ECON5110: Microeconomics

Lecture 3: Sept, 2017

Contents

1 Overview

Having now analyzed the production side of the economy we turn to its converse, the consumption side. As will become apparent, much of the forthcoming analysis is conceptually similar to what we saw on the production side. Just as our analysis of the firm was predicated on a behavioral assumption, profit maximization, we will begin by ascribing a to consumers a criterion for making decisions: utility maximization.

Utility is a nebulous catch all phrase intended to refer numerically to the abstract well being of a consumer. An action or series of actions results in some state-of-affairs which imparts some utility to the consumer, and we assume as a foundational principle that consumers prefer states-of-affairs that provide higher utility to those that provide lower. That is, letting X comprise all possible outcomes, the agent entertains a function $U : X \to \mathbb{R}$ such that she would like to choose actions so as maximize the value of U.

This line of reasoning should sound suspicious if not down right circular, as we must take care to assure our assumptions rest of firm ground. Unlike profit which is generally represented in some tangible method of accounting, we have no ability to view a consumers utility. Even more damningly: while observation dictates that firms do actively try to maximize profit, intuition about our own decision making fails to provide an strong case for the maximization of any function or the existence of a numeric measure of happiness.

Indeed, utility does not exist. What does exist are peoples preferences between alternatives (I prefer mint chip ice cream to strawberry), and their choices (when faced with mint chip or strawberry, which do I pick?). The story of the consumer begins with preferences and choices, and uses these (at least potentially) observable data to construct the notion of utility. Utility becomes another way of representing tangible preference. But, it is a very helpful representation, as it will allow us to make use of many of the same techniques employed in the theory of the firm.

2 Preference

We start the analysis of consumer behavior by considering the consumers **preference**. As a requisite, we must state what are these object over which her preference is defined. Towards this, let X denote the **consumption space**. X represents all of the potential states-of-affairs which might be enacted by the actions of the DM. A simpler, less subtle, interpretation is that X is a set of objects which the agent can literally consume, for example, pork, bananas, holiday packages, flat-shares, etc. For the moment, we will take X to be an abstract set, but we will often add structure. Most commonly, we might fix $X = \mathbb{R}^k$, where each dimension k represents a type of consumption, and the magnitude of the k^{th} dimension represents how much of the k^{th} object is consumed.

A **preference**, denoted by $\succcurlyeq$ is a relation over X—recall this means $\succcurlyeq \subset X \times X$ and we write $x \succcurlyeq y$ iff $(x, y) \in \succcurlyeq$. The interpretation, as the notation suggests, is that $x \succcurlyeq y$ holds if the agent prefers the outcome x to the outcome y. Define $\sim$ and $\succ$ as the symmetric and asymmetric parts of $\succcurlyeq$. While utility is a fable, a preference can, at least in theory, be observed by simply asking the agent about each pair of objects.[1] The goal of the exercise, recall, is to work analytically with a utility function, $U : X \to R$, all the while justifying utility via its consistency with underlying preference. What does consistency mean? A utility function, $U : X \to \mathbb{R}$ **represents** a preference $\succcurlyeq \subset X \times X$ iff

$$U(x) \geq U(y) \iff x \succcurlyeq y.$$

[1] We will dive into more realistic methods of identification in the next section.

Lets examine what it means for a utility function represents a preference relation. It allows us to move back and forth between preferences and utility functions seamlessly: to use the preference paradigm to test our hypotheses with observable data and also use the tools associated with maximization of a function. For example, if B is the set of affordable consumption objects and if $x \in \arg\max_B U$ then we know $x \succcurlyeq y$ for all $y \in B$: according to the agents preferences, x is the best. Recall, from the pervious lecture that if x maximizes U than x maximizes $h \circ U$ for any strictly monotone h. This is really a special case of the following result.

Theorem 1. *Let U represent $\succcurlyeq$. Then V represents $\succcurlyeq$ if and only if there exists a strictly monotone $h : \textsc{im}(U) \to \mathbb{R}$ such that $V = h \circ U$.*

Theorem 1, while mathematically simple, has an important implication to the interpretation of utility. The numerical representation of preferences contains ordinal (i.e., order) information, but does *not* contain cardinal (i.e., magnitude) information. Say X is three elements with $x \succcurlyeq y \succcurlyeq z$. Then $U(x) = 3$, $U(y) = 2$, $U(z) = 1$ represents $\succcurlyeq$ just as faithfully as $U(x) = 3$, $U(y) = 2$, $U(z) = -1000000$. This means, (i) we should not read too much the size of the number, utility is really just another way of talking about $\succcurlyeq$, and (ii) if it is convenient to do so, we can alway take monotone transformations of U.

Recall from past economics classes that an indifference curve traces out all bundles of goods that keep the consumer equally happy. Now, we see this is just $I(x) = \{y \in X \mid y \sim x\}$. Likewise, we can define the set of alternatives that are preferred to, and dis-preferred to, a particular element. Define $UCS, LCS : X \to \mathcal{P}(X)$ as the maps

$$\begin{aligned} UCS :& x \mapsto \{y \in X \mid y \succcurlyeq x\}, \\ LCS :& x \mapsto \{y \in X \mid x \succcurlyeq y\}. \end{aligned}$$

The notation refers to the **upper and lower contour sets** of x.

So, we begin with a preference (theoretically observable data) and wish to end up with a utility function. Which properties of $\succcurlyeq$ need to hold? Clearly, if there is a utility function representing $\succcurlyeq$, then for $x, y \in X$, either $u(x) \geq u(y)$ or $u(y) \geq u(x)$ (or both), implying either $x \succcurlyeq y$ or $y \succcurlyeq x$ (or both). So $\succcurlyeq$ must be **complete**. Furthermore, for $x, y, z \in X$, if $u(x) \geq u(y)$ and $u(y) \geq u(z)$, we must have $u(x) \geq u(z)$. So $\succcurlyeq$ must be **transitive**. A complete and transitive relation is called a **weak order** although in economics it is often referred to as a preference order, for obvious reasons. We call a restriction on $\succcurlyeq$ an **axiom**.

We conclude via the above argument that completeness and transitivity are necessary conditions for the existence of a utility representation. When X is at most countable, they are is also sufficient. Before proving this,

Second, it provides a test for the utility maximization hypothesis. If a consumer's preferences violate transitivity we know there is *no* function that can rationalize the choices they make.

Lemma 1. *So assume $\succcurlyeq$ is a weak order and that $x \succcurlyeq y$. Then $LCS(y) \subseteq LCS(x)$.*

Theorem 2. *Let X be at most countable. Let $\succcurlyeq$ be relation thereover. Then $\succcurlyeq$ is a weak order if and only if there exists a $U : X \to \mathbb{R}$ that represents $\succcurlyeq$.*

Proof when X is finite. We already proved the necessity portion of the proof. So assume $\succcurlyeq$ is a weak order. Then let $U(x) = \#[LCS(x)]$. We need to prove that U represents $\succcurlyeq$. Assume $x \succcurlyeq y$. By Lemma 1, $LCS(y) \subset LCS(x)$. So clearly $U(x) \geq U(y)$. Now assume $x \not\succcurlyeq y$. Then by completeness of $\succcurlyeq$ we have $y \succcurlyeq x$ which implies by the previous line of reasoning that $LCS(x) \subseteq LCS(y)$. Moreover, since $x \not\succcurlyeq y$ we have that $y \notin LCS(x)$, whereas by reflexivity we have that $y \in LCS(y)$. Therefore, $LCS(x) \subsetneq LCS(y)$ or that $U(x) < U(y)$ as desired. ∎

The above proof will not work for countably infinite sets. For example if $X = \mathbb{N}$ and $x \succcurlyeq y$ if and only if $x \leq y$. Then $\#[LCS(n)] = \aleph_0$ for all n. So what to do?

