Chapter 3

The Traditional Approach to Consumer Theory

In the previous section, we considered consumer behavior from a choice-based point of view. That is, we assumed that consumers made choices about which consumption bundle to choose from a set of feasible alternatives, and, using some rather mild restrictions on choices (homogeneity of degree zero, Walras' law, and WARP), were able make predictions about consumer behavior. Notice that our predictions were entirely based on consumer behavior. In particular, we never said anything about why consumers behave the way they do. We only hold that the way they behave should be consistent in certain ways.

The traditional approach to consumer behavior is to assume that the consumer has well-defined preferences over all of the alternative bundles and that the consumer attempts to select the most preferred bundle from among those bundles that are available. The nice thing about this approach is that it allows us to build into our model of consumer behavior how the consumer feels about trading off one commodity against another. Because of this, we are able to make more precise predictions about behavior. However, at some point people started to wonder whether the predictions derived from the preference-based model were in keeping with the idea that consumers make consistent choices, or whether there could be consistent choice-based behavior that was not derived from the maximization of well-defined preferences. It turns out that if we define consistent choice making as homogeneity of degree zero, Walras' law, and WARP, then there are consistent choices that cannot be derived from the preference-based model. But, if we replace WARP with a slightly stronger but still reasonable condition, called the Strong Axiom of Revealed Preference (SARP), then any behavior consistent with these principles can be derived from the maximization of rational preferences.

Next, we take up the traditional approach to consumer theory, often called "neoclassical" consumer theory.

3.1 Basics of Preference Relations

We'll continue to assume that the consumer chooses from among L commodities and that the commodity space is given by $X \subset R_+^L$. The basic idea of the preference approach is that given any two bundles, we can say whether the first is "at least as good as" the second. The "at-least-as-good-as" relation is denoted by the curvy greater-than-or-equal-to sign: $\succeq$. So, if we write $x \succeq y$, that means that "x is at least as good as y."

Using $\succeq$, we can also derive some other preference relations. For example, if $x \succeq y$, we could also write $y \preceq x$, where $\preceq$ is the "no better than" relation. If $x \succeq y$ and $y \succeq x$, we say that a consumer is "indifferent between x and y," or symbolically, that $x \sim y$. The indifference relation is important in economics, since frequently we will be concerned with **indifference sets**. The indifference curve I_y is defined as the set of all bundles that are indifferent to y. That is, $I_y = \{x \in X | y \sim x\}$. Indifference sets will be very important as we move forward, and we will spend a great deal of time and effort trying to figure out what they look like, since the indifference sets capture the trade-offs the consumer is willing to make among the various commodities. The final preference relation we will use is the "strictly better than" relation. If x is at least as good as y and y is not at least as good as x, i.e., $x \succeq y$ and not $y \succeq x$ (which we could write $y \not\succeq x$), we say that $x \succ y$, or x is strictly better than (or strictly preferred to) y.

Our preference relations are all examples of mathematical objects called binary relations. A binary relation compares two objects, in this case, two bundles. For instance, another binary relation is "less-than-or-equal-to," $\leq$. There are all sorts of properties that binary relations can have. The first two we will be interested in are called **completeness** and **transitivity.** A binary relation is complete if, for any two elements x and y in X, either $x \succeq y$ or $y \succeq x$. That is, any two elements can be compared. A binary relation is transitive if $x \succeq y$ and $y \succeq z$ imply $x \succeq z$. That is, if x is at least as good as y, and y is at least as good as z, then x must be at least as good as z.

The requirements of completeness and transitivity seem like basic properties that we would like any person's preferences to obey. This is true. In fact, they are so basic that they form economists' very definition of what it means to be rational. That is, a preference relation $\succeq$ is called **rational** if it is complete and transitive.

When we talked about the choice-based approach, we said that there was implicit in the idea that demand functions satisfy Walras Law the idea that "more is better." This idea is formalized in terms of preferences by making assumptions about preferences over one bundle or another. Consider the following property, called monotonicity:

Definition 5 *A preference relation $\succeq$ is **monotone** if $x \succ y$ for any x and y such that $x_l > y_l$ for $l = 1, ..., L$. It is **strongly monotone** if $x_l \geq y_l$ for all $l = 1, ..., L$ and $x_j > y_j$ for some $j \in \{1, ..., L\}$ implies that $x \succ y$.*

Monotonicity and strong monotonicity capture two different notions of "more is better." Monotonicity says that if every component of x is larger than the corresponding component of y, then x is strictly preferred to y. Strong monotonicity is the requirement that if every component of x is at least as large (but not necessarily strictly larger) than the corresponding component of y and at least one component of x is strictly larger, x is strictly preferred to y.

The difference between monotonicity and strong monotonicity is illustrated by the following example. Consider the bundles $x = (1, 1)$ and $y = (1, 2)$. If $\succeq$ is strongly monotone, then we can say that $y \succ x$. However, if $\succeq$ is monotone but not strongly monotone, then it need not be the case that y is strictly preferred to x. Since preference relations that are strongly monotone are monotone, but preferences that are monotone are not necessarily strongly monotone, strong monotonicity is a more restrictive (a.k.a. "stronger") assumption on preferences.

If preferences are monotone or strongly monotone, it follows immediately that a consumer will choose a bundle on the boundary of the Walrasian budget set. Hence an assumption of some sort of monotonicity must have been in the background when we assumed consumer choices obeyed Walras' Law. However, choice behavior would satisfy Walras' Law even if preferences satisfied the following weaker condition, called local nonsatiation.

Condition 6 *A preference relation $\succeq$ satisfies **local nonsatiation** if for every x and every $\varepsilon > 0$ there is a point y such that $||x - y|| \leq \varepsilon$ and $y \succ x$.*

That is, for every x, there is always a point "nearby" that the consumer strictly prefers to x, and this is true no matter how small you make the definition of "nearby." Local nonsatiation allows for the fact that some commodities may be "bads" in the sense that the consumer would sometimes prefer less of them (like garbage or noise). However, it is not possible for all goods to always be bads if preferences are non-satiated. (Why?)

It's time for a brief discussion about the practice of economic theory. Recall that the object of doing economic theory is to derive testable implications about how real people will behave. But, as we noted earlier, in order to derive testable implications, it is necessary to impose some restrictions on (make assumptions about) the type of behavior we allow. For example, suppose we are interested in the way people react to wealth changes. We could simply assume that people's behavior satisfies Walras' Law, as we did earlier. This allows us to derive testable implications. However, it provides little insight into why they satisfy Walras' Law. Another option would be to assume monotonicity – that people prefer more to less. Monotonicity implies that people will satisfy Walras' Law. But, it rules out certain types of behavior. In particular, it rules out the situation where people prefer less of an object to more of it. But, introspection tells us that sometimes we do prefer less of something. So, we ask ourselves if there is a weaker assumption that allows people to prefer less to more, at least sometimes, that still implies Walras' Law. It turns out that local nonsatiation is just such an assumption. It allows for people to prefer less to more – even to prefer less of everything – the only requirement is that, no matter which bundle the consumer currently selects, there is always a feasible bundle nearby that she would rather have.

By selecting the weakest assumption that leads to a particular result, we accomplish two tasks. First, the weaker the assumptions used to derive a result, the more "robust" it is, in the sense that a greater variety of initial conditions all lead to the same conclusion. Second, finding the weakest possible condition that leads to a particular conclusion isolates just what is needed to bring about the conclusion. So, all that is really needed for consumers to satisfy Walras' Law is for preferences to be locally nonsatiated – but not necessarily monotonic or strongly monotonic.

The assumptions of monotonicity or local nonsatiation will have important implications for the way indifference sets look. In particular, they ensure that $I_x = \{y \in X | y \sim x\}$ are downward sloping and "thin." That is, they must look like Figure 3.1.

If the indifference curves were thick, as in Figure 3.2, then there would be points such as x, where in a neighborhood of x (the dotted circle) all points are indifferent to x. Since there is no strictly preferred point in this region, it is a violation of local-nonsatiation (or monotonicity).

In addition to the indifference set I_x defined earlier, we can also define upper-level sets and lower-level sets. The **upper level set of x** is the set of all points that are at least as good as x, $U_x = \{y \in X | y \succeq x\}$. Similarly, the **lower level set of x** is the set of all points that are no

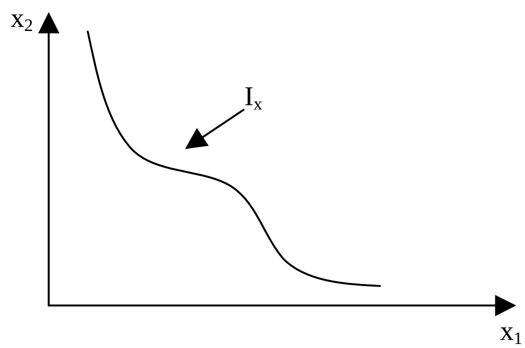

Figure 3.1: Thin Indifferent Sets

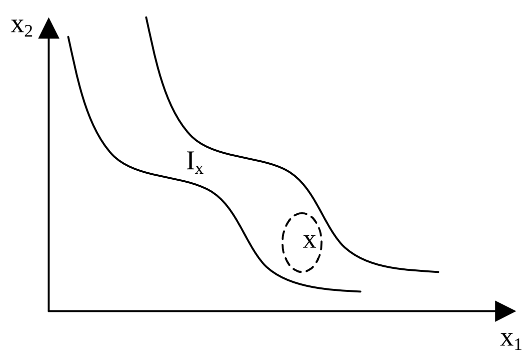

Figure 3.2: Thick Indifference Sets

better than x, $L_x = \{y \in X | x \succeq y\}$. Just as monotonicity told us something about the shape of indifference sets, we can also make assumptions that tell us about the shape of upper and lower level sets.

Recall that a set of points, X, is **convex** if for any two points in the set the (straight) line segment between them is also in the set.[1] Formally, a set X is convex if for any points x and x' in X, every point z on the line joining them, $z = tx + (1-t)\,x'$ for some $t \in [0,1]$, is also in X. Basically, a convex set is a set of points with no holes in it and with no "notches" in the boundary. You should draw some pictures to figure out what I mean by no holes and no notches in the set.

Before we move on, let's do a thought experiment. Consider two possible commodity bundles, x and x'. Relative to the extreme bundles x and x', how do you think a typical consumer feels about an average bundle, $z = tx + (1-t)\,x'$, $t \in (0,1)$? Although not always true, in general, people tend to prefer bundles with medium amounts of many goods to bundles with a lot of some things and very little of others. Since real people tend to behave this way, and we are interested in modeling how real people behave, we often want to impose this idea on our model of preferences.[2]

Exercise 7 *Confirm the following two statements: 1) If $\succeq$ is convex, then if $y \succeq x$ and $z \succeq x$, $ty + (1-t)\,z \succeq x$ as well. (2) Suppose $x \sim y$. If $\succeq$ is convex, then for any $z = ty + (1-t)\,x$, $z \succeq x$.*

Another way to interpret convexity of preferences is in terms of a diminishing **marginal rate of substitution (MRS),** which is simply the slope of the indifference curve. The idea here is that if you are currently consuming a bundle x, and I offer to take some x_1 away from you and replace it with some x_2, I will have to give you a certain amount of x_2 to make you exactly indifferent for the loss of x_1. A diminishing MRS means that this amount of x_2 I have to give you increases the more x_2 that you already are consuming - additional units of x_2 aren't as valuable to you.

The upshot of the convexity and local non-satiation assumptions is that indifference sets have to be thin, downward sloping, and be "bowed upward." There is nothing in the definition of convexity

[1] This is the definition of a **convex set.** It should not be confused with a **convex function**, which is a different thing altogether. In addition, there is such thing as a **concave function.** But, there is **no such thing as a concave set.** I sympathize with the fact that this terminology can be confusing. But, that's just the way it is. My advice is to focus on the meaning of the concepts, i.e., "a set with no notches and no holes."

[2] It is only partly true that when we assume preferences are convex we do so in order to capture real behavior. In addition, the basic mathematical techniques we use to solve our problems often depend on preferences being convex. If they are not (and one can readily think of examples where preferences are not convex), other, more complicated techniques have to be used.

that prevents flat regions from appearing on indifference curves. However, there are reasons why we want to rule out indifference curves with flat regions. Because of this, we strengthen the convexity assumption with the concept of **strict convexity**. A preference relation is strictly convex if for any distinct bundles y and z $(y \neq z)$ such that $y \succeq x$ and $z \succeq x$, $ty + (1-t)\,z \succ x$. Thus imposing strict convexity on preferences strengthens the requirement of convexity (which actually means that averages are at least as good as extremes) to say that averages are strictly better than extremes.

3.2 From Preferences to Utility

In the last section, we said a lot about preferences. Unfortunately, all of that stuff is not very useful in analyzing consumer behavior, unless you want to do it one bundle at a time. However, if we could somehow describe preferences using mathematical formulas, we could use math techniques to analyze preferences, and, by extension, consumer behavior. The tool we will use to do this is called a **utility function.**

A utility function is a function $U(x)$ that assigns a number to every consumption bundle $x \in X$. Utility function $U()$ **represents** preference relation $\succeq$ if for any x and y, $U(x) \geq U(y)$ if and only if $x \succeq y$. That is, function U assigns a number to x that is at least as large as the number it assigns to y if and only if x is at least as good as y. The nice thing about utility functions is that if you know the utility function that represents a consumer's preferences, you can analyze these preferences by deriving properties of the utility function. And, since math is basically designed to derive properties of functions, it can help us say a lot about preferences.

Consider a typical indifference curve map, and assume that preferences are rational. We also need to make a technical assumption, that preferences are continuous. For our purposes, it isn't worth derailing things in order to explain why this is necessary. But, you should look at the example of lexicographic preferences in MWG to see why the assumption is necessary and what can go wrong if it is not satisfied.

The line drawn in Figure 3.3 is the line $x_2 = x_1$, but any straight line would do as well. Notice that we could identify the indifference curve I_x by the distance along the line $x_2 = x_1$ you have to travel before intersecting I_x. Since indifference curves are downward sloping, each I_x will only intersect this line once, so each indifference curve will have a unique number associated with it. Further, since preferences are convex, if $x \succ y$, I_x will lay above and to the right of I_y (i.e. inside I_y), and so I_x will have a higher number associated with it than I_y.

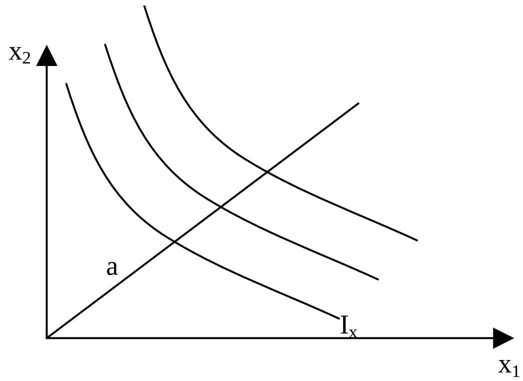

Figure 3.3: Ranking Indifference Curves

We will call the number associated with I_x the utility of x. Formally, we can define a function $u(x_1, x_2)$ such that $u(x_1, x_2)$ is the number associated with the indifference curve on which (x_1, x_2) lies. It turns out that in order to ensure that there is a utility function corresponding to a particular preference relation, you need to assume that preferences are rational and continuous. In fact, this is enough to guarantee that the utility function is a continuous function. The assumption that preferences are rational agrees with how we think consumers should behave, so it is no problem. The assumption that preferences are continuous is what we like to call a **technical assumption,** by which we mean that is that it is needed for the arguments to be mathematically rigorous (read: true), but it imposes no real restrictions on consumer behavior. Indeed, the problems associated with preferences that are not continuous arise only if we assume that all commodities are infinitely divisible (or come in infinite quantities). Since neither of these is true of real commodities, we do not really harm our model by assuming continuous preferences.

3.2.1 Utility is an Ordinal Concept

Notice that the numbers assigned to the indifference curves in defining the utility function were essentially arbitrary. Any assignment of numbers would do, as long as the order of the numbers assigned to various bundles is not disturbed. Thus if we were to multiply all of the numbers by 2, or add 6 to them, or take the square root, the numbers assigned to the indifference curves after the transformation would still represent the same preferences. Since the crucial characteristic of a utility function is the order of the numbers assigned to various bundles, but not the bundles themselves, we say that utility is an ordinal concept.

This has a number of important implications for demand analysis. The first is that if $U(x)$ represents $\succeq$ and $f()$ is a monotonically increasing function (meaning the function is always increasing as its argument increases), then $V(x) = f(U(x))$ also represents $\succeq$. This is very valuable for the following reason. Consider the common utility function $u(x) = x_1^a x_2^{1-a}$, which is called the Cobb-Douglas utility function. This function is difficult to analyze because x_1 and x_2 have different exponents and are multiplied together. But, consider the monotonically increasing function $f(z) = \log(z)$, where "log" refers to the natural logarithm.[3]

$$V(x) = \log[x_1^a x_2^{1-a}] = a \log x_1 + (1-a) \log x_2$$

$V()$ represents the same preferences as $U()$. However, $V()$ is a much easier function to deal with than $U()$. In this way we can exploit the ordinal nature of utility to make our lives much easier. In other words, there are many utility functions that can represent the same preferences. Thus it may be in our interest to look for one that is easy to analyze.

A second implication of the ordinal nature of utility is that the difference between the utility of two bundles doesn't mean anything. For example, if $U(x) - U(y) = 7$ and $U(z) - U(a) = 14$, it doesn't mean that going from consuming z to consuming a is twice as much of an improvement than going from x to y. This makes it hard to compare things such as the impact of two different tax programs by looking at changes in utility. Fortunately, however, we have developed a method for dealing with this, using compensated changes similar to those used in the derivation of the Slutsky matrix in the section on consumer choice.

