, 10)$.

4.3.2 Example: Inflexion

Suppose there is one good, x_1. The utility of the agent is given by $u(x_1) = (x_1 - 2)^3 + 8$. The agent has income $m = 4$ and faces prices $p_1 = 1$, so her budget states that $x_1 \leq 4$.

Figure 11 plots the utility function. Since this is increasing the optimal consumption is clearly $x_1^* = 4$.

Now consider the Lagrangian:

$$\mathcal{L} = (x_1 - 2)^3 + 8 + \lambda[m - p_1 x_1 - p_2 x_2]$$

The FOC is

$$3(x_1 - 2)^2 = \lambda \tag{4.6}$$

There are two solutions to the Kuhn–Tucker conditions. First, suppose $\lambda > 0$. By complimentary slackness, the agent spends her budget. Hence $x_1^* = 4$. Plugging back into (4.6) we see that

$$\lambda = 3(x_1 - 2)^2 = 12 > 0$$

as initially assumed. Hence this solution satisfies the Kuhn–Tucker conditions.

Second, suppose $\lambda = 0$. Equation (4.6) implies that $x_1^* = 2$, and the budget constraint is slack. Again, one can verify this "solution" satisfies the Kuhn–Tucker conditions, even though is is clearly wrong.

The problem is that the utility function has a inflexion at $x_1 = 2$. When the derivative fails to be strictly positive like this, there may be multiple solutions to the Kuhn–Tucker conditions. By part (a) of Theorem 1, one of these is the real solution, but one has to individually check which one.

4.4 Convexity

If preferences fail to be convex, then the solution to the tangency condition may characterise a local maximum or, even worse, a local minimum.

Figure 12 illustrates the problem. Points A, B and C all satisfy the tangency conditions. Point A is the global maximum; point B is a local maximum; point C is a local minimum.

4.4.1 Example: Addiction

Suppose an agent has utility $u(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$. Differentiating, $MU_i = x_i$, so the marginal utility of each good is increasing in the amount of the good consumed. For example, one could imagine the agent becomes addicted to either good.

As shown in Figure 13, these preferences are concave. One can see this formally by showing

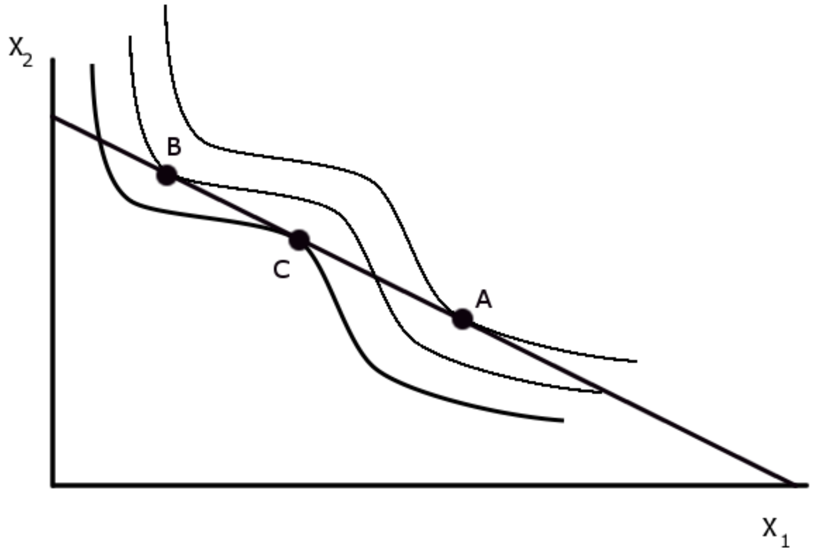

Figure 12: **Failure of Convexity.** Points A, B and C satisfy the tangency condition.

the MRS is increasing in x_1 along an indifference curve. Differentiating,

$$MRS = \frac{MU_1}{MU_2} = \frac{x_1}{x_2} \tag{4.7}$$

The equation of an indifference curve is $x_1^2+x_2^2 = k$. Rearranging, $x_2 = (k-x_1^2)^{1/2}$. Substituting into (4.7),

$$MRS = \frac{x_1}{(k - x_1^2)^{1/2}}$$

which is increasing in x_1.

With this addiction model, the tangency condition yields

$$\frac{x_1}{x_2} = \frac{p_1}{p_2}$$

If $m = 10$, $p_1 = 1$ and $p_2 = 1$, then one would wrongly conclude that $(x_1^*, x_2^*) = (5, 5)$. However, as we can see from figure 13, this is a local minimum.

Looking at figure 13, one can see that the agent's optimal bundle is on the boundary. Intuitively, if the agent becomes addicted, then she wishes to consume only one good. The left and right endpoints of the budget line are $(x_1, x_2) = (m/p_1, 0)$ and $(x_1, x_2) = (0, m/p_2)$ respectively. Comparing these two points, we see that $(x_1^*, x_2^*) = (m/p_1, 0)$ if $p_1 \leq p_2$ and $(x_1^*, x_2^*) = (0, m/p_2)$ if $p_1 \geq p_2$.

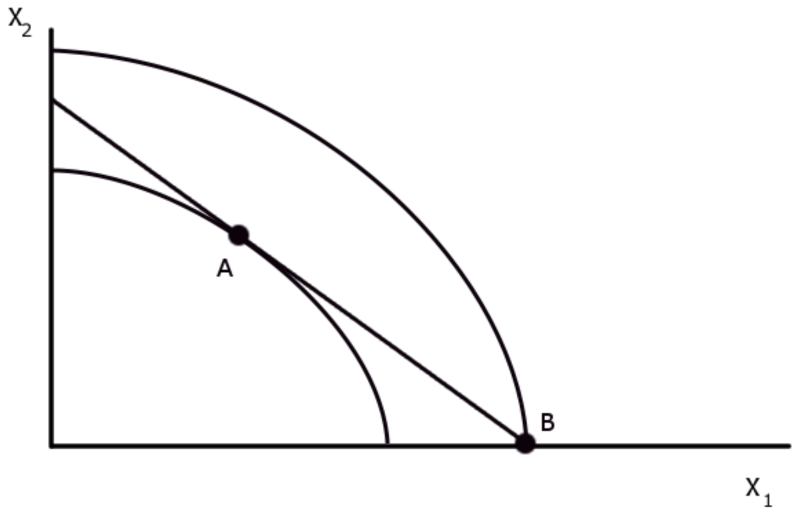

Figure 13: **Addictive preferences.** Point A satisfies the tangency condition but is not optimal. The agent's optimal bundle is point B.

4.5 Boundary Solutions

In Section 3 we assumed that the solution to the agent's UMP is internal. In practice, there are many goods that a typical person chooses not to buy.

Figure 14 provides an illustration. In this example,

$$MRS(x_1, x_2) > \frac{p_1}{p_2} \tag{4.8}$$

for all internal (x_1, x_2). Since the indifference curve is always steeper than the budget line, the agent's optimal bundle is $(x_1^*, x_2^*) = (m/p_1, 0)$. The intuition behind this is straightforward: rearranging, (4.8) we see that

$$\frac{MU_1}{p_1} > \frac{MU_2}{p_2}$$

which means that the bang–per–buck from x_1 is always bigger than that from x_2. As a result the agent consumes only x_1.

There are two ways to solve problems where boundary constraints may bind. First, one can insert Lagrange multipliers for the boundary constraints as in equation (4.2). One can then use the Kuhn–Tucker conditions to derive the agent's optimal bundle.

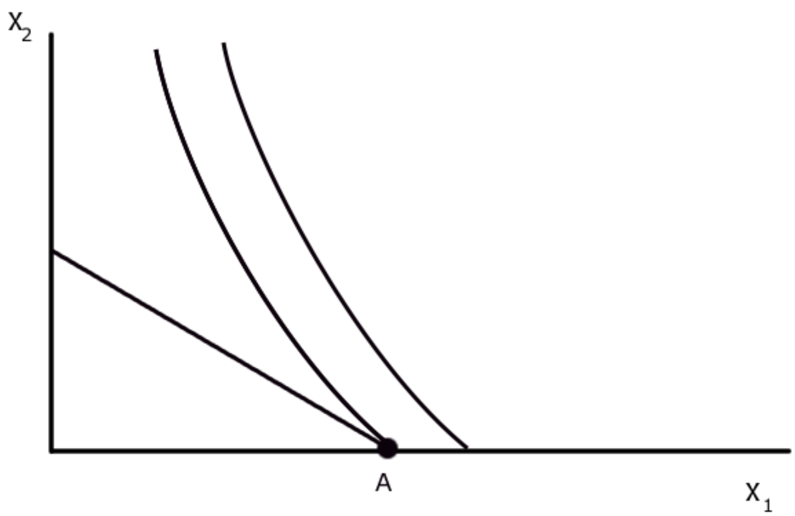

Figure 14: **Boundary Solutions.** In this figure the indifference curves are always steeper than the budget line. As a result, the solution is at point A, on the boundary.

Second, if preferences are convex, one can simply ignore the boundary constraints. If one finds that $x_i^* < 0$, then set $x_i^* = 0$ and resolve.

For a certain class of preferences, boundary problems will never be an issue. An indifference curve implicitly defines x_2 as a function of x_1. Let this function be denoted $x_2(x_1)$, so that $u(x_1, x_2(x_1)) = k$. The **Inada conditions** state that:

$$\lim_{x_1 \to 0} MRS(x_1, x_2(x_1)) = \infty \qquad \text{and} \qquad \lim_{x_1 \to \infty} MRS(x_1, x_2(x_1)) = 0 \tag{4.9}$$

Under these assumptions, the agent places a huge value on the first unit of both goods, and so will consume a positive amount of each. Notably, these Inada conditions are satisfied by Cobb Douglas preferences, $u(x_1, x_2) = x_1^\alpha x_2^\beta$, where

$$MRS = \frac{MU_1}{MU_2} = \frac{\alpha x_1^{\alpha-1} x_2^\beta}{\beta x_1^\alpha x_2^{\beta-1}} = \frac{\alpha x_2}{\beta x_1} \tag{4.10}$$

Along an indifference curve $x_1^\alpha x_2^\beta = k$, so $x_2 = k^{1/\beta} x_1^{-\alpha/\beta}$ and substituting into (4.10),

$$MRS = \frac{\alpha x_2}{\beta x_1} = \frac{\alpha}{\beta} k^{1/\beta} x_1^{-(\alpha+\beta)/\beta}$$

which satisfies the Inada conditions (4.9).

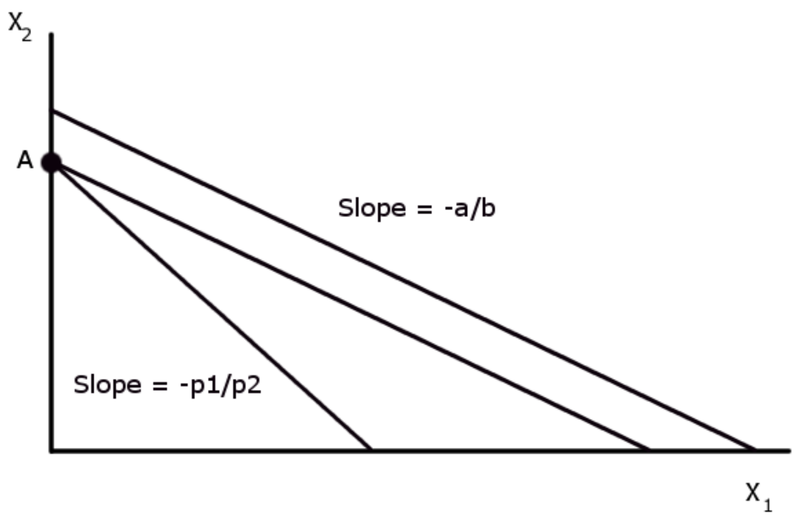

Figure 15: **Perfect Substitutes.** The slope of the agent's indifference curve is $-\alpha/\beta$, while the slope of the agent's budget line is $-p_1/p_2$. In this figure $\alpha/\beta < p_1/p_2$, so the bang–per–buck is higher from good 2 than good 1 and the agent chooses point A on the boundary.

4.5.1 Example: Perfect Substitutes

Suppose an agent has a utility function

$$U(x_1, x_2) = \alpha x_1 + \beta x_2$$

The indifference curves are straight lines with slope $-\alpha/\beta$. The budget line has slope $-p_1/p_2$. As shown in figure 15, the agent's optimal choice will occur at one of the endpoints. If

$$\frac{\alpha}{p_1} > \frac{\beta}{p_2}$$

then the bang–per–buck from good 1 exceeds that from good 2, so the agent's optimal bundle is $(x_1^*, x_2^*) = (m/p_1, 0)$. If

$$\frac{\alpha}{p_1} < \frac{\beta}{p_2}$$

then the bang–per–buck from good 2 exceeds that from good 1, so the agent's optimal bundle is $(x_1^*, x_2^*) = (0, m/p_2)$. If

$$\frac{\alpha}{p_1} = \frac{\beta}{p_2}$$

then the agent is indifferent between all points on the budget line.

5 General Properties of Solutions

It is useful to know some general properties of Marshallian demand (1.1) and indirect utility (1.2).

The Marshallian demand is homogenous of degree zero in prices and income. That is,

$$x_i^*(p_1, p_2, m) = x_i^*(\alpha p_1, \alpha p_2, \alpha m) \qquad \text{for } \alpha > 0$$

This means that a doubling of prices and income has no affect on the agent's demand. Intuitively, the agent buys the same goods whether the currency is denominated in Euros or Dollars.

The indirect utility function has three attractive properties.

1. The indirect utility function is homogenous of degree zero in prices and income. That is,

$$v(p_1, p_2, m) = v(\alpha p_1, \alpha p_2, \alpha m) \qquad \text{for } \alpha > 0$$

The agent's utility depends on what she buys, which is unaffected by the form of the currency (see above).

2. The indirect utility function is increasing in income and decreasing in prices. An increase in the agent's income expands the budget set and thereby increases the agent's utility from her most preferred choice. Conversely, an increase in a price contracts the agent's budget set and thereby decreases the agent's utility from her most preferred choice.

3. Roy's Identity:[4]

$$\frac{\partial v(p_1, p_2, m)}{\partial p_i} = -x_i^*(p_1, p_2, m)\frac{\partial v(p_1, p_2, m)}{\partial m} \tag{5.1}$$

Suppose p_1 increases by 1¢. Then there is a direct and indirect effect on the agent's utility. The direct effect is that, holding demand constant, the agent's effective income falls by $x_1^* \times 1¢$. The indirect effect is that as relative prices change, the agent rebalances her optimal choice. This indirect effect, however, is small since the agent's initial choice was optimal under the initial price, so is almost optimal under the new price. Putting this together, we see that the effect of a 1¢ price rise on utility equals x_1^* times the effect of a 1¢ drop in income. This is exactly what

[4] Advanced.

equation (5.1) says.

Roy's identity is useful since it enables us to calculate the agent's Marshallian demand from her indirect utility function. To illustrate this result, consider the symmetric Cobb–Douglas model. The agent has utility $u(x_1, x_2) = x_1 x_2$, income m and faces prices p_1 and p_2. Using the results in Section 3.2 we know that the agent has indirect utility

$$v(p_1, p_2, m) = \frac{m^2}{4p_1p_2}$$

Using Roy's identity (5.1), we therefore have,

$$\begin{aligned} x_1 * (p_1, p_2, m) &= -\left[\frac{\partial v(p_1, p_2, m)}{\partial p_i}\right]\left[\frac{\partial v(p_1, p_2, m)}{\partial m}\right]^{-1} \\ &= -\left[-\frac{m^2}{4p_1^2 p_2}\right]\left[\frac{2m}{4p_1p_2}\right]^{-1} \\ &= \frac{m}{2p_1} \end{aligned}$$

which is the optimal demand we found in Section 3.2.

The formal proof is in two steps. First, we show that λ measures the marginal utility of income, i.e. the bang–per buck. Observe that an agent's indirect utility is defined by

$$v(p_1, p_2, m) = u(x_1^*(p_1, p_2, m), x_2^*(p_1, p_2, m)) \tag{5.2}$$

Differentiating with respect to m,

$$\begin{aligned} \frac{\partial v(p_1, p_2, m)}{\partial m} &= \frac{\partial u(x_1^*, x_2^*)}{\partial x_1}\frac{\partial x_1^*(p_1, p_2, m)}{\partial m} + \frac{\partial u(x_1^*, x_2^*)}{\partial x_2}\frac{\partial x_2^*(p_1, p_2, m)}{\partial m} \\ &= \lambda p_1 \frac{\partial x_1^*(p_1, p_2, m)}{\partial m} + \lambda p_2 \frac{\partial x_2^*(p_1, p_2, m)}{\partial m} \end{aligned} \tag{5.3}$$

where the second line uses the FOCs (3.9) and (3.10). At the optimum the agent's budget holds:

$$p_1 x_1^*(p_1, p_2, m) + p_1 x_2^*(p_1, p_2, m) = m \tag{5.4}$$

Differentiating (5.4) with respect to m,

$$p_1 \frac{\partial x_1^*(p_1, p_2, m)}{\partial m} + p_2 \frac{\partial x_2^*(p_1, p_2, m)}{\partial m} = 1$$

Substituting into (5.3) we obtain,

$$\frac{\partial v(p_1, p_2, m)}{\partial m} = \lambda \tag{5.5}$$

as required.

We can now prove Roy's identity. Differentiating (5.2) with respect to p_1,

$$\begin{aligned} \frac{\partial v(p_1, p_2, m)}{\partial p_1} &= \frac{\partial u(x_1^*, x_2^*)}{\partial x_1} \frac{\partial x_1^*(p_1, p_2, m)}{\partial p_1} + \frac{\partial u(x_1^*, x_2^*)}{\partial x_2} \frac{\partial x_2^*(p_1, p_2, m)}{\partial p_1} \\ &= \lambda p_1 \frac{\partial x_1^*(p_1, p_2, m)}{\partial p_1} + \lambda p_2 \frac{\partial x_2^*(p_1, p_2, m)}{\partial p_1} \end{aligned} \tag{5.6}$$

where the second line uses the FOCs (3.9) and (3.10). Differentiating the budget constraint (5.4) with respect to p_1,

$$x_1^*(p_1, p_2, m) + p_1 \frac{\partial x_1^*(p_1, p_2, m)}{\partial p_1} + p_2 \frac{\partial x_2^*(p_1, p_2, m)}{\partial p_1} = 0$$

Substituting into (5.6),

$$\frac{\partial v(p_1, p_2, m)}{\partial p_1} = -\lambda x_1^*(p_1, p_2, m) \tag{5.7}$$

Using (5.7) we obtain,

$$\frac{\partial v(p_1, p_2, m)}{\partial p_1} = -x_1^*(p_1, p_2, m) \frac{\partial v(p_1, p_2, m)}{\partial m}$$

as required.

6 Comparative Statics

In this Section we consider how demand for good 1 is affected by the agent's income, the price of good 1 and the price of good 2. These effects are further analysed in the EMP notes. First, we introduce the notion of elasticities.

6.1 Elasticities

Suppose we are interested in how the price of a good, p, affects the demand for that good, $x(p)$. We can measure this change in absolute terms or percentage terms.

First, we might be interested in the impact of a 1¢ increase in the price. This is measured by the derivative $dx(p)/dp$. This is simple to calculate but has the disadvantage that the measure depends on the currency we are using. For example, suppose a consumer is buying a good in US dollars and has demand:

$$x(p) = 10 - 10p \tag{6.1}$$

Differentiating, $dx/dp = -10$. Suppose we now change the currency to British pounds, and suppose \$2 = £1. The demand function becomes:

$$x(p) = 10 - 20p \tag{6.2}$$

Differentiating, $dx/dp = -20$. We see that, while nothing fundamental has changed, the sensitivity of demand to price has doubled, simply because we have relabelled the currency.