Proof when X is infinite. Since X is denumerable, there exists a bijection $i : X \to \mathbb{N}$. Let

$$U : x \mapsto \sum_{y \in LCS(x)} \frac{1}{2^{i(y)}}.$$

∎

3 Consumption in $\mathbb{R}^n_+$

We will now take consumption to be represented by $\mathbb{R}^n_+$. The common interpretation here is that each dimension is a particular *type* of consumption, with a point in $\mathbb{R}^n_+$ specifying the amount of each of the n goods to be consumed. What is meant by type? The obvious is that there are different consumption goods—wine and beer being an oft cited example for some odd reason. The (x_w, x_b) represents the state of affairs in which the agent consumes x_w worth of wine and x_b worth of beer. But there are many other interpretations. Each dimension could be a point in time: (x_1, x_2) being the amount of wine at time 1 and at time 2; or a point and time and a particular consumption good: $(x_{1w}, x_{1b}, x_{2w}, x_{2b})$; or, an interpretation that will receive special attention in the future, each dimension is a hypothetical state upon which consumption can be conditioned: (x_r, x_s) being the amount of wine when it rains and is sunny, respectively.

$\mathbb{R}^n_+$ is not countable, so Theorem 2 does not apply. It is clear that completeness and transitivity are still necessary, but alas, they are no longer sufficient. The standard example is $\mathbb{R}^2$ where $(x_1, x_2) \succcurlyeq (y_1, y_2)$ if and only if $x_1 > y_1$ or if $x_1 = y_1$ and $x_2 \geq y_2$. These are called lexicographic preferences, and they are how the alphabet is ordered.

There are several different additional axiomatic restrictions that will guarantee a representation, but the standard/modern approach is *topological*. A preference relation over $\mathbb{R}^n_+$ is **continuous** if $UCS(\boldsymbol{x})$ and $LCS(\boldsymbol{x})$ are closed for all $\boldsymbol{x}$. (Recall $B \subseteq \mathbb{R}^n_+$ is closed if for all sequences $\{\boldsymbol{x}_n\}_{n \in \mathbb{N}} \subseteq B$ if $\lim_{n \to \infty} \boldsymbol{x}_n$ exists than it is contained in B; that is, a set is closed if it contains all of its limit points).

Remark 3. *If $\succcurlyeq$ is continuous if and only if for all all sequences $\{\boldsymbol{x}_n\}_{n \in \mathbb{N}}$ such that $\boldsymbol{x}_n \succcurlyeq \boldsymbol{y}$ for all n and $\boldsymbol{x}_n \to \boldsymbol{x}$ we also have $\boldsymbol{x} \succcurlyeq \boldsymbol{y}$.*

The continuity of $\succcurlyeq$, in tandem with $\succcurlyeq$ being a weak-order, will guarantee that there exists a utility representation of the preference. In fact, it is too strong: there exists non-continuous preferences that admit representation. Continuity implies that not only the existence of a representation, but, as betrayed by terminology, that there exists a *continuous* representation.[2]

Theorem 4. *Let $\succcurlyeq$ be a relation over X. Then $\succcurlyeq$ is a continuous weak order if and only if there exists a continuous $U : X \to \mathbb{R}$ that represents $\succcurlyeq$.*

The proof of Theorem 4 is rather involved and is outside of the scope of this class. It should be noted that there exist non-continuous U that represent a continuous $\succcurlyeq$, a result that can most easily seen in light of Theorem 1. This is yet another instance of the general principal that preferences only contain ordinal information. We might find a representation of of $\succcurlyeq$ with a particular preference but this is just one of many ways of representing the same preference.

Shapes of Utility functions. We will end the discussion on preferences by considering a few more axioms which hold normative appeal in some situations. First, if we interpret each dimension of $\mathbb{R}^n$ as a type of consumption, as we assume that consumption is good, then it seems reasonable to assume that more of any particular dimension is a positive. We say that

[2]For perspicuity: a function $f : \mathbb{R}^n \to \mathbb{R}$ is **continuous** if for every sequence $\{\boldsymbol{x}_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^n$ such that $\boldsymbol{x}_n \to \boldsymbol{x}$ then $f(\boldsymbol{x}_n) \to f(\boldsymbol{x})$. Take heed, however, this is not the definition of continuity in more general settings.

$\succcurlyeq$ is **weakly monotone** if $\boldsymbol{x} \geq \boldsymbol{y}$ implies $\boldsymbol{x} \succcurlyeq \boldsymbol{y}$ and **strictly monotone** if $\boldsymbol{x} \geq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$ implies $\boldsymbol{x} \succ \boldsymbol{y}$. Hopefully, by this point, the following result is pretty obvious.

Theorem 5. *Let U represent $\succcurlyeq$. Then $\succcurlyeq$ is weakly monotone iff U is non-decreasing ($\boldsymbol{x} \geq \boldsymbol{y}$ implies $U(\boldsymbol{x}) \geq U(y)$), and $\succcurlyeq$ is strictly montone iff U is strictly increasing ($\boldsymbol{x} \geq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$ implies $U(\boldsymbol{x}) > U(\boldsymbol{y})$).*

The above theorem is proven essentially by definition. A weaker condition—but one often sufficient for our results in consumer theory—is local non-satiation. A preference $\succcurlyeq$ is **locally non-satiated** for every $\epsilon > 0$ and every $\boldsymbol{x}$ there is a $\boldsymbol{y}$ such that $|x_i - y_i| \leq \epsilon$ and $\boldsymbol{y} \succ \boldsymbol{x}$. Local non-satiation implies that no matter what consumption bundle the agent currently has, she would be better off by moving slightly in some direction (notice, this direction could be a decrease in some dimension). Also, local non-satiation does not mean that the agent cannot have become satiated in some dimension, so long as there is another that by slightly altering would make her better off.

Example 1. *Let $X = \mathbb{R}^2_+$ and assume that $U(x, y) = x + y$. Then x is strictly monotone, weakly monotone and non-satiated. $U(x, y) = \min\{x, y\}$ (Leontief) on the other hand is weakly but not strictly monotone. To see this: if $x > y$ then $U(x + \epsilon, y) = U(x, y)$. Finally, notice $U(x, y) = x - y$ is not monotone at all, but is locally non-satiated, since increasing x always increase utility (as does decreasing y when $y \neq 0$).*

A final word on the shape of preferences: call $\succcurlyeq$ **convex** if for all $x, y, z \in X$ where $y \succcurlyeq x$ and $z \succcurlyeq x$, and for every $\alpha \in [0, 1]$ we have

$$\alpha y + (1 - \alpha)z \succcurlyeq x.$$

Call $\succcurlyeq$ strictly convex if the consequent holds with *strict* preference, whenever $y \neq z$ and $\alpha \neq 0, 1$.

Remark 6. *$\succcurlyeq$ is convex if and only if $UCS(\boldsymbol{x})$ is convex for all $\boldsymbol{x}$.*

When might convexity of preference be an appropriate assumption? Consider the case where different dimensions are different time periods: then convexity says: if the agent prefers 2 wines today to the bundle $\boldsymbol{x}$ and also prefers 2 wines tomorrow to the bundle $\boldsymbol{x}$ then she must also prefer 1 wine today and 1 wine tomorrow to $\boldsymbol{x}$. Convexity indicates that the agent prefers to smooth consumption: to mix bundles together rather than concentrate on a single dimension. What does convexity say about the representing function?

Theorem 7. *If U represents $\succcurlyeq$ and U is (strictly) concave then $\succcurlyeq$ is (strictly) convex.*

Concave utility functions (particularly over a single dimension, like money) are often employed because they exhibit diminishing marginal returns.[3] Theorem 7 therefore says that diminishing marginal returns implies convex preference.

The converse of Theorem 7 does not hold. So what can we definitively say about convex preferences.