3.2.2 Some Basic Properties of Utility Functions

If preferences are convex, then the indifference curves will be convex, as will the upper level sets. When a function's upper-level sets are always convex, we say that the function is (sorry about this) **quasiconcave**. The importance of quasiconcavity will become clear soon. But, I just want to drill into you that quasiconcave means convex upper level sets. Keep that in mind, and things will be much easier.

For example, consider a special case of the Cobb-Douglas utility function I mentioned earlier.

$$u(x_1, x_2) = x_1^{\frac{1}{4}} x_2^{\frac{1}{4}}.$$

Figure 3.4 shows a three-dimensional (3D) graph of this function.

[3]Despite what you are used to, economists always use log to refer to the natural log, ln, since we don't use base 10 logs at all.

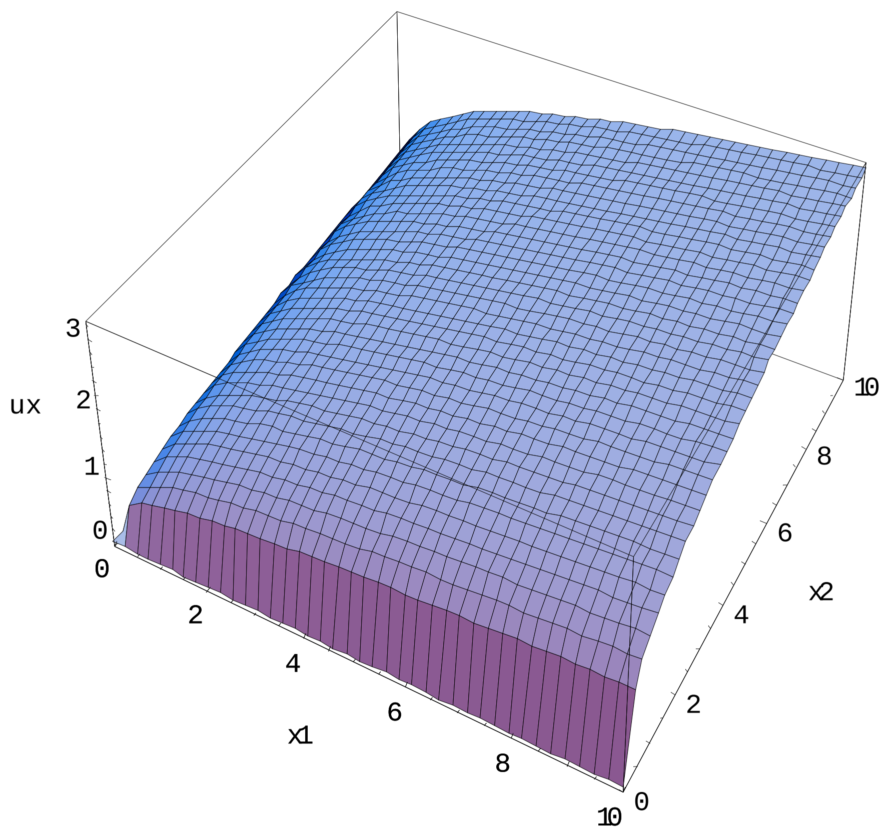

Figure 3.4: Function $u(x)$

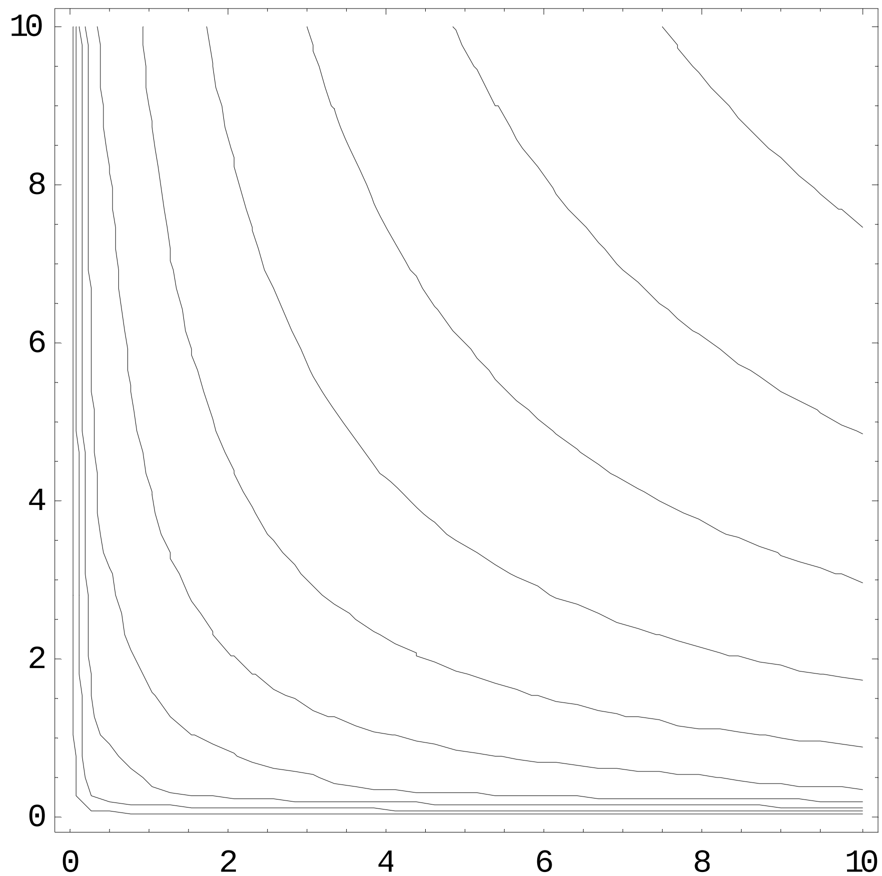

Figure 3.5: Level sets of $u(x)$

Notice the curvature of the surface. Now, consider Figure 3.5, which shows the level sets (I_x) for various utility levels. Notice that the indifference curves of this utility function are convex. Now, pick an indifference curve. Points offering more utility are located above and to the right of it. Notice how the contour map corresponds to the 3D utility map. As you move up and to the right, you move "uphill" on the 3D graph.

Quasiconcavity is a weaker condition than concavity. Concavity is an assumption about how the numbers assigned to indifference curves change as you move outward from the origin. It says that the increase in utility associated with an increase in the consumption bundle decreases as you move away from the origin. As such, it is a cardinal concept. Quasiconcavity is an ordinal concept. It talks only about the shape of indifference curves, not the numbers assigned to them. It can be shown that concavity implies quasiconcavity but a function can be quasiconcave without being concave (can you draw one in two dimensions). It turns out that for the results on utility

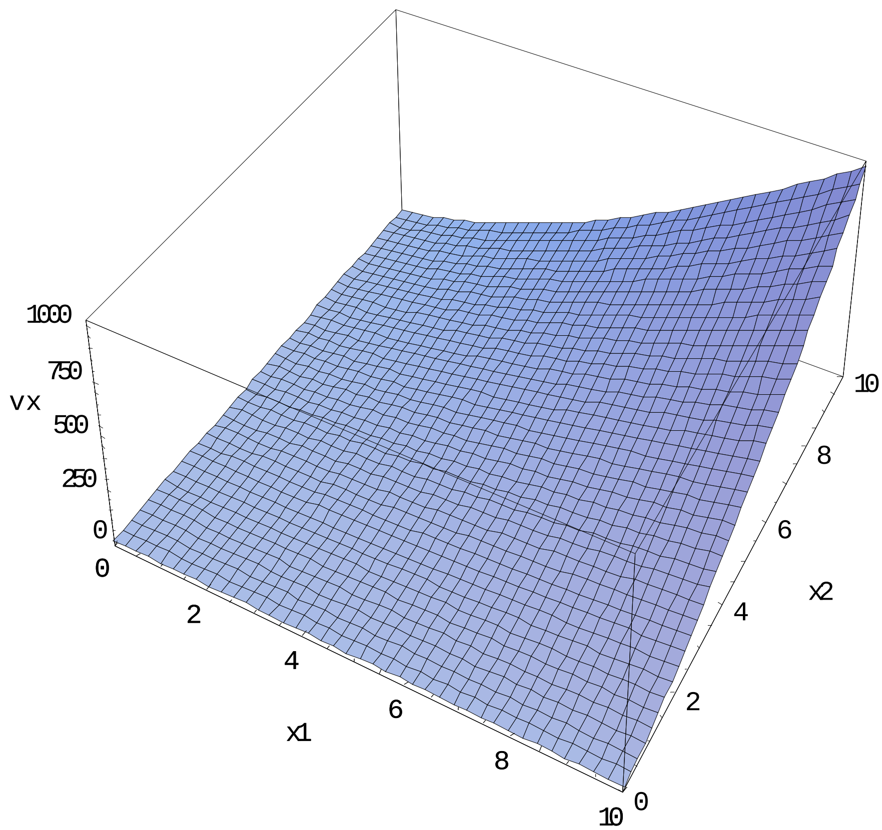

Figure 3.6: Function $v(x)$

maximization we will develop later, all we really need is quasiconcavity. Since concavity imposes cardinal restrictions on utility (which is an ordinal concept) and is stronger than we need for our maximization results, we stick with the weaker assumption of quasiconcavity.[4]

Here's an example to help illustrate this point. Consider the following function, which is also of the Cobb-Douglas form:

$$v(x) = x_1^{\frac{3}{2}} x_2^{\frac{3}{2}}.$$

Figure 3.6 shows the 3D graph for this function. Notice that $v(x)$ is "curved upward" instead of downward like $u(x)$. In fact, $v(x)$ is a not a concave function, while $u(x)$ is a concave function.[5] But, both are quasiconcave. We already saw that $u(x)$ was quasiconcave by looking at its level

[4] As in the case of convexity and strict convexity, a strictly quasiconcave function is one whose upper level sets are strictly convex. Thus a function that is quasiconcave but not strictly so can have flat parts on the boundaries of its indifference curves.

[5] See Simon and Blume or Chiang for good explanations of concavity and convexity in multiple dimensions.

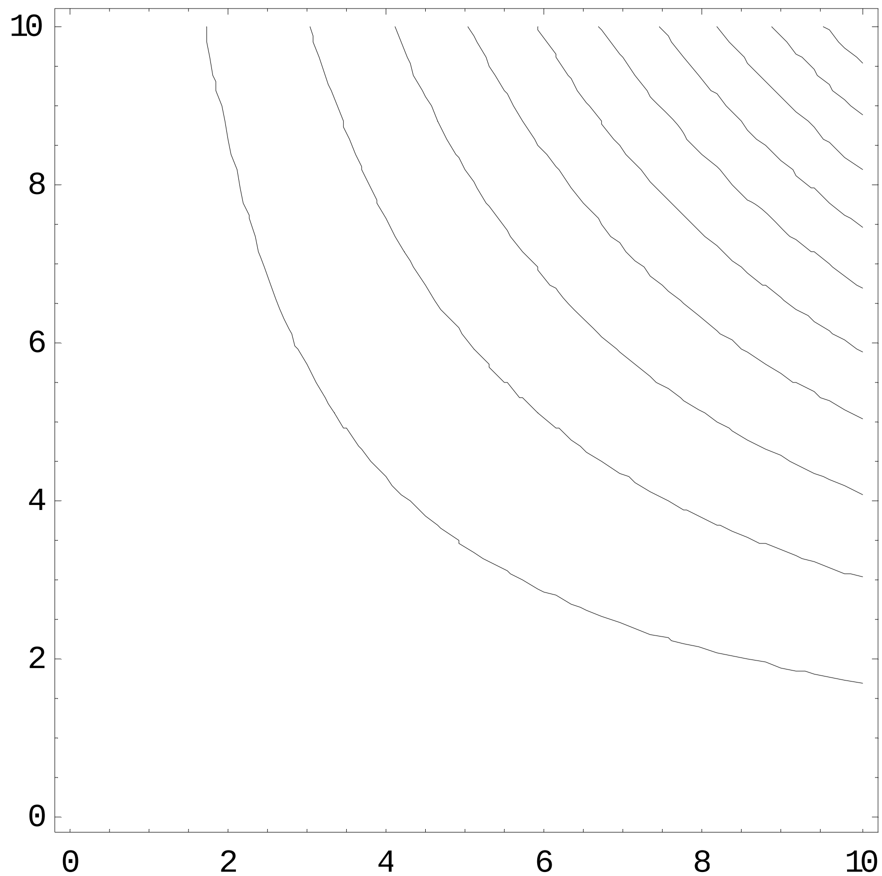

Figure 3.7: Level sets of $v(x)$.

sets.[6] To see why $v(x)$ is quasiconcave, let's look at the level sets of $v(x)$ in Figure 3.7. Even though $v(x)$ is curved in the other direction, the level sets of $v(x)$ are still convex. Hence $v(x)$ is quasiconcave. The important point to take away here is that quasiconcavity is about the shape of level sets, not about the curvature of the 3D graph of the function.

Before going on, let's do one more thing. Recall $u(x) = x_1^{\frac{1}{4}} x_2^{\frac{1}{4}}$ and $v(x) = x_1^{\frac{3}{2}} x_2^{\frac{3}{2}}$. Now, consider the monotonic transformation $f(u) = u^6$. We can rewrite $v(x) = x_1^{\frac{6}{4}} x_2^{\frac{6}{4}} = \left(x_1^{\frac{1}{4}} x_2^{\frac{1}{4}}\right)^6 = f(u(x))$. Hence utility functions $u(x)$ and $v(x)$ actually represent the same preferences! Thus we see that utility and preferences have to do with the shape of indifference curves, not the numbers assigned to them. Again, utility is an ordinal, not cardinal, concept.

Now, here's an example of a function that is not quasiconcave.

$$h(x) = \left(x^2 + y^2\right)^{\frac{1}{4}} \left(2 + \frac{1}{4}\left(\sin\left(8 \arctan\left(\frac{y}{x}\right)\right)\right)^2\right)$$

[6] That isn't a proof, just an illustration.

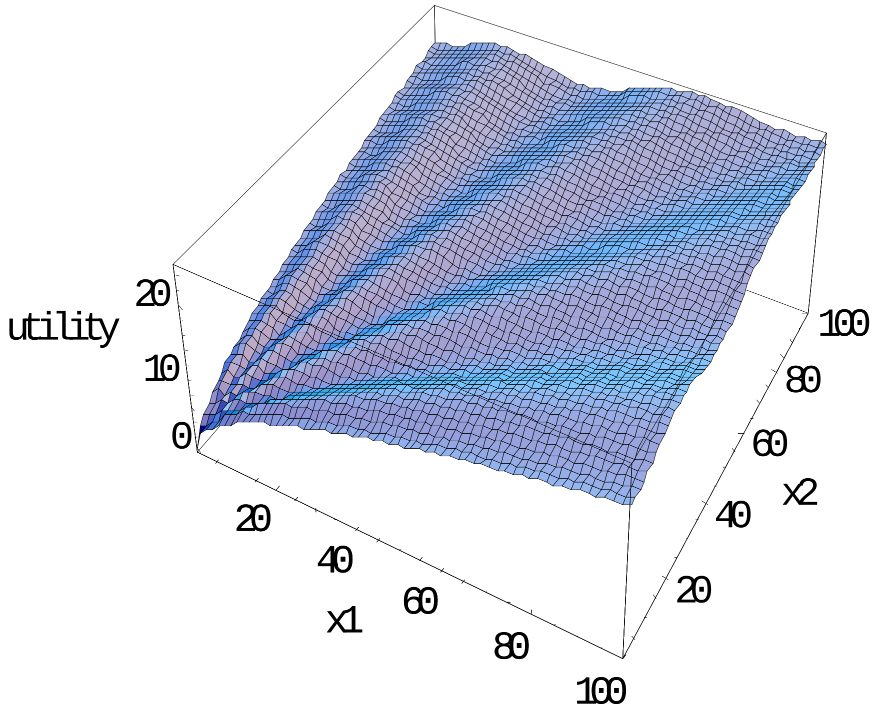

Figure 3.8: Function $h(x)$

Don't worry about where this comes from. Figure 3.8 shows the 3D plot of $h(x)$.

Figure 3.9 shows the isoquants for this utility function. Notice that the level sets are not convex. Hence, function $h(x)$ is not quasiconcave. After looking at the mathematical analysis of the consumer's problem in the next section, we'll come back to why it is so hard to analyze utility functions that look like $h(x)$.

3.3 The Utility Maximization Problem (UMP)

Now that we have defined a utility function, we are prepared to develop the model in which consumers choose the bundle they most prefer from among those available to them.[7] In order to

[7]Notice that in the choice model, we never said why consumers make the choices they do. We only said that those choices must appear to satisfy homogeneity of degree zero, Walras' law, and WARP. Now, we say that the consumer acts to maximize utility with certain properties.