In order to overcome this problem, we can measure the effect of a 1% increase in price. Define the **price elasticity of demand** to equal the percentage increase in demand caused by a 1% increase in the price. Since the percentage change in x equals the absolute change, Δx, divided by the level, the elasticity is given by

$$\epsilon_{x,p} = \frac{\Delta x/x}{\Delta p/p} = \frac{\Delta x}{\Delta p}\frac{p}{x}$$

In differential terms, we have

$$\epsilon_{x,p} = \frac{dx(p)}{dp}\frac{p}{x} \tag{6.3}$$

In example (6.1) we therefore have

$$\epsilon_{x,p} = (-10)\frac{p}{10-10p} = -\frac{p}{1-p}$$

Observe that we obtain exactly the same number if the price is denominated in pounds (6.2) since percentage changes are independent of the currency. We can also write the elasticity (6.3) as

$$\epsilon_{x,p} = \frac{d \ln x(p)}{d \ln p}$$

The proof is as follows:

$$\frac{d \ln x(p)}{d \ln p} = \frac{d}{d \ln p} \ln x(e^{\ln p}) = \frac{dx(e^{\ln p})}{dp} \frac{e^{\ln p}}{x(e^{\ln p})} = \frac{dx(p)}{dp} \frac{p}{x(p)}$$

There are different kinds of elasticities we can define. In the two good model, the agent's demand for good 1, $x_1^*(p_1, p_2, m)$, depends on her income and the price of both goods. We can correspondingly define the **income elasticity of demand**:

$$\epsilon_{x_1,m} = \frac{dx_1^*(p_1, p_2, m)}{dm} \frac{m}{x_1^*(p_1, p_2, m)}$$

The **own–price elasticity of demand**:

$$\epsilon_{x_1,p_1} = \frac{dx_1^*(p_1, p_2, m)}{dp_1} \frac{p_1}{x_1^*(p_1, p_2, m)}$$

And the **cross–price elasticity of demand**:

$$\epsilon_{x_1,p_2} = \frac{dx_1^*(p_1, p_2, m)}{dp_2} \frac{p_2}{x_1^*(p_1, p_2, m)}$$

The fact that demand is homogenous of degree zero (see Section 5) implies that a 1% increase in income and both prices does not affect the agent's demand. Hence

$$\epsilon_{x_1,m} + \epsilon_{x_1,p_1} + \epsilon_{x_1,p_2} = 0.$$

6.2 Income Effects

Suppose an agent's income increases. Figure 16 shows that her budget constraint shifts outwards. The line linking her optimal bundles for different levels of income is called the **income offer curve** or the income expansion path.[5]

Figure 17 show how the consumption of one particular good varies with the agent's income. This is called the **Engel curve**.

[5] If the income offer curve is linear, preferences are said to be **homothetic**. Perfect substitutes, perfect complements and Cobb Douglas preferences are all examples of homothetic preferences.

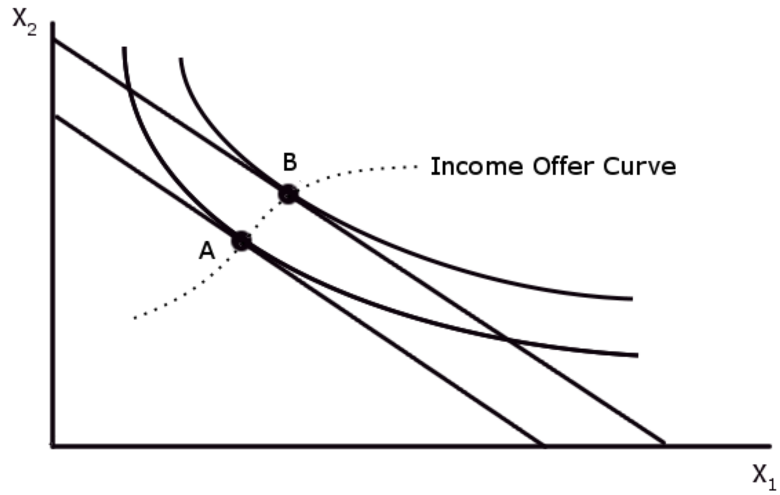

Figure 16: **Income Effects: A Normal Good.** This figure shows the effect of an increase in the agent's income on her demand for both goods. Her choice moves from point A to point B, increasing her consumption of both goods.

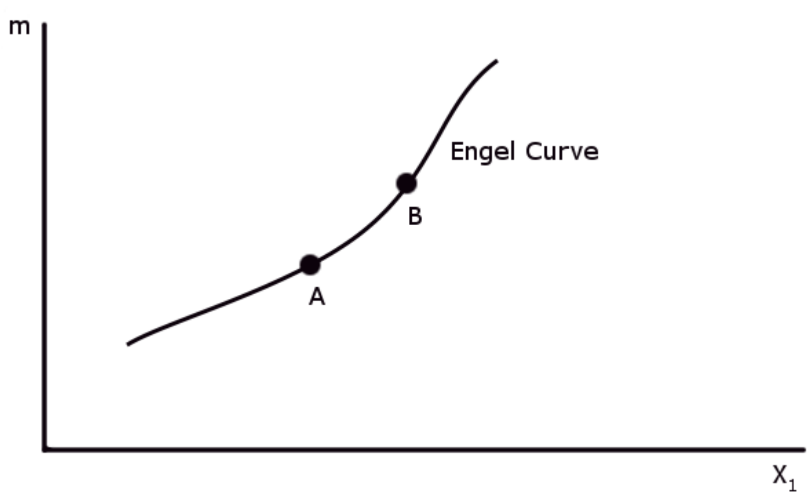

Figure 17: **Engel Curve.** This figure shows the effect of an increase in the agent's income on her demand for good 1. Points A and B correspond to those in figure 16.

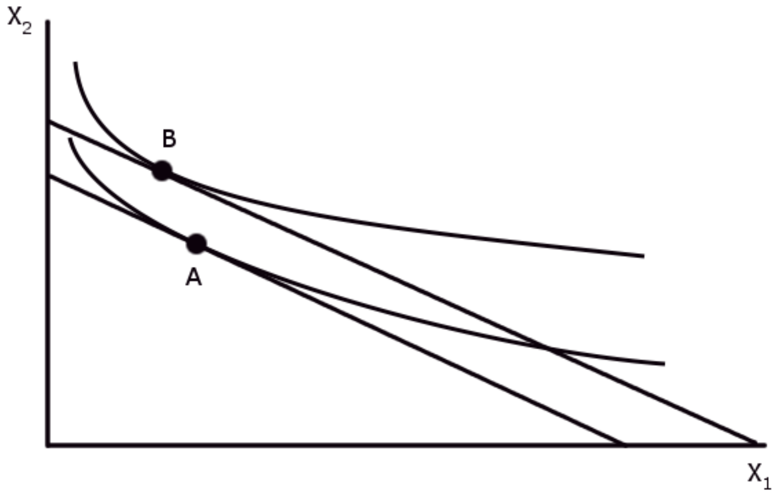

Figure 18: **Income Effects: An Inferior Good.** This figure shows the effect of an increase in the agent's income on her demand for both goods. Her choice moves from point A to point B, increasing her consumption of x_2 but reducing her consumption of x_1. Good 1 is therefore inferior.

When an agent's income rises her demand may rise or fall. If her demand rises with income, the good is **normal** (see figure 16). If her demand falls with income, the good is **inferior** (see figure 18). Many goods are normal for some ranges of income and inferior for others. For example, for very poor people in China, rice consumption increases in income; for well off people in China, rice consumption falls in income as people substitute towards meat.

The normal/inferior distinction concerns the effect of income on the absolute consumption of the good. One may also wonder about the effect of income on the budget share of a good. For example, housing often accounts for around 30% of people's spending, independent of their income. A good is a **luxury** if a 1% increase in income leads to a more than 1% increase in consumption. That is, $\epsilon_{x_1,m} > 1$. A good is a **necessity** if a 1% increase in income leads to a less than 1% increase in consumption. That is, $\epsilon_{x_1,m} < 1$.

6.3 Own Price Effects

Suppose the price of good 1 increases. Figure 19 shows that the budget line pivots inwards and, in this case, leads the agent to consume less x_1 and more x_2. The line linking her optimal bundles for different levels of p_1 is called the **price offer curve**.

Figure 19: **Price Effects.** This figure shows the effect of an increase in the price of good 1 on the agent's demand for both goods. As p_1 rises, the agent's choice moves from A to B. As a result, the consumption of x_1 falls while the consumption of x_2 rises.

Figure 20 shows how the consumption of good 1 varies with p_1. This is, surprisingly enough, called the **demand curve**. Good 1 is called **ordinary** if an the demand curve is downward sloping, so an increase in p_1 causes a reduction in x_1^*. Good 1 is a **Giffen good** if the demand is locally upward sloping, so an increase in p_1 causes an increase in x_1^*.

To understand how price affects demand, note that we can decompose the impact of an increase in p_1 into two effects. First, holding the agent's purchasing power constant, there is a change in change in relative prices causing good 1 to become more expensive relative to good 2. This is called the **substitution effect** and causes the demand for good 1 to fall. Second, holding relative prices fixed, the increase in p_1 reduces the agent's purchasing power. This is called the **income effect** and causes the demand for good 1 to fall if it is normal, and rise if it is inferior. With a Giffen good, the increase in p_1 causes demand for good 1 to fall a little via the substitution effect and causes demand for good 1 to rise a lot via the income effect. For example, when the price of rice rises in China some poor people experience a cut in their purchasing power, can no longer afford meat and consequently consume more rice.[6] For more on income and substitution effects see the EMP notes.

We end with some more jargon. If a 1% increase in price changes demand by less than 1%, demand is called **inelastic**. That is, $-1 < \epsilon_{x_1,p_1} < 1$. If a 1% increase in price changes demand

[6]See Jensen and Miller (2009), "Giffen Behaviour and Subsistence Consumption", *American Economic Review.*

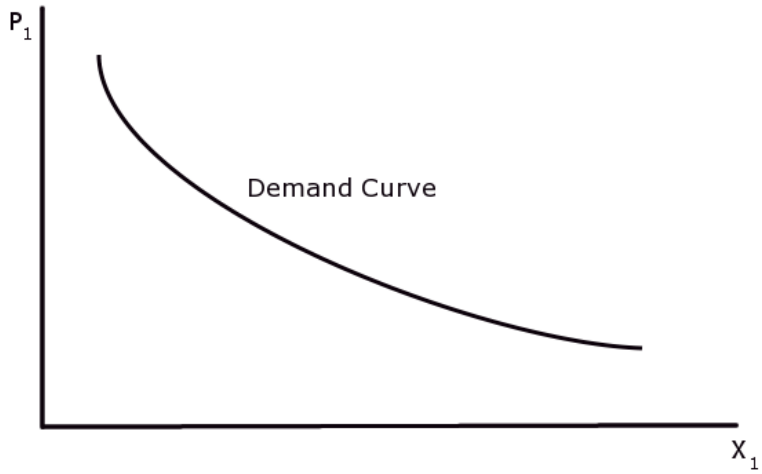

Figure 20: **Demand Curve for Good 1.** As the price of good 1 falls, the demand for good 1 increases. Hence the good is ordinary.

by more than 1%, demand is called **elastic**. That is, $\epsilon_{x_1,p_1} < -1$ or $\epsilon_{x_1,p_1} > 1$.

Exercise: The spending on good 1 is given by $p_1 x_1^*(p_1, p_2, m)$. Show that an increase in p_1 reduces the agent's spending on the good if and only if the good is ordinary and demand is elastic.

6.4 Cross Price Effects

Suppose the price of good 1 increases. Such an increase may cause the demand for good 2 to rise (as in figure 19) or fall (as in figure 21).

Goods 1 and 2 are **gross substitutes** if an increase in the price of one increases the demand for the other. Mathematically,

$$\frac{\partial x_1^*}{\partial p_2} > 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial p_1} > 0$$

For example, when the price of pizza rises, the demand for hamburgers goes up (and vice versa). Hence pizza and hamburgers are gross substitutes.

Goods 1 and 2 are **gross complements** if an increase in the price of one decreases the demand

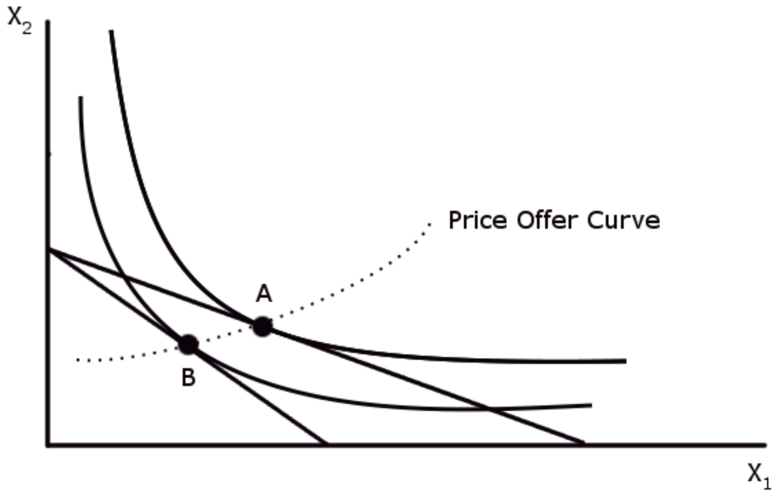

Figure 21: **Price Effects.** This figure shows the effect of an increase in the price of good 1 on the agent's demand for both goods. As p_1 rises, the agent's choice moves from A to B. As a result, the consumption of both goods rises.

for the other. Mathematically,

$$\frac{\partial x_1^*}{\partial p_2} < 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial p_1} < 0$$

For example, when the price of buns rises, the demand for hamburgers goes down (and vice versa). Hence buns and hamburgers are gross complements.

There is a problem with the idea of gross substitutes/complements: the cross derivatives may have different signs. For example, suppose x_1 are domestic flights and x_2 are international flights. We may have the following scenario. An increase in p_1 causes the agent to become poorer, since she often flies home to see her parents, leading her to cut back on international holidays and reducing x_2^*. An increase in p_2 causes the agent to take fewer international holidays and more domestic holidays, leading to an increase in x_1^*. Hence we have

$$\frac{\partial x_1^*}{\partial p_2} > 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial p_1} < 0$$

It is unclear whether these goods are complements or substitutes. In the EMP notes we introduce the idea of net substitutes, where we can never have problems like this.

Expenditure Minimisation Problem

Simon Board

This Version: September 20, 2009
First Version: October, 2008.

The expenditure minimisation problem (EMP) looks at the reverse side of the utility maximisation problem (UMP). The UMP considers an agent who wishes to attain the maximum utility from a limited income. The EMP considers an agent who wishes to find the cheapest way to attain a target utility. This approach complements the UMP and has several rewards:

- It enables us to analyse the effect of a price change, holding the utility of the agent constant.
- It enables us to decompose the effect of a price change on an agent's Marshallian demand into a substitution effect and an income effect. This decomposition is called the Slutsky equation.
- It enables us to calculate how much we need to compensate a consumer in response to a price change if we wish to keep her utility constant.

1 Model

We make several assumptions:

1. There are N goods. For much of the analysis we assume $N = 2$ but nothing depends on this.
2. The agent takes prices as exogenous. We normally assume prices are linear and denote them by $\{p_1, \ldots, p_N\}$.

3. Preferences satisfy completeness, transitivity and continuity. As a result, a utility function exists. We normally assume preferences also satisfy monotonicity (so indifference curves are well behaved) and convexity (so the optima can be characterised by tangency conditions).

The expenditure minimisation problem is

$$\min_{x_1,\ldots,x_N} \sum_{i=1}^{N} p_i x_i \qquad \text{subject to} \qquad u(x_1,\ldots,x_N) \geq \overline{u} \tag{1.1}$$
$$x_i \geq 0 \quad \text{for all } i$$

The idea is that the agent is trying to find the cheapest way to attain her target utility, $\overline{u}$. The solution to this problem is called the **Hicksian demand** or compensated demand. It is denoted by

$$h_i(p_1,\ldots,p_N,\overline{u})$$

The money the agent must spend in order to attain her target utility is called her expenditure. The **expenditure function** is therefore given by

$$e(p_1,\ldots,p_N,\overline{u}) = \min_{x_1,\ldots,x_N} \sum_{i=1}^{N} p_i x_i \qquad \text{subject to} \qquad u(x_1,\ldots,x_N) \geq \overline{u}$$
$$x_i \geq 0 \quad \text{for all } i$$

Equivalently, the expenditure function equals the amount the agent spends on her optimal bundle,

$$e(p_1,\ldots,p_N,\overline{u}) = \sum_{i=1}^{N} p_i h_i(p_1,\ldots,p_N,\overline{u})$$

1.1 Example

Suppose there are two goods, x_1 and x_2. Table 1 shows how the agent's utility (the numbers in the boxes) varies with the number of x_1 and x_2 consumed.

To keep things simple, suppose the agent faces prices $p_1 = 1$ and $p_2 = 1$ and wishes to attain utility $\overline{u} = 12$. The agent can attain this utility by consuming $(x_1, x_2) = (6, 2)$, $(x_1, x_2) = (4, 3)$, $(x_1, x_2) = (3, 4)$ or $(x_1, x_2) = (2, 6)$. Of these, the cheapest is either $(x_1, x_2) = (4, 3)$ or $(x_1, x_2) = (3, 4)$. In either case, her expenditure is $4 + 3 = 7$.

$x_1 \backslash x_2$	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

Table 1: **Utilities from different bundles.**

Now suppose the agent faces prices $p_1 = 1$ and $p_2 = 3$ and still wishes to attain utility $\overline{u} = 12$. The combinations of (x_1, x_2) that attain this utility remain unchanged, however the price of these bundles is different. Now the cheapest is $(x_1, x_2) = (6, 2)$, and the agent's expenditure is $6 + 2 \times 3 = 12$.

While this "table approach" can be used to illustrate the basic idea, one can see that it quickly becomes hard to solve even simple problems. Fortunately, calculus comes to our rescue.

2 Solving the Expenditure Minimisation Problem

2.1 Graphical Solution

We can solve the problem graphically, as with the UMP. The components are also similar to that problem.

First, we need to understand the constraint set. The agent can choose any bundle where (a) the agent attains her target utility, $u(x_1, x_2) \geq \overline{u}$; and (b) the quantities are positive, $x_1 \geq 0$ and $x_2 \geq 0$. If preferences are monotone, then the bundles that meet these conditions are exactly the ones that lie above the indifference curve with utility $\overline{u}$. See figure 1.

Second, we need to understand the objective. The agent wishes to pick the bundle in the constraint set that minimises her expenditure. Just like with the UMP, we can draw the level curves of this objective function. Define an iso–expenditure curve by the bundles of x_1 and x_2 that deliver constant expenditure:

$$\{(x_1, x_2) : p_1 x_1 + p_2 x_2 = const\}$$

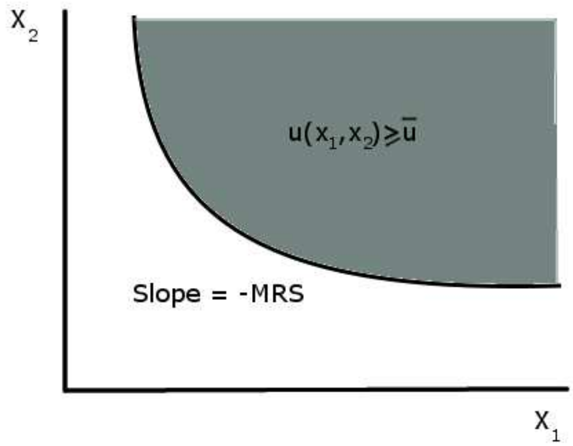

Figure 1: **Constraint Set.** The shaded area shows the bundles that yield utility $\overline{u}$ or more.