Theorem 8. *If U represents $\succcurlyeq$ then U is (strictly) quasi-concave iff $\succcurlyeq$ is (strictly) convex.*

Proof. We will prove sufficiency. Assume that $\succcurlyeq$ is convex. Consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n_+$. Without loss of generality assume that $\boldsymbol{x} \succcurlyeq \boldsymbol{y}$ so that $U(\boldsymbol{x}) \geq U(\boldsymbol{y})$. Then $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \succcurlyeq \boldsymbol{y}$ by convexity, which implies $U(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq U(\boldsymbol{y})$. Hence U is quasi-concave. ■

Two quick mathy tidbits. First, notice that, once proven, the necessity of Theorem 8 will prove Theorem 7 immediately, since every concave function is quasi-concave. Second, Theorem 8 states that *every* representing U is quasi-concave, hence it *must* be that quasi-concavity is an ordinal preference.

4 Choice

While preferences are clearly a more tangible object than utility, it is still not clear that we can observe an agent's preference. Short of actually asking the agent her ranking (and taking the proper measure to ensure this procedure is incentivized) we, as modelers, only have access to the agent's preferences (and therefore utility) via the actions she takes. This section quickly outlines how one might move from observing the choices of a consumer to understanding her preference.

The set up here assumes that we observe the agent making choices in a variety of decision problems. A **decision problem** is a set of possible outcomes, and therefore, a subset of X. The set $A \subseteq X$ represents the feasible consumption outcomes. Given a decision problem, A, we assume that we observe which element the agent chooses. To make matters simple and delineate the frontier of what is identifiable, we assume that we observe the agent's choice from every possible decision problem. This information is encoded by a **choice function**, $\mathscr{C} : \mathcal{P}(X) \to \mathcal{P}(X)$, such that $\mathscr{C}(A) \subseteq A$.

[3] Use this line of reasoning at your own peril: diminishing marginal returns, and indeed concavity in general, is a *cardinal* notion.

Example 2. *Let* $X = \{a, b, c\}$*. The agent is willing to choose either* b *or* c *so long as both are available, but otherwise chooses only* a *(all 1 element menus are dictated by definition):* $\mathscr{C}(\{a, b\}) = \mathscr{C}(\{a, c\}) = \{a\}$ *and* $\mathscr{C}(\{b, c\}) = \mathscr{C}(\{a, b, c\}) = \{b, c\}$*.*

While the choice function in the previous example is perfectly valid, it is a little bit strange. Because a is not in $\mathscr{C}(X)$, we can conclude that the agent prefers some other element. But in each two element menu, the agent picks a. Thus this choice functions seems inconsistent with any underlying preference.

Towards making this formal, let $\succcurlyeq$ be a relation over X. Then we say a choice function, $\mathscr{C}$, is **rationalized** by $\succcurlyeq$ if for every A we have $\mathscr{C}(A) = \{x \in A \mid x \succcurlyeq y, \forall y \in A\}$. That is, if the agent's choice from any decision problem is the set of all $\succcurlyeq$-maximal elements.

Example 3. *Let* $X = \{a, b, c\}$ *and let* $a \succ b \sim c$*:* $\mathscr{C}(\{a, b\}) = \mathscr{C}(\{a, c\}) = \mathscr{C}(\{a, b, c\}) = \{a\}$ *and* $\mathscr{C}(\{b, c\}) = \{b, c\}$*.*

Examples 2 and 3 are similar (we only changed the value of $\mathscr{C}(X)$, but, we argued the former is not compatible with an underlying preference relation, whereas the later obviously is. Given our priority in eventually finding a utility function to describe consumer behavior, we are now tasked with finding conditions on $\mathscr{C}$ such that it is rationalized by a weak order.

The first requirement is rather straight-forward. (Let $\succcurlyeq$ be a weak order). First, notice that if A is a finite set, then there must be some $\succcurlyeq$-maximal element. Hence, $\mathscr{C}$ must be non-empty (we ignore the empty decision problem). Call a choice function **finitely non-empty** if it satisfies this condition. Both example 2 and 3 are finitely non-empty.

The next requirement is a bit more subtle. We need to ensure consistency between the choices in different decision problems. What goes wrong in Example 2? It is that a was *revealed* to be preferred to b when it was the unique choice form $\{a, b\}$. But then this preference was violated when b was later chosen from a menu when a *could have* been chosen but was not. How do we rule this out?

Say a choice rule satisfied the **weak axiom of revealed preference, or WARP** if whenever $x, y \in A \cap B$ and $x \in \mathscr{C}(A)$ and $y \in \mathscr{C}(B)$ then $x \in \mathscr{C}(B)$. By switching the roles of A and B we see that, symmetrically, $y \in \mathscr{C}(A)$. Intuitively, if $x \in \mathscr{C}(A)$ and $y \in A$ then given any rationalizing relation, we have $x \succcurlyeq y$. Likewise $y \in \mathscr{C}(B)$ and $x \in B$ imply $y \succcurlyeq x$. So, $x \sim y$, indicating whenever either x or y is chosen, the other must also be chosen. In Example 2, $a, b \in \{a, b\} \cap \{a, b, c\}$, $a \in \mathscr{C}(\{a, b\})$ and $b \in \mathscr{C}(\{a, b, c\})$ but $a \notin \mathscr{C}(\{a, b, c\})$. So this choice function violates WARP.

Theorem 9. *A choice function $\mathscr{C}$ is rationalizable iff it is finitely non-empty and satisfies WARP.*

Proof. We will prove sufficiency. Let $\mathscr{C}$ be as dictated. Define $\succcurlyeq$ over X as $x \succcurlyeq y$ if and only if $x \in \mathscr{C}(\{x, y\}$. We must show (i) $\succcurlyeq$is complete and transitive and (ii) that $\succcurlyeq$ rationalizes $\mathscr{C}$. Since $\mathscr{C}$ is finitely non-empty either x or y (or both) is in $\mathscr{C}(\{x, y\})$. So $\succcurlyeq$ is complete. Now, to show transitivity it suffices to show that if $x \in \mathscr{C}(\{x, y\})$ and $y \in \mathscr{C}(\{y, z\})$ it must also be that $x \in \mathscr{C}(\{x, z\})$. Assume the antecedent. There are 3 cases to consider: Case (1): if $x \in \mathscr{C}(\{x, y, z\}$. Then $x \in \mathscr{C}(\{x, z\}$ since otherwise we would violate WARP (since then, $x, z \in \{x, z\} \cap \{x, y, z\}$, $z \in \mathscr{C}(\{x, z\})$ and $x \in \mathscr{C}(\{x, y, z\})$ but $x \notin \mathscr{C}(\{x, z, \})$. Case (2): if $y \in \mathscr{C}(\{x, y, z\}$. Applying WARP to $\{x, y\}, \{x, y, z\}$ we see that also $x \in \mathscr{C}(\{x, y, z\})$. Apply case (1). Case (3): if $z \in \mathscr{C}(\{x, y, z\}$. Applying WARP to $\{y, z\}, \{x, y, z\}$ we see that also $y \in \mathscr{C}(\{x, y, z\})$. Apply case (2).

Now, assume by way of contradiction that $\succcurlyeq$ did not rationalize $\mathscr{C}$. Then there is some decision problem B such that $\mathscr{C}(B) \neq \{x \in B \mid x \succcurlyeq y, \forall y \in B\}$. There are 2 cases. Case (1) there is a $y \in C(B)$ such that $y \notin \{x \in B \mid x \succcurlyeq y, \forall y \in B\}$. This second requirement implies there is an $x \in B$ such that $x \succ y$, which by definition implies $y \notin \mathscr{C}(\{x, y\})$. Applying WARP to $\{x, y\}$ and B delivers that $y \in C(\{x, y\})$, a contradiction. The other case, that there $y \in \{x \in B \mid x \succcurlyeq y, \forall y \in B\}$ such that $y \notin C(B)$ is similar. ■

5 Demand Theory

Now, with some foundational understanding of utility functions, we will contemplate how utility translates into behavior within a more specific environment. While preferences, utility functions, and choice functions are all defined over the entire domain of potentially feasible alternatives, consumers are most often constrained by their budgets. Which bundles of consumption are possible depends of which bundles are *affordable*. It is within this specific set of constraints that we will further develop demand theory.