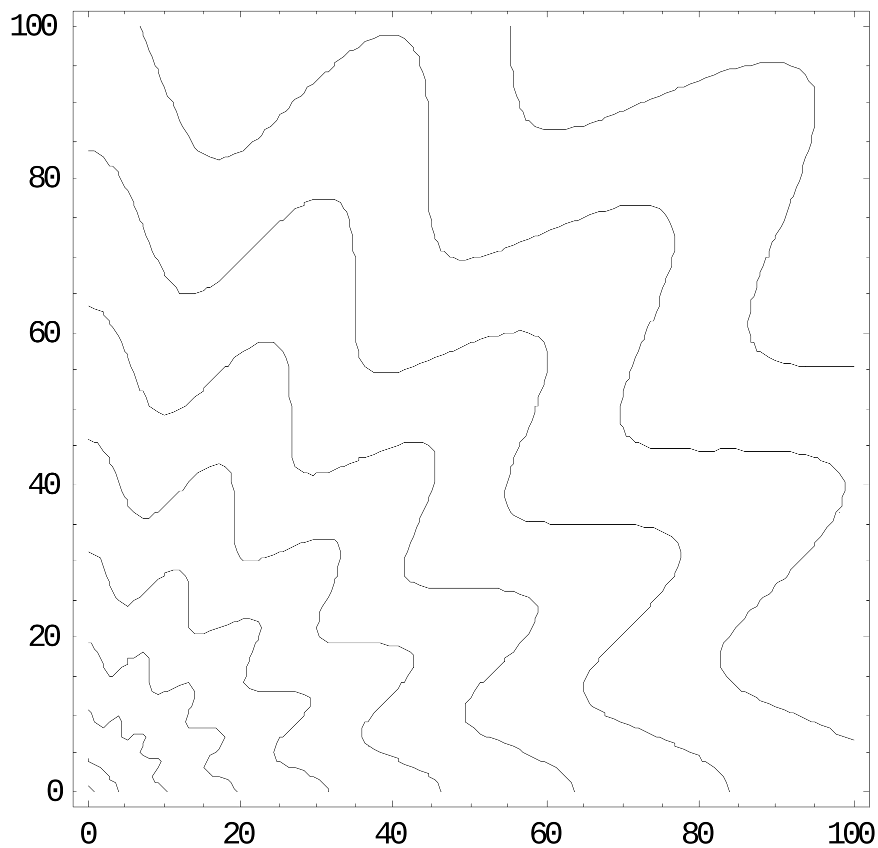

Figure 3.9: Level Sets of $h(x)$

ensure that the problem is "well-behaved," we will assume that preferences are rational, continuous, convex, and locally nonsatiated. These assumptions imply that the consumer has a continuous utility function $u(x)$, and the consumer's choices will satisfy Walras' Law. In order to use calculus techniques, we will assume that $u()$ is differentiable in each of its arguments. Thus, in other words, we assume indifference curves have no "kinks."

The consumer's problem is to choose the bundle that maximizes utility from among those available. The set of available bundles is given by the Walrasian budget set $B_{p,w} = \{x \in X | p \cdot x \leq w\}$. We will assume that all prices are strictly positive (written $p >> 0$) and that wealth is strictly positive as well. The consumer's problem can be written as:

$$\begin{aligned} &\max_{x \geq 0} u(x) \\ s.t. \quad : \quad & p \cdot x \leq w. \end{aligned}$$

The first question we should ask is: Does this problem have a solution? Since $u(x)$ is a continuous function and $B_{p,w}$ is a closed and bounded (i.e., compact) set, the answer is yes by the Weierstrass theorem - a continuous function on a compact set achieves its maximum. How do we find the solution? Since this is a constrained maximization problem, we can use Lagrangian methods. The Lagrangian can be written as:

$$L = u(x) + \lambda (w - p \cdot x)$$

Which implies Kuhn-Tucker first-order conditions (FOC's):

$$\begin{aligned} u_i(x^*) - \lambda^* p_i &\leq 0 \text{ and } x_i (u_i(x^*) - \lambda^* p_i) = 0 \text{ for } i = (1, ..., L) \\ w - p \cdot x^* &\geq 0 \text{ and } \lambda^* (w - p \cdot x^*) = 0 \end{aligned}$$

Note that the optimal solution is denoted with an asterisk. This is because the first-order conditions don't hold everywhere, only at the optimum. Also, note that the value of the Lagrange multiplier λ is also derived as part of the solution to this problem.

Now, we have a system with $L+1$ unknowns. So, we need $L+1$ equations in order to solve for the optimum. Since preferences are locally non-satiated, we know that the consumer will choose a consumption bundle that is on the boundary of the budget set. Thus the constraint must bind. This gives us one equation.

The conditions on x_i are complicated because we must allow for the possibility that the consumer chooses to consume $x_i^* = 0$ for some i at the optimum. This will happen, for example, if the relative price of good i is very high. While this is certainly a possibility, "corner solutions" such as these are not the focus of the course, so we will assume that $x_i^* > 0$ for all i for most of our discussion. But, you should be aware of the fact that corner solutions are possible, and if you come across a corner solution, it may appear to behave strangely.

Generally speaking, we will just assume that solutions are interior. That is, that $x_i^* > 0$ for all commodities i. In this case, the optimality condition becomes

$$u_i\left(x_i^*\right) - \lambda^* p_i = 0. \tag{3.1}$$

Solving this equation for λ^* and doing the same for good j yields:

$$-\frac{u_i\left(x_i^*\right)}{u_j\left(x_j^*\right)} = -\frac{p_i}{p_j} \text{ for all } i, j \in \{1, ..., L\}.$$

This turns out to be an important condition in economics. The condition on the right is the slope of the budget line, projected into the i and j dimensions. For example, if there are two commodities, then the budget line can be written $x_2 = -\frac{p_1}{p_2}x_1 + \frac{w}{p_2}$. The left side, on the other hand, is the slope of the utility indifference curve (also called an **isoquant** or **isoutility curve**). To see why $-\frac{u_i\left(x_i^*\right)}{u_j\left(x_j^*\right)}$ is the slope of the isoquant, consider the following identity: $u\left(x_1, x_2\left(x_1\right)\right) \equiv k$, where k is an arbitrary utility level and $x_2\left(x_1\right)$ is defined as the level of x_2 needed to guarantee the consumer utility k when the level of commodity 1 consumed is x_1. Differentiate this identity with respect to x_1:[8]

$$\begin{aligned} u_1 + u_2\frac{dx_2}{dx_1} &= 0 \\ \frac{dx_2}{dx_1} &= -\frac{u_1}{u_2} \end{aligned}$$

So, at any point (x_1, x_2), $-\frac{u_1(x_1,x_2)}{u_2(x_1,x_2)}$ is the slope of the implicitly defined curve $x_2\left(x_1\right)$. But, this curve is exactly the set of points that give the consumer utility k, which is just the indifference curve. As mentioned earlier, we call the slope of the indifference curve the **marginal rate of substitution (MRS):** $MRS = -\frac{u_1}{u_2}$.

Thus the optimality condition is that at the optimal consumption bundle, the MRS (the rate that the consumer is willing to trade good x_2 for good x_1, holding utility constant) must equal the ratio of the prices of the two goods. In other words, the slope of the utility isoquant is the same as the slope of the budget line. Combine this with the requirement that the optimal bundle be on

[8]Here, we adopt the common practice of using subscripts to denote partial derivatives, $\frac{\partial u(x)}{\partial x_i} = u_i$.

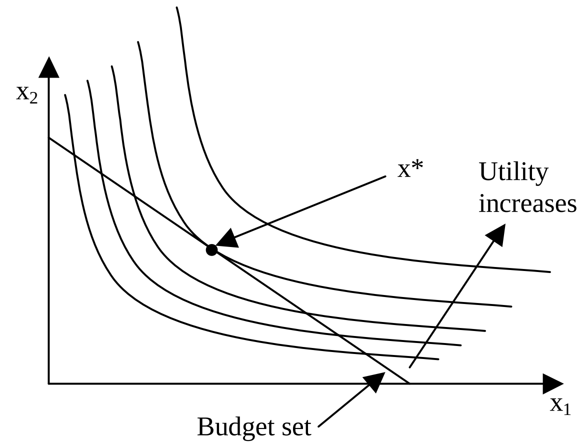

Figure 3.10: Tangency Condition

the budget line, and this implies that the utility isoquant will be tangent to the budget line at the optimum.

In Figure 3.10, x^* is found at the point of tangency between the budget set and one of the utility isoquants. Notice that because the level sets are convex, there is only one such point. If the level sets were not convex, this might not be the case. Consider, for example, Figure 3.9. Here, for any budget set, there will be many points of tangency between utility isoquants and the budget set. Some will be maximizers, and some won't. The only way to find out whether a point is a maximizer is to go through the long and unpleasant process of checking the second-order conditions. Further, even once a maximizer is found, it may behave strangely. We discuss this point further in Section 3.3.1 below.

So, we are looking for the point of tangency between the budget set and a utility isoquant. One way to do this would be to use the following procedure:

1. Choose a point on the budget line, call it z. Find its upper level set U_z. Find $U_z \cap B_{p,w}$. This gives you the set of points that are feasible and at least as good as z.

2. If this set contains only z, you are done: z is the utility maximizing point. If this set contains more than just z, choose an arbitrary point on the budget line that is also inside U_z and repeat the process. Keep going until the only point that is in both the upper level set and the budget set is that point itself. This point is the optimum.

The problem with this procedure is that it could potentially take a very long time to find the optimal point. The calculus approach allows us to do it much faster by finding the point along the budget line that has the same slope as the indifference curve. This is a much easier task, but it turns out that it is really just a shortcut for the procedure outlined above.

3.3.1 Walrasian Demand Functions and Their Properties

So, suppose that we have found the utility maximizing point, x^*. What have we really found? Notice that if the prices and wealth were different, the utility maximizing point would have been different. For this reason, we will write the endogenous variable x^* as a function of prices and wealth, $x(p,w) = (x_1(p,w), x_2(p,w) ..., x_L(p,w))$. This function gives the utility maximizing bundle for any values of p and w. We will call $x(p,w)$ the consumer's **Walrasian demand function**, although it is also sometimes called the Marshallian or ordinary demand function. This is to distinguish it from another type of demand function that we will study later.

As a consequence of what we have done, we can immediately derive some properties of the Walrasian demand function:

1. **Walras' Law:** $p \cdot x(p,w) \equiv w$ for all p and w. This follows from local nonsatiation. Recall the definition of local non-satiation: For any $x \in X$ and $\varepsilon > 0$ there exists a $y \in X$ such that $||x - y|| < \varepsilon$ and $y \succ x$. Thus the only way for x to be the most preferred bundle is if there the nearby point that is better is not in the budget set. But, this can only happen if x satisfies $p \cdot x(p,w) \equiv w$.

2. **Homogeneity of degree zero in** (p,w). The definition of homogeneity is the same as always. $x(\alpha p, \alpha w) = x(p,w)$ for all p, w and $\alpha > 0$. Just as in the choice based approach, the budget set does not change: $B_{p,w} = B_{\alpha p, \alpha w}$. Now consider the first-order condition:

$$-\frac{u_i(x_i^*)}{u_j\left(x_j^*\right)} = -\frac{p_i}{p_j} \text{ for all } i,j \in \{1,...,L\}.$$

   Suppose we multiply all prices by $\alpha > 0$. This makes the right hand side $-\frac{\alpha p_i}{\alpha p_j} = -\frac{p_i}{p_j}$, which is just the same as before. So, since neither the budget constraint nor the optimality condition are changed, the optimal solution must not change either.

3. **Convexity of** $x(p,w)$**.** Up until now we have been assuming that $x(p,w)$ is a unique point. However, it need not be. For example, if preferences are convex but not strictly convex, the

isoquants will have flat parts. If the budget line has the same slope as the flat part, an entire region may be optimal. However, we can say that if preferences are convex, the optimal region will be a convex set. Further, we can add that if preferences are strictly convex, so that $u()$ is strictly quasiconcave, then $x(p,w)$ will be a single point for any (p,w). This is because strict quasiconcavity rules out flat parts on the indifference curve.

A Note on Optimization: Necessary Conditions and Sufficient Conditions

Notice that we derived the first-order conditions for an optimum above. However, while these conditions are necessary for an optimum, they are not generally sufficient - there may be points that satisfy them that are not maxima. This is a technical problem that we don't really want to worry about here. To get around it, we will assume that utility is quasiconcave and monotone (and some other technical conditions that I won't even mention). In this case we know that the first-order conditions are sufficient for a maximum.

In most courses in microeconomic theory, you would be very worried about making sure that the point that satisfies the first-order conditions is actually a maximizer. In order to do this you need to check the second order conditions (make sure the function is "curved down"). This is a long and tedious process, and, fortunately, the standard assumptions we will make, strict quasiconcavity and monotonicity, are enough to make sure that any point that satisfies the first-order conditions is a maximizer (at least when the constraint is linear). Still, you should be aware that there is such things as second-order conditions, and that you either need to check to make sure they are satisfied or make assumptions to ensure that they are satisfied. We will do the latter, and leave the former to people who are going to be doing research in microeconomic theory.

A Word on Nonconvexities

It is worthwhile to spend another moment on what can happen if preferences are not convex, i.e. utility is not quasiconcave. We already mentioned that with nonconvex preferences it becomes necessary to check second-order conditions to determine if a point satisfying the first-order conditions is really a maximizer. There can also be other complications. Consider a utility function where the isoquants are not convex, shown in Figure 3.11.

When the budget line is given by line 1, the optimal point will be near x. When the budget line is line 2, the optimal points will be either x or y. But, none of the points between x and y on line

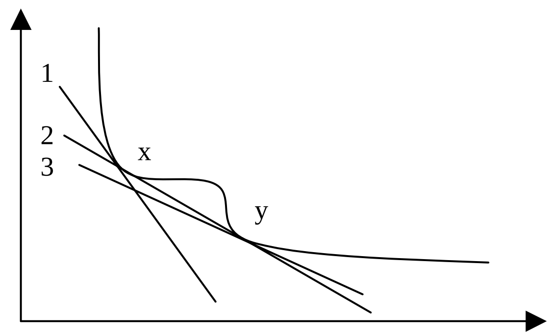

Figure 3.11: Nonconvex Isoquants

2 are as good as x or y (a violation of the idea that averages are better than extremes). Finally, if the budget line is given by line 3, the only optimal point will be near y. Thus the optimal point jumps from x to y without going through any intermediate values.

Now, lines 1, 2, and 3 can be generated by a series of compensated decreases in the price of good 1 (plotted on the horizontal axis). And, intuitively, it seems like people's behavior should change by a small amount if the price changes by a small amount. But, if the indifference curves are non-convex, behavior could change a lot in response to small changes in the exogenous parameters. Since non-convexities result in predictions that do not accord with how we feel consumers actually behave, we choose to model consumers as having convex preferences. In addition, non-convexities add complications to solving and analyzing the consumer's maximization problem that we are very happy to avoid, so this provides another reason why we assume preferences are convex.

Actually, the same sort of problem can arise when preferences are convex but not strictly convex. It could be that behavior changes a lot in response to small changes in prices (although it need not do so). In order to eliminate this possibility and ensure a unique maximizing bundle, we will generally assume that preferences are strictly convex and that utility is strictly quasiconcave (i.e., has strictly convex upper level sets).

3.3.2 The Lagrange Multiplier

You may recall that the optimal value of the Lagrange multiplier is the shadow value of the constraint, meaning that it is the increase in the value of the objective function resulting from a slight relaxation of the constraint. If you don't remember this, you should reacquaint yourself with the point by looking in the math appendix of your favorite micro text or "math for economists" book.[9] If you still don't believe this is true, I present you with the following derivation. In addition to showing this fact about the value of λ^*, it also illustrates a common method of proof in economics.

Consider the utility function $u(x)$. If we substitute in the demand functions, we get

$$u(x(p,w))$$

which is the utility achieved by the consumer when she chooses the best bundle she can at prices p and wealth w. The constraint in the problem is:

$$p \cdot x \leq w.$$

So, relaxing the constraint means increasing w by a small amount. If this is unfamiliar to you, think about why it is so: The budget set $B_{p,w+dw}$ strictly includes the budget set $B_{p,w}$, and so any bundle that could be chosen before the wealth increase could also be chosen after. Since there are more feasible points, the constraint after the wealth increase is a relaxation of the constraint before. We can analyze the effect of this by differentiating $u(x(p,w))$ with respect to w :

$$\begin{aligned}
\frac{d}{dw} u(x(p,w)) &= \sum_{i=1}^{L} \frac{du_i}{dx_i} \frac{dx_i}{dw} \\
&= \sum_{i=1}^{L} \lambda p_i \frac{dx_i}{dw} \\
&= \lambda \sum_{i=1}^{L} p_i \frac{dx_i}{dw} \\
&= \lambda.
\end{aligned}$$

The transition from the first line to the second line is accomplished by substituting in the first-order condition: $\frac{du_i}{dx_i} - \lambda p_i = 0$. The transition from the second line to the third line is trivial (you can factor out λ since it is a constant). The transition from the third line to the fourth line comes from the comparative statics of Walras' Law that we derived in the choice section. Since $p \cdot x(p,w) \equiv w$, $\sum p_i \frac{dx_l}{dw} = 1$ (you could rederive this by differentiating the identity with respect to w if you want).

3.3.3 The Indirect Utility Function and Its Properties

The Walrasian demand function $x(p,w)$ gives the commodity bundle that maximizes utility subject to the budget constraint. If we substitute this bundle into the utility function, we get the utility

[9] If you don't have a favorite, I recommend "Mathematics for Economists" by Simon and Blume.

that is earned when the consumer chooses the bundle that maximizes utility when prices are p and wealth is w. That is, define the function $v(p,w)$ as:

$$v(p,w) \equiv u(x(p,w)).$$

We call $v(p,w)$ the **indirect utility function**. It is *indirect* because while utility is a function of the commodity bundle consumed, x, indirect utility function $v(p,w)$ is a function of p and w. Thus it is indirect because it tells you utility as a function of prices and wealth, not as a function of commodities. You can think of it this way. Given prices p and wealth w, $x(p,w)$ is the commodity bundle chosen and $v(p,w)$ is the utility that results from consuming $x(p,w)$. But, if I know $v(p,w)$, then given any prices and wealth I can calculate utility without first having to solve for $x(p,w)$.

Just as $x(p,w)$ had certain properties, so does $v(p,w)$. In fact, most of them are inherited from the properties of $x(p,w)$. Suppose that preferences are locally nonsatiated. The indirect utility function corresponding to these preferences $v(p,w)$ has the following properties:

1. Homogeneity of degree zero: Since $x(p,w) = x(\alpha p, \alpha w)$ for $\alpha > 0$,

$$v(\alpha p, \alpha w) = u(x(\alpha p, \alpha w)) = u(x(p,w)) = v(p,w).$$

   In other words, since the bundle you consume doesn't change when you scale all prices and wealth by the same amount, neither does the utility you earn.