These iso–expenditure curves are just like budget curves and so have slope $-p_1/p_2$. See figure 2

The aim of the agent is to choose the bundle (x_1, x_2) in the constraint set that is on the lowest iso–expenditure curve and hence minimises her expenditure. Ignoring boundary problems and kinks, the solution has the feature that the iso–expenditure curve is tangent to the target indifference curve. As a result, their slopes are identical. The *tangency condition* can thus be written as

$$MRS = \frac{p_1}{p_1} \tag{2.1}$$

This is illustrated in figure 3.

The intuition behind (2.1) is as follows. Using the fact that $MRS = MU_1/MU_2$,[1] equation (2.1) implies that

$$\frac{MU_1}{MU_2} = \frac{p_1}{p_1} \tag{2.2}$$

[1]Recall: $MU_i = \partial U/\partial x_i$ is the marginal utility from good i.

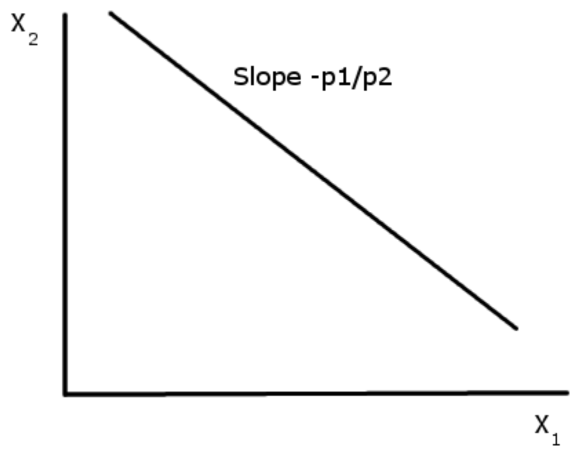

Figure 2: **Iso-Expenditure Curve.** This figure shows the bundles that induce constant expenditure.

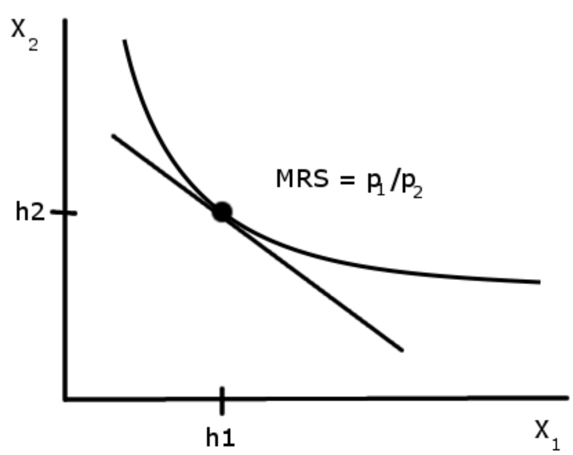

Figure 3: **Optimal Bundle.** This figure shows how the cheapest bundle that attains the target utility satisfies the tangency condition.

Rewriting (2.2) we find

$$\frac{p_1}{MU_1} = \frac{p_2}{MU_1}$$

The ratio p_i/MU_i measures the cost of increasing utility by one util, or the "cost–per–bang". At the optimum the agent equates the cost–per–bang of the two goods. Intuitively, if good 1 has a higher cost–per–bang 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 utility at a lower cost.

If preferences are monotone, then the constraint will bind,

$$u(x_1, x_2) = \overline{u}, \tag{2.3}$$

The tangency equation (2.2) and constraint equation (2.3) can then be used to solve for the two Hicksian demands.

If there are N goods, the agent will equalise the cost–per–bang from each good, giving us $N-1$ equations. Using the constraint equation (2.3), we can solve for the agent's Hicksian demands.

The tangency condition (2.2) is the same as that under the UMP. This is no coincidence. We discuss the formal equivalence in Section 4.2.

2.2 Example: Symmetric Cobb Douglas

Suppose $u(x_1, x_2) = x_1 x_2$. The tangency condition yields:

$$\frac{x_2}{x_1} = \frac{p_1}{p_2} \tag{2.4}$$

Rearranging, $p_1 x_1 = p_2 x_2$.

The constraint states that $\overline{u} = x_1 x_2$. Substituting (2.4) into this yields,

$$\overline{u} = \frac{p_1}{p_2} x_1^2$$

Solving for x_1, the Hicksian demand is given by

$$h_1(p_1, p_2, \overline{u}) = \left(\frac{p_2}{p_1}\overline{u}\right)^{1/2} \tag{2.5}$$

Similarly, we can solve for the Hicksian demand for good 2,

$$h_2(p_1, p_2, \overline{u}) = \left(\frac{p_1}{p_2}\overline{u}\right)^{1/2}$$

We can now calculate the agent's expenditure

$$\begin{aligned} e(p_1, p_2, \overline{u}) &= p_1 h_1(p_1, p_2, \overline{u}) + p_2 h_2(p_1, p_2, \overline{u}) \\ &= 2(\overline{u} p_1 p_2)^{1/2} \end{aligned} \tag{2.6}$$

2.3 Lagrangian Solution

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

$$\mathcal{L} = p_1 x_1 + p_2 x_2 + \lambda[\overline{u} - u(x_1, x_2)]$$

As with the UMP, the term in brackets can be thought as the penalty for violating the constraint. That is, the agent is punished for falling short of the target utility.

The FOCs with respect to x_1 and x_2 are

$$\frac{\partial L}{\partial x_1} = p_1 - \lambda \frac{\partial u}{\partial x_1} = 0 \tag{2.7}$$

$$\frac{\partial L}{\partial x_2} = p_2 - \lambda \frac{\partial u}{\partial x_2} = 0 \tag{2.8}$$

If preferences are monotone then the constraint will bind,

$$u(x_1, x_2) = \overline{u} \tag{2.9}$$

These three equations can then be used to solve for the three unknowns: x_1, x_2 and λ.

Several remarks are in order. First, this approach is identical to the graphical approach. Dividing (2.7) by (2.8) yields

$$\frac{\partial u/\partial x_1}{\partial u/\partial x_2} = \frac{p_1}{p_2}$$

which is the same as (2.2). Moreover, the Lagrange multiplier is

$$\lambda = \frac{p_1}{MU_1} = \frac{p_2}{MU_2}$$

is exactly the cost–per–bang.

Second, if preferences are not monotone, the constraint (2.9) 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 goods. In this case, one obtains N first order conditions and the constraint equation (2.9).

3 General Results

3.1 Properties of Expenditure Function

The expenditure function exhibits four important properties.

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

$$e(p_1, p_2, \overline{u}) = e(\alpha p_1, \alpha p_2, \overline{u})$$

for $\alpha > 0$. Intuitively, if the prices of x_1 and x_2 double, then the cheapest way to attain the target utility does not change. However, the cost of attaining this utility doubles.

2. The expenditure function is increasing in $(p_1, p_2, \overline{u})$. If we increase the target utility $\overline{u}$, then the constraint becomes harder to satisfy and the cost of attaining the target increases. If we increase p_1 then it costs more to buy any bundle of goods and it costs more to attain the target utility.

3. The expenditure function is concave in prices (p_1, p_2). Fix the target utility $\overline{u}$ and prices $(p_1, p_2) = (p'_1, p'_2)$. Solving the EMP we obtain Hicksian demands $h'_1 = h_1(p'_1, p'_2, \overline{u})$ and $h'_2 = h_2(p'_1, p'_2, \overline{u})$. Now suppose we fix demands and change p_1, the price of good 1. This gives us a pseudo–expenditure function

$$\eta_{h'_1,h'_2}(p_1) = p_1 h'_1 + p'_2 h'_2$$

This pseudo–expenditure function is linear in p_1 which means that, if we keep demands constant, then expenditure rises linearly with p_1. Of course, as p_1 rises the agent can reduce her expenditure by rebalancing her demand towards the good that is cheaper. This means that

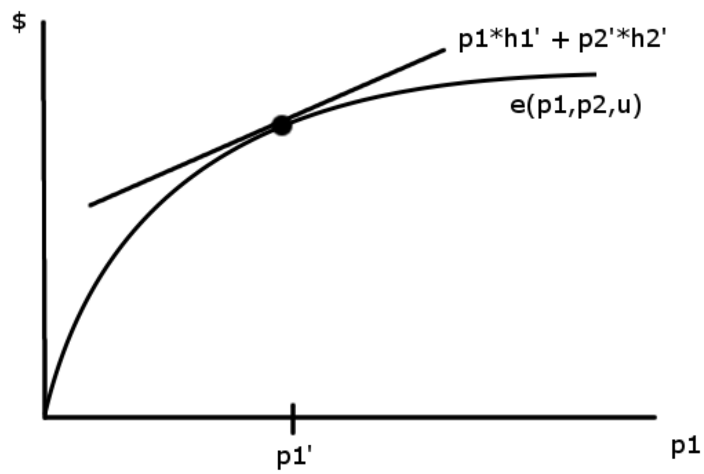

Figure 4: **Expenditure Function.** This figure shows how the expenditure function lies under the pseudo–expenditure function.

real expenditure function lies below the pseudo–expenditure function and is therefore concave. See figure 4.

More formally, the expenditure function is given by the lower envelope of the pseudo-expenditure functions. That is, for any bundle (x_1, x_2), the cost of this bundle at prices (p_1, p_2) is given by

$$\eta_{x_1,x_2}(p_1, p_2) = p_1 x_1 + p_2 x_2$$

The expenditure function is then the minimum of these pseudo–expenditure functions given the bundle (x_1, x_2) attains the target utility. Mathematically,

$$e(p_1, p_2, \overline{u}) = \min\{p_1 x_1 + p_2 x_2 : u(x_1, x_2) = \overline{u}\} \tag{3.1}$$

Thus the expenditure function is the lower minimum of a collection of linear functions, and is therefore concave.[2] See figure 5.

4. Sheppard's Lemma: The derivative of the expenditure function equals the Hicksian demand. That is,

$$\frac{\partial}{\partial p_1} e(p_1, p_2, \overline{u}) = h_1(p_1, p_2, \overline{u}) \tag{3.2}$$

[2]Exercise: Show that the minimum of two concave functions is concave.

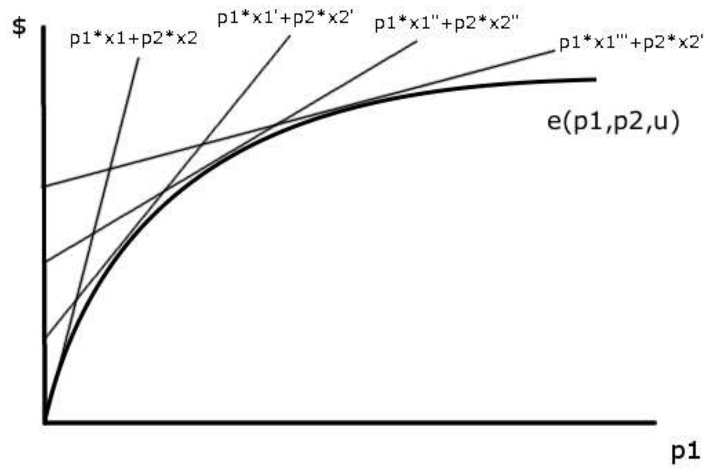

Figure 5: **Envelope Property of Expenditure Function.** This figure shows the expenditure function equals the lower envelope of the pseudo expenditure functions.

The idea behind this result can be seen from figure 4. At $p_1 = p_1'$ the expenditure function is tangential to the pseudo–expenditure function. The pseudo–expenditure is linear in p_1 with slope $h_1(p_1', p_2', \overline{u})$. Hence the expenditure function also has slope $h_1(p_1', p_2', \overline{u})$.

The intuition behind Sheppard's Lemma is as follows. Suppose an agent wishes to attain target utility $\overline{u} = 25$ and faces prices $p_1 = \$1$ and $p_2 = \$1$. Furthermore, suppose that the cheapest way to attain the target utility is by consuming $h_1 = 5$ and $h_2 = 5$. Next, consider an increase in p_1 of 1¢. This change has a direct and indirect effect. The direct effect is that, holding demand constant, the agent's spending rises by $h_1 \times 1¢ = 5¢$; the indirect effect is that the agent will change her demands. However, the tangency condition illustrated in figure 3 shows that the agent is close to indifferent between choosing the optimal quantity and nearby quantities, so the rebalancing demand will will have a very small impact on her expenditure. We thus conclude that $\Delta e = h_1 \Delta p_1$, Rewriting,

$$\frac{\Delta e}{\Delta p_1} = h_1$$

This is the discrete version of equation (3.2).

Here is a formal proof of Sheppard's Lemma. By definition of the expenditure function,

$$e(p_1, p_2, \overline{u}) = p_1 h_1(p_1, p_2, \overline{u}) + p_2 h_2(p_1, p_2, \overline{u})$$

Differentiating with respect to p_1 yields

$$\frac{\partial}{\partial p_1}e(p_1,p_2,\overline{u}) = h_1(p_1,p_2,\overline{u}) + p_1\frac{\partial h_1(p_1,p_2,\overline{u})}{\partial p_1} + p_2\frac{\partial h_2(p_1,p_2,\overline{u})}{\partial p_1} \tag{3.3}$$

As discussed above, we have decomposed the effect of the price change into a direct effect (the first term) and an indirect effect (the second and third terms). We now wish to show the indirect effect is zero. From the agent's minimisation problem in Section 2.3, the FOCs are

$$p_i = \lambda\frac{\partial u(h_1,h_2)}{\partial x_i}$$

We also know that the agent's constraint binds:

$$u(h_1(p_1,p_2,\overline{u}),h_2(p_1,p_2,\overline{u})) = \overline{u} \tag{3.4}$$

Substituting the FOCs into (3.3)

$$\frac{\partial}{\partial p_1}e(p_1,p_2,\overline{u}) = h_1(p_1,p_2,\overline{u}) + \lambda\left[\frac{\partial u(h_1,h_2)}{\partial x_1}\frac{\partial h_1(p_1,p_2,\overline{u})}{\partial p_1} + \frac{\partial u(h_1,h_2)}{\partial x_2}\frac{\partial h_2(p_1,p_2,\overline{u})}{\partial p_1}\right] \tag{3.5}$$

Differentiating (3.4) with respect to p_1 yields

$$\frac{\partial u(h_1,h_2)}{\partial x_1}\frac{\partial h_1(p_1,p_2,\overline{u})}{\partial p_1} + \frac{\partial u(h_1,h_2)}{\partial x_2}\frac{\partial h_2(p_1,p_2,\overline{u})}{\partial p_1} = 0 \tag{3.6}$$

Substituting (3.6) into (3.5) yields Sheppard's Lemma.

3.2 Properties of Hicksian Demand

Hicksian demand has three important properties. These follow from the properties of the expenditure function derived above.

1. Hicksian demand is homogenous of degree zero in prices. That is,

$$h_1(p_1,p_2,\overline{u}) = h_1(\alpha p_1,\alpha p_2,\overline{u})$$

for $\alpha > 0$. Intuitively, doubling both prices does not alter the cheapest way to obtain the target utility $\overline{u}$.

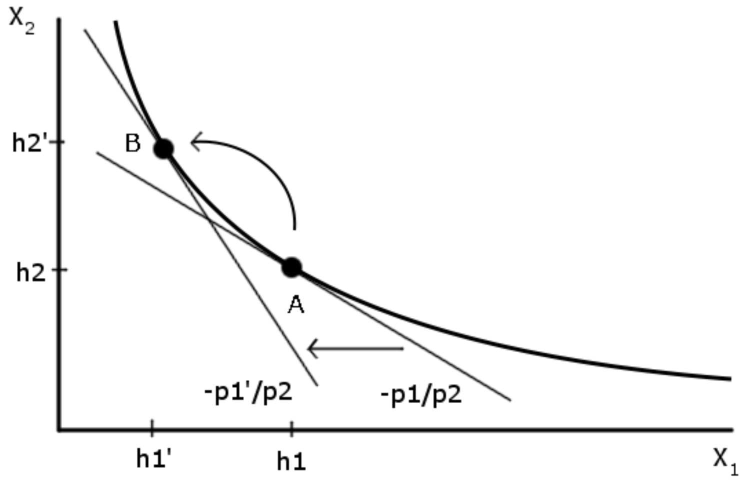

Figure 6: **Hicksian Demand and Own Price Effects.** This figure shows the effect of an increase in p_1, from p_1 to p_1'. The optimal bundle moves from A to B.

2. The Law of Hicksian Demand: The Hicksian demand for good i is decreasing in p_i. That is,

$$\frac{\partial}{\partial p_i} h_i(p_1, p_2, \overline{u}) \leq 0$$

Intuitively, when p_1 rises the relative prices become tilted in favour of good 2. The cheapest way to attain the target utility then consists of less of good 1 and more of good 2. Graphically this can be seen from figure 6. As p_1 rises to p_1', the iso–expenditure function becomes steeper and the optimal bundle involves less of good 1 and more of good 2.[3]

A formal proof of this result uses the properties of the expenditure function:

$$\frac{\partial}{\partial p_1} h_1(p_1, p_2, \overline{u}) = \frac{\partial^2}{\partial p_1^2} e(p_1, p_2, \overline{u}) \leq 0$$

where the equality comes from Sheppard's Lemma and the inequality follows from the concavity of the expenditure function.

This result highlights a big difference between Hicksian demand and Marshallian demand. An increase in p_1 always reduces the Hicksian demand for good 1 but may, in the case of a Giffen good, increase the Marshallian demand. This is because the effect of a price change on Marshallian demand has two effects: a substitution effect (a change in relative prices) and an

[3]The fact that the demand for good 2 always rises is an artifact of there only being 2 goods.

income effect (a change in the consumer's purchasing power). In comparison, the change in Hicksian demand isolates the substitution effect.

3. Hicksian demand has symmetric cross derivatives. That is,

$$\frac{\partial}{\partial p_2} h_1(p_1, p_2, \overline{u}) = \frac{\partial}{\partial p_1} h_2(p_1, p_2, \overline{u})$$

The proof of this result also uses the properties of the expenditure function.

$$\frac{\partial}{\partial p_2} h_1(p_1, p_2, \overline{u}) = \frac{\partial}{\partial p_2}\left[\frac{\partial}{\partial p_1} e(p_1, p_2, \overline{u})\right] = \frac{\partial}{\partial p_1}\left[\frac{\partial}{\partial p_2} e(p_1, p_2, \overline{u})\right] = \frac{\partial}{\partial p_1} h_2(p_1, p_2, \overline{u})$$

The first and third equalities come from Sheppard's Lemma and the second from Young's theorem.

We say goods x_1 and x_2 are **net substitutes** if

$$\frac{\partial}{\partial p_2} h_1(p_1, p_2, \overline{u}) > 0 \qquad \text{and} \qquad \frac{\partial}{\partial p_1} h_2(p_1, p_2, \overline{u}) > 0$$

We say goods x_1 and x_2 are **net complements** if

$$\frac{\partial}{\partial p_2} h_1(p_1, p_2, \overline{u}) < 0 \qquad \text{and} \qquad \frac{\partial}{\partial p_1} h_2(p_1, p_2, \overline{u}) < 0$$

The symmetry of the cross derivatives means that we cannot have one cross–derivative positive negative and the opposite cross–derivative negative, as with gross substitutes and complements.[4]

4 Income and Substitution Effects

We are often interested in how price changes affect Marshallian demand. This matters to firms when choosing prices, to government when choosing tax rates and to economists when making forecasts. For example: how much will demand for ethanol increase if we lower the price by $10?

We saw with the UMP that an increase in p_1 may lead to a large decrease in demand (if demand is elastic), may lead to a small decrease in demand (if demand is inelastic) or may lead to an increase in demand (in the case of a Giffen good). One major issue is that an increase in the

[4]Exercise: Suppose there are two goods. Show they must be net substitutes.