What is affordable depends on the prices of goods and the wealth of the consumer. We continue to employ the assumption that $X = \mathbb{R}^n_+$. Given a vector of prices, $\boldsymbol{p}$, and a wealth, w, the corresponding **budget set**, $B(\boldsymbol{p}, w) \subset \mathbb{R}^n$, is

$$B(\boldsymbol{p}, w) = \{\boldsymbol{x} \in X \mid \boldsymbol{p} \cdot \boldsymbol{x} \leq w\}.$$

Compare this constraint to the constraint of a cost minimization problem; the same implicit assumptions are in play: prices are linear and goods are perfectly divisible.

Given a budget set (i.e., given prices and a wealth level), what does the consumer choose? She will choose the bundle that maximize her preference. Assume that a consumer can be described a preference, $\succcurlyeq$. Then the **walrasian demand** (often called the Marshallian demand) of the consumer is a correspondence $\boldsymbol{x}^\star : \mathbb{R}^n_{++} \times \mathbb{R} \to \mathbb{R}^n_+$ such that

$$\boldsymbol{x}^\star(\boldsymbol{p}, w) = \{\boldsymbol{x} \in B(\boldsymbol{p}, w) \mid \boldsymbol{x} \succcurlyeq \boldsymbol{y}, \forall \boldsymbol{y} \in B(\boldsymbol{p}, w)\}.$$

Setting $\mathscr{C}$ as the choice function rationalized by $\succcurlyeq$, is equivalent to

$$\boldsymbol{x}^\star(\boldsymbol{p}, w) = \mathscr{C}(B(\boldsymbol{p}, w)).$$

Finally if U represents $\succcurlyeq$, then also,

$$\boldsymbol{x}^\star(\boldsymbol{p}, w) = \operatorname*{arg\,max}_{\boldsymbol{y} \in \mathbb{R}^n_+} \; U(\boldsymbol{y}) \quad \text{subject to } \boldsymbol{p} \cdot \boldsymbol{y} \leq w.$$

Given that preferences, choice, and utility are three different representations of the same underling idea, it should be clear that each definition of $\boldsymbol{x}^\star$ is equivalent. Generally, however, the most convenient way to describe $\boldsymbol{x}^\star$ is via a utility function. This is principally because we know how to solve for the maximum of an function subject to an inequality constraint. The utility definition of walrasian demand should being giving you serious flashbacks to profit maximization! This is especially true when U is a differentiable function, in which case, calculus makes our lives particularly easy.[4]

What can we say about walrasian demand? Just as before, when talking about production, a constant scale increase in prices should not affect allocations, in this case consumption.

Theorem 10. *Walrasian demand is h.d.0.*

Proof. This follows directly form the fact that $B(\boldsymbol{p}, w) = B(t\boldsymbol{p}, tw)$. ■

How much will the consumer spend? Whenever she is locally non-satiated, she will spend her entire wealth. This is super obvious when $\succcurlyeq$ is monotone, less so with vanilla LNS. Whats going on? Well LNS rules out the possibility that all goods are bad, there must be some direction to move in (in the consumption space) that increases utility. If we aren't on the boundary, we must be able to increase utility. Formally:

Theorem 11 (Walras' Law)**.** *Let $\succcurlyeq$ be locally non-satiated. Then $\boldsymbol{p} \cdot \boldsymbol{x}^\star(\boldsymbol{p}, w) = w$.*

[4]A word of caution though, differentiability is an cardinal property—we have not delineated any condition of $\succcurlyeq$ to ensure the existence of a differentiable representation.

Proof. Suppose not. Then there exists an $\boldsymbol{x} \in \boldsymbol{x}^\star(\boldsymbol{p}, w)$ such that $\boldsymbol{p} \cdot \boldsymbol{x} < w$. Then there exists an $\epsilon > 0$ such that for all $\boldsymbol{y} \in X$ such that $|\boldsymbol{x} - \boldsymbol{y}| \leq \epsilon$, we have $\boldsymbol{p} \cdot \boldsymbol{y} \leq w$ (i.e., the inner product is continuous). By local non-satiation there is some $\boldsymbol{y}'$ in this neighborhood with $\boldsymbol{y}' \succ \boldsymbol{x}$. Since $\boldsymbol{y}'$ is affordable, this is a contradiction. ■

In general, there maybe many different utility maximizing and affordable bundles. However, when $\succcurlyeq$ is convex (so U is quasi-concave, by Theorem, 8), the *set* of maximizers is convex.

Theorem 12. *If $\succcurlyeq$ is convex then $\boldsymbol{x}^\star(\boldsymbol{p}, w)$ is convex.*

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{x}^\star(\boldsymbol{p}, w)$. Since $\boldsymbol{x} \in \boldsymbol{x}^\star(\boldsymbol{p}, w)$ and $\boldsymbol{y} \in B(\boldsymbol{p}, w)$, we have $\boldsymbol{x} \succcurlyeq \boldsymbol{y}$. Let $t \in [0, 1]$. Then we have $t\boldsymbol{x} + (1-t)\boldsymbol{y} \in B(\boldsymbol{p}, w)$ by the convexity of $B(\boldsymbol{p}, w)$ and $t\boldsymbol{x} + (1-t)\boldsymbol{y} \succcurlyeq \boldsymbol{y}$ by the convexity of $\succcurlyeq$. So $t\boldsymbol{x} + (1-t)\boldsymbol{y} \succcurlyeq \boldsymbol{y} \succcurlyeq \boldsymbol{z}$ for all $\boldsymbol{z} \in B(\boldsymbol{p}, w)$ (where the second preferential statement follows from $\boldsymbol{y} \in \boldsymbol{x}^\star(\boldsymbol{p}, w)$). ■

Theorem 13. *If $\succcurlyeq$ is strictly convex then $\boldsymbol{x}^\star(\boldsymbol{p}, w)$ is a singleton.*

Demand Theory with Calculus. We will now take the additional assumption that $\succcurlyeq$ is represented by a continuously differentiable utility function. While this is a very common assumption in applied theory, we should acknowledge that it is an assumption made directly on the intangible U. With this additional structure, the investigation into Walrasian demand boils down to solving a constrained maximization problem via the KKT method of constrained optimization. In light of Theorem 11, we can further simplify the inequality constraints to equality constraints, so long as the preferences are LNS.

Example 4. *Fix $(\boldsymbol{p}, w)$ and let $U(x, y) = (x^r + y^r)^{\frac{1}{r}}$. The associated Lagrangian to the utility max problem is*

$$\mathcal{L} = (x^r + y^r)^{\frac{1}{r}} - \lambda(p_x x + p_y y - w),$$

where we can use the Lagrangian method (rather than KKT) since U is monotone, hence, LNS. The FOCs are

$$\frac{1}{r}((x^r + y^r)^{\frac{1-r}{r}} r x^{r-1} = \lambda p_x \tag{$\mathcal{L}_x$}$$

$$\frac{1}{r}((x^r + y^r)^{\frac{1-r}{r}} r y^{r-1} = \lambda p_y \tag{$\mathcal{L}_y$}$$

$$p_x x + p_y y = w \tag{$\mathcal{L}_\lambda$}$$

Dividing $\mathcal{L}_x$ by $\mathcal{L}_y$ we see that $(\frac{x}{y})^{r-1} = \frac{p_x}{p_y}$, or that $x = (\frac{p_x}{p_y})^{\frac{1}{r-1}} y$. Plugging back into $\mathcal{L}_\lambda$ we have

$$p_x(\frac{p_x}{p_y})^{\frac{1}{r-1}} y + p_y y = p_x^{\frac{r}{r-1}} p_y^{\frac{-1}{r-1}} y + p_y^{\frac{r}{r-1}} p_y^{\frac{-1}{r-1}} y = w$$

or that $y^\star = \frac{wp_y^{\frac{1}{r-1}}}{p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}}$. Often we will denote $\sigma = \frac{1}{1-r}$ so that $y^\star = \frac{wp_y^{\sigma-1}}{p_x^\sigma + p_y^\sigma}$.