2. $v(p,w)$ is strictly increasing in w and non-increasing in p_l. If $x_l > 0$, $v(p,w)$ is strictly decreasing in p_l. Indirect utility is strictly increasing in w by local non-satiation. If $x(p,w)$ is optimal and preferences are locally non-satiated, there must be a point just on the other side of the budget line that the consumer strictly prefers. If w increases a little bit, this point will become feasible, and the consumer will earn higher utility. Indirect utility is non-increasing in p_l since an increase in p_l shrinks the feasible region. After p_l increases, the budget line lies inside of the old budget set. Since the consumer could have chosen these points but didn't, the consumer can be no better off than before the price increase. Note that the consumer will do strictly worse unless it is the case that $x_l(p,w) = 0$ and the consumer still chooses to consume $x_l = 0$ after the price increase. In this case, the consumer's consumption bundle does not change, so neither does her utility. This is a subtle point having to do with corner solutions. But, a carefully drawn picture should make it all clear (See Figure 3.12).

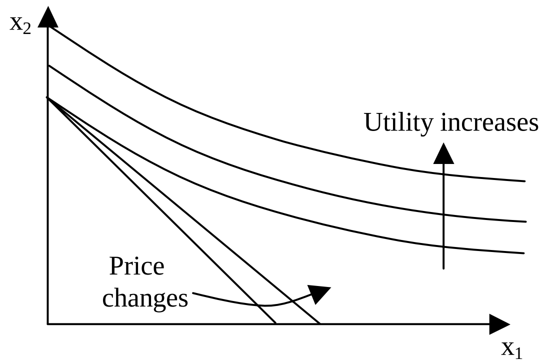

Figure 3.12: $V(p,w)$ is decreasing in p

3. $v(p,w)$ is quasiconvex in (p,w). In other words, the set $\{(p,w)\,|v(p,w) \leq \bar{v}\}$ is convex for all $\bar{v}$. Consider two distinct price-wealth vectors (p',w') and (p'',w'') such that $v(p',w') = v(p'',w'')$. Since the consumer chooses her most preferred consumption bundle at each price, $x(p',w')$ is preferred to all other bundles in $B_{p',w'}$ and $x(p'',w'')$ is preferred to all other bundles in $B_{p'',w''}$. Now, consider the budget set formed at an average price $p^a = ap' + ap''$ and wealth $w^a = aw' + (1-a)w''$. Every bundle in B_{p^a,w^a} will lie within either $B_{p',w'}$ or $B_{p'',w''}$. Hence the utility of any bundle in B_{p^a,w^a} can be no larger than the utility of the chosen bundle, $v(p',w)$. Thus $v(p',w) \geq v(p_a,w)$, which proves the result. Note: see the diagram on page 57 of MWG.[10] A question for you: What is the intuitive meaning of this?

4. $v(p,w)$ is continuous in p and w. Small changes in p and w result in small changes in utility. This is especially clear in the case where indifference curves are strictly convex and differentiable.

3.3.4 Roy's Identity

Consider the indirect utility function: $v(p,w) = \max_{x \in B_{p,w}} u(x)$. The function $v(p,w)$ tells you how much utility the consumer earns when prices are p and wealth is w. Thanks to a very clever bit of mathematics, we can exploit this in order to figure out the relationship between the indirect utility function and the demand functions $x(p,w)$.

[10] This is a slightly informal argument. Formally, we must show for any (p',w') and (p'',w''), $v(p^a,w^a) \leq \max\{v(p',w'), v(p'',w'')\}$. But, the same intuition continues to hold.

The definition of the indirect utility function implies that the following identity is true:

$$v(p,w) \equiv u(x(p,w)).$$

Differentiate both sides with respect to p_l :

$$\frac{\partial v}{\partial p_l} = \sum_{i=1}^{L} \frac{\partial u}{\partial x_i} \frac{\partial x_i}{\partial p_l}.$$

But, based on the first-order conditions for utility maximization, which we know hold when $u()$ is evaluated at the optimal x, $x(p,w)$ (see equation (3.1)):

$$\frac{\partial u}{\partial x_i} = \lambda p_i.$$

And, we also know (from Section 3.3.2) that the Lagrange multiplier is the shadow price of the constraint: $\lambda = \frac{\partial v}{\partial w}$. Hence:

$$\begin{aligned} \frac{\partial v}{\partial p_l} &= \sum_{i=1}^{L} \frac{\partial u}{\partial x_i} \frac{\partial x_i}{\partial p_l} = \sum_{i=1}^{L} \lambda p_i \frac{\partial x_i}{\partial p_l} \\ &= \lambda \sum_{i=1}^{L} p_i \frac{\partial x_i}{\partial p_l} = \frac{\partial v}{\partial w} \sum_{i=1}^{L} p_i \frac{\partial x_i}{\partial p_l}. \end{aligned}$$

Now, recall the comparative statics result of Walras' Law with respect to a change in p_l :

$$x_l(p,w) = -\sum_{i=1}^{L} p_i \frac{\partial x_i}{\partial p_l}.$$

Substituting this in yields:

$$\begin{aligned} \frac{\partial v}{\partial p_l} &= -\frac{\partial v}{\partial w} x_l(p,w) \\ x_l(p,w) &= -\frac{\frac{\partial v}{\partial p_l}}{\frac{\partial v}{\partial w}}. \end{aligned}$$

The last equation, known as Roy's identity, allows us to derive the demand functions from the indirect utility function. This is useful because in many cases it will be easier to estimate an indirect utility function and derive the direct demand functions via Roy's identity than to derive $x(p,w)$ directly. Estimating Roy's identity involves estimating a single equation. Estimating $x(p,w)$, on the other hand, amounts to finding for every value of p and w the solution to a set of $L+1$ first-order equations, which themselves may have unknown parameters.

3.3.5 The Indirect Utility Function and Welfare Evaluation

Consider the situation where the price of good 1 increases from p_1 to p_1'. What is the impact of this price change on the consumer? One way to measure it is in terms of the indirect utility function:

$$impact = v\left(p', w\right) - v\left(p, w\right).$$

That is, the impact of the price change is equal to the difference in the consumer's utility at prices p' and p. While this is certainly a measure of the impact of the price change, it is essentially useless. There are a number of reasons why, but they all hinge on the fact that utility is an ordinal, not a cardinal concept. As you recall, the only meaning of the numbers assigned to bundles by the utility function is that $x \succ y$ if and only if $u(x) > u(y)$. In particular, if $u(x) = 2u(y)$, this does not mean that the consumer likes bundle x twice as much as bundle y. Also, if $u(x) - u(z) > u(s) - u(t)$, this doesn't mean that the consumer would rather switch from bundle z to x than from t to s. Because of this, there is really nothing we can make of the numerical value of $v(p', w) - v(p, w)$. The only thing we can say is that if this difference is positive, the consumer likes (p', w) more than (p, w). But, we can't say how much more.

Another problem with using the change in the indirect utility function as a measure of the impact of a policy change is that it cannot be compared across consumers. Comparing the change in two different utility functions is even more meaningless than comparing the change in a single person's utility function. This is because even if both utility functions were cardinal measures of the benefit to a consumer (which they aren't), there would still be no way to compare the scales of the two utility functions. This is the "problem of interpersonal comparison of utility," which arises in many aspects of welfare economics.

As a possible solution to the problem, consider the following thought experiment. Initially, prices and wealth are given by (p, w). I am interested in measuring the impact of a change in prices to p'. So, I ask you the following question: By how much would I have to change your wealth so that you are indifferent between (p', w) and (p, w')? That is, for what value w' does

$$v\left(p', w\right) = v\left(p, w'\right).$$

The change in wealth, $w' - w$, in essence gives a monetary value for the impact of this change in price. And, this monetary value is a better measure of the impact of the price change than the utility measurement, because it is, at least to a certain extent, comparable.[11] You can compare

[11] I say to a certain extent, because the value of an additional Δw dollars of wealth will depend on the initial state.

the impact of two different changes in prices by looking at the associated changes in wealth needed to compensate the consumer. Similarly, you can compare the impact of price changes on different consumers by comparing the changes in wealth necessary to leave them just as well off.[12]

Finally, although using the amount of money needed to compensate the consumer is an imperfect measure of the impact of a policy decision, it has one huge benefit for the neoclassical economist, and that is that it is observable, at least in principle. There is nothing we can do to observe utility scales. However, we can often elicit from people the amount of money they would find equivalent to a certain policy change, either through experiments, surveys, or other estimation techniques.

3.4 The Expenditure Minimization Problem (EMP)

In the previous section, I argued that a good measure of the impact of a change in prices was the change in wealth necessary to make the consumer as well off at the old prices and new wealth as she was at the new prices and old wealth. However, this is not an easy exercise when all you have to work with is the indirect utility function. If we had a function that tells you how much wealth you would need to have in order to achieve a certain level of utility, then we would be able to do this much more efficiently. There is such a function. It is called the **expenditure function,** and in this section we will develop it.

The **expenditure minimization problem** (EMP) asks the question, if prices are p, what is the minimum amount the consumer would have to spend to achieve utility level u? That is:

$$\begin{aligned} & \min_x p \cdot x \\ s.t. \quad : \quad & u(x) \geq u. \end{aligned}$$

Before we go on, let's take a moment to figure out what the endogenous and exogenous variables are here. The exogenous variables are prices p and the reservation (or target) utility level u. The endogenous variable is x, the consumption bundle. So, in words, the expenditure minimization bundle amounts to finding the bundle x that minimizes the cost of achieving utility u when prices are p.

The Lagrangian for this problem is given by:

$$L_{EMP} = p \cdot x - \lambda \left(u(x) - u \right).$$

For instance, poor people presumably value the same wealth increment more than rich people.

[12] Again, this measure is imperfect because it assumes that the two consumers have the same marginal utility of wealth.

Assuming an interior solution, the first-order conditions are given by:[13]

$$
\begin{aligned}
p_i - \lambda u_i(x) &= 0 \text{ for } i \in \{1, ..., L\} \\
\lambda(u(x) - u) &= 0
\end{aligned} \tag{3.2}
$$

If $u()$ is well behaved (e.g., quasiconcave and increasing in each of its arguments), then the constraint will bind, and the second condition can be written as $u(x) = u$. Further, a unique solution to this problem will exist for any values of p and u. We will denote the value of the solution to this problem by $h(p, u) \in X$. That is, $h(p, u)$ is an L dimensional vector whose l^{th} component, $h_l(p, u)$ gives the amount of commodity l that is consumed when the consumer minimizes the cost of achieving utility u at prices p. The function $h(p, u)$ is known as the **Hicksian (or compensated) demand function.**[14] It is a demand function because it specifies a consumption bundle. It differs from the Walrasian (or ordinary) demand function in that it takes p and u as its arguments, whereas the Walrasian demand function takes p and w as its arguments.

In other words, $h(p, u)$ and $x(p, w)$ are the answers to two different but related problems. Function $x(p, w)$ answers the question, "Which commodity bundle maximizes utility when prices are p and wealth is w?" Function $h(p, u)$ answers the question, "Which commodity bundle minimizes the cost of attaining utility u when prices are p?" We'll return to the difference between the two types of demand shortly.

Since $h(p, u)$ solves the EMP, substitution of $h(p, u)$ into the first-order conditions for the EMP yields the identities (assuming the constraint binds):

$$
\begin{aligned}
p_i - \lambda u_i(h(p, u)) &\equiv 0 \text{ for } i \in \{1, ..., L\} \\
u(h(p, u)) - u &\equiv 0
\end{aligned}
$$

Further, just as we defined the indirect utility function as the value of the objective function of the UMP, $u(x)$, evaluated at the optimal consumption bundle, $x(p, w)$, we can also define such a function for the EMP. The **expenditure function,** denoted $e(p, u)$, is defined by:

$$
e(p, u) \equiv p \cdot h(p, u)
$$

and is equal to the minimum cost of achieving utility u, for any given p and u.

[13] Again, remember that if $f(y)$ is a function with a vector y as its argument, the notation f_i will frequently be used as shorthand notation for $\frac{\partial f}{\partial y_i}$. Thus u_i denotes $\frac{\partial u}{\partial x_i}$.

[14] We'll return to why $h(p, u)$ is called the compensated demand function in a while.

3.4.1 A First Note on Duality

Consider the first-order conditions (from (3.2)) for x_i and x_j. Solving each for λ yields:

$$\begin{aligned} \frac{p_i}{u_i} &= \lambda = \frac{p_j}{u_j} \\ \frac{u_i}{u_j} &= \frac{p_i}{p_j}. \end{aligned} \tag{3.3}$$

Recall that this is the same tangency condition we derived in the UMP. What does this mean? Consider a price vector p and wealth w. The bundle that solves the UMP, $x^* = x(p, w)$ is found at the point of tangency between the budget line and the consumer's utility isoquant. The consumer's utility at this point is given by $u^* = u(x^*)$. Thus x^* is the point of tangency between the line $p \cdot x = w$ and the curve $u(x) = x^*$.

Now, consider the EMP when the target utility level is given by u^*. The bundle that solves the EMP is the bundle that achieves utility u^* at minimum cost. This is located by finding the point of tangency between the curve $u(x) = u^*$ and a budget line (which is what (3.3) says). But, we already know from the UMP that the curve $u(x) = u^*$ is tangent to the budget line $p \cdot x = w$ at x^* (and is tangent to no other budget line). Hence x^* must solve the EMP problem when the target utility level is u^*! Further, since x^* lies on the budget line, $p \cdot x^* = w$. So the minimum cost of achieving utility u^* is w. Thus the UMP and the EMP pick out the same point.

Let me restate what I've just argued. If x^* solves the UMP when prices are p and wealth is w, then x^* solves the EMP when prices are p and the target utility level is $u(x^*)$. Further, maximal utility in the UMP is $u(x^*)$ and minimum expenditure in the EMP is w. This result is called the "duality" of the EMP and the UMP.

The UMP and the EMP are considered dual problems because the constraint in the UMP is the objective function in the EMP and vice versa. This is illustrated by looking at the graphical solutions to the two problems. In the UMP, shown in Figure 3.13, you keep increasing utility until you find the one that is tangent to the budget line.In the EMP, on the other hand, shown in Figure 3.14, you keep decreasing expenditure (which is like shifting a budget line toward the origin) until you find the expenditure line that is tangent to $u(x) = u^*$. Although the process of finding the optimal point is different in the UMP and EMP, they both pick out the same point because they are looking for the same basic relationship, as expressed in equation (3.3).

The duality relationship between the EMP and the UMP is captured by the following identities,

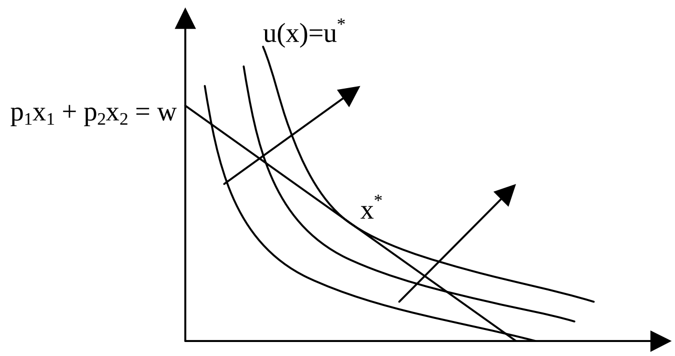

Figure 3.13: The Utility Maximization Problem

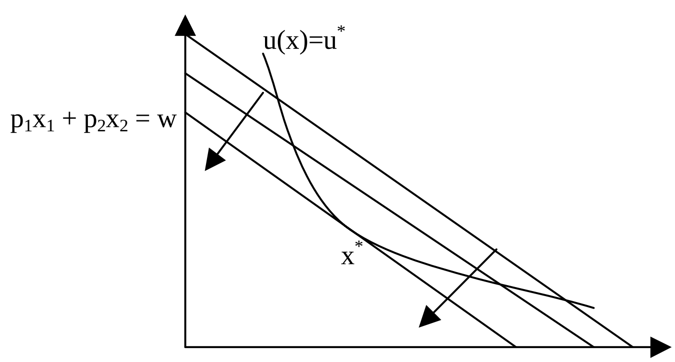

Figure 3.14: The Expenditure Minimization Problem

to which we will return later:

$$\begin{aligned} h(p, v(p, w)) &\equiv x(p, w) \\ x(p, e(p, u)) &\equiv h(p, u). \end{aligned}$$

These identities restate the principles discussed previously. The first says that the commodity bundle that minimizes the cost of achieving the maximum utility you can achieve when prices are p and wealth is w is the bundle that maximizes utility when prices are p and wealth is w. The second says that the bundle that maximizes utility when prices are p and wealth is equal to the minimum amount of wealth needed to achieve utility u at those prices is the same as the bundle that minimizes the cost of achieving utility u when prices are p.

Similar identities can be written using the indirect utility function and expenditure function:

$$\begin{aligned} u &\equiv v(p, e(p, u)) \\ w &\equiv e(p, v(p, w)). \end{aligned}$$

Note to MWG readers: There is a mistake in Figure 3.G.3. The relationships on the horizontal line connecting $v(p, w)$ and $e(p, u)$ **should be the ones written directly above.**

The main implication of the previous analysis is this: The expenditure function contains the exact same information as the indirect utility function. And, since the indirect utility function can be used (by Roy's identity) to derive the Walrasian demand functions, which can, in turn, be used to recover preferences, **the expenditure function contains the exact same information as the utility function.** This means that if you know the consumer's expenditure function, you know her utility function, and vice versa. No information is lost along the way. This is another expression of what people mean when they say that the UMP and EMP are dual problems - they contain exactly the same information.

3.4.2 Properties of the Hicksian Demand Functions and Expenditure Function

In this section, we refer both to function $u(x)$ and to a particular level of utility, u. In order to be clear, let's put a bar over the u when we are talking about a level of utility, i.e., $\bar{u}$. Just as we derived the properties of $x(p, w)$ and $v(p, w)$, we can also derive the properties of the Hicksian demand functions $h(p, \bar{u})$ and expenditure function $e(p, \bar{u})$. Let's begin with $h(p, \bar{u})$. We will assume that $u()$ is a continuous utility function representing a locally non-satiated preference relation.