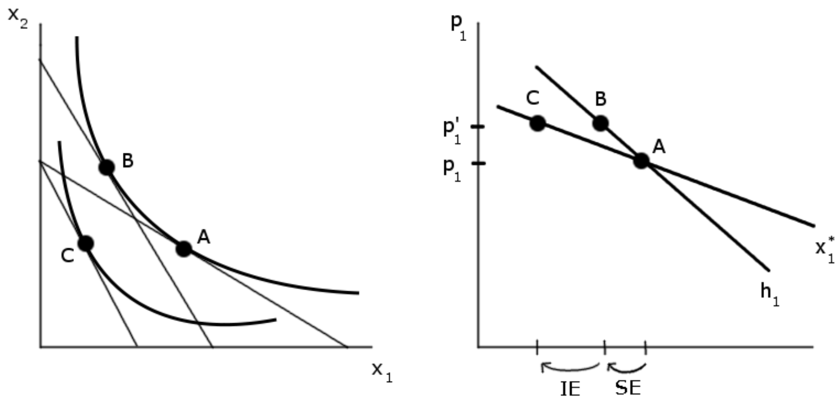

Figure 7: **Substitution and Income Effects with Normal Good.** With a normal good, both substitution effect (SE) and income effect (IE) are negative.

price of good 1 has two effects: it both makes good 1 relatively more expensive (the substitution effect) and reduces the agent's purchasing power (the income effect). This section will separate these effects. In Section 4.1 we do this graphically. In Section 4.3 we do this mathematically.

4.1 Pictures

Suppose we start at point A in figures 7 and 8. When p_1 increases, the budget line pivots around it's left end and demand falls from A to C. We can decompose this change into two effects.

1. A change in relative prices, keeping utility constant. This is the shift from A to B, and is called the **substitution effect**. This equals the change in Hicksian demand and, appealing to the Law of Hicksian Demand, is negative.
2. A change in income, keeping relative prices constant. This is the shift from B to C, and is called the **income effect**. This effect is positive if the good is normal, and negative if the good is inferior.

Exercise: draw the equivalent picture for a Giffen good.

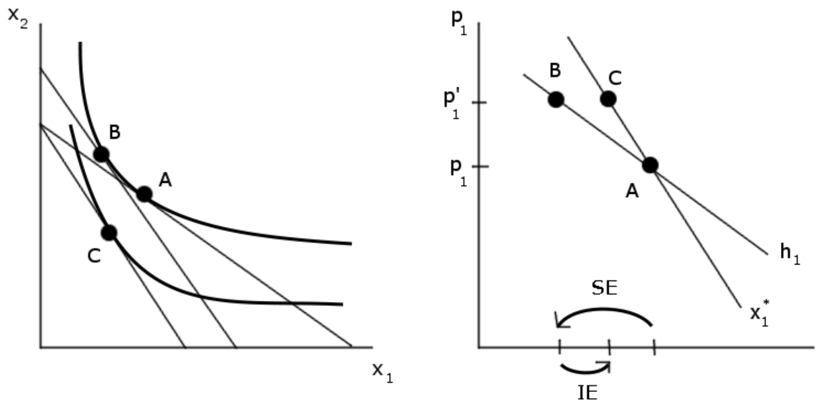

Figure 8: **Substitution and Income Effects with Normal Good.** With an inferior good, substitution effect (SE) is negative while the income effect (IE) is positive.

4.2 Relation between the UMP and EMP

The EMP and UMP are closely related. To illustrate, suppose the agent has \$10 to spend on two goods. Suppose her utility is maximised when $(x_1, x_2) = (5, 5)$ and she can attain 25 utils.[5] What is the cheapest way for the agent to attain 25 utils? Given this information, the answer must be $(x_1, x_2) = (5, 5)$. Moreover, her expenditure is \$10. The reason is as follows. First, we know that the agent can obtain 25 utils from \$10, so the cheapest way to obtain 25 utils is at most \$10. That is, $e \leq \$10$. Now suppose, by contradiction, that the agent can obtain 25 utils for, say, \$8. Then, if preferences are monotone, she will be able to obtain strictly more than 25 utils with \$10, contradicting our initial assumptions.

We can state this result formally. Fix prices (p_1, p_2) and income m. Marshallian demand is given by $x_i^*(p_1, p_2, m)$ and indirect utility is $v(p_1, p_2, m)$. Consider the EMP:

$$\min_{x_1, x_2} p_1 x_1 + p_2 x_2 \qquad \text{subject to} \qquad u(x_1, x_2) \geq v(p_1, p_2, m)$$

The induced Hicksian demand is given by $h_i(p_1, p_2, v(p_1, p_2, m))$ while the expenditure function is $e(p_1, p_2, v(p_1, p_2, m))$. Then using the reasoning above, one can show that

$$e(p_1, p_2, v(p_1, p_2, m)) = m \tag{4.1}$$

$$h_i(p_1, p_2, v(p_1, p_2, m)) = x_i^*(p_1, p_2, m) \tag{4.2}$$

[5]These numbers come from assuming $p_1 = 1$, $p_2 = 1$ and $u(x_1, x_2) = x_1 x_2$.

Suppose we start with income m. Equation (4.1) says that the minimum expenditure required to reach $v(p_1, p_2, m)$, the most utility from m, is just m. Equation (4.2) says that an agent who wishes to maximise her utility from m and one who wishes to find the cheapest way to attain $v(p_1, p_2, m)$ will buy the same goods. Intuitively, in both cases, they will spend m and will do so by equating the bang–per–buck from each good.

Equation (4.1) is practically useful. Fixing prices and omitting them from the arguments, it says that $e(v(m)) = m$. Since the expenditure function is increasing in $\overline{u}$, we can invert it and obtain:

$$v(m) = e^{-1}(m) \tag{4.3}$$

Hence the indirect utility function equals the inverse of the expenditure function. To illustrate this result, suppose $u(x_1, x_2) = x_1 x_2$. From equation (2.6), we know that

$$e(\overline{u}) = 2\sqrt{\overline{u}p_1p_2}$$

We invert this equation by letting $m = e(\overline{u})$ and $v(m) = \overline{u}$, and solving for $v(m)$. This yields

$$v(m) = \frac{m^2}{4p_1p_2}$$

One can verify that this indeed the indirect utility function.

We can also state a second, closely related, result. Fix prices (p_1, p_2) and target utility $\overline{u}$. Hicksian demand is given by $h_i(p_1, p_2, \overline{u})$ and the expenditure function is $e(p_1, p_2, \overline{u})$. Consider the UMP:

$$\max_{x_1, x_2} u(x_1, x_2) \qquad \text{subject to} \qquad p_1x_1 + p_2x_2 \leq e(p_1, p_2, \overline{u})$$

The induced Marshallian demand is given by $x_i^*(p_1, p_2, e(p_1, p_2, \overline{u}))$ while the indirect utility is $v(p_1, p_2, e(p_1, p_2, \overline{u}))$. One can show that

$$v(p_1, p_2, e(p_1, p_2, \overline{u})) = \overline{u} \tag{4.4}$$

$$x_i^*(p_1, p_2, e(p_1, p_2, \overline{u})) = h_i(p_1, p_2, \overline{u}) \tag{4.5}$$

Suppose we start with target utility $\overline{u}$. Equation (4.4) says that the most utility the agent can get from $e(p_1, p_2, \overline{u})$, the money required to reach $\overline{u}$, is just $\overline{u}$. Equation (4.5) says that an agent who wishes to find the cheapest way to attain $\overline{u}$ and one who wishes to maximise her utility from $e(p_1, p_2, \overline{u})$ will buy the same goods. Intuitively, in both cases, they will attain utility $\overline{u}$

and will do so by equating the bang–per–buck from each good.

Fixing prices and omitting them from the arguments, equation (4.4) says that $v(e(\overline{u})) = \overline{u}$. Since the indirect function is increasing in m, we can invert it and obtain:

$$e(\overline{u}) = v^{-1}(\overline{u}) \tag{4.6}$$

Hence the expenditure function equals the inverse of the indirect utility function. Together, equations (4.3) and (4.6) mean we can move back and forwards between the expenditure function and indirect utility function.

4.3 Slutsky Equation: Own Price Effects

Suppose p_1 increases by Δp_1. There are two effects:

1. Fixing the agent's utility, relative prices change causing demand to rise by $\frac{\partial h_1}{\partial p_1}\Delta p_1$. Since $\frac{\partial h_1}{\partial p_1} < 0$, this effect causes demand to fall. This is the substitution effect.
2. Fixing relative prices, the agent's income falls by $x_1^*\Delta p_1$. As a result, her demand falls by $x_1^*\frac{\partial x_1^*}{\partial m}\Delta p_1$. This is the income effect.

Putting these effects together, we have

$$\Delta x_1^* = \frac{\partial h_1}{\partial p_1}\Delta p_1 - x_1^*\frac{\partial x_1^*}{\partial m}\Delta p_1$$

Dividing by Δp_1 yields the Slutsky equation.

Theorem 1 (Own–Price Slutsky Equation)**.** *Fix prices* (p_1, p_2) *and income* m*, and let* $\overline{u} = v(p_1, p_2, m)$ *be the indirect utility. Then*

$$\frac{\partial}{\partial p_1}x_1^*(p_1, p_2, m) = \frac{\partial}{\partial p_1}h_1(p_1, p_2, \overline{u}) - x_1^*(p_1, p_2, m)\frac{\partial}{\partial m}x_1^*(p_1, p_2, m) \tag{4.7}$$

A formal proof is reasonably straightforward. Using equation (4.5),

$$h_i(p_1, p_2, \overline{u}) = x_i^*(p_1, p_2, e(p_1, p_2, \overline{u}))$$

Differentiating with respect to p_1 yields

$$\begin{aligned}\frac{\partial}{\partial p_1}h_1(p_1,p_2,\overline{u}) &= \frac{\partial x_1^*(p_1,p_2,e(p_1,p_2,\overline{u}))}{\partial p_1} + \frac{\partial x_1^*(p_1,p_2,e(p_1,p_2,\overline{u}))}{\partial m}\frac{\partial e(p_1,p_2,\overline{u})}{\partial p_1} \\ &= \frac{\partial x_1^*(p_1,p_2,e(p_1,p_2,\overline{u}))}{\partial p_1} + \frac{\partial x_1^*(p_1,p_2,e(p_1,p_2,\overline{u}))}{\partial m}x_1^*(p_1,p_2,e(p_1,p_2,\overline{u})) \quad (4.8)\end{aligned}$$

where the second line comes from Sheppard's Lemma. Using the definition of $\overline{u}$ and equation (4.1),

$$e(p_1,p_2,\overline{u}) = e(p_1,p_2,v(p_1,p_2,m)) = m \quad (4.9)$$

Substituting (4.9) into (4.8) and rearranging yields (4.7), as required.

4.4 Slutsky Equation: Cross Price Effects

Equation (4.7) analyses the effect of a change in p_1 on the demand for good 1. We can use the same approach to analyse the effect of a change in p_2 on the demand for good 1.

Suppose p_2 increases by Δp_2. As before, there are two effects:

1. Fixing the agent's utility, relative prices change causing demand to rise by $\frac{\partial h_1}{\partial p_2}\Delta p_2$. Recall that $\frac{\partial h_1}{\partial p_2} > 0$ if the goods are net substitutes and $\frac{\partial h_1}{\partial p_2} < 0$ are net complements.
2. Fixing relative prices, the agent's income falls by $x_2^*\Delta p_2$. As a result, her demand falls by $x_2^*\frac{\partial x_1^*}{\partial m}\Delta p_2$.

Putting these effects together, we have

$$\Delta x_1^* = \frac{\partial h_1}{\partial p_2}\Delta p_2 - x_2^*\frac{\partial x_1^*}{\partial m}\Delta p_2$$

Dividing by Δp_2 yields the Slutsky equation for cross–price effects.

Theorem 2 (Cross–Price Slutsky Equation)**.** *Fix prices (p_1,p_2) and income m, and let $\overline{u} = v(p_1,p_2,m)$ be the indirect utility. Then*

$$\frac{\partial}{\partial p_2}x_1^*(p_1,p_2,m) = \frac{\partial}{\partial p_2}h_1(p_1,p_2,\overline{u}) - x_2^*(p_1,p_2,m)\frac{\partial}{\partial m}x_1^*(p_1,p_2,m) \quad (4.10)$$

The proof is almost identical to that of (4.7). Using equation (4.5),

$$h_i(p_1, p_2, \overline{u}) = x_i^*(p_1, p_2, e(p_1, p_2, \overline{u}))$$

Differentiating with respect to p_2 yields

$$\begin{aligned}\frac{\partial}{\partial p_2} h_1(p_1, p_2, \overline{u}) &= \frac{\partial x_1^*(p_1, p_2, e(p_1, p_2, \overline{u}))}{\partial p_2} + \frac{\partial x_1^*(p_1, p_2, e(p_1, p_2, \overline{u}))}{\partial m} \frac{\partial e(p_1, p_2, \overline{u})}{\partial p_2} \\ &= \frac{\partial x_1^*(p_1, p_2, e(p_1, p_2, \overline{u}))}{\partial p_2} + \frac{\partial x_1^*(p_1, p_2, e(p_1, p_2, \overline{u}))}{\partial m} x_2^*(p_1, p_2, e(p_1, p_2, \overline{u})) \quad (4.11)\end{aligned}$$

where the second line comes from Sheppard's Lemma. Using the definition of $\overline{u}$ and equation (4.1),

$$e(p_1, p_2, \overline{u}) = e(p_1, p_2, v(p_1, p_2, m)) = m \tag{4.12}$$

Substituting (4.12) into (4.11) and rearranging yields (4.10).

4.5 Slutsky Equation: Example

We illustrate the Slutsky equation with our running example. Let $u(x_1, x_2) = x_1 x_2$. From the UMP we know that

$$\begin{aligned} x_1^*(p_1, p_2, m) &= \frac{m}{2p_1} \\ v(p_1, p_2, m) &= \frac{m^2}{4p_1 p_2} \end{aligned}$$

From the EMP (see Section 2.2) we know that

$$\begin{aligned} h_1(p_1, p_2, \overline{u}) &= \left(\overline{u}\frac{p_2}{p_1}\right)^{1/2} \\ e(p_1, p_2, \overline{u}) &= 2(\overline{u} p_1 p_2)^{1/2} \end{aligned}$$

The left hand side of the Slutsky equation states

$$\frac{\partial}{\partial p_1} x_1^*(p_1, p_2, m) = -\frac{1}{2} m p_1^{-2} \tag{4.13}$$

The right hand side is

$$\begin{aligned}\frac{\partial h_1}{\partial p_1} - x_1^* \frac{\partial x_1^*}{\partial m} &= -\frac{1}{2}\overline{u}^{1/2} p_1^{-1/2} p_2^{1/2} - \frac{1}{4} m p_1^{-2} \\ &= -\frac{1}{4} m p_1^{-2} - \frac{1}{4} m p_1^{-2} \qquad (4.14)\end{aligned}$$

where the second line uses $\overline{u} = v(p_1, p_2, m)$.

Observe that (4.13) equals (4.14) as we would hope. Moreover, the two terms in equation (4.14) are identical. This means that the substitution and income effects are of equal size: both account for 50% of the fall in demand.

5 Consumer Surplus

It is often important to put a monetary value on the effect of a price change on an agent's utility. For example, the government may wish to evaluate the impact of a tax change; or a court may wish to evaluate the negative effect of collusion on consumers.

To gain some intuition, suppose the consumer has monetary valuations for each unit of the good. In particular suppose their valuations are given by table 2.

Unit	Valuation $
1	10
2	8
3	6
4	4
5	2

Table 2: **Agent's Valuations**

Suppose the price of the good is initially $p_1 = 3$. Since the agent buys a unit if and only if her valuation exceeds the price, she will buy 4 units. Her consumer surplus, the difference between her willingness to pay and the price she pays, equals

$$\text{CS} = (10-3) + (8-3) + (6-3) + (4-3) = \$16$$

Suppose the price rises to $p_1 = 7$. The agent then consumes 2 units and her consumer surplus is

$$\text{CS} = (10-7) + (8-7) = \$4.$$

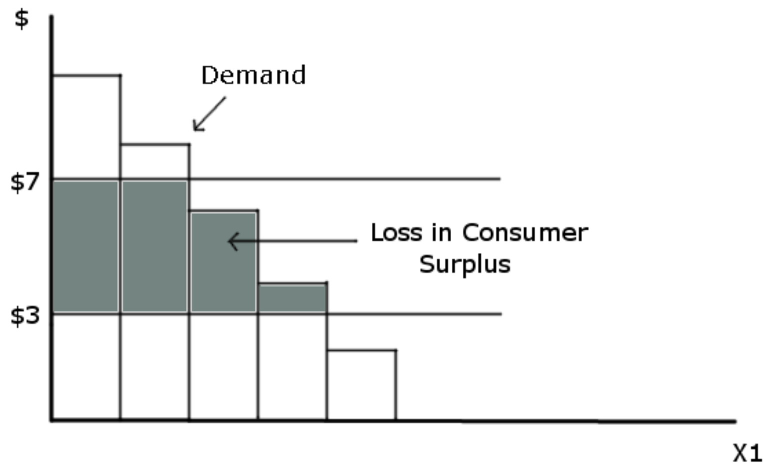

Figure 9: **Consumer Surplus with Quasilinear Demand.** The figure shows the agent's demand curve. The shaded area is the loss in CS due to the price increase.

Hence the agent would need to be compensated \$12 for this price increase. This is shown in figure 9.

This exercise is familiar from introductory economics courses: consumer surplus is the area under the agent's Marshallian demand curve. However this approach assumes the agent has quasilinear utility, allowing us to associate a monetary value to each unit demanded by the agent. In Section 5.1 we show that the welfare effect of a price change is determined by the area under the Hicksian demand curve rather than the Marshallian demand. In Section 5.2 we see that, when utility is quasi–linear then Hicksian demand and Marshallian demand coincide, justifying the approach taken above.

5.1 Compensating Variation

Suppose prices and income are initially (p_1, p_2, m), and that p_1 increases to p_1'. The **compensating variation** is defined by

$$CV = e(p_1', p_2, \overline{u}) - e(p_1, p_2, \overline{u})$$

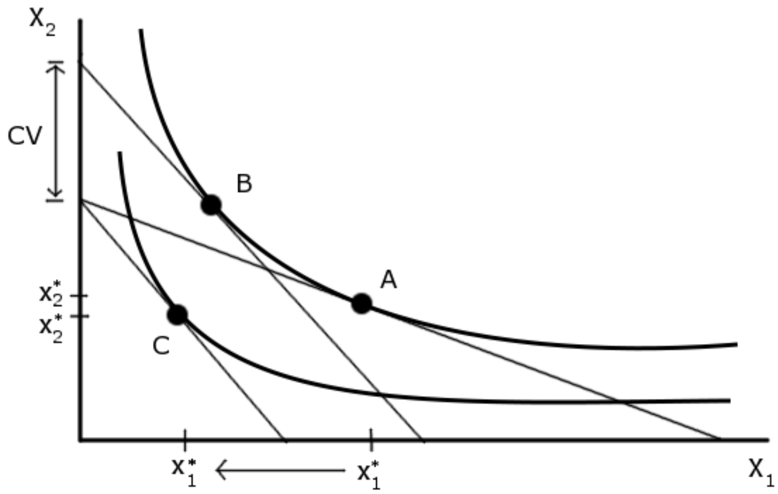

Figure 10: **Compensating Variation and Indifference Curves.** This figure shows the effect on an increase in p_1. The Marshallian demand falls from A to C. The Hicksian demand moves from A to B. The compensating variation equals the difference between the consumer's original income and the income she would need to attain $\overline{u}$.

The CV is thus the extra spending needed to keep the agent at their original utility level. That is, an increase in income of CV completely compensates the agent for the price increase.[6] This is shown in figure 10.