One important fact about walrasian demand, which holds generally is that

$$\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}} = \frac{p_x}{p_y}.$$

The left hand side of this equation is referred to as the **marginal rate of substitution**. At the optimum, it must be that the ratio of marginal utilities is equal to the ratio of prices. (Recall the same thing held in cost minimization with respect to the marginal rate of *technical* substitution). If this were not true then the consumer could "sell" one unit of x to "buy" one unit of y, which would increase her utility.

One might worry that, given our previous fixation on the cardinality of properties, that this condition is dependent on the choice of utility function. However, if U and V both represent x then there is a strictly increasing h such that $h \circ U = V$. Then

$$\frac{\frac{\partial V}{\partial x}}{\frac{\partial V}{\partial y}} = \frac{\frac{\partial h}{\partial U}\frac{\partial U}{\partial x}}{\frac{\partial h}{\partial U}\frac{\partial U}{\partial y}} = \frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}}.$$

Of corse, this assumes that h is differentiable. This is just a statement that indifference curves must lie tangent to the budget line (of course this will be violated if the solution is a boundary solution. For example, work out $U(x, y) = 2x + y$).

When the consumer allocates consumption optimally, what is her utility? This is found using the **indirect utility function**: $v : \mathbb{R}^n_{++} \times \mathbb{R} \to \mathbb{R}$ as

$$v(\boldsymbol{p}, w) = U(\boldsymbol{x}^\star(\boldsymbol{p}, w)),$$

which is reminiscent of the profit function of a firm.

Example 4. *Then we have the indirect utility function as:*

$$\begin{aligned}
v(\boldsymbol{p}, w) &= \big((\frac{w p_y^{\frac{1}{r-1}}}{p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}})^r + (\frac{w p_x^{\frac{1}{r-1}}}{p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}})^r \big)^{\frac{1}{r}} \\
&= w \big(\frac{p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}}{\big(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}\big)^r} \big)^{\frac{1}{r}} \\
&= w \big(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}} \big)^{\frac{1-r}{r}}
\end{aligned}$$

We can use the envelope theorem and v to give a nice interpretation to λ? Notice that $\frac{\partial v}{\partial w} = \lambda$. So the Lagrange multiplier is the marginal utility of an additional dollar (i.e., pound) of income. While we will not get into the details, it can be shown that, so long as utility is a continuous function, then $\boldsymbol{x}^\star$ is upper hemi continuous (a generalization of continuity to set valued functions) and v is continuous. This is important for applications, since it states that if the parameters are measured with a slight imprecision, the error in predictions will be correspondingly small.

There are some other things we can show about the indirect utility function, all of which are importations of the results for profit maximization and cost minimization. The proof techniques are identical.

Theorem 14. *The indirect utility function $v(\boldsymbol{p}, w)$ is non-increasing in $\boldsymbol{p}$ and non-decreasing in w, homogeneous of degree 0, and quasi-convex in $\boldsymbol{p}$.*

Expenditure Minimization. The dual problem to utility maximization ask the guarantee a minimum level of utility, and to do so in the cheapest possible way. The problem is formally, as a function of the parameters $(\boldsymbol{p}, u)$:

$$\min_{\boldsymbol{x} \in X} \boldsymbol{p} \cdot \boldsymbol{x} \quad \text{subject to } U(\boldsymbol{x}) \geq u.$$

We call the correspondence $\boldsymbol{h}^\star : \mathbb{R}^n_{++} \times \mathbb{R} \to \mathbb{R}^n_+$, which specifies the set of minimizers to the above problem, the **Hicksian demand** of the consumer. The minimized expenditure function: $e : \mathbb{R}^n_{++} \times \mathbb{R} \to \mathbb{R}$, with $e(\boldsymbol{p}, u) = \boldsymbol{p} \cdot \boldsymbol{h}^\star(\boldsymbol{p}, u)$. Hicksian demand is often called compensated demand, because it assumes, in the face of price changes, that the consumer can spend any amount of money so as to attain the same level of utility. Therefore, she must be compensated in the face of price changes.

Hicksian demand and the expenditure minimization are *literally* the same as conditional factor demand and cost minimization. They specify a mathematically identical problem, so we can just go ahead an import the following result:

Theorem 15. *The expenditure function $e(\boldsymbol{p}, u)$ is non-decreasing in $\boldsymbol{p}$, homogeneous of degree 1 in p, concave in $\boldsymbol{p}$, continuous in $\boldsymbol{p}$, and $\boldsymbol{h}^\star(\boldsymbol{p}, u) = \nabla_{\boldsymbol{p}} e(\boldsymbol{p}, u)$*

Proof. See Theorems 8, 9, 10, and 11 of Lecture 2. ■

Just as in producer theory, we saw that the supply function was upward sloping, we now see the corresponding *law of compensated demand.* When the prices of goods increase, the compensated (i.e., hicksian) demand cannot increase; compensated demand is downward sloping. This result follows from the concavity of e. A similar result does *not* hold for Walrasian demand.

Example 4. *Fix $(\boldsymbol{p}, u)$ and let $U(x, y) = (x^r + y^r)^{\frac{1}{r}}$. The associated Lagrangian to the utility minimization problem is*

$$\mathcal{L} = p_x x + p_y y - \lambda((x^r + y^r)^{\frac{1}{r}} - u),$$

The FOCs are

$$p_x = \lambda \frac{1}{r}(x^r + y^r)^{\frac{1-r}{r}} r x^{r-1} \tag{$\mathcal{L}_x$}$$

$$p_y = \lambda \frac{1}{r}(x^r + y^r)^{\frac{1-r}{r}} r y^{r-1} \tag{$\mathcal{L}_y$}$$

$$(x^r + y^r)^{\frac{1}{r}} = u \tag{$\mathcal{L}_\lambda$}$$

Dividing the conditions, we get, again,

$$x = y(\frac{p_x}{p_y})^{\frac{1}{r-1}}.$$

Plugging and chugging, we get $u = \big(y^r(\frac{p_x}{p_y})^{\frac{r}{r-1}} + y^r\big)^{\frac{1}{r}} = y\big((\frac{p_x}{p_y})^{\frac{r}{r-1}} + (\frac{p_y}{p_y})^{\frac{r}{r-1}})\big)^{\frac{1}{r}}$. Which gives us:

$$\begin{aligned} y &= \frac{u}{((\frac{p_x}{p_y})^{\frac{r}{r-1}} + (\frac{p_y}{p_y})^{\frac{r}{r-1}})^{\frac{1}{r}}} \\ &= \frac{u p_y^{\frac{1}{r-1}}}{(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}})^{\frac{1}{r}}} \end{aligned}$$

Hence the minimum expenditure is:

$$\begin{aligned} e(\boldsymbol{p},u) &= \frac{u p_y^{\frac{1}{r-1}} p_y}{(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}})^{\frac{1}{r}}} + \frac{u p_x^{\frac{1}{r-1}} p_x}{(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}})^{\frac{1}{r}}} \\ &= \frac{u(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}})}{(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}})^{\frac{1}{r}}} \\ &= u(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}})^{\frac{r-1}{r}} \end{aligned}$$

Duality. Recall from producer theory that there was a weak form of duality with regard to cost minimization and profit maximization. Profit maximization implied cost minimization, hence if a firm was asked to minimized costs while producing the profit maximizing quantity, they would demand the same allocation as if they were simply asked to maximize profit. Consumer theory will exhibit a stronger form of duality between utility maximization and expenditure minimization. We have

Theorem 16. *If U is continuous and satisfies LNS then:*

1. $e(\boldsymbol{p}, v(\boldsymbol{p}, w)) = w$
2. $v(\boldsymbol{p}, e(\boldsymbol{p}, u)) = u$
3. $\boldsymbol{h}^\star(p, v(\boldsymbol{p}, w)) = \boldsymbol{x}^\star(\boldsymbol{p}, w)$
4. $\boldsymbol{x}^\star(\boldsymbol{p}, e(\boldsymbol{p}, u)) = \boldsymbol{h}^\star(\boldsymbol{p}, u)$

The most obvious proof is a proof by picture. But we can also do things analytically.