Properties of the Hicksian Demand Functions

The Hicksian demand functions have the following properties:

1. Homogeneity of degree zero in $p : h(\alpha p, \bar{u}) \equiv h(p, \bar{u})$ for $p, \bar{u}$, and $\alpha > 0$. NOTE: THIS IS HOMOGENEITY IN P, NOT HOMOGENEITY IN P AND U! Homogeneity of degree zero is best understood in terms of the graphical presentation of the EMP. The solution to the EMP is the point of tangency between the utility isoquant $u(x) = \bar{u}$ and one of the budget lines. This is determined by the slope of the expenditure lines (lines of the form $p \cdot x = k$, where k is any constant). Any change that doesn't affect the slope of the budget lines should not affect the cost-minimizing bundle (although it will affect the expenditure on the cost minimizing bundle). Since the slope of the expenditure line is determined by relative prices and since scaling all prices by the same amount does not affect relative prices, the solution should not change. More formally, the EMP at prices αp is

$$\begin{aligned} &\min_x \alpha p \cdot x \\ s.t. \quad : \quad & u(x) \geq \bar{u}. \end{aligned}$$

But, this problem is formally equivalent to:

$$\min \alpha (p \cdot x) : s.t. : u(x) \geq x$$

which is equivalent to:

$$\alpha \min_x p \cdot x : s.t. : u(x) \geq x$$

which is just the same as the EMP when prices are p, except that total expenditure is multiplied by α, which doesn't affect the cost minimizing bundle.

2. No excess utility: $u(h(p, \bar{u})) = \bar{u}$. This follows from the continuity of $u()$. Suppose $u(h(p, \bar{u})) > \bar{u}$. Then consider a bundle h' that is slightly smaller than $h(p, \bar{u})$ on all dimensions. By continuity, if h' is sufficiently close to $h(p, \bar{u})$, then $u(h') > \bar{u}$ as well. But, then h' is a bundle that achieves utility $\bar{u}$ at lower cost than $h(p, \bar{u})$, which contradicts the assumption that $h(p, \bar{u})$ was the cost minimizing bundle in the first place.[15] From this we can conclude that the constraint always binds in the EMP.

[15] This type of argument - called "Proof by Contradiction" - is quite common in economics. If you want to show a implies b, assume that b is false and show that if b is false then a must be false as well. Since a is assumed to be true, this implies that b must be true as well.

3. If preferences are convex, then $h(p,\bar{u})$ is a convex set. If preferences are strictly convex (i.e. $u()$ is strictly quasiconcave), then $h(p,\bar{u})$ is single valued.

Properties of the Expenditure Function

Based on the properties of $h(p,\bar{u})$, we can derive properties of the expenditure function, $e(p,\bar{u})$.

1. Function $e(p,\bar{u})$ is **homogeneous of degree one in** p: Since $h(p,\bar{u})$ is homogeneous of degree zero in p, this means that scaling all prices by $\alpha > 0$ does not affect the bundle demanded. Applying this to total expenditure:

$$e(\alpha p,\bar{u}) = \alpha p \cdot h(\alpha p,\bar{u}) = \alpha p \cdot h(p,\bar{u}) = \alpha e(p,\bar{u}).$$

   In words, if all prices change by a factor of α, the same bundle as before achieves utility level $\bar{u}$ at minimum cost, only it now costs you twice as much as it used to. This is exactly what it means for a function to be homogeneous of degree one.

2. Function $e(p,\bar{u})$ is **strictly increasing in** $\bar{u}$ and **non-decreasing in** p_l for any l. I'll give the argument here to show that $e(p,\bar{u})$ cannot be decreasing in $\bar{u}$. There are a few more details to show that it cannot stay constant either, but most of the intuition of the argument is contained in showing that $e(p,\bar{u})$ cannot be strictly decreasing. The argument is by contradiction. Suppose that for $\bar{u}' > \bar{u}$, $e(p,\bar{u}) > e(p,\bar{u}')$. But, then $h(p,\bar{u}')$ satisfies the constraint $u(x) \geq \bar{u}$ and does so at lower cost than $h(p,\bar{u})$, which contradicts the assumption that $h(p,\bar{u})$ is the cost minimizing bundle that achieves utility level $\bar{u}$. The argument that $e(p,\bar{u})$ is nondecreasing in p_l uses another method which is quite common, a method I call "feasible but not optimal." Let p and p' differ only in component l, and let $p'_l > p_l$. From the definition of the expenditure function, $e(p',\bar{u}) = p' \cdot h(p',\bar{u}) \geq p \cdot h(p',\bar{u}) \geq e(p,\bar{u})$. The first equality follows from the definition of the expenditure function, the first $\geq$ follows from the fact that $p' > p$ (note: $p' \cdot h(p',\bar{u}) > p \cdot h(p,\bar{u})$ if $h_l(p',\bar{u}) > 0$), and the second $\geq$ follows from the fact that $h(p',\bar{u})$ achieves utility level $\bar{u}$ but does not necessarily do so at minimum cost (i.e. $h(p',\bar{u})$ is feasible in the EMP for $(p,\bar{u})$ but not necessarily optimal).

3. Function $e(p,\bar{u})$ is **concave in** p.[16] Consider two price vectors p and p', and let $p^a =$

[16] Recall the definition of concavity. Consider y and y' such that $y \neq y'$. Function $f(y)$ is concave if, for any $a \in [0,1]$, $f(ay + (1-a)y') \geq af(y) + (1-a)f(y')$.

$ap + (1-a)\,p'$.

$$\begin{aligned} e\left(p^a, \bar{u}\right) &= p^a \cdot h\left(p^a, \bar{u}\right) \\ &= ap \cdot h\left(p^a, \bar{u}\right) + (1-a)\,p \cdot h\left(p^a, \bar{u}\right) \\ &\geq ap \cdot h\left(p, \bar{u}\right) + (1-a)\,p' \cdot h\left(p', \bar{u}\right) \end{aligned}$$

where the first line is the definition of $e(p, \bar{u})$, the second follows from the definition of p^a, and the third follows from the fact that $h(p^a, \bar{u})$ is feasible but not optimal in the EMP for $(p, \bar{u})$ and $(p', \bar{u})$.

The following heuristic explanation is also helpful in understanding the concavity of $e(p, \bar{u})$. Suppose prices change from p to p'. If the consumer continued to consume the same bundle at the old prices, expenditure would increase linearly:

$$\Delta\text{expenditure} = \left(p' - p\right) \cdot h\left(p, \bar{u}\right).$$

But, in general the consumer will not continue to consume the same bundle after the price change. Rather, he will rearrange his bundle in order to minimize the cost of achieving $\bar{u}$ at the new prices, p'. Since this will save the consumer some money, total expenditure will decrease at less than a linear rate. And, an alternate definition of concavity is that the function always increases at less than a linear rate. In other words, $f(x)$ is concave if it always lies below its tangent lines.[17]

3.4.3 The Relationship Between the Expenditure Function and Hicksian Demand

Just as there was a relationship between the indirect utility function $v(p, w)$ and the Walrasian demand functions $x(p, w)$, there is also a relationship between the expenditure function $e(p, \bar{u})$ and the Hicksian demand function $h(p, \bar{u})$. In fact, it is even more straightforward for $e(p, \bar{u})$ and $h(p, \bar{u})$. Let's start with the derivation

$$e\left(p, \bar{u}\right) \equiv p \cdot h\left(p, \bar{u}\right)$$

Since this is an identity, differentiate it with respect to p_i:

$$\frac{\partial e}{\partial p_i} \equiv h_i\left(p, \bar{u}\right) + \sum_j p_j \frac{\partial h_j}{\partial p_i}.$$

[17] This explanation will be clearer once we show that $h_l(p, u) = \frac{\partial e(p,u)}{\partial p_l}$, i.e. that $h_l(p, u)$ is exactly the rate of increase in expenditure if p_l increases by a small amount. Thus $e(p', u) - e(p, u) \leq h(p, u) \cdot (p' - p)$ is exactly the definition of concavity.

Now, substitute in the first-order conditions, $p_j = \lambda u_j$

$$\frac{\partial e}{\partial p_i} \equiv h_i\left(p, \bar{u}\right) + \lambda \sum_j u_j \frac{\partial h_j}{\partial p_i}. \tag{3.4}$$

Since the constraint binds at any optimum of the EMP,

$$u\left(h\left(p, \bar{u}\right)\right) \equiv \bar{u}$$

Differentiate with respect to p_i :

$$\sum_j u_j \frac{\partial h_j}{\partial p_i} = 0$$

and substituting this into (3.4) yields:

$$\frac{\partial e}{\partial p_j} \equiv h_j\left(p, \bar{u}\right). \tag{3.5}$$

That is, the derivative of the expenditure function with respect to p_j is just the Hicksian demand for commodity j.

The importance of this result is similar to the importance of Roy's identity. Frequently, it will be easier to measure the expenditure function than the Hicksian demand function. Since we are able to derive the Hicksian demand function from the expenditure function, we can derive something that is hard to observe from something that is easier to observe.

From (3.5) we can derive several additional properties (assuming $u()$ is strictly quasiconcave and $h()$ is differentiable):

1. (a) $\frac{\partial h_i}{\partial p_j} = \frac{\partial^2 e}{\partial p_i \partial p_j}$. This one follows directly from the fact that (3.5) is an identity. Let $D_p h\left(p, \bar{u}\right)$ be the matrix whose i^{th} row and j^{th} column is $\frac{\partial h_i}{\partial p_j}$. This property is thus the same as saying that $D_p h\left(p, \bar{u}\right) \equiv D_p^2 e\left(p, \bar{u}\right)$, where $D_p^2 e\left(p, \bar{u}\right)$ is the matrix of second derivatives (Hessian matrix) of $e\left(p, \bar{u}\right)$.

   (b) $D_p h\left(p, \bar{u}\right)$ is a negative semi-definite (n.s.d.) matrix. This follows from the fact that $e\left(p, \bar{u}\right)$ is concave, and concave functions have Hessian matrices that are n.s.d. The main implication is that the diagonal elements are non-positive, i.e., $\frac{\partial h_i}{\partial p_i} \leq 0$.

   (c) $D_p h\left(p, \bar{u}\right)$ is symmetric. This follows from Young's Theorem (that it doesn't matter what order you take derivatives in): $\frac{\partial h_i}{\partial p_j} = \frac{\partial^2 e}{\partial p_i \partial p_j} = \frac{\partial h_j}{\partial p_i}$. The implication is that the cross-effects are the same – the effect of increasing p_j on h_i is the same as the effect of increasing p_i on h_j.

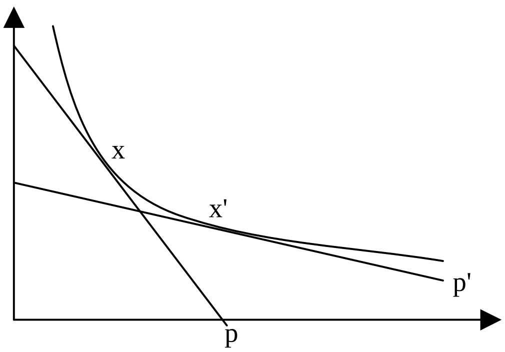

Figure 3.15: Compensated Demand

(d) $\sum_j \frac{\partial h_i}{\partial p_j} p_j = 0$ for all i. This follows from the homogeneity of degree zero of $h(p, \bar{u})$ in p. Consider the identity:

$$h(ap, \bar{u}) \equiv h(p, \bar{u}).$$

Differentiate with respect to a and evaluate at $a = 1$, and you have this result.

The Hicksian demand curve is also known as the **compensated demand curve**. The reason for this is that implicit in the definition of the Hicksian demand curve is the idea that following a price change, you will be given enough wealth to maintain the same utility level you did before the price change. So, suppose at prices p you achieve utility level $\bar{u}$. The change in Hicksian demand for good i following a change to prices p' is depicted in Figure 3.15.

When prices are p, the consumer demands bundle x, which has total expenditure $p \cdot x = w$. When prices are p', the consumer demands bundle x', which has total expenditure $p' \cdot x' = w'$. Thus implicit in the definition of the Hicksian demand curve is the idea that when prices change from p to p', the consumer is compensated by changing wealth from w to w' so that she is exactly as well off in utility terms after the price change as she was before.

Note that since $\frac{\partial h_i}{\partial p_i} \leq 0$, this is another statement of the compensated law of demand (CLD). When the price of a good goes up and the consumer is compensated for the price change, she will not consume more of the good. The difference between this version and the previous version we saw (in the choice based approach) is that here, the compensation is such that the consumer can achieve the same utility before and after the price change (this is known as Hicksian substitution), and in the previous version of the CLD the consumer was compensated so that she could just afford the same bundle as she did before (this is known as Slutsky compensation). It turns out that the two types of compensation yield very similar results, and, in fact, for differential changes in price, they are identical.

3.4.4 The Slutsky Equation

Recall that the whole point of the EMP was to generate concepts that we could use to evaluate welfare changes. The purpose of the expenditure function was to give us a way to measure the impact of a price change in dollar terms. While the expenditure function does do this (you can just look at $e(p', u) - e(p, u)$), it suffers from another problem. The expenditure function is based on the Hicksian demand function, and the Hicksian demand function takes as its arguments prices and the target utility level u. The problem is that while prices are observable, utility levels certainly are not. And, while we can generate some information by asking people over and over again how they compare certain bundles, this is not a very good way of doing welfare comparisons.

To summarize our problem: The Walrasian demand functions are based on observables (p, w) but cannot be used for welfare comparisons. The Hicksian demand functions, on the other hand, can be used to make welfare comparisons, but are based on unobservables.

The solution to this problem is to somehow derive $h(p, u)$ from $x(p, w)$. Then we could use our observations of p and w to derive $h(p, u)$, and use $h(p, u)$ for welfare evaluation. Fortunately, we can do exactly this. Suppose that $u(x(p, w)) = u$ (which implies that $e(p, u) = w$), and consider the identity:

$$h_i(p, u) \equiv x_i(p, e(p, u)).$$

Differentiate both sides with respect to p_j :

$$\begin{aligned}
\frac{\partial h_i}{\partial p_j} &\equiv \frac{\partial x_i(p, e(p, u))}{\partial p_j} + \frac{\partial x_i(p, e(p, u))}{\partial e(p, u)} \frac{\partial e(p, u)}{\partial p_j} \\
&\equiv \frac{\partial x_i(p, w)}{\partial p_j} + \frac{\partial x_i(p, w)}{\partial w} h_j(p, u) \\
&\equiv \frac{\partial x_i(p, w)}{\partial p_j} + \frac{\partial x_i(p, w)}{\partial w} x_j(p, e(p, u)) \\
&\equiv \frac{\partial x_i(p, w)}{\partial p_j} + \frac{\partial x_i(p, w)}{\partial w} x_j(p, w).
\end{aligned}$$

The equation

$$\frac{\partial h_i(p, v(p, w))}{\partial p_j} \equiv \frac{\partial x_i(p, w)}{\partial p_j} + \frac{\partial x_i(p, w)}{\partial w} x_j(p, w)$$

is known as the Slutsky equation. Note that it provides the link between the Walrasian demand functions $x(p, w)$ and the Hicksian demand functions, $h(p, u)$. Thus if we estimate the right-hand side of this equation, which is a function of the observables p and w, then we can derive the value of the left-hand side of the equation, even though it is based on unobservable u.

Recall that implicit in the idea of the Hicksian demand function is the idea that the consumer's wealth would be adjusted so that she can achieve the same utility after a price change as she did before. This idea is apparent when we look at the Slutsky equation. It says that the change in demand when the consumer's wealth is adjusted so that she is as well off after the change as she was before is made up of two parts. The first, $\frac{\partial x_i(p,w)}{\partial p_j}$, is equal to how much the consumer would change demand if wealth were held constant. The second, $\frac{\partial x_i(p,w)}{\partial w} x_i(p,w)$, is the additional change in demand following the compensation in wealth.

For example, consider an increase in the price of gasoline. If the price of gasoline goes up by one unit, consumers will tend to consume less of it, if their wealth is held constant (since it is not a Giffen good). However, the fact that gasoline has become more expensive means that they will have to spend more in order to achieve the same utility level. The amount by which they will have to be compensated is equal to the change in price multiplied by the amount of gasoline the consumer buys, $x_i(p,w)$. However, when the consumer is given $x_i(p,w)$ more units of wealth to spend, she will adjust her consumption of gasoline further. Since gasoline is normal, the consumer will increase her consumption. Thus the compensated change in demand (sometimes called the **pure substitution effect**), $\frac{\partial h_i(p,v(p,w))}{\partial p_j}$, will be the sum of the uncompensated change (also known as the substitution effect), $\frac{\partial x_i(p,w)}{\partial p_j}$, and the wealth effect, $\frac{\partial x_i(p,w)}{\partial w} x_i(p,w)$.[18]

In order to make this clear, let's rearrange the Slutsky equation and go through the intuition again.

$$\frac{\partial x_i(p,w)}{\partial p_j} \equiv \frac{\partial h_i(p,v(p,w))}{\partial p_j} - \frac{\partial x_i(p,w)}{\partial w} x_j(p,w)$$

Here, we are interested in explaining an uncompensated change in demand in terms of the compensated change and the wealth effect. Think about the effect of an increase in the price of bananas on a consumer's Walrasian demand for bananas. If the price of bananas were to go up, and my wealth were adjusted so that I could achieve the same amount of utility before and after the change, I would consume fewer bananas. This follows directly from the CLD: $\frac{\partial h_i}{\partial p_i} \leq 0$. However, the change in compensated demand assumes that the consumer will be compensated for the price change. Since an increase in the price of bananas is a bad thing, this means that $\frac{\partial h_i}{\partial p_i}$ has built into it the idea

[18] This latter term is often called the "income effect," which is not quite right. Variable w stands for total wealth, which is more than just income. When people call this the income effect (as I sometimes do), they are just being sloppy.

that income will be increased in order to compensate the consumer. But, in reality consumers are not compensated for price changes, so we are interested in the uncompensated change in demand $\frac{\partial x_i}{\partial p_i}$. This means that we must remove from the compensated change in demand the effect of the compensation. Since $\frac{\partial h_i}{\partial p_i}$ assumed an increase in wealth, we must impose a decrease in wealth, which is just what the terms $-\frac{\partial x_i(p,w)}{\partial w} x_i(p, w)$ are. The decrease in wealth is given by $-x_i(p, w)$, and the effect of this decrease on demand for bananas is given by $\frac{\partial x_i}{\partial w}$.[19]

3.4.5 Graphical Relationship of the Walrasian and Hicksian Demand Functions

Demand functions are ordinarily graphed with price on the vertical axis and quantity on the horizontal axis, even though this is technically "backward." But, we will follow with tradition and draw our graphs this way as well.