The compensating variation can be related to the Hicksian demand curve. Applying the fundamental theorem of calculus,[7]

$$\begin{aligned} CV &= \int_{p_1}^{p_1'} \frac{\partial}{\partial p_1} e(\tilde{p}_1, p_2, \overline{u})\, d\tilde{p}_1 \\ &= \int_{p_1}^{p_1'} h_1(\tilde{p}_1, p_2, \overline{u})\, d\tilde{p}_1 \qquad (5.1) \end{aligned}$$

where the second equation follows from Sheppard's Lemma. Equation (5.1) says that the lost welfare from the price change equals the area under the Hicksian demand curve. See figure 11.

[6]There is a closely related measure of welfare called the equivalent variation. We will not discuss this here.
[7]The fundamental theorem of calculus says that $f(b) - f(a) = \int_a^b f'(x)dx$.

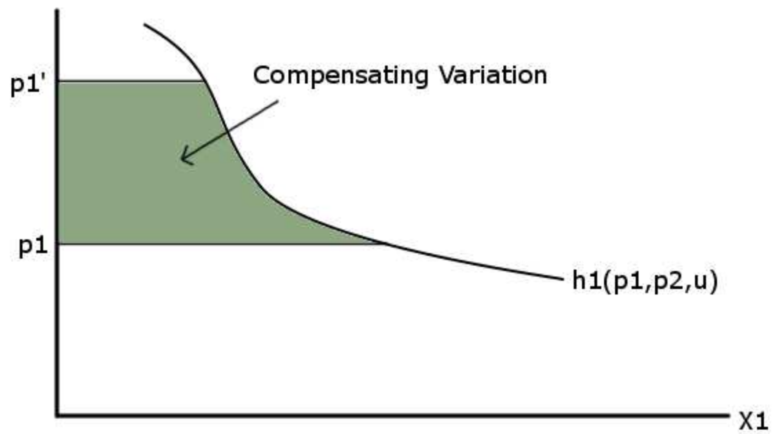

Figure 11: **Compensating Variation and Hicksian Demand.** This figure shows that CV equals the area under the demand curve.

5.2 Quasilinear Utilities

While we may wish to calculate the area under the Hicksian demand, it is often easier to calculate the area under the Marshallian demand curve. For example, in empirical applications, it is easy to estimate the Marshallian demand by looking at how much people buy at different prices.

Suppose utility is **quasilinear** in that it can be represented by a utility function of the form

$$u(x_1, x_2) = v(x_1) + x_2$$

where we assume $v(\cdot)$ is increasing and concave. Under this specification, the marginal utility of the second good is constant. For example, x_2 could be a general aggregate good or cash.

When utility is quasilinear we can think of an agent's utility in terms of dollar valuations, as at the start of this section. The argument is as follows. The agent's problem is to maximise her utility subject to her budget constraint, $p_1x_1 + p_2x_2 \leq m$. Since utility is monotone, the budget constraint will bind. Using the substitution method, the budget constraint becomes

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}x_1$$

Substituting this into the utility function, the agent maximises

$$v(x_1) - \frac{p_1}{p_2}x_1 + \frac{m}{p_2} \tag{5.2}$$

Notice the last term is a constant and can be ignored. If x_2 is interpreted as cash, we can normalise $p_2 = 1$. The agent then chooses x_1 to maximise

$$v(x_1) - p_1 x_1$$

The agent's choice is independent of m, so she acts as if she values x_1 units of good 1 at $v(x_1)$, independent of the units of x_2 being consumed. We can then think of $v'(x_1)$ as her valuation of the marginal unit of x_1.

Under quasilinear utility, the Hicksian and Marshallian demands coincide. Ignoring boundary problems, the Marshallian demand is derived by maximising (5.2). The first–order condition implies that Marshallian demand is implicitly given by

$$v'(x_1^*(p_1, p_2, m)) = \frac{p_1}{p_2} \tag{5.3}$$

Turning to the EMP, the agent minimises

$$\mathcal{L} = p_1 x_1 + p_2 x_2 + \lambda[\overline{u} - v(x_1) - x_2]$$

The first first–order conditions are

$$\begin{aligned} p_1 &= \lambda v'(x_1) \\ p_2 &= \lambda \end{aligned}$$

Looking at the ratio of these two equations, Hicksian demand is implicitly given by

$$v'(h_1(p_1, p_2, \overline{u})) = \frac{p_1}{p_2} \tag{5.4}$$

From equations (5.3) and (5.4) we see that Marshallian demand and Hicksian demand coincide. Hence the compensating variation is given by

$$\begin{aligned} CV &= \int_{p_1}^{p_1'} h_1(\tilde{p}_1, p_2, \overline{u})\, d\tilde{p}_1 \\ &= \int_{p_1}^{p_1'} x_1^*(\tilde{p}_1, p_2, m)\, d\tilde{p}_1 \end{aligned}$$

This result provides a foundation for the classical measure of consumer surplus.

6 Endowments of Goods

In the UMP we assume that agents are endowed with income m and use it to maximise their utility. While this is a useful model to address demand for retail products, it is sometimes more accurate to assume agents are endowed with goods which they can sell on the open market. There are two reasons for analysing this model:

- The model is important for understanding practical problems such as a worker's choice of labour supply (Section 6.1), and an agent's decision to smooth consumption over time (Section 6.2).
- When we analyse the entire economy, we will want to close the model. Hence we wish the agents who demand goods to also work for firms that make goods.

Suppose there are N goods and the agent starts with endowments $\{\omega_1, \ldots, \omega_N\}$, where $\omega_i \geq 0$ for all i. The consumer can sell these goods at market prices $\{p_1, \ldots, p_N\}$. For example, an agent may own a farm which produces vegetables and may sell the produce to buy meat. The agent has income

$$m = \sum_{i=1}^{N} p_i \omega_i \tag{6.1}$$

Given equation (6.1) the agent's problem is the same as that studied so far. We can derive her Marshallian demand and indirect utility (see figure 12). We can also derive her Hicksian demand and expenditure function (since these are independent of income)

The one major difference from the model with exogenous income is that a price change now affects the agent's income as well as the goods she buys. We study this in Section 6.3. We first consider two applications.

6.1 Labour Supply

Suppose an agent has utility $u(x_1, x_2) = x_1 x_2$ over leisure x_1 and a general consumption good x_2. The agent has exogenous income m and can also work at wage w. She has T hours which

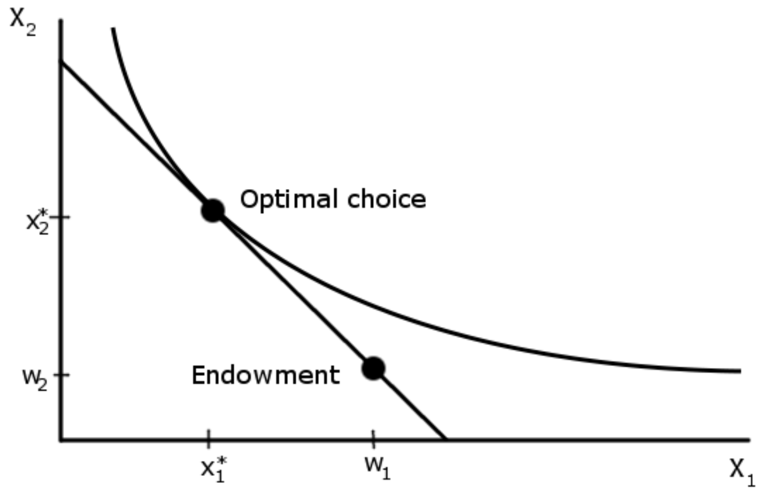

Figure 12: **Marshallian Demand with Endowments** This figure shows the optimal choice when the agent has endowments of the two goods.

she can allocate to either work or leisure. We normalise the price of x_2 to $p_2 = 1$.

The agent's budget constraint is

$$x_2 = w(T - x_1) + m$$

The left hand side equals the agent's spending on consumption; the right hand side equals her income. As a thought experiment, one can imagine the agent selling all T units of her labour and then buying x_1 units of it back at price w to be consumed as leisure. We can thus rewrite the budget constraint as

$$wx_1 + x_2 = wT + m$$

The left hand side is the goods consumed (including leisure, consumed at price w). The right hand side is the agent's endowment income, as in equation (6.1).

Ignoring boundary constraints,[8] her problem is

$$\max_{x_1,x_2} \mathcal{L} = x_1x_2 + \lambda[m + w(T - x_1) - x_2]$$

[8]With this problem the boundary constraints are slightly different to normal since the agent cannot consume more that T units of leisure. We thus have $T \geq x_1 \geq 0$ and $x_2 \geq 0$.

The FOCs are

$$x_2 = \lambda w$$
$$x_1 = \lambda$$

Taking the ratio of these FOCs, we see that

$$\frac{x_2}{x_1} = w$$

As before, the left hand side is the MRS, while the right hand side is the price ratio. Using the budget constraint the agent's demands are given by

$$x_1^* = \frac{1}{2w}[wT + m] \tag{6.2}$$
$$x_2^* = \frac{1}{2}[wT + m] \tag{6.3}$$

Equations (6.2) and (6.3) show that the consumer splits her endowment income of $wT + m$ equally between leisure and consumption. This is just like the solution to the Cobb Douglas problem without endowments (see UMP notes), where we found that

$$x_1^* = \frac{1}{2p_1}m \qquad \text{and} \qquad x_2^* = \frac{1}{2p_2}m \tag{6.4}$$

We can now evaluate an effect of a change in wages. Differentiating (6.2) and (6.3),

$$\frac{\partial x_1^*}{\partial w} = \frac{1}{2w^2}wT - \frac{1}{2w^2}[wT + m] = -\frac{1}{2w^2}m \tag{6.5}$$
$$\frac{\partial x_2^*}{\partial w} = \frac{1}{2}T \tag{6.6}$$

From (6.5), we see an increase in the wage reduces the amount of leisure the agent consumes. There are two effects here: an increase in the wage raises the relative price of leisure and reduces demand (the substitution effect); it also makes the agent richer and increases the demand for leisure (the income effect). In this case the substitution effect dominates the income effect: we analyse this formally in Section 6.4.

From (6.6), we see an increase in the wage increase the amount of x_2 the agent consumes. This is because an increase in wages increase the value of the agent's endowment; in comparison, without endowments, equation (6.4) shows that x_2^* is independent of p_1.

6.2 Intertemporal Optimisation

Suppose an agent allocates consumption (e.g. money) across two periods. Let the consumption in period 1 and 2 be x_1 and x_2 respectively. The agent's utility is

$$u(x_1, x_2) = \ln(x_1) + (1+\beta)^{-1}\ln(x_2) \tag{6.7}$$

where $\beta \geq 0$ is the agent's discount rate.

In periods 1 and 2 the agent is endowed with income m_1 and m_2, respectively. The agent can save at interest rate $r \geq 0$, so that \$1 in period 1 is worth \$(1+r) in period 2. As a result, the agent's budget constraint is

$$m_1 + (1+r)^{-1}m_2 = x_1 + (1+r)^{-1}x_2 \tag{6.8}$$

The left hand side of (6.8) is the agent's lifetime income in terms of period 1 dollars. The right and side is the agent's lifetime spending. We say they are borrowing when $x_1 > m_1$ and saving when $x_1 < m_1$.

We can solve this problem just as we would solve a regular utility maximisation problem, where $p_1 = 1$ and $p_2 = (1+r)^{-1}$. See figure 13. Using (6.7) and (6.8) the tangency condition, $MRS = p_1/p_2$, becomes

$$(1+\beta)\frac{1/x_1}{1/x_2} = (1+r)$$

Rearranging,

$$x_1^* = \frac{1+\beta}{1+r}x_2^* \tag{6.9}$$

Equation (6.2) immediately implies that if $r = \beta$ then the agent consumes the same in each period, $x_1^* = x_2^*$. Intuitively, since the agent's per–period utility is concave, she wishes to smooth her consumption across time. If $r = \beta$, then the agent is just as impatient as the market, so she will perfectly smooth her consumption across the two periods. If $\beta > r$ then the agent is more impatient than the market and she consumes more in the first period, $x_1^* > x_2^*$.

Using the budget constraint, demand is given by

$$x_1^* = \frac{1+\beta}{2+\beta}[m_1 + (1+r)^{-1}m_2] \qquad \text{and} \qquad x_2^* = \frac{1+r}{2+\beta}[m_1 + (1+r)^{-1}m_2]$$

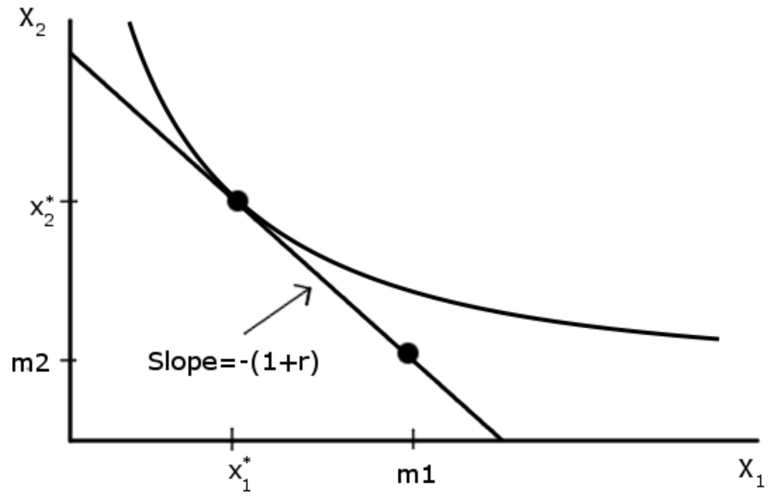

Figure 13: **Intertemporal Optimisation** This figure shows an agent who has a high income in period 1 and a low income in period 2. At the optimum, she saves in period 1.

6.3 Own Price Effects

Suppose there is an increase in p_1. As with a fixed income m, the budget line becomes steeper. However, since the value of the endowment changes, it is no longer true that the budget set shrinks. Rather, the budget line pivots around the endowment: see figure 14.

As in Section 4, we can decompose the price change into a substitution and income effect. However, the income effect has to be adjusted for the change in the value of the endowment. Suppose p_1 increases by Δp_1. Then there are two effects:

1. Fixing the agent's utility, relative prices change causing demand to rise by $\frac{\partial h_1}{\partial p_1}\Delta p_1$. Since $\frac{\partial h_1}{\partial p_1} < 0$, this effect causes demand to fall. This is the substitution effect.
2. Fixing relative prices, the agent's income rises by $(\omega_1 - x_1^*)\Delta p_1$. This means that the agent's income rises if she is a net seller of the good (as in the labour example), and falls if she is a net buyer of the good. As a result, her demand rises by $(\omega_1 - x_1^*)\frac{\partial x_1^*}{\partial m}\Delta p_1$. This is the income effect.

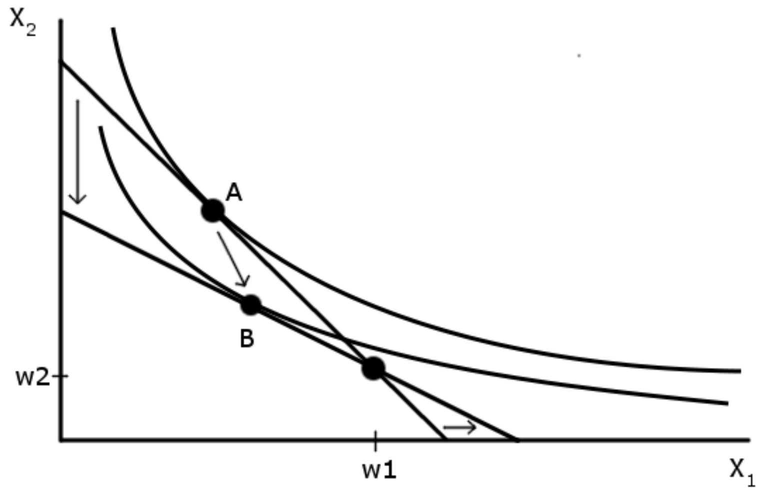

Figure 14: **Own–Price Effects with Endowments.** This figure shows the effect on a *decrease* in p_1 when the agent is endowed with $\{\omega_1, \omega_2\}$. Note the income looks like it goes down, even though prices fall. This is because, at point A, the agent owns more of good 1 than she buys, $\omega_1 > x_1^*$. Hence a decrease in p_1 reduces her purchasing power.

Putting these effects together, we have

$$\Delta x_1^* = \frac{\partial h_1}{\partial p_1}\Delta p_1 + (\omega_1 - x_1^*)\frac{\partial x_1^*}{\partial m}\Delta p_1$$

Dividing by Δp_1 yields the Slutsky equation.

Theorem 3 (Own–Price Slutsky Equation with Endowments)**.** *Fix prices* (p_1, p_2)*, income* m *and endowments* (ω_1, ω_2)*, and let* $\overline{u} = v(p_1, p_2, m)$ *be the indirect utility. Then*

$$\frac{\partial}{\partial p_1} x_1^*(p_1, p_2, m) = \frac{\partial}{\partial p_1} h_1(p_1, p_2, \overline{u}) + (\omega_1 - x_1^*(p_1, p_2, m))\frac{\partial}{\partial m} x_1^*(p_1, p_2, m) \tag{6.10}$$

This result follows from the regular Slutsky equation (4.7). All we need to do is define net demand for good 1 by $z_1^*(p_1, p_2, m) = x_1^*(p_1, p_2, m) - \omega_1$. We can then apply the regular Slutsky equation to the agent's net demand:

$$\frac{\partial}{\partial p_1} z_1^*(p_1, p_2, m) = \frac{\partial}{\partial p_1} h_1(p_1, p_2, \overline{u}) - z_1^*(p_1, p_2, m)\frac{\partial}{\partial m} z_1^*(p_1, p_2, m) \tag{6.11}$$

Since z_1^* and x_1^* differ by a constant term, we can put equation (6.11) back in terms of $x_i^*(p_1, p_2, m)$, yielding equation (6.10).

6.4 Labour Supply and the Slutsky Equation

We now apply the Slutsky equation (6.10) to the labour supply problem in Section 6.1. From equation (6.2), the Marshallian demand is

$$x_1^*(p_1, p_2, m) = \frac{1}{2w}[wT + m] \tag{6.12}$$

Using the (6.2) and (6.3) the indirect utility is

$$v(p_1, p_2, m) = x_1^* x_2^* = \frac{1}{4w}[wT + m]^2 \tag{6.13}$$

From (2.5) and using $p_1 = w$ and $p_2 = 1$, the Hicksian demand is

$$h_1(p_1, p_2, \overline{u}) = \overline{u}^{1/2} w^{-1/2} \tag{6.14}$$

We now have all the elements we need.

The left hand side of the Slutsky equation is

$$\frac{\partial}{\partial p_1} x_1^* = -\frac{1}{2w^2} m$$

The right hand side of the Slutsky equation is

$$\begin{aligned}
\frac{\partial h_1}{\partial p_1} + (\omega_1 - x_1^*)\frac{\partial x_1^*}{\partial m} &= -\frac{1}{2} w^{-3/2} \overline{u}^{1/2} + \left[T - \frac{1}{2w}[wT + m]\right] \frac{1}{2w} \\
&= -\frac{1}{2} w^{-3/2} \frac{1}{2} w^{-1/2} [wT + m] + \frac{1}{4w^2}[wT - m] \\
&= -\frac{1}{4w^2}[wT + m] + \frac{1}{4w^2}[wT - m] \\
&= -\frac{1}{2w^2} m
\end{aligned}$$

where the first line uses (6.12) and (6.14), and the second uses $\overline{u} = v(p_1, p_2, m)$ and equation (6.13). We can therefore see that the substitution effect outweighs the income effect, and as m becomes smaller these two effects grow closer in magnitude. In the limit, as $m \to 0$, leisure demand (and hence labour supply) are independent of the wage.

Firm's Problem

Simon Board*

This Version: September 20, 2009
First Version: December, 2009.