Proof that utility max implies expenditure min. Fix $(\boldsymbol{p}, w)$ and let $x^\star$ solve the utility maximization problem. Define $u = v(\boldsymbol{p}, w) = U(\boldsymbol{x}^\star)$. Then we must show that $\boldsymbol{x}^\star$ is the cheapest bundle that attains utility level u; that is, it solves the expenditure minimization problem for $(\boldsymbol{p}, u)$. Assume this was not the case, then there is some $\boldsymbol{y}$ such that $\boldsymbol{p} \cdot \boldsymbol{y} < \boldsymbol{p} \cdot \boldsymbol{x}^\star \leq w$. By LNS we can find an affordable bundle, $\boldsymbol{z}$ that such that $U(\boldsymbol{z}) > U(\boldsymbol{y}) \geq u = U(\boldsymbol{x}^\star)$. A contradiction to definition of $\boldsymbol{x}^\star$. ■

Example 4. *For the CES function, we had $v(\boldsymbol{p}, w) = w\big(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}\big)^{\frac{r-1}{r}}$. Rather than re-deriving the expenditure minimization function, we can rely of the duality. We know that if set $u = v(\boldsymbol{p}, w)$ then, $e(\boldsymbol{p}, u) = w$ so, solving for w in the above gives us: $w = \frac{u}{\big(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}\big)^{\frac{1-r}{r}}}$ or $e(\boldsymbol{p}, u) = w = u\big(p_x^{\frac{r}{r-1}} + p_y^{\frac{r}{r-1}}\big)^{\frac{r-1}{r}}$, which was the cost minimization function solved for prior.*

One more quick identity.

Theorem 17 (Roy's Identity)**.** *Given u is continuous and everything is well defined:*

$$\boldsymbol{x}_i^\star(\boldsymbol{p}, w) = -\frac{\frac{\partial v(\boldsymbol{p},w)}{\partial p_i}}{\frac{\partial v(\boldsymbol{p},w)}{\partial w}}.$$

Proof. Define $u = v(\boldsymbol{p}, w)$. Differentiate the identity $v(\boldsymbol{p}, e(\boldsymbol{p}, u)) = u$, which gives us

$$\frac{\partial v(\boldsymbol{p}, w)}{\partial p_i} + \frac{\partial v(\boldsymbol{p}, w)}{\partial w}\frac{\partial e(\boldsymbol{p}, w)}{\partial p_i} = 0 \tag{1}$$

we also have, from Theorem 16 that $\boldsymbol{x}^\star(\boldsymbol{p}, w) = \boldsymbol{h}^\star(\boldsymbol{p}, u)$. Moreover from Theorem 15, we know $\boldsymbol{h}_i^\star(\boldsymbol{p}, u) = \frac{\partial e(\boldsymbol{p},w)}{\partial p_i}$. Plugging this into (1) and rearranging delivers the proof. ■

Slutsky Decomposition. Hicksian demand is downward sloping, but a similar result does not hold for Walrasian demand. Why? Consider what happens when a consumer, on a fixed budget, faces an increase in the price good i. There are two effects: first the relative cost of all other goods has fallen, leading the consumer to increase her demand for substitute goods. Second, the consumer must spend more of her budget to consume the same amount of good i, so she is relativity poorer than before. She may want to make further substitutions between goods because of the changes in her income.

The first effect, called the **substitution effect**, is always negative. The second, the **income effect**, can go either way. So when combined, it is possible that the income effect is more positive than the substitution effect is negative, leading to an overall increase in demand. Such goods are called **giffen** goods.

> "As Mr.Giffen has pointed out, a rise in the price of bread makes so large a drain on the resources of the poorer labouring families and raises so much the marginal utility of money to them, that they are forced to curtail their consumption of meat and the more expensive farinaceous foods: and, bread being still the cheapest food which they can get and will take, they consume more, and not less of it."
>
> —Alfred Marshall (Namesake of Marshallian Demand)

We can make all of this pristinely formal:

Theorem 18 (Slutsky Decomposition)**.** *If u is continuous, locally non-satiated, and strictly* *quasi-concave, and h is differentiable, then:*

$$\frac{\partial \boldsymbol{x}_i^\star(\boldsymbol{p},w)}{\partial p_j} = \frac{\partial h(\boldsymbol{p},v(\boldsymbol{p},w))}{\partial p_j} - \frac{\partial \boldsymbol{x}_i^\star(\boldsymbol{p},w)}{\partial w}\boldsymbol{x}_j^\star(p,w)$$

Proof. From Theorem 16 we have $\boldsymbol{h}_i^\star(p,u) = x_i^\star(\boldsymbol{p}, e(\boldsymbol{p},u))$. Differentiating with respect to p_j provides:

$$\frac{\partial \boldsymbol{h}_i^\star(\boldsymbol{p},u)}{\partial p_j} = \frac{\partial \boldsymbol{x}_i^\star(\boldsymbol{p},w)}{\partial p_j} + \frac{\partial \boldsymbol{x}_i^\star(\boldsymbol{p},w)}{\partial w}\frac{\partial e(\boldsymbol{p},u)}{\partial p_j}.$$

Now notice, from Theorem 15, we have $\frac{\partial e(\boldsymbol{p},u)}{\partial p_j} = \boldsymbol{h}_j^\star(\boldsymbol{p},u)$. Appealing again to the duality, we know that $u = v(\boldsymbol{p},w)$ we have $\boldsymbol{h}_j^\star(\boldsymbol{p},u) = \boldsymbol{x}_j^\star(\boldsymbol{p},w)$. Rearranging, yields the conclusion. ■

The decomposition relates changes in compensated demand with changed in uncompensated demand. The change in Hicksian demand is just the substitution effect. The second term is the income effect, this is the change in consumption of x_i because of the change in income. But the change in income is proportional to the amount of the x_j that was consumed.

The relation between this result and the above discussion comes when we consider the effect of a change in p_i on the consumption of x_i. We know from (a corollary to) Theorem 15 that Hicksian demand is downward sloping in price. So the substitution effect is always negative. The income effect, however, can go both ways. If $\frac{\partial \boldsymbol{x}_i^\star(\boldsymbol{p},w)}{\partial w} > 0$, the x_i is called a **normal** good. Increasing the consumers wealth increases their consumption of x_i. In this case the income and substitution effects reinforce each other: the overall change in consumption of x_i with respect to an increase in p_i is negative. Goods which exhibit a negative income effect are called **inferior**. Here, the two effects work against one another. If the income effect outweighs the substitution effect, so that $\frac{\partial \boldsymbol{x}_i^\star(\boldsymbol{p},w)}{\partial p_i} > 0$, then we have a giffen good on our hands.