The difference between the compensated demand response to a price change and the uncompensated demand response to a price change is equal to the wealth effect:

$$\frac{\partial h_i(p, v(p, w))}{\partial p_j} \equiv \frac{\partial x_i(p, w)}{\partial p_j} + \frac{\partial x_i(p, w)}{\partial w} x_j(p, w)$$

Since $\frac{\partial h_i}{\partial p_j}$ is negative, when the wealth effect is positive (i.e., good i is normal) this means that the Hicksian demand curve will be steeper than the Walrasian demand curve at any point where they cross.[20] If, on the other hand, the wealth effect is negative (i.e. good i is inferior), this means that the Hicksian demand curve will be less steep than the Walrasian demand curve (see MWG Figure 3.G.1).

Let's go into a bit more detail in working out the relative slopes of the Walrasian and Hicksian demand curves and determining how changes in u shift the Hicksian demand curve (depending on whether the good is normal or inferior).

In this subsection we discuss how changes in the exogenous parameters, p and u, affect the Hicksian Demand curve when it is drawn on the typical P-Q axes.

[19] We are interested in the total change in consumption of bananas when the price of bananas goes up. In the real world, we don't compensate people when prices change. But, the Slutsky equation tells us that the total (uncompensated) effect of a change in the price of bananas is a combination of the substitution effect (compensated effect) and the wealth effect. This result should be familiar to you from your intermediate micro course. If it isn't, you may want to take a look at the (less abstract) treatment of this point in an intermediate micro text, such as Varian's Intermediate Microeconomics. Test of understanding: If bananas are a normal good, could demand for bananas ever rise when the price increases (i.e. could bananas be a Giffen good)? Answer using the Slutsky equation.

[20] Remember that the graphs are backwards, so a less negative slope $\frac{\partial h_i}{\partial p_j}$ is actually steeper.

Part 0: Recall that $\frac{\partial h_i}{\partial p_i} \leq 0$ because the Slutsky matrix is negative semi-definite. We'll assume, as is typical, that $\frac{\partial h_i}{\partial p_i} < 0$. To keep things simple, I'll omit the subscripts for the rest of the subsection, since we're always talk about a single good.

Part 1: Relationship between Hicksian and Walrasian Demand.

By duality, we know that through any point on Walrasian demand there is a Hicksian demand curve through that point. How do their slopes compare? This is given by the Slutsky equation. But, keep in mind two things. First, since we put p on the vertical axis and x on the horizontal axis, when we draw a graph, the derivatives of the demand functions aren't slopes. They're inverse slopes. That is, the slope of the Walrasian demand is $\frac{1}{\partial x / \partial p}$ and the slope of the Hicksian is $\frac{1}{\partial h / \partial p}$. Second, these derivatives are usually negative. So, we have to be a bit careful about thinking about quantities that are larger (i.e., further to the right on the number line) and quantities that are larger in magnitude (i.e., are further from zero on the number line). Since slopes are negative, a "larger" slope corresponds to a flatter curve. You'lls ee why this is important in a minute.

To figure out whether $x(p, w)$ or $h(p, u)$ through a point is steeper, use the Slutsky equation.

$$\frac{\partial h}{\partial p} = \frac{\partial x}{\partial p} + \frac{\partial x}{\partial w} x.$$

The answer will depend on whether x is normal or inferior. So, begin by considering a normal good. In this case, $\frac{\partial x}{\partial p} x > 0$, so:

$$\begin{aligned} \frac{\partial h}{\partial p} &> \frac{\partial x}{\partial p} \\ \left| \frac{\partial h}{\partial p} \right| &< \left| \frac{\partial x}{\partial p} \right| \\ \frac{1}{\left| \frac{\partial h}{\partial p} \right|} &> \frac{1}{\left| \frac{\partial x}{\partial p} \right|} \end{aligned}$$

The first line comes from the Slutsky equation and the fact that the income effect is positive. The second comes from the fact that, for a normal good, both sides are negative, and hence if $\frac{\partial h}{\partial p} > \frac{\partial x}{\partial p}$, $\frac{\partial h}{\partial p}$ is smaller in magnitude (absolute value) than $\frac{\partial x}{\partial p}$. The third follows from the second sinxe if $x < y$ and both are positve, then $1/x > 1/y$. Hence, for normal goods, the Hicksian Demand through a point is steeper than the Walrasian Demand through that point.

For an inferior good, things reverse. To simplify, suppose x is inferior but not Giffen (so that

$\frac{\partial x}{\partial p} < 0$ – you can do the Giffen case on your own). In this case $\frac{\partial x}{\partial p} x < 0$, so:

$$\begin{aligned} \frac{\partial h}{\partial p} &< \frac{\partial x}{\partial p} \\ \left|\frac{\partial h}{\partial p}\right| &> \left|\frac{\partial x}{\partial p}\right| \\ \frac{1}{\left|\frac{\partial h}{\partial p}\right|} &< \frac{1}{\left|\frac{\partial x}{\partial p}\right|} \end{aligned}$$

and so Hicksian demand is flatter than Walrasian Demand.

Part 2: Dependence of Hicksian demand on u.

How does changing u shift the Hicksian Demand curve? Again, the answer depends on whether the good is normal or inferior. To see how, use duality:

$$h(p,u) \equiv x(p, e(p,u)),$$

and differentiate both sides with respect to u :

$$\frac{\partial h}{\partial u} \equiv \frac{\partial x}{\partial w}\frac{\partial e}{\partial u}.$$

By the properties of the expenditure function, we know that $\frac{\partial e}{\partial u} > 0$ (see MWG Prop 3.E.2, p. 59), so that $\frac{\partial h}{\partial u}$ has the same sign as $\frac{\partial x}{\partial w}$. Hence, when the good is normal, increasing u increases Hicksian demand **for any price.** Thus, increasing u shifts the Hicksian demand curve to the right. Similarly, when the good is inferior, increasing u decreases Hicksian demand for any price, and thus increasing u shifts the Hicksian demand for an inferior good to the left. The intuition is that in order to achieve a higher utility level, the consumer must spend more, and consumption increases with expenditure for a normal good and decreases with expenditure for an inferior good.

A couple of pictures. These pictures depict Walrasian and Hicksian demand before and after a price decrease for a normal good and for an inferior good. Note that $p^1 < p^0$ so that $u^1 > u^0$. For the normal good, Hicksian demand is steeper than Walrasian, and shifts to the right when the price decreases. For the inferior good, Hicksian demand is flatter than Walrasian and shifts to the left when the price decreases.

Substitutes and Complements Revisited Remember when we studied the UMP, we said that goods i and j were gross complements or substitutes depending on whether $\frac{\partial x_i}{\partial p_j}$ was negative or positive? Well, notice that we could also classify goods according to whether $\frac{\partial h_i}{\partial p_j}$ is negative or positive. In fact, we will call goods i and j complements if $\frac{\partial h_i}{\partial p_j} < 0$ and substitutes if $\frac{\partial h_i}{\partial p_j} > 0$. That

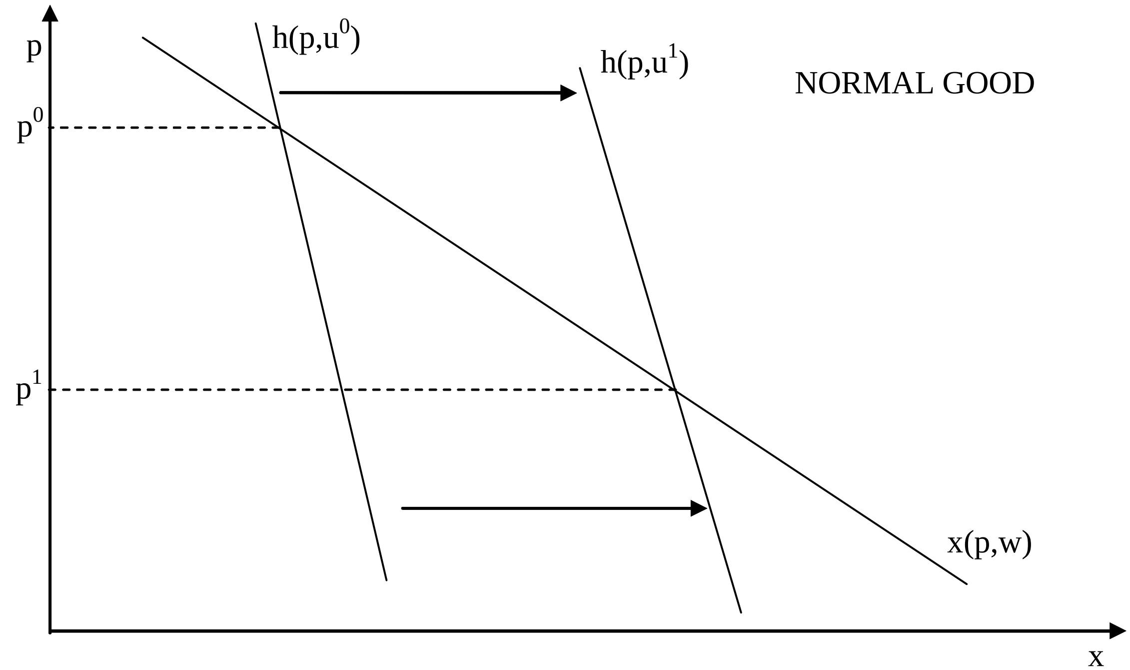

Figure 3.16:

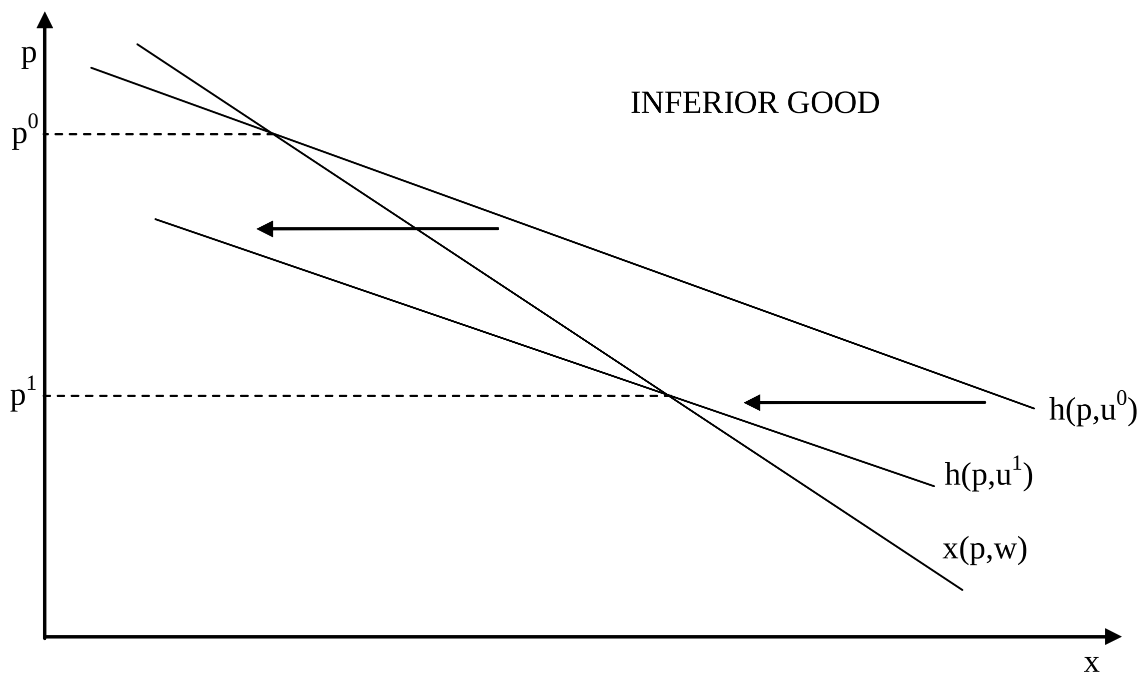

Figure 3.17:

is, we drop the "gross" when talking about the Hicksian demand function.[21] In many ways, the Hicksian demand function is the proper function to use to talk about substitutes and complements since it separates the question of wealth effects and substitution effects. For example, it is possible that good j is a gross complement for good i while good i is a gross substitute for good j (if good i is normal and good j is inferior), but no such thing is possible when talking about (just plain) complements or substitutes since $\frac{\partial h_i}{\partial p_j} = \frac{\partial h_j}{\partial p_i}$.

3.4.6 The EMP and the UMP: Summary of Relationships

The relationships between all of the parts of the EMP and the UMP are summarized in Figure 3.G.3 of MWG and similar figures appear in almost any other micro theory book. So, I urge you to look it over (with the proviso about the typo that I mentioned earlier).

Here, I'll do it in words. Start with the UMP.

$$\begin{aligned} &\max u(x) \\ s.t &: \quad p \cdot x \leq w. \end{aligned}$$

The solution to this problem is $x(p,w)$, the Walrasian demand functions. Substituting $x(p,w)$ into $u(x)$ gives the indirect utility function $v(p,w) \equiv u(x(p,w))$. By differentiating $v(p,w)$ with respect to p_i and w, we get Roy's identity, $x_i(p,w) \equiv -\frac{v_{p_i}}{v_w}$.

Now the EMP.

$$\begin{aligned} &\min p \cdot x \\ s.t. &: \quad u(x) \geq u. \end{aligned}$$

The solution to this problem is the Hicksian demand function $h(p,u)$, and the expenditure function is defined as $e(p,u) \equiv p \cdot h(p,u)$. Differentiating the expenditure function with respect to p_j gets you back to the Hicksian demand, $h_j(p,u) \equiv \frac{\partial e(p,u)}{\partial p_j}$.

The connections between the two problems are provided by the duality results. Since the same bundle that solves the UMP when prices are p and wealth is w solves the EMP when prices are p and the target utility level is $v(p,w)$, we have that

$$\begin{aligned} x(p,w) &\equiv h(p, v(p,w)) \\ h(p,u) &\equiv x(p, e(p,u)). \end{aligned}$$

[21] You can think of the 'gross' as referring to the fact that $\frac{\partial x_i}{\partial p_j}$ captures the effect of the price change before adding in the effect of compensation, sort of like how gross income is sales before adding in the effect of expenses.

Applying these identities to the expenditure and indirect utility functions yields more identities:

$$\begin{aligned} v(p, e(p, u)) &\equiv u \\ e(p, v(p, w)) &\equiv w. \end{aligned}$$

Note: These last equations are where the mistake is in the book. Finally, from the relationship between $x(p, w)$ and $h(p, u)$ we can derive the Slutsky equation:

$$\frac{\partial h_i(p, v(p, w))}{\partial p_j} \equiv \frac{\partial x_i(p, w)}{\partial p_j} + \frac{\partial x_i(p, w)}{\partial w} x_j.$$

If you are really interested in such things, there is also a way to recover the utility function from the expenditure function (see a topic in MWG called "integrability"), but I'm not going to go into that here.

3.4.7 Welfare Evaluation

Underlying our approach to the study of preferences has been the ultimate goal of developing a tool for the welfare evaluation of policy changes. Recall that:

1. The UMP leads to $x(p, w)$ and $v(p, w)$, which are at least in principle observable. However, $v(p, w)$ is not a good tool for welfare analysis.
2. The EMP leads to $h(p, u)$ and $e(p, u)$, which are based on unobservables (u) but provide a good measure for the change in a consumer's welfare following a policy change.
3. The Slutsky equation provides the link between the observable concepts, $x(p, w)$, and the useful concepts, $h(p, u)$.

In this section, we explore how these tools can be used for welfare analysis. The neoclassical preference-based approach to consumer theory gives us a measure of consumer well-being, both in terms of utility and in terms of the wealth needed to achieve a certain level of well-being. It turns out that this is crucial for welfare evaluation.

We will consider a consumer with "well-behaved" preferences (i.e. a strictly increasing, strictly quasiconcave utility function). The example we will focus on is the welfare impact of a price change.

Consider a consumer who has wealth w and faces initial prices p^0. Utility at this point is given by

$$v\left(p^0, w\right).$$

If prices change to p^1, the consumer's utility at the new prices is given by:

$$v\left(p^1, w\right).$$

Thus the consumer's utility increases, stays constant, or decreases depending on whether:

$$v\left(p^1, w\right) - v\left(p^0, w\right)$$

is positive, equal to zero, or negative.

While looking at the change in utility can tell you whether the consumer is better off or not, it cannot tell you how much better off the consumer is made. This is because utility is an ordinal concept. The units that utility is measured in are arbitrary. Thus it is meaningless to compare, for example, $v\left(p^1, w\right) - v\left(p^0, w\right)$ and $v\left(p_2, w\right) - v\left(p_3, w\right)$. And, if $v\left(\right)$ and $y\left(\right)$ are the indirect utility functions of two people, it is also meaningless to compare the change in v to the change in y.

However, suppose we were to compare, instead of the direct utility earned at a particular price-wealth pair, the wealth needed to achieve a certain level of utility at a given price-wealth pair. To see how this works, let

$$\begin{aligned} u^1 &= v\left(p^1, w\right) \\ u^0 &= v\left(p^0, w\right). \end{aligned}$$

We are interested in comparing the expenditure needed to achieve u^1 or u^0. Of course, this will depend on the particular prices we use. It turns out that we have broad latitude to choose whichever set of prices we want, so let's call the reference price vector p^{ref}, and we'll assume that it is strictly greater than zero on all components.