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 $q \in \Re_+$. 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 $z_i \in \Re_+$ 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,

$$\text{Isoquant} = \{(z_1, z_2) \in \Re^2_+ | f(z_1, z_2) = const.\}$$

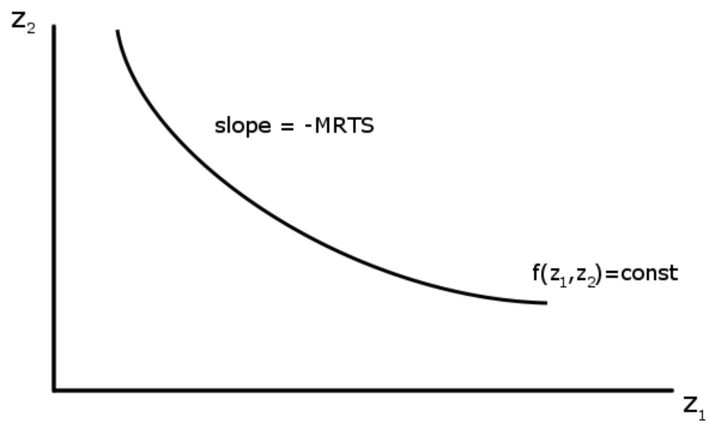

Figure 1: **Isoquant.** This figure shows two isoquants. Each curve depicts the bundles that yield constant output.

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

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

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

$$MRTS = \frac{MP_1}{MP_2} \tag{1.4}$$

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 $z' = (z'_1, z'_2)$,

$$\left.\begin{array}{l} z_i \geq z'_i \text{ for each } i \\ z_i > z'_i \text{ for some } i \end{array}\right\} \quad \text{implies } f(z_1, z_2) > f(z'_1, z'_2).$$

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

$$f(tz + (1-t)z') \geq f(z') \qquad \text{for all } t \in [0,1] \tag{1.5}$$

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

$$f(tz_1, tz_2) \leq tf(z_1, z_2) \qquad \text{for } t \geq 1 \tag{1.6}$$

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

$$f(tz_1, tz_2) = tf(z_1, z_2) \qquad \text{for } t \geq 1$$

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

$$f(tz_1, tz_2) \geq tf(z_1, z_2) \qquad \text{for } t \geq 1$$

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 $z' = (z'_1, z'_2)$,

$$f(tz + (1-t)z') \geq tf(z) + (1-t)f(z') \qquad \text{for } t \in [0,1] \tag{1.7}$$

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

$$f(z_1, z_2) = z_1^{\alpha} z_2^{\beta} \qquad \text{for } \alpha \geq 0 \text{ and } \beta \geq 0$$

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.

$$f(tz_1, tz_2) = (tz_1)^{\alpha}(tz_2)^{\beta}t = t^{\alpha+\beta} z_1^{\alpha} z_2^{\beta} = t^{\alpha+\beta} f(z_1, z_2)$$

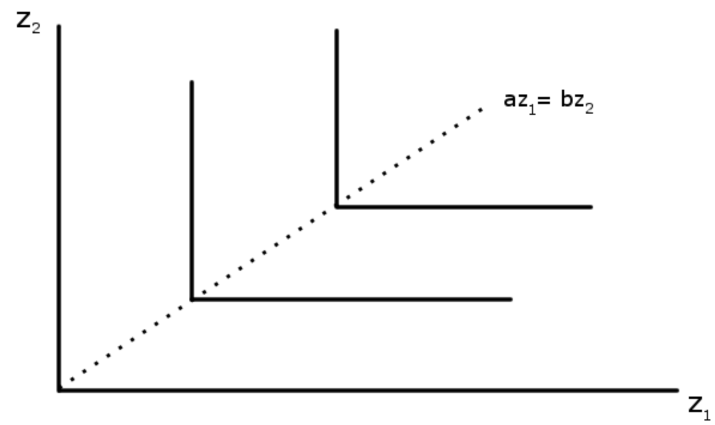

Figure 2: **Isoquants for Leontief Technology.** The isoquants are L–shaped, with the kink along the line $\alpha z_1 = \beta z_2$.

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.

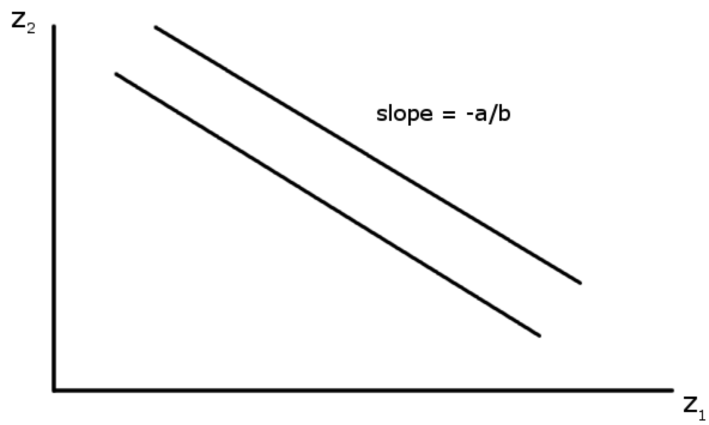

Figure 3: **Isoquants for Perfect Substitutes.** The isoquants are straight line with slope $-\alpha/\beta$.

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

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

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 **input demands** which are denoted by

$$z_i^*(r_1,\dots,r_N,q)$$

The money the firm must spend in order to attain its target output is its cost. The **cost function** is therefore

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

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

$$c(r_1,\dots,r_N,q) = \sum_{i=1}^{N} r_i z_i^*(r_1,\dots,r_N,q) \tag{3.2}$$

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 **average cost** is,

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

The **marginal cost** equals the cost of each additional unit,

$$MC(r_1,r_2,q) = \frac{dc(r_1,r_2,q)}{dq}$$

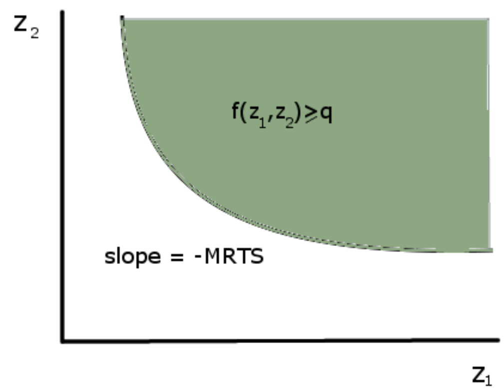

Figure 4: **Constraint Set.** This figure shows the set of inputs that deliver the target output, q.

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:

$$\{(z_1, z_2) : r_1 z_1 + r_2 z_2 = const.\}$$

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.

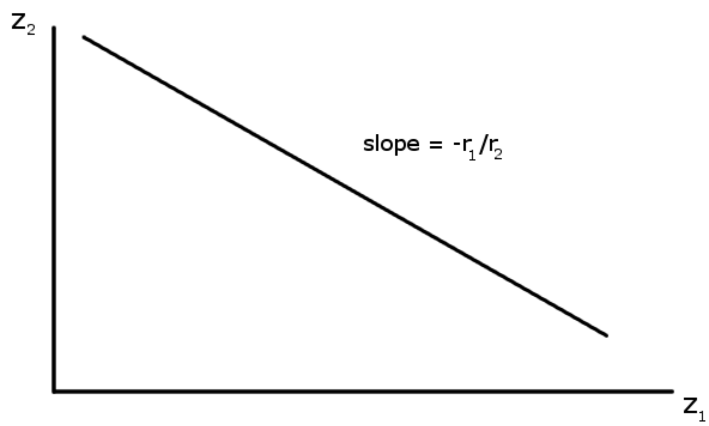

Figure 5: **Isocost.** The isocost function shows the set of inputs which cost the same amount of money.

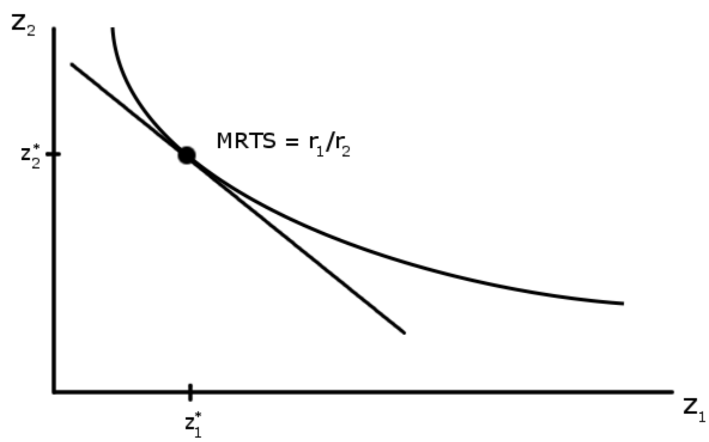

Figure 6: **Tangency.** This figure shows that, at the optimal input combination, the isocost curve is tangent to the isoquant.

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

$$\frac{MP_1}{MP_2} = \frac{r_1}{r_1} \tag{3.4}$$

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,

$$f(z_1, z_2) = q. \tag{3.5}$$

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 $r_1 z_1 = r_2 z_2$, so the firm spends the same on both its inputs.

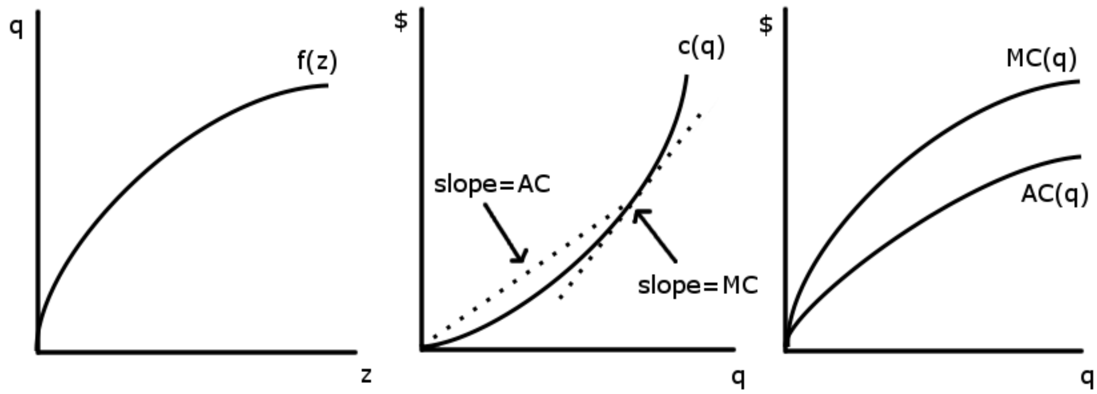

Figure 7: **Cost curves.** This figure shows the cost, average cost and marginal cost curves for the Cobb–Douglas example.

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

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

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

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

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 *minimising* 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

$$\frac{\partial L}{\partial z_1} = r_1 - \lambda \frac{\partial u}{\partial z_1} = 0 \tag{3.6}$$

$$\frac{\partial L}{\partial z_2} = r_2 - \lambda \frac{\partial u}{\partial z_2} = 0 \tag{3.7}$$

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

$$f(z_1, z_2) = q \tag{3.8}$$

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

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,

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

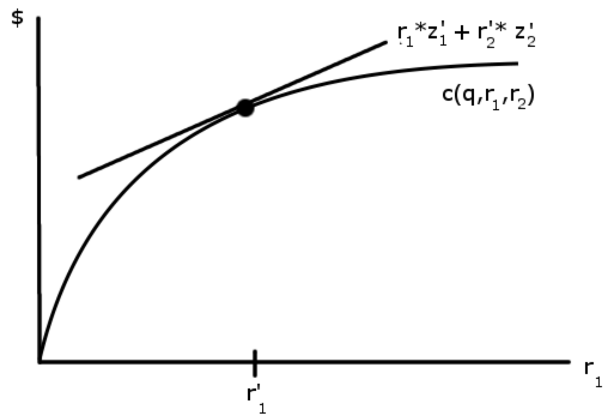

Figure 8: **Concavity of Cost Function in Input Prices.** This figure shows how the cost function lies under the pseudo–cost function.

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,

$$\frac{\partial}{\partial r_1} c(r_1, r_2, q) = z_1^*(r_1, r_2, q) \tag{3.9}$$

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 Δr_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:

$$\frac{d}{dz_i} MP_i(z_1, z_2) = \frac{d^2}{dz_i^2} f(z_1, z_2) \le 0$$

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 $MC(q) \ge AC(q)$, is flat when $MC(q) = AC(q)$ and is and decreasing when $MC(q) \le AC(q)$. Suppose the firm currently produces n units of output, and that the marginal cost of the $(n+1)^{\text{st}}$ unit is higher than the average cost of the first n. 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 $c'(q)q \ge c(q)$. Rearranging, this condition is just $MC(q) \ge AC(q)$, 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

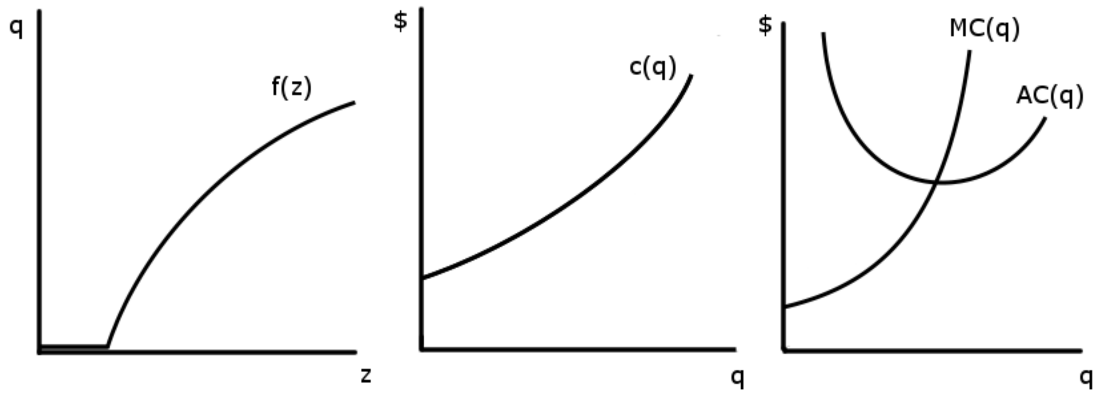

Figure 9: **Cost Curves for a Nonconcave Production Function I: Fixed Cost.** This figure shows the cost, average cost and marginal cost curves when the firm must pay a fixed cost.

positive quantities but requires a fixed cost needed to initiate production.[2] 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.[3] 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.[4]

[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".

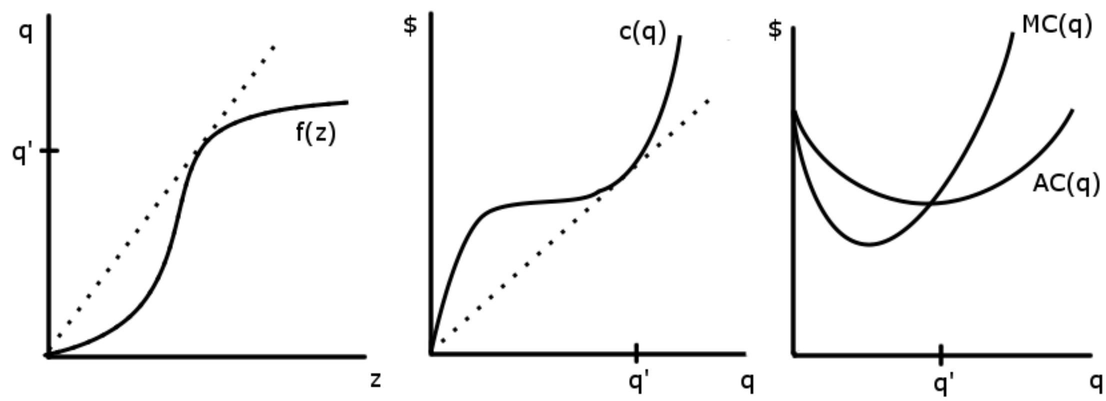

Figure 10: **Cost Curves for a Nonconcave Production Function II.** This figure shows the cost, average cost and marginal cost curves.

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

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

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

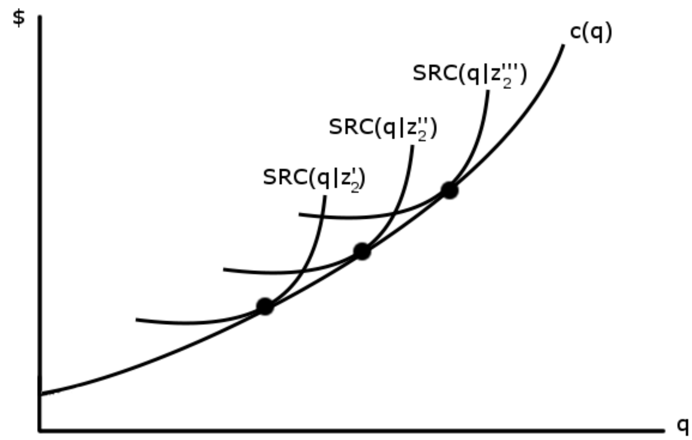

Figure 11: **Long Run and Short Run Costs.** This figure shows the long run cost curve and the short run cost curves corresponding to three levels of the second input.

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 $r_1 z_1' + r_1 z_2'$, 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 it's 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

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

The first–order conditions are

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

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

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 $f = z_1^{1/3} z_2^{1/3}$. Profit is given by

$$\pi = p z_1^{1/3} z_2^{1/3} - r_1 z_1 - r_2 z_2$$

The FOCs are

$$\frac{1}{3} p z_1^{-2/3} z_2^{1/3} = r_1$$
$$\frac{1}{3} p z_1^{1/3} z_2^{-2/3} = r_2$$

Solving these two equations yields input demands:

$$z_1^*(p, r_1, r_2) = \frac{1}{27} \frac{p^3}{r_1^2 r_2} \qquad \text{and} \qquad z_2^*(p, r_1, r_2) = \frac{1}{27} \frac{p^3}{r_1 r_2^2}$$

The optimal supply is

$$q^*(p, r_1, r_2) = (z_1^*)^{1/3} (z_2^*)^{1/3} = \frac{1}{9} \frac{p^2}{r_1 r_2}$$

The profit function is

$$\pi^*(p, r_1, r_2) = pq^* - r_1 z_1^* - r_2 z_2^* = \frac{1}{27} \frac{p^3}{r_1 r_2}$$

4.3 Two–Step Solution

Step 1. Find the cheapest way to attain output q. Recall the cost function is given by

$$c(q, r_1, r_2) = \min_{z_1, z_2} r_1 z_1 + r_2 z_2 \qquad \text{subject to} \qquad \begin{aligned} & f(z_1, z_2) \geq q \\ & z_i \geq 0 \quad \text{for all } i \end{aligned}$$

Step 2. Find the profit–maximising output. Given a cost function, the firm's problem is

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

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,

$$p = MC(q, r_1, r_2) \tag{4.4}$$

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,

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

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,

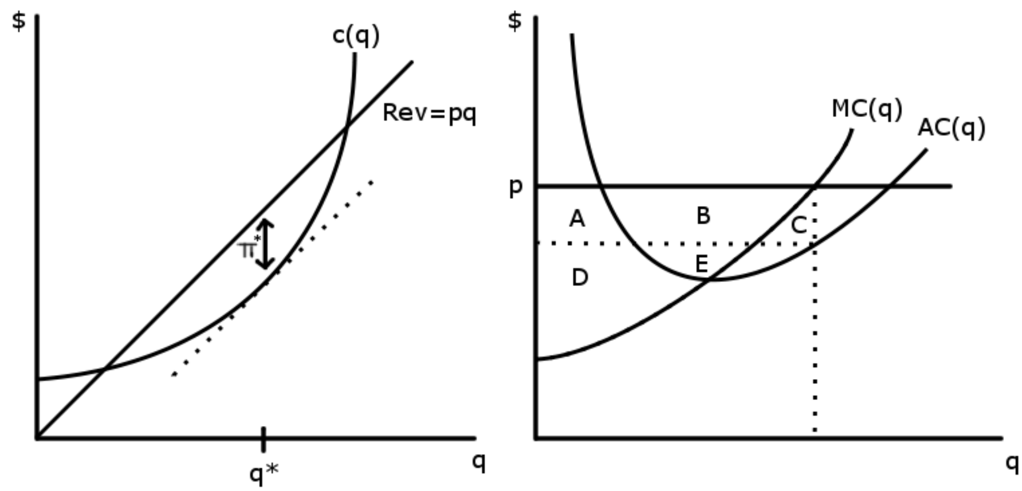

Figure 12: **Profit maximisation.** The left panel shows that profit is maximised when the revenue line is parallel to the cost line. The vertical gap, is then equal to the firm's profit. The right panel shows that profit is maximised when the price equals to marginal cost. Profit then equals A+B+C.