6 Risk

Until now, we have interpreted the consumption objects as specific entities that deliver a certain utility when consumed. A can of beer, a new sweater, a Boeing B-17 "Flying Fortress" Bomber, or a bag of apples are specific things; it is reasonable (although by no means the end of the story) to assume these goods impart an exact and certain utility when consumed. What about a lottery ticket, a stock, or a an insurance plan? These consumption objects are **risky**, there is an intrinsic element of chance that will determine different final utility levels. This section seeks to expand our analysis of consumer behavior to risky prospects.

Some motivating clarification: a lottery ticket *could* be treated just like a can of beer. The consumer has some preference relation relating lottery tickets to beer, and all of the above analysis goes through without a hitch. In other words, we are not claiming that the previous results don't work with the introduction of risk, but rather, that they disregard the *additional* structure embodied by the riskiness of situation at large.

So what is this additional structure? How do we represent risk? As always, let X denote a set of consumption objects. The difference here, is that we will assume X represents *final* consumption objects, those things that are the to be consumed when all risk has been resolved. Then, we consider the set of **simple lotteries** over X. A simple lottery is a probability distribution on X such that only a finite number of elements obtain with positive probability.[5] Formally a simple lottery p is a function $p : X \to [0, 1]$ such that $\{x \in X \mid p(x) > 0\}$ is finite and $\sum_X p(x) = 1$ (the latter condition is well defined by the former).

Let P denote set set of all simple probabilities on X. When X is finite P can be identified with the $|X| - 1$ dimensional simplex. Therefore, even when X is finite, P is uncountable. There are infinitely many lotteries that assign just 2 goods, one for each $\alpha \in [0, 1]$. However, X need not be finite. Indeed, we could have $X = \mathbb{R}^n_+$, so that we completely subsume the previous analysis. Then, lotteries are random allocations of vectors of consumption. For each $x \in X$, we also let x denote the lottery that assigns x with probability 1: that is, we let x represent p such that $p(x) = 1$ and $p(y) = 0$ for all $y \neq x$.

A **compound** lottery is a lottery of lotteries. For example with probability α the consumer received the lottery p and with remaining probability she receives the lottery q: $\alpha p + (1 - \alpha)q$ where $\alpha \in [0, 1]$ and $p, q \in P$. The **reduction** of this lottery is the lottery that yields the same distribution of final outcomes. Formally $\alpha p+(1-\alpha)q : X \to [0, 1]$ is given by $(\alpha p+(1-\alpha)q)(x) = \alpha p(x) + (1 - \alpha)q(x)$.

It is straightforward to show that $\alpha p + (1 - \alpha)q$, as defined above, is itself a simple lottery. We will make the assumption that we can associate compound lotteries with their reductions. What does this assumption entail? That the timing of the resolution of risk does not change the consumer's preferences. A compound lottery has to stages of resolution, whereas the reduction has only 1. By identifying the two we are tacitly assuming (tacit because it is baked into the structure, rather than imposed as an axiom), the consumer does not care when risk gets resolved. In light of this assumption, however, we can denote the lottery p such that $p(x) = \alpha$

[5] To keep things at the appropriate level of rigor, we will have to put up with some impression. Those inducted into the world of formal probability or measure theory will be annoyed, but, rest assured, everything can be redone in a totally rigorous way. To wit: if X is uncountable we should really say that the **support** of the lottery is finite.

and $p(y) = (1-\alpha)$ as $\alpha x + (1-\alpha)y$.

We will assume the decision makers preference, $\succcurlyeq$, is over P. How should the consumer evaluate the utility of a lottery? Say the decision maker is playing a game, in which she might win, lose, or draw. She prefers winning to drawing to loosing. So we know that whatever representation we find must have $U(w) > U(d) > U(l)$. Lets go ahead and assume $U(w) = 1$ and $U(l) = 0$. The decision maker had two strategies, the first yields a draw for sure (the action d) and the second a win and a loss with equal probability (the action $\frac{1}{2}w + \frac{1}{2}l$. Now, what is we learn that the decision maker is indifferent between these two actions? $d \sim \frac{1}{2}w + \frac{1}{2}l$. What should this, intuitively, imply about $U(d)$? It seems natural to assume that $U(d) = \frac{1}{2}$.

The lottery $\frac{1}{2}w + \frac{1}{2}l$ will deliver a utility of 1 with probability $\frac{1}{2}$ and 0 with probability $\frac{1}{2}$. So the *average* utility the decision maker receives when playing the action will be $\frac{1}{2}$. Since d is indifferent to this action, the utility of d must also be $\frac{1}{2}$. (Must be?!, heavens no, that's an overstatement. But it is a reasonable behavioral model as that it be!). More generally we are interested in the **Expected Utility** representation:

$$U(p) = \sum_X p(x)u(x) \tag{EU}$$

where $u : X \to \mathbb{R}$ is a **utility index**. An expected utility maximizer evaluates a lottery by first evaluating the utility that each possible resolution of risk would bear, then takes the average (or expectation) according to the probability of each outcome.

There is myriad evidence that people do not behave according to expected utility. So why study it? (1) It is normatively appealing. One can argue that it is the philosophically correct way of making decisions under risk. (2) It is simple, and useful in applications. Often the deviations from the theory are not detrimental to the results of an application, so cautiously using EU can simply analysis. (3) Deviations are often systematic, and therefore EU gives a baseline model to compare other more complicated behavioral models. In other words it provides a common reference to equate new models.

Notice that we have assumed (again, tacitly, via the structure of the objects in question) that all risk is *objective*. That is, the probabilities of outcomes are unquestionably given in the setup of the problem. Our game playing decision maker, however, might believe that her opponent is quite dull, and that she has a high chance of winning. A different decision maker, facing the same set of actions, might believe her opponent is quick witted, and that her chances of winning are slim. Thus the probabilities of winning and loosing associated with different actions might be a *subjective* component of preference, just like the ranking over sure outcomes.

Savage, if a seminal and bar setting work, dealt with the issue of jointly identifying both the subjective probability of events and the utility associated with outcomes. We will not deal with it, but it is something to be aware of.

Expected Utility Axioms. First off, we know that since U represents $\succcurlyeq$, it must be that $\succcurlyeq$ is a weak order. If we want U to be continuous, the of course $\succcurlyeq$ must also be continuous. But, here we will actually provide a weaker notion of continuity which will get the job done. (This is a game mathematicians, and by proxy theorists, like to play. Find the weakest axioms to achieve the result). Say that $\succcurlyeq$ is **Archemedean** if for all $p, q, r \in P$ with $p \succ q \succ r$ there exists an $\alpha, \beta \in (0, 1)$ such that $\alpha p + (1 - \alpha)r \succ q \succ \beta p + (1 - \beta)r$. The Archemedean axiom has a similar flavor to continuity, but is a little weaker (try to prove a continuous $\succcurlyeq$ satisfies it). It effectively states that there cannot be an r so bad (or a p so good) that no mixture can overturn the preference with q.

The second new axiom is what will ensure the expectation like structure to the representation. Imagine there are two different strategies of a game, the first results in a win half the time and a draw half the time, the second a lose have the time and a draw half the time. Since the (expected utility maximizing) player prefers winning to losing, she must prefer the first strategy to the second. Why? When the utility associated with a_1 is $U(a_1) = \frac{1}{2}u(w) + \frac{1}{2}u(d)$ and with a_2 is $U(a_2) = \frac{1}{2}u(l) + \frac{1}{2}u(d)$. Clearly, we have just added a constant to each side of the inequality $u(w) > u(l)$. Generally, the independence axiom ensures such a relation is always respected: call $\succcurlyeq$ **independent** if for all $p, q, r \in P$ and $\alpha \in (0, 1)$ we have $p \succcurlyeq q$ if and only if $\alpha p + (1 - \alpha)r \succcurlyeq \alpha q + (1 - \alpha)r$. The axiom is so named because it says the consumers ranking between lotteries is independent of the common parts. Going back to the example, the player ignores the common $\frac{1}{2}$ probability of drawing, basing her preference only on the aspects of the actions which differ.

Theorem 19 (von Neumann and Morgenstern)**.** *The preference relation $\succcurlyeq$ is an Archemedean, dependent, weak order if and only if there exists a bounded $u : X \to \mathbb{R}$ such that*

$$U(p) = \sum_X p(x)u(x) \tag{EU}$$

represents $\succcurlyeq$. Moreover, $V(p) = \sum_X p(x)v(x)$ also represents $\succcurlyeq$ for some $v : X \to \mathbb{R}$ if and only if $v = au + b$ for some $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$.

We will not prove this theorem here, as it is rather involved. There are many good resources online, ripe for the googling.

We will however, prove the uniqueness claim:

Proof. We will first show that if $v = au + b$ then V represents $\succcurlyeq$. We have

$$\begin{aligned} V(p) \geq V(q) &\iff \sum_X p(x)(au(x)+b) \geq \sum_X q(x)(au(x)+b) \\ &\iff a\sum_X p(x)u(x) + b \geq a\sum_X q(x)u(x) + b \\ &\iff \sum_X p(x)u(x) \geq \sum_X q(x)u(x) \\ &\iff p \succcurlyeq q \end{aligned}$$

Ok, so now, we must show that if V also represents $\succcurlyeq$ it must be that $v = au + b$. Assume this was not the case for all such a, b. Choose any $x, y, z \in X$ such that $x \succ y \succ z$. Let $a = \frac{v(x)-v(z)}{u(x)-u(z)}$ (which is necessarily strictly positive) and $b = v(z) - bu(z)$. (Notice, if no 3 element satisfy this, we are done trivially, as any 2 points are collinear). It is easy to check that $v(x) = au(x) + b$ and $v(z) = au(z) + b$. So our assumption entails that $v(y) \neq au(y) + b$. Let $\alpha \in (0, 1)$ be the unique number such that $\alpha x + (1-\alpha)z \sim y$ (exists by the Archemedean axiom, unique by representation). Then

$$\begin{aligned} v(y) &= V(\alpha x + (1-\alpha)z) \\ &= \alpha v(x) + (1-\alpha)v(z) \\ &= \alpha(au(x)+b) + (1-\alpha)(au(z)+b) \\ &= a(\alpha u(x) + (1-\alpha)u(z)) + b \\ &= aU(\alpha x + (1-\alpha)z) + b = au(y) + b. \end{aligned}$$

This contradicts our assumption, so it must be that $v = au + b$. ∎

We will often say that the representation U is unique upto affine transformations or that U is cardinally unique. Notice two quick points about this claim. (i) The functional representation, U, is itself affine. Recall, this means $U(\alpha p + (1-\alpha)q) = \alpha U(p) + (1-\alpha)U(q)$. We could verify this directly, or (at least in the case when $|X|$ is finite) appeal back to the math notes, Remark 6: a function is affine if and only if can be written as the dot-product with a vector. Here the vector is $(u(x_1)\dots u(x_n))$. (ii) In light of the cardinal uniqueness of u and the previous point, we can see immediately that if U and V both represent $\succcurlyeq$ *and both are affine* then $V = aU + b$. Hey, those math notes are finally coming in handy. Is anyone actually reading this?

Risk Aversion. Given our basic understanding of how to represent preferences over risky domains, we can now turn our attention risk attitudes, or, how to quantify a consumer's tolerance for risk. Generally, we think that consumers are risk *averse*, that they prefer less risk, keeping expected consumption levels constant (we will make this precise later). The insurance industry exists entirely to reduce exposure to risk; stocks and other risky securities must pay a premium to entice investors, etc. Of course there are other, albeit less universal, examples of risk seeking behavior: gambling key among them.

To simplify matters we will consider preferences over money, restricting out attention to a single dimensional prize space. We could view this as a reduced form representation of the more general problem. Assuming a fixed set of prices, we can define $u(w) = v(\boldsymbol{p}, w)$, where v is an indirect utility function. We will assume that the underlying utility index over money is continuous and strictly monotone (these follow from the addition of the continuity and monotonicity axioms, respectively).

For any lottery $p \in P$, there is a unique $x \in \mathbb{R}$ such that $u(x) = U(p) = \sum_{\mathbb{R}} p(x)u(x)$. Since u is continuous, we can apply the intermediate value theorem to obtain existence, and strict monotonicity delivers uniqueness. We will call such an amount of money the **certainty equivalent**, and denote it by c_p. The certainty equivalent of p, c_p, is the amount of money such that the consumer is indifferent between receiving the risky lottery p or c_p with certainty. It is clear that $c_p = u^{-1}(U(p))$, which when considering degenerate lotteries provides, as a sanity check, $c_x = u^{-1}(U(x)) = u^{-1}(u(x)) = x$: a degenerate lottery is its own certainty equivalent.

Notice that both the underlying prize space, $X = \mathbb{R}$ and the set of all lotteries thereover, P, are convex. Because of this that we can consider, for any $p \in P$ the **expected payoff** of p: $e_p = \sum_{\mathbb{R}} p(x)x$. The expected payoff of a lottery is exactly what it sounds like, it is the amount of money the DM can expect to receive on average.

What does it intuitively mean to be risk averse? We said above that a risk averse agent prefers less risk, keeping expected consumption levels constant. One obvious place where we can make such a comparison is between a lottery p and its expected payoff e_p. These two alternatives provide the same expected consumption level, but the latter is risk free. So we say an consumer is **risk averse** if $U(p) \leq U(e_p)$ for all p. She is strictly risk averse of the inequality is strict, and risk seeking if the inequality is reversed. A risk neutral consumer is both risk seeking and risk averse, so $U(p) = U(e_p)$. Equivalently, we could have define risk aversion using certainty equivalents: a consumer is risk averse if and only if $c_p \leq e_p$.

Example 5. *A risk averse consumer has a utility index over wealth given by* $u(x) = x^{\frac{1}{2}}$. *The* *consumer currently has a wealth of* 100*. She might suffer a loss (say her house is on fire) of* 64 *with probability* $\frac{1}{2}$*. How much is she willing to pay to insure herself fully against the loss?*

If she does not insure herself her expected utility is $\frac{1}{2}100^{\frac{1}{2}} + \frac{1}{2}(100-64)^{\frac{1}{2}} = \frac{10}{2} + \frac{6}{2} = 8$*. If she does insure herself, at cost* c*, her expected utility is* $(100-c)^{\frac{1}{2}}$*. Setting these equal and solving for* c *we see that the consumer is willing to pay* $c = 36$*. This is more than the expected loss:* $\frac{1}{2}64 = 32$*.*

What if the DM had linear preferences? How does c *change? Then the solution would be to set* $\frac{1}{2}100 + \frac{1}{2}(100-64) = 68$ *equal to* $100 - c$ *so that* $c = 32$*. Notice this is exactly the expected loss.*

The above example hints at two things. First, the shape of the utility function relates to the underlying risk attitude in a systematic way. Curvature of the utility function (i.e., concavity) is what drives risk aversion.

Theorem 20. *If a consumer has a concave preference over* $\mathbb{R}$ *then she is risk averse, if she has linear preferences she is risk neutral.*

Proof. The definition of concavity/linearity delivers this immediately for lotteries with 2 elements. For a general proof, appeal to Jensen's inequality. ∎

The second thing is that if there are two consumers, one of whom is more risk averse than the other, both can be made better off by *trading risk*. In other words, the less risk averse consumer insures the more risk averse one (for some premium) and makes both better. This is how insurance markets work. The insurance firm is presumably much less risk averse than an individual consumer because it has many different contracts, smoothing risk out across its many commitments.