The expenditure needed to achieve utility level u at prices p^{ref} is just

$$e\left(p^{ref}, u\right).$$

Thus, if we want to compare the expenditure needed to achieve utility u^0 and u^1, this is given by:

$$\begin{aligned} & e\left(p^{ref}, u^1\right) - e\left(p^{ref}, u^0\right) \\ & e\left(p^{ref}, v\left(p^1, w\right)\right) - e\left(p^{ref}, v\left(p^0, w\right)\right). \end{aligned}$$

This expression will be positive whenever it takes more wealth to achieve utility u^1 at prices p^{ref} than to achieve u^0. Hence this expression will also be positive, zero, or negative depending on whether $u^1 > u^0$, $u^1 = u^0$, or $u^1 < u^0$. However, the units now have meaning. The difference is measured in dollar terms. Because of this, $e\left(p^{ref}, v(p,w)\right)$ is often called a **money metric indirect utility function.**

We can construct a money metric indirect utility function using virtually any strictly positive price as the reference price p^{ref}. However, there are two natural candidates: the original price, p^0, and the new price, p^1. When $p^{ref} = p^0$, the change in expenditure is equal to the change in wealth such that the consumer would be indifferent between the new price with the old wealth and the old price with the new wealth. Thus it asks what change in wealth would be equivalent to the change in price. Formally, define the **equivalent variation**, $EV\left(p^0, p^1, w\right)$, as

$$EV\left(p^0, p^1, w\right) = e\left(p^0, v\left(p^1, w\right)\right) - e\left(p^0, v\left(p^0, w\right)\right) = e\left(p^0, v\left(p^1, w\right)\right) - w.$$

since $e\left(p^0, v\left(p^0, w\right)\right) = w$. Equivalent variation is illustrated in MWG Figure 3.I.2, panel a. Notice that the compensation takes place at the old prices – the budget line shifts parallel to the one for $\left(p^0, w\right)$.

Since $w = e\left(p^1, v\left(p^1, w\right)\right)$, an alternative definition of EV would be:

$$EV\left(p^0, p^1, w\right) = e\left(p^0, v\left(p^1, w\right)\right) - e\left(p^1, v\left(p^1, w\right)\right).$$

In this form, EV asks how much more money does it take to achieve utility level $v\left(p^1, w\right)$ at p^0 than at p^1. Note: if $EV < 0$, this means that it takes less money to achieve utility $v\left(p^1, w\right)$ at p^0 than p^1 (which means that prices have gone up to get to p^1, at least on average).

When considering the case where the price of only one good changes, EV has a useful interpretation in terms of the Hicksian demand curve. Applying the fundamental theorem of calculus and the fact that $\frac{\partial e(p,u)}{\partial p_i} = h_i(p,u)$, if only the price of good 1 changes, we have:[22]

$$e\left(p^0, v\left(p^1, w\right)\right) - e\left(p^1, v\left(p^1, w\right)\right) = \int_{p_1^1}^{p_1^0} h_1\left(s, p_{-1}^0, v\left(p^1, w\right)\right) ds$$

Thus the absolute value of EV is given by the area to the left of the Hicksian demand curve between p_1^0 and p_1^1. If $p_1^0 < p_1^1$, EV is negative - a welfare loss because prices went up. If $p_1^0 > p_1^1$, EV is

[22] Often when we are interested in a particular component of a vector - say, the price of good i - we will write the vector as (p_i, p_{-i}), where p_{-i} consists of all the other components of the price vector. Thus, (p_i^*, p_{-i}) stands for the vector $(p_1, p_2, ..., p_{i-1}, p_i^*, p_{i+1}, ..., p_L)$. It's just a shorthand notation.

Another notational explanation - in an expression such as p_1^0, the superscript refers to the timing of the price vector (i.e. new or old prices), and the subscript refers to the commodity. Thus, p_1^0 is the old price of good 1.

positive - a welfare gain because prices went down. The relevant area is depicted in MWG Figure 3.I.3, panel a.

The other case to consider is the one where the new price is taken as the reference price. When $p^{ref} = p^1$, the change in expenditure is equal to the change in wealth such that the consumer is indifferent between the original situation (p^0, w) and the new situation $(p^1, w + \Delta w)$. Thus it asks how much wealth would be needed to compensate the consumer for the price change. Formally, define the **compensating variation** (depicted in MWG Figure 3.I.2, panel b)

$$CV\left(p^0, p^1, w\right) = e\left(p^1, v\left(p^1, w\right)\right) - e\left(p^1, v\left(p^0, w\right)\right) = w - e\left(p^1, v\left(p^0, w\right)\right).$$

Again, when only one price changes, we can readily interpret CV in terms of the area to the left of a Hicksian demand curve. However, this time it is the Hicksian demand curve for the old utility level, u^0. To see why, note that $w = e\left(p^0, v\left(p^0, w\right)\right)$, and so (again assuming only the price of good 1 changes):

$$CV\left(p^0, p^1, w\right) = e\left(p^0, v\left(p^0, w\right)\right) - e\left(p^1, v\left(p^0, w\right)\right) = \int_{p_1^1}^{p_1^0} h_1\left(s, p_{-1}^0, v\left(p^0, w\right)\right) ds,$$

which is positive whenever $p_1^0 > p_1^1$ and negative whenever $p_1^0 < p_1^1$. The relevant area is illustrated in MWG Figure 3.I.3, panel b.

Recall that whenever good i is a normal good, increasing the target utility level u shifts $h_i(p_i, \bar{p}_{-i}, u)$ to the right in the (x_i, p_i) space. This is because in order to achieve higher utility the consumer will need to spend more wealth, and if the good is normal and the consumer spends more wealth, more of the good will be consumed. Thus when the good is normal, $EV \geq CV$. On the other hand, if the good is inferior, then increasing u shifts $h_i(p_i, \bar{p}_{-i}, u)$ to the left, and $CV \geq EV$. When there is no wealth effect on the good, i.e., $\frac{\partial x_i(p,w)}{\partial w} = 0$, then $CV = EV$.

Figure 3.I.3 also shows the Walrasian demand curve. In fact, it shows it crossing $h\left(p_1, p_{-1}^0, v\left(p^1, w\right)\right)$ at p_1^1 and $h_1\left(p_1, p_{-1}^0, v\left(p^0, w\right)\right)$ at p_1^0. This results from the duality of utility maximization and expenditure minimization. Formally, we have the equalities

$$\begin{aligned} h_1\left(p^0, v\left(p^0, w\right)\right) &= x\left(p^0, w\right) \\ h_1\left(p^1, v\left(p^1, w\right)\right) &= x\left(p^1, w\right), \end{aligned}$$

which each arise from the identity $h_i(p, v(p, w)) \equiv x_i(p, w)$. The result of this is that the Walrasian demand curve crosses the Hicksian demand curves at the two points mentioned above, and that the area to the left of the Walrasian demand curve lies somewhere between the EV and CV. There are a number of comments that must be made on this topic:

1. Although the area to the left of the Hicksian demand curve is equal to the change in the expenditure function, the area to the left of the Walrasian demand function has no ready interpretation.

2. The area to the left of the Walrasian demand curve is called the change in Marshallian consumer surplus, ΔCS, and is probably the notion of welfare change that you are used to from your intermediate micro courses.

3. Unfortunately, the change in Marshallian consumer surplus is a meaningless measure (see part 1) except for:

   (a) If there are no wealth effects on the good whose price changes, then $EV = CV = \Delta CS$.

   (b) Since ΔCS lies between EV and CV, it can sometimes be a good approximation of the welfare impact of a price change. This is especially true if wealth effects are small.

4. Some might argue that ΔCS is a useful concept because it is easier to compute than EV or CV since it does not require estimation of the Hicksian demand curves. But, if you know about the Slutsky equation (which you do), this isn't such a problem.

So Which is Better, EV or CV?

Both EV and CV provide dollar measures of the impact of a price change on consumer welfare, and there are circumstances in which each is the appropriate measure to use. EV does have one advantage over CV, though, and that is that if you want to consider two alternative price changes, EV gives you a meaningful measure, while CV does not (necessarily). For example, consider initial price p^0 and two alternative price vectors p^a and p^b. The quantities $EV\left(p^0, p^a, w\right)$ and $EV\left(p^0, p^b, w\right)$ are both measured in terms of wealth at prices p^0 and thus they can be compared. On the other hand, $CV\left(p^0, p^a, w\right)$ is in terms of wealth at prices p^a and $CV\left(p^0, p^b, w\right)$ is in terms of wealth needed at prices p^b, which cannot be readily compared.

This distinction is important in policy issues such as deciding which commodity to tax. The impact of placing a tax on gasoline vs. the impact of placing a tax on electricity needs to be measured with respect to the same reference price if we want to compare the two in a meaningful way. This means using EV.

Example: Deadweight Loss of Taxation.

Suppose that the government is considering putting a tax of $t > 0$ dollars on commodity 1. The current price vector is p^0. Thus the new price vector is $p^1 = \left(p_1^0 + t, p_2^0, ..., p_L^0\right)$.

After the tax is imposed, consumers purchase $h_1\left(p^1, u^1\right)$ units of the good, where $u^1 = v\left(p^1, w\right)$. The tax revenue raised by the government is therefore $T = th_1\left(p^1, u^1\right)$. However, in order to raise this T dollars, the government must increase the effective price of good 1. This makes consumers worse off, and the amount by which it makes consumers worse off is given by:

$$EV\left(p^0, p^1, w\right) = \int_{p_1^1}^{p_1^0} h_1\left(s, p_{-1}^0, u^1\right) ds.$$

Since $p_1^1 > p_1^0$, EV is negative and gives the amount of money that consumers would be willing to pay in order to avoid the tax. Thus consumers are made worse off by $EV\left(p^0, p^1, w\right)$ dollars. Since the tax raises T dollars, the net impact of the tax is

$$-EV\left(p^0, p^1, w\right) - T.$$

The previous expression, known as the **deadweight loss (DWL)** of taxation, gives the amount by which consumers would have been better off, measured in dollar terms, if the government had just taken T dollars away from them instead of imposing a tax. To put it another way, consumers see the tax as equivalent to losing EV dollars of income. Since the tax only raises T dollars of income, $-EV - T$ is the dollar value of the consumers' loss that is not transferred to the government as tax revenue. It simply disappears.

Well, it doesn't really disappear. Consumers get utility from consuming the good. In response to the tax, consumers decrease their consumption of the good, and this decreases their utility and is the source of the deadweight loss. On the other hand, a tax that does not distort the price consumers must pay for the good would not change their compensated demand for the good. Consequently, it would not lead to a deadweight loss. This is one argument for lump-sum taxes instead of per-unit taxes. Lump-sum taxes (each consumer pays T dollars, regardless of the consumption bundle each one purchases) do not distort consumers' purchases, and so they do not lead to deadweight losses. However, lump-sum taxes have problems of their own. First, they are regressive, meaning that they impact the poor more than the rich, since everybody must pay the same amount. Second, lump-sum taxes do not charge the users of commodities directly. So, there is some question whether, for example, money to pay for building and maintaining roads should be raised by charging everybody the same amount or by charging a gasoline tax or by charging drivers a toll each time they use the road. The lump-sum tax is non-distortionary, but it must be paid by people who don't drive, even people who can't afford to drive. The gasoline tax is paid by all drivers, including people who don't use the particular roads being repaired, and it is distortionary in the sense that people will generally reduce their driving in response to the tax, which induces a deadweight loss. Charging a toll to those who use the road places the burden of paying for repairs on exactly those who are benefiting from having the roads. But like the gasoline tax, it is also distortionary (since people will tend to avoid toll roads). And, since the tolls are focused on relatively few consumers, the tolls may have to be quite high in order to raise the necessary funds, imposing a large burden on those people who cannot avoid using the toll roads. These are just some of the issues that must be considered in deciding which commodities should be taxed and how.

3.4.8 Bringing It All Together

Recall the basic dilemma we faced. The UMP yields solution $x(p,w)$ and value function $v(p,w)$ that are based on observables but not useful for doing welfare evaluation since utility is ordinal. The EMP yields solution $h(p,u)$ and value function $e(p,u)$, which can be used for welfare evaluation but are based on u, which is unobservable. As I have said, the link between the two is provided by the Slutsky equation

$$\frac{\partial h_i(p, v(p,w))}{\partial p_j} = \frac{\partial x_i(p,w)}{\partial p_j} + \frac{\partial x_i(p,w)}{\partial w} x_i(p,w).$$

We now illustrate how this is implemented. Suppose the price of good 1 changes. EV is given by:

$$EV\left(p^0, p^1, w\right) = \int_{p_1^1}^{p_1^0} h_1\left(s, p_{-1}^0, u^1\right) ds.$$

We can approximate $h_1\left(s, p_{-1}^0, u^1\right)$ using a first-order Taylor approximation.

Recall, a first-order Taylor approximation for a function $f(x)$ at point x_0 is given by:

$$\tilde{f}(x) \cong f(x_0) + f'(x_0)(x - x_0).$$

This gives a linear approximation to $f(x)$ that is tangent to $f(x)$ at x_0 and a good approximation for x that are not too far from x_0. But, the further x is away from x_0, the worse the approximation will be. See Figure 3.18.

Now, the first-order Taylor approximation to $h\left(s, p_{-1}^0, u^1\right)$ is given by:

$$h_1\left(s, p_{-1}^0, u^1\right) \cong h_1\left(p_1^1, p_{-1}^0, u^1\right) + \frac{\partial h_1(p, v(p,w))}{\partial p_1}\left(s - p_1^1\right) \tag{3.6}$$

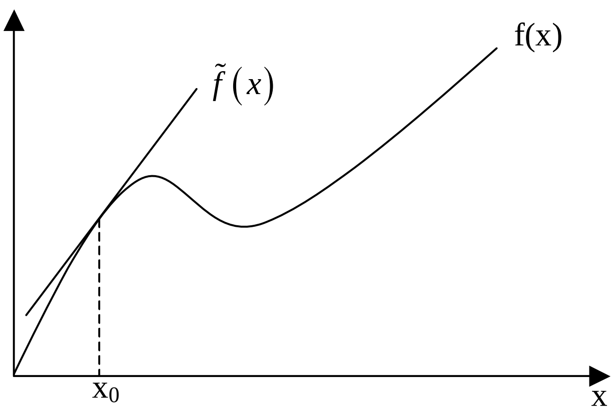

Figure 3.18: A First-Order Taylor Approximation

Note that we have taken as our original point $p = \left(p_1^1, p_{-1}^0\right)$. That is, the price vector after the price change? Why do we do this? The reason is that we are using the Hicksian demand curve for u^1, the utility level after the price change. Because of that, we also want to use the price after the price change. We know that, at $p = \left(p_1^1, p_{-1}^0\right)$, $h\left(p_1^1, p_{-1}^0, u^1\right) = x_1\left(p_1^1, p_{-1}^0, w\right)$. This fact, along with the Slutsky equation, allows us to rewrite (3.6) as:

$$\begin{aligned} \tilde{h}_1\left(s, p_{-1}^0, u^1\right) &\cong x_1\left(p_1^1, p_{-1}^0, w\right) \\ &+ \left(\frac{\partial x_1\left(p_1^1, p_{-1}^0, w\right)}{\partial p_1} + \frac{\partial x_1\left(p_1^1, p_{-1}^0, w\right)}{\partial w} x_1\left(p_1^1, p_{-1}^0, w\right)\right)\left(s - p_1^1\right). \end{aligned} \tag{3.7}$$

The last equation provides an approximation for the Hicksian demand curve based only on observable quantities. That is, we have eliminated the need to know the (unobservable) target utility level. Finally, note that demand $x_1\left(p_1^1, p_{-1}^0, w\right)$ and derivatives $\frac{\partial x_1\left(p_1^1, p_{-1}^0, w\right)}{\partial p_1}$ and $\frac{\partial x_1\left(p_1^1, p_{-1}^0, w\right)}{\partial w}$ can be observed or approximated using econometric techniques. Note that the difference between this approximation and one based on the Walrasian demand curve is the addition of the wealth-effect term, $\frac{\partial x_1\left(p_1^1, p_{-1}^0, w\right)}{\partial w} x_1\left(p_1^1, p_{-1}^0, w\right)$.

Figure 3.19 illustrates the first-order Taylor approximation to EV. Since the "original point" in our estimate to the Hicksian demand function is $p = \left(p_1^1, p_{-1}^0\right)$, estimated Hicksian demand $\tilde{h}_1$ is coincides with and it tangent to the actual Hicksian demand h_1 at this point. As you move to prices that are further away from p_1^1, the approximation is less good. True EV is the area left of h_1. Thus, the estimation "error", the difference between true EV and estimated EV, is given by the area between $\tilde{h}_1$ and h_1 between prices p_0^1 and p_1^1.