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

$$q^*(p, r_1, r_2) = \frac{1}{9} \frac{p^2}{r_1 r_2}$$

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 $c(q) = 20q^2 - 8q^3 + q^4$.

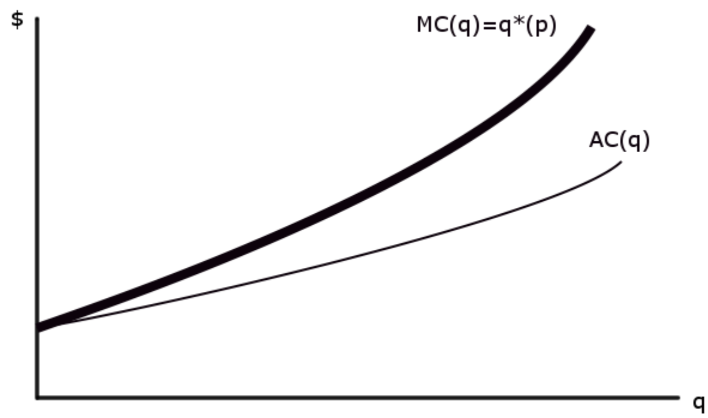

Figure 13: **Supply Curve with Convex Costs.** This figure shows how the supply curve coincides with the marginal cost curve.

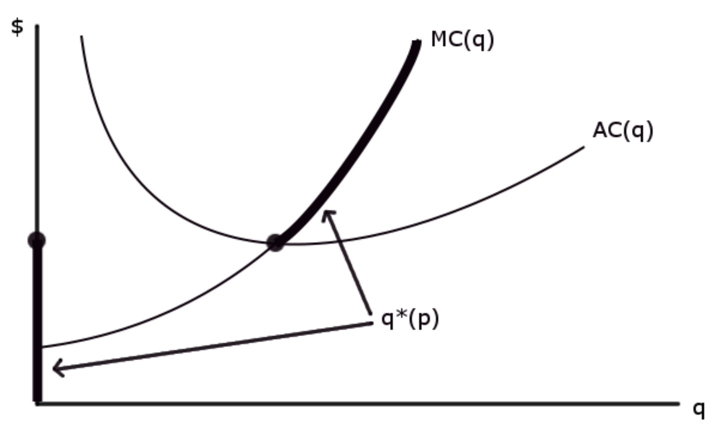

Figure 14: **Supply Curve with Nonconvex Costs I: Fixed Costs.** This figure shows how the supply curve coincides with the marginal cost curve when it lies above the average cost.

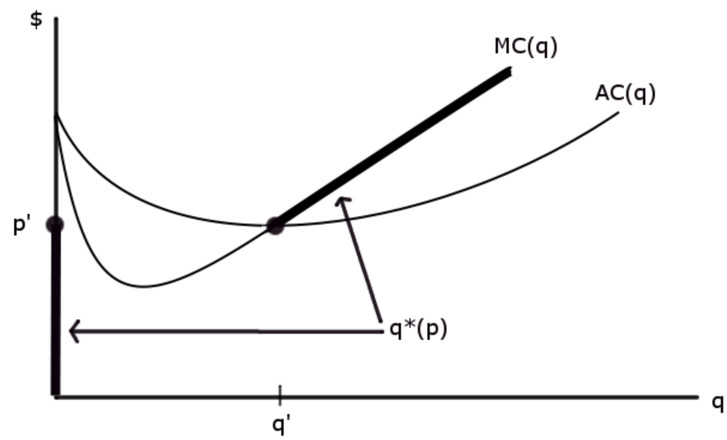

Figure 15: **Supply Curve with Nonconvex Costs II: U–Shaped Marginal Cost.** This figure shows how the supply curve coincides with the marginal cost curve when it lies above the average cost.

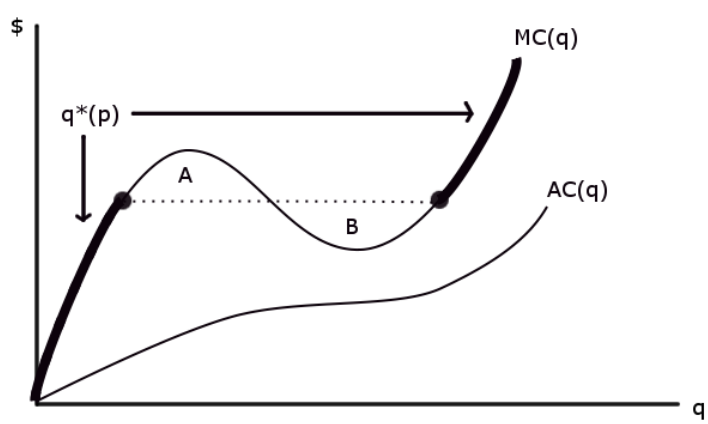

Figure 16: **Supply Curve with Nonconvex Costs III.** This figure shows how the supply curve coincides with the marginal cost curve when it lies above the average cost.

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,

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

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

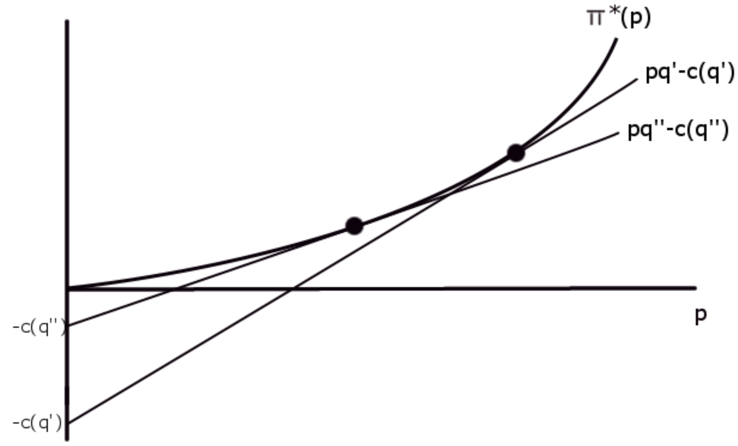

Figure 17: **Convexity of Profit Functions** This figure shows how the profit function equals the upper envelope of the pseudo–profit functions, $pq - c(q)$.

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

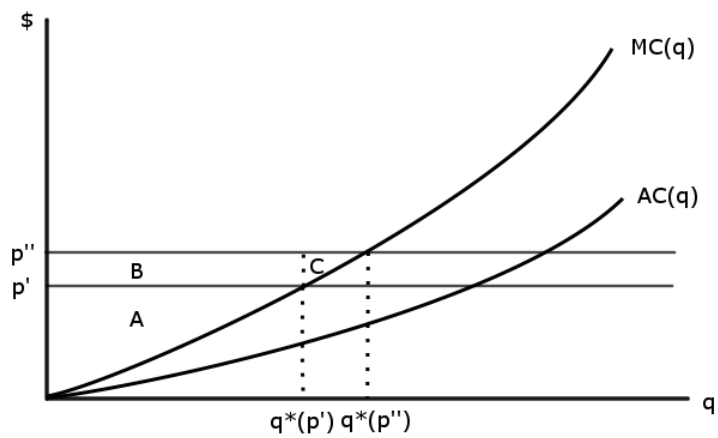

Figure 18: **Convexity of Profit Functions** This figure shows how the profit function is convex in the price and that the derivative equals the current supply.

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

Partial Equilibrium: Positive Analysis

Simon Board*

This Version: November 28, 2009
First Version: December 1, 2008.

In this Chapter we consider consider the interaction between different agents and firms, and solve for equilibrium prices and quantities.

Section 1 introduces the idea of partial equilibrium. Section 2 looks at how we aggregate agent's demand curves and firm's supply curves to form market demand and market demand supply. Section 3 defines an equilibrium, and discusses basic properties thereof. Section 4 looks at short-run equilibrium, where entry and exit are not possible. Finally, Section 5 considers long-run equilibrium, where entry and exit are possible.

1 Partial Equilibrium

When studying partial equilibrium, we consider the equilibrium in one market, taking as exogenous prices in other markets and agents' incomes, as well as preferences and technology. The main advantage of this model is simplicity: the equilibrium price is found by equating supply and demand. The model can also be used for welfare analysis, evaluating the effect of tax changes or the introduction of tariffs. However, the assumption that we can analyse one market independently of others can be dubious in some cases.

As an alternative, economists sometimes study general equilibrium. In this model, we fix preferences and technology and suppose agents are endowed with goods and shares. We then

*Department of Economics, UCLA. http://www.econ.ucla.edu/sboard/. Please email suggestions and typos to sboard@econ.ucla.edu.

solve for all prices simultaneously, equating supply and demand in each market. While this approach is far more general (hence the name), it is harder to analyse.

To illustrate the difference between partial and general equilibrium consider the worldwide market for cars.[1] A partial equilibrium analysis would add up the world's demand for cars to form a market demand curve. It would also add up the different firms' supply curves to form a market supply curve. The price of cars can then be found where demand equals supply. Intuitively, if the price were lower the would be excess demand and the price would be bid up; if the price were higher there would be excess supply and competition among suppliers would drive the price down. An exogenous increase in demand from China, due to Government construction of highways, would then shift up the demand curve, raising equilibrium price and quantity. In a general equilibrium model, there would be many other effects. First, the value of car firms would rise, increasing the income of their shareholders. Second, the increased demand for cars would push up the price of complements, such as oil. Third, there would be an increase in demand for inputs, such as steel, so commodity prices would rise. Ultimately, there is no single correct model: rather, there is a tradeoff between complexity and realism which depends on the markets at hand and the questions one is interested in.

1.1 Partial Equilibrium Model

We make the following assumptions:

1. We are interested in market 1. The price of this good, p_1, is to be determined.[2] We assume that all consumers and firms face the same prices (the law of one price), and that they are all price takers.
2. On the demand side, there are J agents who desire good 1. Each agent j has income m^j and utility $u^j(x_1, \dots, x_N)$. The consumer spends her income on N outputs which have prices $\{p_1, p_2 \dots, p_N\}$, where $\{p_2 \dots, p_N\}$ are exogenous. For simplicity, we often take $N = 2$.
3. On the supply side, there are K firms who sell good 1. Each firm k has a production technology $f^k(z_1, \dots, z_M)$, where the M input prices are given exogenously by $\{r_1 \dots, r_M\}$. For simplicity, we often take $M = 2$.

[1] We assume that there is one type of car. This assumption is highly stylised but, recall, the purpose of a model is not to describe the real world exactly, but to make useful abstractions.

[2] For simplicity, we will sometimes denote the price of this good by p.

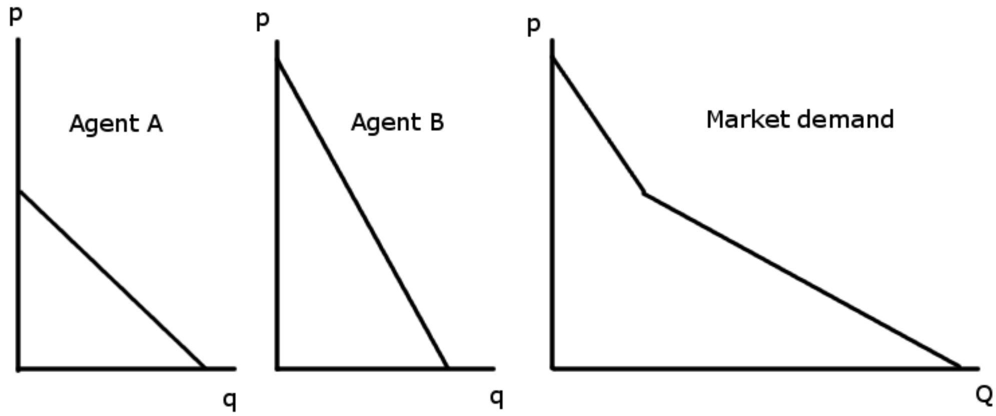

Figure 1: **Summing Demand.** Market demand is the horizontal sum of individual demands.

2 Market Demand and Supply

2.1 Market Demand

Solving the consumer's utility maximisation problem, we can derive agent j's Marshallian demand for good 1,

$$x_1^j(p_1, \ldots, p_N; m^j)$$

We can then form market demand by summing the individual agent's demands:

$$X_1(p_1, \ldots, p_N; m^1, \ldots, m^J) = \sum_{j=1}^{I} x_1^j(p_1, \ldots, p_N; m^j)$$

Market demand depends on the tastes of the agents, the prices of goods, the number of agents and the distribution of income in the economy. This means that if we redistribute money from every agent to one special agent (think Bill Gates), then this will change the demand for high-value items like yachts.

Figure 1 shows that when we sum demand, we are effectively adding the demand curves horizontally.

2.2 Market Supply

Solving the firm's profit maximisation problem, we can derive firm k's supply of good 1,

$$q^k(p_1, r_1 \ldots, r_M)$$

We can then form market supply by summing the individual firm's supply functions:

$$Q(p_1, r_1 \ldots, r_M) = \sum_{k=1}^{I} q^k(p_1, r_1 \ldots, r_M)$$

Note that market demand depends on the technologies of the firms, the price of good 1, the prices of the inputs and the number of firms. This means that if the number of firms increases then the market supply curve will shift out; similarly, if the number of firms decreases then the supply curve will shift in. We will see examples of this below.

3 Equilibrium

The equilibrium price of good 1, p_1^*, is found by equating supply and demand,

$$X_1(p_1^*, p_2; m^1, \ldots, m^I) = Q(p_1^*, r_1, r_2) \tag{3.1}$$

where, for simplicity, we assume there are 2 output goods and 2 input goods. Figure 2 shows the classic demand and supply picture.

The idea behind equation (3.1) is that if there is excess demand then consumers will bid the market price up, and if there is excess supply firms will compete for customers and reduce the market price.[3]

We now consider two general questions concerning the equilibrium price. First, does there exist a price p_1^* that satisfies (3.1)? Second, is there one one price p_1^* that satisfies (3.1) or many?

[3] This process is called tatonnement. One problem with this intuition is that it contradicts our assumption of price taking: how can firms undercut each other when they cannot affect the price? This suggests that the correct interpretation of price taking is not that firms cannot offer a different price, but that they have no incentive to do so.

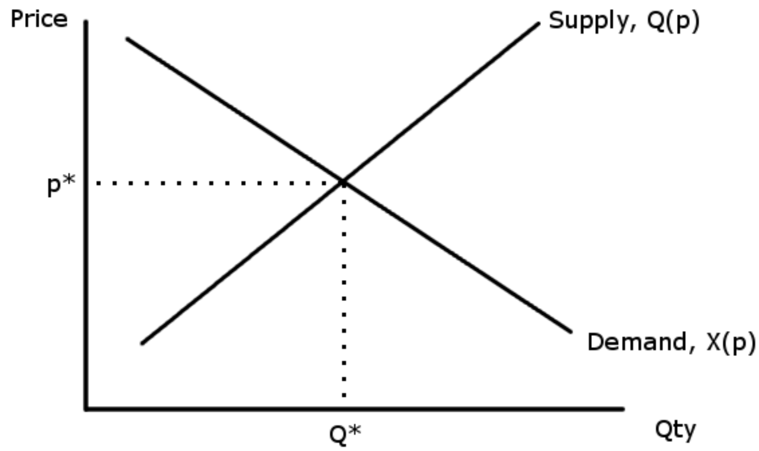

Figure 2: **Equilibrium.** In equilibrium, supply equals demand.

3.1 Existence

Does a price, p_1^*, exist that equates supply and demand? We can ensure existence if three conditions hold, all of which are satisfied in Figure 2.

1. For high prices, supply exceeds demand. This is trivially satisfied since demand must fall to zero as the price rises to infinity.
2. For low prices, demand exceeds supply. This is easy to satisfy since one would expect supply to fall to zero as prices converge to zero.
3. The supply and demand functions are continuous in p_1.

Figure 3 shows an example with one firm and one consumer where an equilibrium does not exist. The problem in this example is that at price p' the consumer wants 2 units, while the firm wishes to supply either 0 or 3 units. This problem is caused by the fixed cost: if the firm's production function is concave then its supply curve will be continuous. However, even if the individual firms' supply functions contain jumps, this is not a problem if the market is sufficiently large. In the example, if there are 3 identical agents then we can have two firms producing $q = 3$. The total demand then equals $2 + 2 + 2 = 6$, while the total supply equals

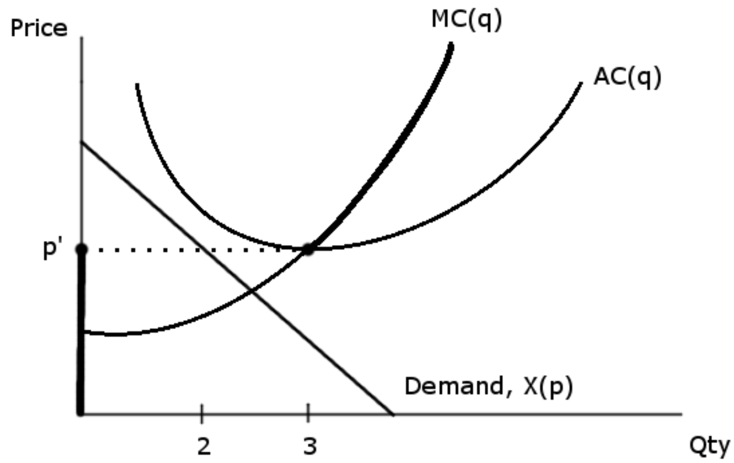

Figure 3: **Nonexistence of Equilibrium.** Due to a fixed cost, supply jumps at price p'. As a result, no equilibrium exists.

$3 + 3 = 6$, and we have an equilibrium.[4]

3.2 Uniqueness

In Figure 2 there is a unique equilibrium price. Can there be more? From the law of supply we know that the supply curve is upward sloping. If the good is ordinary for each agent then the market demand is downward sloping and there can at most be one equilibrium. In contrast, if the good is a Giffen good, then there may be multiple equilibria, as shown in Figure 4.

4 Short-Run Equilibrium

In Section 3 we discussed equilibrium in general. It is useful to differentiate between the following time periods:[5]

[4] Without fixed costs, equilibrium may also fail to exist if the marginal cost curve is first increasing, then decreasing, and increasing again. This causes a discontinuity in the supply function, as shown in the Production Chapter. Again, this is not a problem if the market is sufficiently large.

[5] In the Production Chapter, we considered a fourth case when some factors are flexible, but others are not. We omit this here for simplicity.

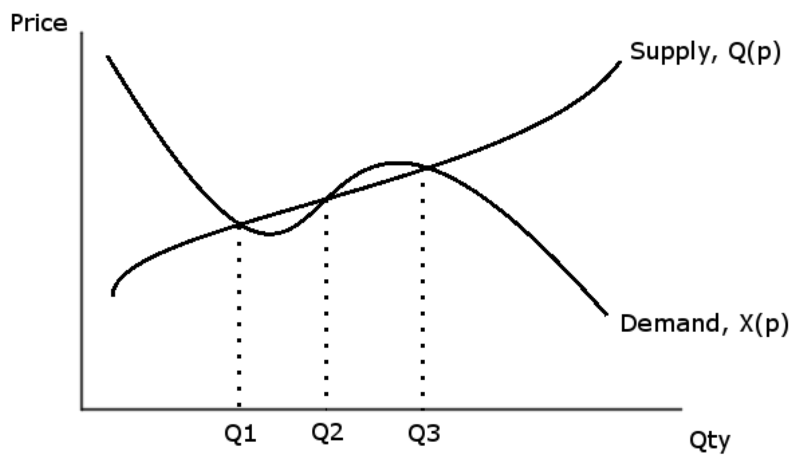

Figure 4: **Multiple Equilibria.** In this figure, the demand curve has an upward sloping component. As a result, there are three equilibria with quantity Q_1, Q_2 and Q_3 being sold.

1. In the **very short run** all the factors of production are fixed, and output is fixed.
2. In the **short run** all factors are flexible, but fixed costs are sunk. Firms cannot enter or exit.
3. In the **long run** all factors are flexible and fixed costs are not sunk. Hence firms can enter and exit freely.

In the very short-run, the analysis is straightforward since output is fixed. We solve a numerical example in Section 5.3.

In this Section we analyse short-run equilibrium, assuming the number of firms is fixed. In this case, each firm operates on its individual supply curve, and market supply equals the sum of individual firm supply.

4.1 Example

Suppose there are 15 agents, ten of whom have income $m_j = 10$ and five of whom have $m_j = 20$. Each agent has symmetric Cobb-Douglas utility,

$$u^j(x_1, x_2) = x_1 x_2$$

Solving the agent's utility maximisation problem, the demand function of each agent for good 1 is

$$x_1^j = \frac{m_j}{2p_1}$$

Summing over demand curves, market demand is given by

$$X_1 = \frac{200}{2p_1} = \frac{100}{p_1}$$

Next, suppose there are 9 firms with production function

$$f^k(z_1, z_2) = (z_1 - 1)^{1/3}(z_2 - 1)^{1/3}$$

Solving the firm's cost maximisation problem, its cost function is

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

Solving the profit-maximisation problem, its supply function is given by

$$q^k = \frac{1}{9}\frac{p_1^2}{r_1 r_2}$$

Summing over firms, market supply is thus

$$Q = \frac{p_1^2}{r_1 r_2}$$

Equating supply and demand,

$$p_1^* = (100 r_1 r_2)^{1/3}$$

4.2 Shifts in Supply and Demand

We are often interested in how shifts in supply or demand affects prices and quantities. A shift in demand may be due to changes in income (e.g. a recession), changes in the price of substitutes or complements, changes in the quality of the goods, and so forth. Similarly, a shift in the supply curve may be due to changes in input prices, the number of producers or changes in technology.

Figure 5 shows the effect of an increase in demand. The left-hand side shows the market from the firm's perspective, while the right-hand side shows the entire market. The increase in demand causes the equilibrium price to rise and each firm to move up its supply function. As

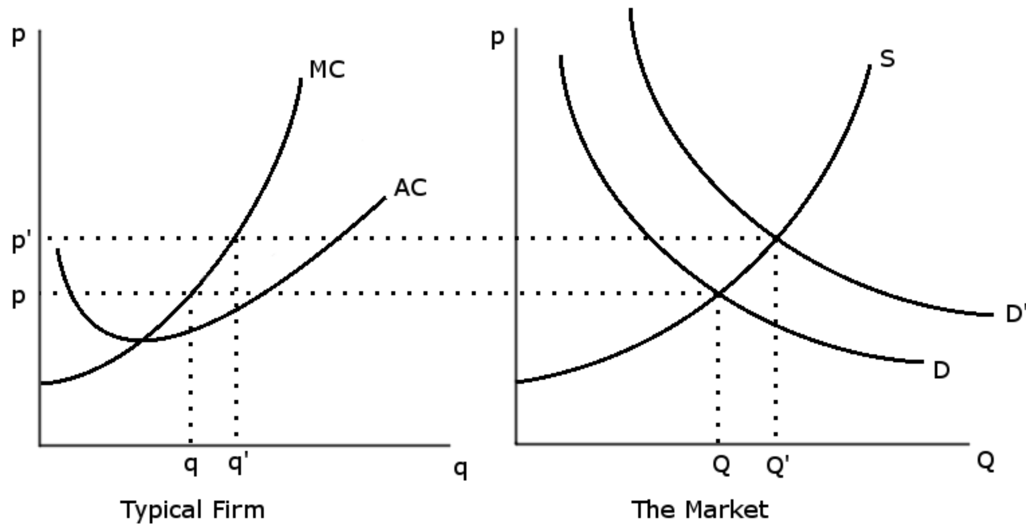

Figure 5: **Increase in Demand.** This figure shows the effect on an increase in demand from the firm's perspective (left), and the market's perspective (right). The original demand is D, with price p and quantities q and Q. When demand shifts up to D', the price rises to p' and quantities rise to q' and Q'. Note that, in this picture, entry and exit are impossible so the firm's supply curve coincides with its marginal cost curve.

a result, each firm's profit increases.

Figure 6 shows the effect of a shift in demand for different supply functions. When supply is inelastic there is a small change in quantity and a large change in price. Conversely, when supply is elastic, there is a large change in quantity and a small change in price.

Exercise: How does a shift in the supply curve affect price and quantity when the demand curve is elastic/inelastic?

5 Long-Run Equilibrium

In the long run, we assume there are many identical potential entrants. This means that entry will occur if there exists a quantity q such that $p > AC(q)$. Conversely, exit will occur if $p < AC(q)$. As a result price must be driven down to the minimum average cost

$$p = \min_{q \geq 0} AC(q).$$

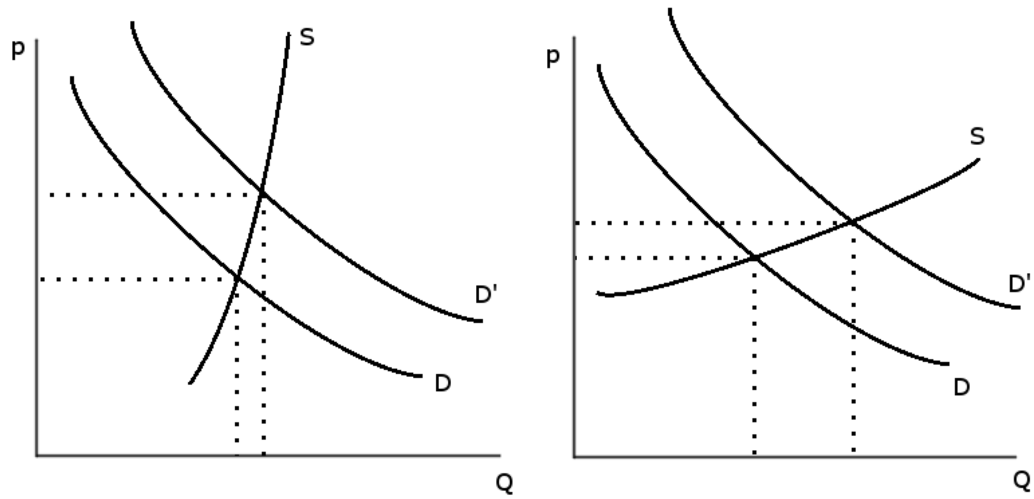

Figure 6: **Shifts in Demand.** This picture shows the effect of an increase in demand when supply is inelastic (left) or elastic (right). When supply is inelastic, the price rises substantially, but there is little extra output. When supply is inelastic, the price rises a little, and there is a lot of extra output.

We can thus think of the **long-run supply function** as a horizontal line equal to the minimum average cost. Note: this does not contradict the first-order condition of each firm since $AC(q) = MC(q)$ when $AC(q)$ is minimised. Hence we have $p = AC(q) = MC(q)$. See Figure 7.

In this class, we adopt a simple model of "the long-run", where there are an infinite number of identical firms. One could imagine an alternative model, where news firms are not as efficient as current firms, implying that the long-run supply function is upward sloping.[6] As an extension, one might suppose there is a difference between demand in the short-run and the long-run. For example, consider the market for oil. In the short-run, demand is very inelastic as consumers have fixed commutes. In the long-run consumers can change their cars, organise car-pools and even change jobs, in reaction to a change in the price of oil, making demand much more inelastic.

5.1 Example

Here we continue the example from Section 4.1. The algebra can get a little involved, so I urge you to work with example out, on your own, with $r_1 = r_2 = 1$.

[6]This would also occur if new firms bid up the price of common inputs.

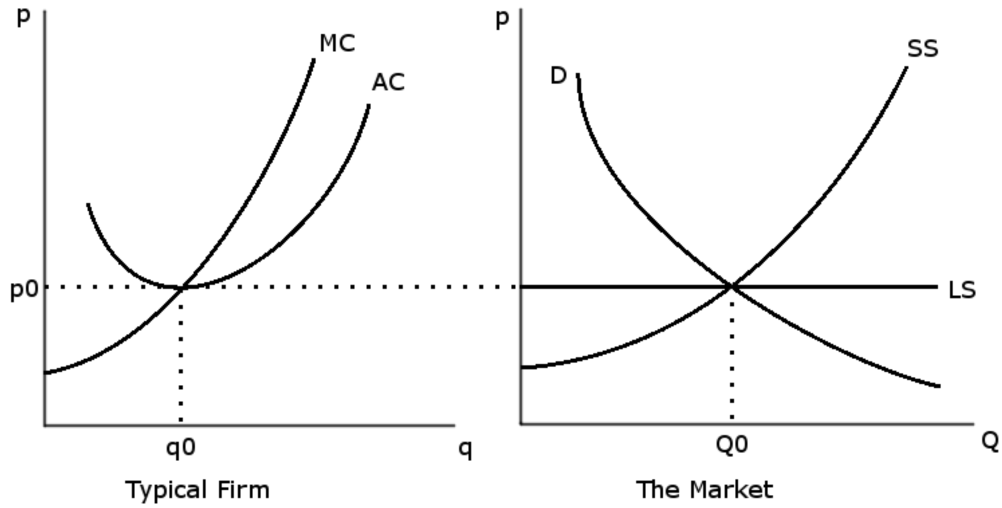

Figure 7: **Long-Run Equilibrium.** In the long-run equilibrium, each firm produces at the minimum average-cost, and the firms enter until aggregate supply equals aggregate demand. In this figure, LS is the long-run supply curve, while SS is the short-run supply corresponding to output Q_0.

The firm's cost function is

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

Dividing by q, average cost is

$$AC(q) = c(q)/q = 2(r_1 r_2)^{1/2} q^{1/2} + (r_1 + r_2) q^{-1}$$

As argued above, average cost is minimised in the long-run. Differentiating,

$$\frac{dAC(q)}{dq} = (r_1 r_2)^{1/2} q^{-1/2} - (r_1 + r_2) q^{-2}$$

Setting this equal to zero and rearranging, the AC-minimising quantity is

$$q^* = (r_1 r_2)^{-1/3} (r_1 + r_2)^{2/3}$$

The long-run supply curve is a horizontal line at the minimum average cost. Substituting q^* into the average cost, the price in the long-run is

$$p^* = AC(q^*) = 3(r_1 r_2)^{1/3} (r_1 + r_2)^{1/3}$$

We now switch to the market level. Since we know the price level, we can calculate the market demand. Using the demand function from Section 4.1,

$$X(p^*) = 100/p^* = \frac{100}{3}(r_1 r_2)^{-1/3}(r_1 + r_2)^{-1/3}$$

In equilibrium, aggregate supply equals aggregate demand. Since each firm produces q^*, we know the number of firms is

$$K^* = X(p^*)/q^* = \frac{100}{3}(r_1 + r_2)^{-1}$$

Note that K^* may not be an integer, but we won't worry about this.

5.2 Shifts in Demand

Suppose we start in long-run equilibrium, as shown in figure 7. Suppose there in an increase in demand (e.g. due to government expenditure).

In the very short run, quantity is constant, and the market price rises a lot. This is shown in figure 8.

In the short run, firms increase their output in response to the increase in demand. Each firm moves up its individual supply curve (i.e. their marginal cost curve) and the market moves up its aggregate supply curve. As a result, firms now make positive profits, as shown in figure 9.

In the long-run the positive industry profits attracts entrants. This causes the industry short-run supply curve to shift out, and the price level to move back to the minimum average cost. The new aggregate quantity is given by the intersection of the long-run supply function and the demand function. Each individual firm returns to operating at minimum average cost, and making zero profits. See figure 10.

5.3 Example

Suppose market demand is linear:

$$X(p) = 1500 - 50p.$$

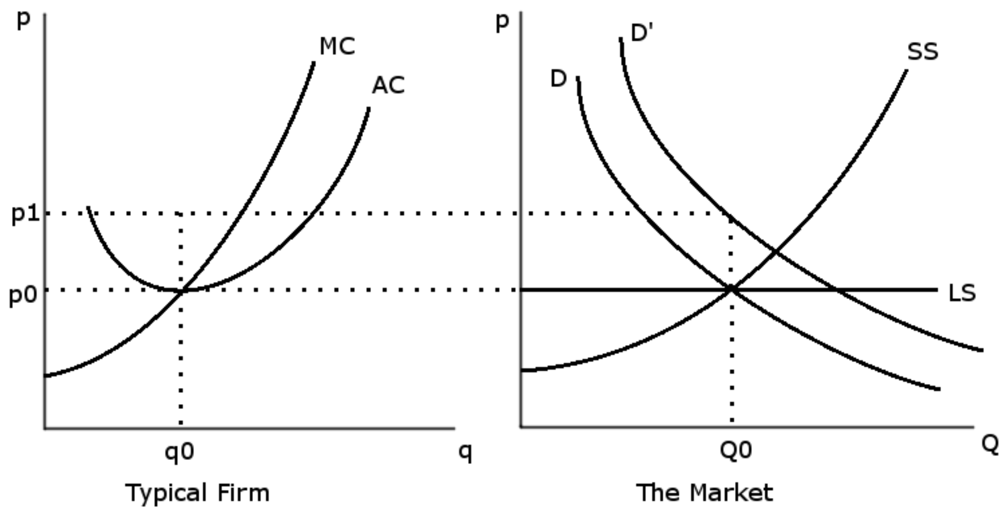

Figure 8: **Very-Short Run.** In the very-short run, each firm produces the same quantity q_0 and no firms enter, so market quantity is given by Q_0. The price rises to p_1 to clear the market.

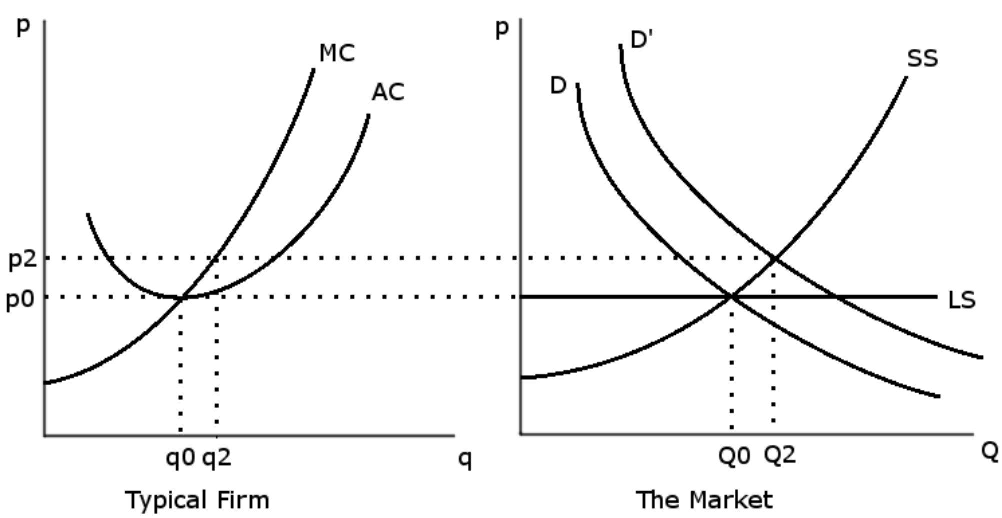

Figure 9: **Short Run.** In the short run, the individual firms increase their quantity to q_2. As a result, the market moves up its short-run supply curve and the equilibrium is given by p_2.

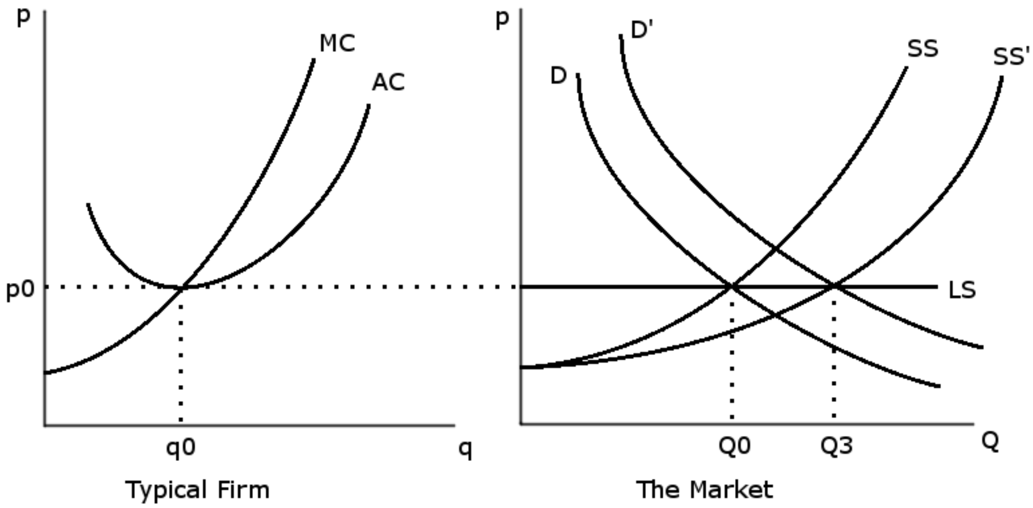

Figure 10: **Long Run.** In the long run, new firms enter, the industry reverts to the long run supply function and the short run supply function shifts to SS'. At this point, each firm again produces at long-run minimum average cost, q_0.

There are many potential firms, each with cost function

$$c(q) = 100 + q^2/4.$$

Let's calculate the initial long-run equilibrium. The average cost is

$$AC(q) = c(q)/q = 100q^{-1} + q/4$$

In the long-run equilibrium, firms operate at minimum average cost. Differentiating,

$$\frac{dAC(q)}{dq} = -100q^{-2} + 1/4 = 0$$

Rearranging, each firm produces

$$q^* = 20$$

Substituting into the average cost, the market price is then

$$p^* = AC(q^*) = 10$$

Market demand is

$$X(p^*) = 1500 - 50p^* = 1000$$

which equals market supply. Hence the number of firms is

$$K^* = X(p^*)/q^* = 1000/20 = 50.$$

Suppose demand falls (e.g. the price of a complement rises). In particular, new demand is given by

$$X(p) = 1200 - 50p.$$

In the **very short run**, each firm's output is fixed. Hence market output is fixed at 1000. Equating supply and demand,

$$1200 - 50p = 1000.$$

Hence the equilibrium price is $p^* = 4$.

In the **short run**, each firm reduces its output due to the reduction in demand. Each firm maximises profits:

$$\pi(q) = pq - c(q) = pq - 100 - q^2/4$$

Taking the first-order condition, the optimal output is $q^*(p) = 2p$. This gives this individual firm's supply curve. There are 50 firms, so the market supply curve is

$$Q(p) = 100p$$

Equating supply and demand,

$$1200 - 50p = 100p$$

Rearranging, the short-run equilibrium price is $p^* = 8$. Each firm produces $q = 2p^* = 16$, and makes profits

$$\pi = pq - c(q) = 8 \times 16 - 100 - 16^2/4 = -36.$$

In the **long run**, the negative profits cause firms to exit the industry. In the long-run, the price returns to $p^* = 10$, and each active firm produces $q^* = 20$. Given the price, market demand is

$$X(p^*) = 1200 - 50p^* = 700$$

which also equals market supply. Thus the number of firms is given by

$$K^* = X(p^*)/q^* = 700/20 = 35$$

which is less than the 50 firms we started with.