In the case of CV, CV is computed as the area left of the Hicksian demand curve at the original

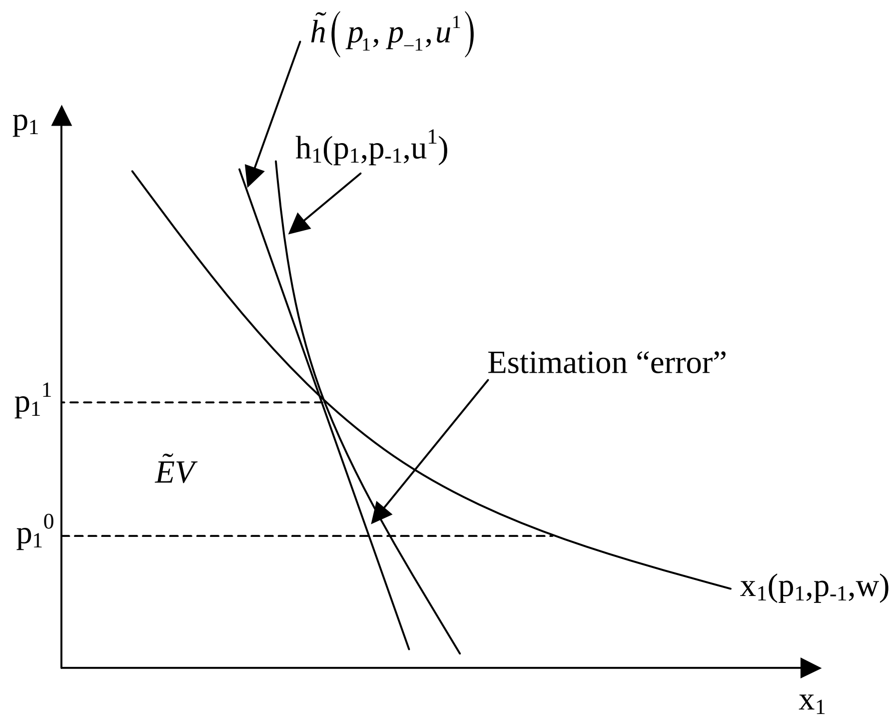

Figure 3.19: The First-Order Taylor Approximation to EV

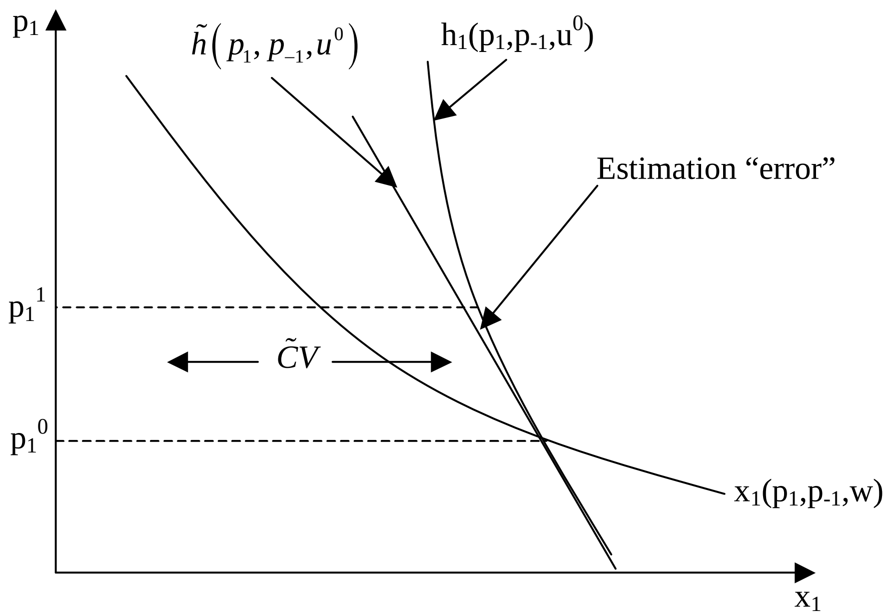

Figure 3.20: The First-Order Taylor Approximation to CV

utility level, $h_1\left(p_1, p_{-1}, u^0\right)$. Because of this, we must use the original price as the "original point" in the Taylor approximation. Thus, for the purposes of CV, estimated Hicksian demand is given by:

$$\begin{aligned} h_1\left(s, p_{-1}^0, u^1\right) &\cong h_1\left(p_1^0, p_{-1}^0, u^1\right) + \frac{\partial h_1\left(p, v\left(p, w\right)\right)}{\partial p_1}\left(s - p_1^0\right) \\ &\cong x_1\left(p_1^0, p_{-1}^0, w\right) \\ &\quad + \left(\frac{\partial x_1\left(p_1^0, p_{-1}^0, w\right)}{\partial p_1} + \frac{\partial x_1\left(p_1^0, p_{-1}^0, w\right)}{\partial w} x_1\left(p_1^0, p_{-1}^0, w\right)\right)\left(s - p_1^0\right). \end{aligned}$$

The diagram for the Taylor approximation to CV corresponding to Figure 3.19 therefore looks like Figure 3.20:

Welfare Evaluation: An Example

Data: $x_1^0 = 100$, $p_1^0 = 10$, $\frac{\partial x}{\partial p} = -4$, $\frac{\partial x}{\partial w} = 0.02$. Note: there is no information on w.

Suppose the price of good 1 increases to $p' = 12.5$. How much should a public assistance program aimed at maintaining a certain standard of living be increased to offset this price increase?

To answer this question, we are looking for the CV of the price change. To compute this, we need to approximate the Hicksian demand curve for the original utility level, $h_1\left(p_1, p_{-1}, u^0\right)$.

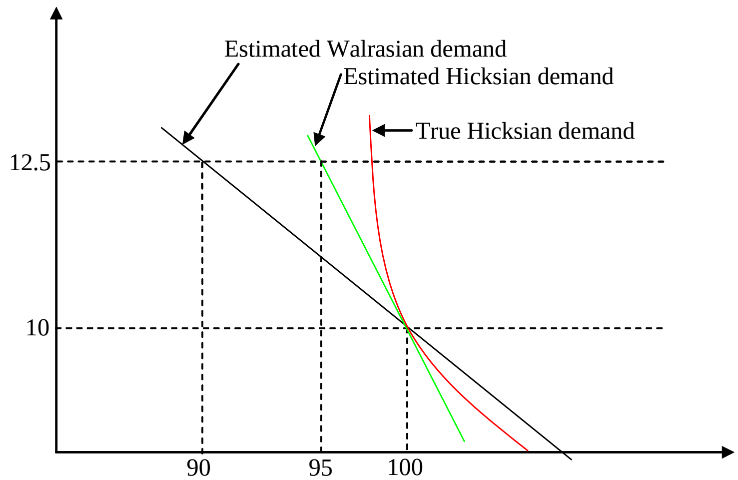

Figure 3.21:

1. We know that $h_1\left(10, p_{-1}, u^0\right) = x_1\left(10, p_{-1}, w\right)$.

2. The slope of the $h_1\left(p_1, p_{-1}, u^0\right)$ can be approximated using the data and the Slutsky equation.

$$\begin{aligned}\frac{\partial h}{\partial p} &= \frac{\partial x}{\partial p} + \frac{\partial x}{\partial w}x \\ &= -4 + 0.02\,(100) \\ &= -2\end{aligned}$$

3. So, at price 12.5, Hicksian demand is given by

$$\begin{aligned}h &= 100 + \frac{\partial h}{\partial p}dp \\ &= 100 + (-2)\,2.5 = 95\end{aligned}$$

4. To compute CV, compute the area of a trapezoid (or the area of a rectangle plus a triangle):

$$|CV| = (2.5)\left(\frac{95+100}{2}\right) = 243.75.$$

Since the price is increasing, we know that $CV < 0$, so $CV = -243.75$.

We could also estimate the change in Marshallian Consumer Surplus. This is just the area to the left of the Walrasian demand curve between the two prices. Hence $\Delta CS = -(2.5)\left(\frac{90+100}{2}\right) = -237.5$. Hence if we were to use the Marshallian consumer surplus in this case, we would not compensate the consumer enough for the price increase.

Another thing we could do is figure that the harm done to the consumer is just the change in price times the original consumption of this good, i.e., $2.5\,(100) = 250$. However, if we gave the consumer 250 additional dollars, we would be overcompensating for the price increase.

3.4.9 Welfare Evaluation for an Arbitrary Price Change

The basic analysis of welfare change using CV and EV considers the case of a single price change. However, what should we do if the policy change is not a single price change? For changes in multiple prices, we can just compute the CV for each of the changes (i.e., changing prices one by one and adding the CV (or EV) from each of the changes along a "path" from the original price to the new price). If price and wealth change, we can add the change in wealth to the CV (or EV) from the price changes (see below). But, what if the policy change involves something other than prices and wealth, such a change in environmental quality, roads, etc. How do we value such a change?

The answer is that, if we have good estimates of Walrasian demand, we can always represent the change as a change in a budget set. After doing so, we can compute the CV is the usual way.

Part 1: Any arbitrary policy change can be thought of as a simultaneous change in p and w.

To illustrate, suppose that we have a good estimate of consumers' demand functions (i.e., we fit a flexible functional form for demand using high-quality data). Let $x\,(p, w)$ denote demand. Suppose that initially prices and wealth are $\left(p^0, w^0\right)$ and the consumer chooses bundle $x\left(p^0, w^0\right)$. Now, suppose that "something happens" that leads the consumer to choose bundle x' instead of x^0. What is the CV (or EV) of this change?

The first step is to note that, if demand is quasiconcave, there is some price-wealth vector for which x^0 and x' are optimal choices. You can find these price-wealth vectors, which we'll call $\left(p^0, w^0\right)$ and (p', w'), by solving the equations $x^0 = \left(p^0, w^0\right)$ and $x' = x\,(p', w')$. (In reality you probably already know $\left(p^0, w^0\right)$ and have an observation of x' or estimate of.) Remember, we have a good estimate of $x\,(p, w)$. Once we find (p', w'), then we know that the change in the consumer's utility in going from x^0 to $x^{'}$ is just $v\,(p', w') - v\left(p^0, w^0\right)$, and so the impact of the policy change reduces to computing the EV or CV for this simultaneous change in p and w. Let $v\,(p', w') = u'$ and $v\left(p^0, w^0\right) = u^0$.

Part 2: Compute the EV or CV for a simultaneous change in p and w.

So, we've recast the policy change as a change from $\left(p^0, w^0\right)$ to (p', w'), letting u^0 and u' denote the utility levels before and after the change. To compute EV, return to the definition of EV we used before.

$$EV = e\left(p^0, u'\right) - e\left(p^0, u^0\right)$$

Adding and subtracting $e(p', u')$, we get:

$$EV = \left[e\left(\ p^0, u'\right) - e\left(\ p', u'\right)\right] + \left[e\left(p', u'\right) - e\left(p^0, u^0\right)\right].$$

But, note that $e(p', u') = w'$ and $e\left(p^0, u^0\right) = w^0$, so

$$EV = \left[e\left(\ p^0, u'\right) - e\left(\ p', u'\right)\right] + w' - w^0, \qquad (*)$$

and note that $\left[e\left(p^0, u'\right) - e(p', u')\right]$ is as in the definition of EV when only a price changes. So, if only the price of good 1 changes, EV can be written as:

$$EV = \int_{p_1'}^{p_1^0} h_1\left(s, p_{-1}, u'\right) ds + \left(w' - w^0\right),$$

and this can be estimated in the usual way from the estimated Walrasian demand curve.

If multiple prices change, we change them one by one and add up the integral from each change, and then we add the change in wealth. That is, if prices change from $\left(p_1^0, p_2^0, ..., p_L^0\right)$ to $(p_1', p_2', ..., p_L')$ and wealth changes from w^0 to w', the EV is:

$$EV = \int_{p_1'}^{p_1^0} h_1\left(s, p_2^0, ..., p_L^0\right) ds + \int_{p_2'}^{p_2^0} h_2\left(p_1', s, p_3^0, ..., p_L^0\right) ds + ... + \int_{p_L'}^{p_L^0} h_2\left(p_1', p_2', ..., p_{L-1}', s\right) ds + w' - w^0.$$

If you replace each Hicksian demand with an estimate based on Marshallian demand and the Slutsky equation, you can estimate this using only observables. It is tedious, but certainly possible.

This is a diagram that illustrates the whole thing. Suppose a policy change moves the consumer's consumption bundle from x^0 to x'. To compute the EV, the first thing you do is find the (p, w) for which $x^0 = x\left(p^0, w^0\right)$. This budget set is labeled $B\left(p^0, w^0\right)$. Then, you find the (p, w) for which x' is optimal, which we call (p', w'). This budget set (red) is labeled $B(p', w')$. Denote the initial utility level u^0 and the final utility level u', and note that neither the utility levels nor the indifference curves (which are drawn in as dotted lines for illustration) are observed.

Next, we decompose the change from $\left(p^0, w^0\right)$ to (p', w') into two parts. Part 1 is a change in wealth holding prices fixed at p^0. Let y denote the point the consumer chooses at $\left(p^0, w'\right)$, and let u^y denote the utiltiy earned. This point and the associated budget set are in blue. Note that moving from budget set $B\left(p^0, w^0\right)$ to budget set $B\left(p^0, w'\right)$ is just like losing $w' - w^0$ dollars (since prices don't change this is, in fact, exactly what happens). This is where the $\left(w' - w^0\right)$ term comes from in expression (*) above. Distance $w' - w^0$ is denoted on the left. (Note that in the diagram, these distances are scaled by p_2, since we are showing them on the x_2-axis.)

Part 2 of the decomposition is the change in prices from p^0 to p' when wealth is w'. But, note that this just the kind of EV we computed in the simple case. That is, prices change but wealth remains constant. Let z denote the point that offers the same utility as x' but is chosen at prices p'. That is, $z = x\left(p^0, w' + EV\left(p^0, p'.w'\right)\right)$. The budget line supporting z is denoted in green, and the EV for the price change from p^0 to p' at wealth w' is just the distance that the budget shifts up from blue $B\left(p^0, w'\right)$ to green $B\left(p^0, w' + EV\left(p^0, p'.w'\right)\right)$ denoted $EV\left(p^0, p', w'\right)$ on the left. Since

$$\begin{aligned} EV\left(p^0, p', w'\right) &= e\left(p^0, u'\right) - e\left(p^0, u^y\right) \\ &= e\left(p^0, u'\right) - w' \\ &= e\left(p^0, u'\right) - e\left(p', u'\right), \end{aligned}$$

this is just like the EV's we computed when only prices changed. This is where the $\left[e\left(p^0, u'\right) - e\left(p', u'\right)\right]$ comes from in expression (*) above.

The total EV is the sum of these two parts. The distance is denoted Total EV in the diagram. Note that since the consumer ends up worse off overall, the total EV should be negative.

You could also do something similar for CV.

$$\begin{aligned} CV &= e\left(p', u'\right) - e\left(p', u^0\right) \\ &= \left[e\left(p', u'\right) - e\left(p^0, u^0\right)\right] + \left[e\left(p^0, u^0\right) - e\left(p', u^0\right)\right] \\ &= w' - w^0 + \left[e\left(p^0, u^0\right) - e\left(p', u^0\right)\right], \end{aligned}$$

and note that once again $\left[e\left(p^0, u^0\right) - e\left(p', u^0\right)\right]$ is as in our original definition of CV. So, this term can be rewritten in terms of integrals of Hicksian demand curves at utility level u^0.

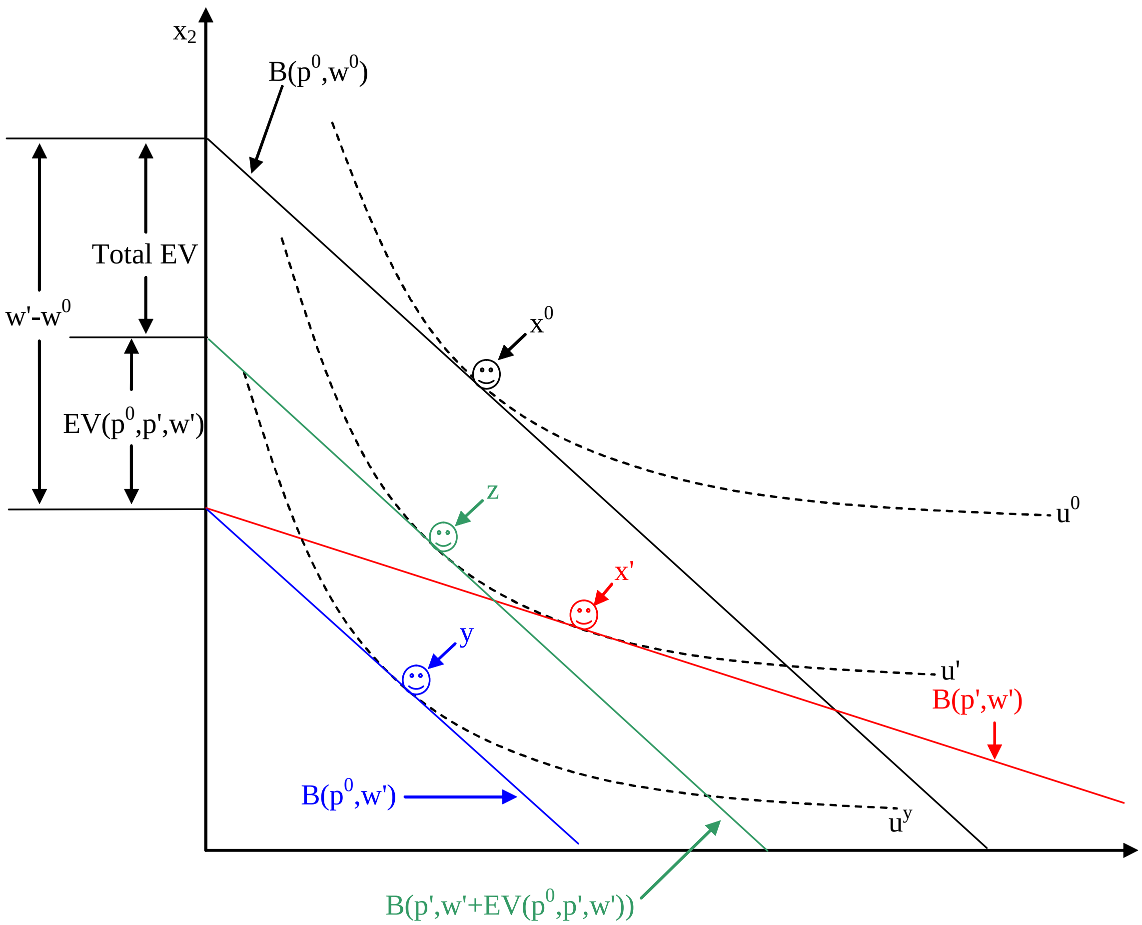

Figure 3.22: