Choice, Preferences and Utility

Mark Dean

Lecture Notes for Fall 2018 PhD Class in Behavioral Economics - Columbia University

1 Introduction

The first topic that we are going to cover is the relationship between choice, preferences and utility maximization. It is worth thinking about these issues in some detail as utility maximization is *the* canonical model of behavior within economics. Even a lot of 'behavioral' models start with the assumption that people maximize some sort fixed preference relation.

In your first year classes, you proved two fundamental results: Just as a reminder

Definition 1 *Let $\mathcal{A} \subseteq 2^X/\varnothing$ be a collections of choice sets, and $C : \mathcal{A} \rightarrow 2^X/\varnothing$ be a choice correspondence. We say that a complete preference relation $\succeq$ rationalizes C if, for any $A \in \mathcal{A}$*

$$C(A) = \{x \in A | x \succeq y \ \forall \ y \in A\}$$

Theorem 1 *For any finite set X and complete choice correspondence $C : 2^X/\varnothing \rightarrow 2^X/\varnothing$, there exists a complete preference relation $\succeq$ that rationalizes that choice correspondence if and only if C satisfies property α and β.*

where

Axiom 1 (Property α) *If $x \in B \subseteq A$ and $x \in C(A)$, then $x \in C(B)$*

Axiom 2 (Property β) *If $x, y \in C(A)$, $A \subseteq B$ and $y \in C(B)$ then $x \in C(B)$*

Theorem 2 *Let X be a finite set. A binary relation $\succeq$ on X has a utility representation if and only if it is a complete preference relation.*

These results are very powerful - they tell us what the observable implications of utility maximization are. However, they are also somewhat limited, in that they rely on quite strong assumptions. Specifically, they assume that

- We observe a complete choice correspondence - that is choices from every non-empty subset of X
- We observe a choice correspondence - so we get to see all the elements that a DM thinks is best in a given set.
- X is finite.

On can easily think of cases in which all three of these assumptions may fail. So the first thing that we are going to do in this chapter is extend the results of theorems 1 and 2 to cover cases in which we do not observe choices from every choice set, observe only a single choice, and in which X may not be finite (this last case is where things will get a little bit technical).

Here is a quick guide to some of the source materiel for this lecture if you want to learn more:

- "Notes on the Theory of Choice" by David Kreps gives a good, non-technical introduction to the relationship between choice, preferences and utility in chapters 2 and 3
- "Real Analysis with Economic Applications "by Efe Ok gives an extremely technical (but very readable) introduction to the same topic, and is the source you want if you are interested in the details of the extension of our results to infinite spaces. Section A1 gives a good introduction to order theory, while section B4 discusses ordinal utility theory
- "Lectures in Microeconomic Theory" by Ariel Rubinstein lectures 1 to 3 covers these issues well, and also discusses some of the common failures of rationality
- The article "Consistency, Choice and Rationality" by Walter Bossert and Kotaro Suzumura (available on the web) goes through many of the issues in this chapter in extraordinary detail
- "Utility Theory for Decision Making" by Peter Fishburn is tough going, but contains almost all the key results in utility theory

2 What If We Do Not Observe Choices from Every Choice Set?

Theorem 1 assumes that we observe choices from every subset of the set X. This is an extremely strong assumption, as the number of choices that we have to observe gets very large very quickly as the size of X increases - if X has n elements, then we need to observe $2^n - 1$ choices. Given that we are often not going to have a data set that includes a complete choice correspondence, then a natural question is whether we can drop the word 'complete' from the statement of the theorem. In other words, if we observe a choice correspondence on an arbitrary subset S of the power set of X does theorem 1 still hold? The answer is no, as the following example shows.

Example 1 *Let* $X = \{x, y, z\}$ *and say we observe the following (incomplete) choice correspondence*

$$\begin{aligned} C(\{x,y\}) &= \{x\} \\ C(\{z,y\}) &= \{y\} \\ C(\{x,z\}) &= \{z\} \end{aligned}$$

This choice correspondence satisfies properties α *and* β *trivially.* α *is satisfied because we do not observe any choices from sets that are subsets of each other.* β *is satisfied because we never see two objects chosen from the same set. However, there is no way that we can rationalize these choices with a complete preference relation. The first observation implies that* $x \succ y$*, the second that* $y \succ z$ *and the third that* $z \succ x$[1]*. Thus, any binary relation that would rationalize these choices would be intransitive.*

In fact, in order for theorem 1 to hold, we don't have to observe choices from all subsets of X, but we do have to need at least all subsets of X that contain two and three elements (you should go back and look at the proof of theorem 1 and check that you agree with this statement.)

But this condition is still too strong. In many cases we will not observe choices from all such subsets. What can we say in this case? The key here is the principle of revealed preference. The logic of the concept of revealed preference is as follows. Let us begin by assuming that our DM is a preference maximizer - their choices are the result of maximizing some set of preferences. If this is the case, then their choices reveal something about those preferences. In particular, the

[1] Note that I am using $\succ$ in the sense that $x \succ y$ if $x \succeq y$ but not $y \succeq x$

things that the DM chooses from as set must be the best elements in the set, according to the preferences. Thus, if we see an object being chosen from a set, then we know that it is being *revealed preferred* to all the other objects in that set. Note here that what we really mean is that it has been **directly revealed preferred,** in the sense that it is at least as good as all the other objects in the set. Furthermore, if we want our preferences to be transitive, then the fact that x has been revealed directly preferred to y and y revealed directly preferred to z is enough to conclude that x is preferred to z. Finally, if x is chosen from some set while y was available from the same set and was not chosen, we would want to conclude that x is strictly preferred to y.

Definition 2 (The Principle of Revelaed Preference) *Let C be a choice correspondence on a set X. We say that x is revealed directly preferred to y if, for some $\{x, y\} \subset A$ $x \in C(A)$, in which case we write $xR^D y$. We say that x is revealed preferred to y if there exists some sequence $\{x_1,\ x_2, ...x_n\} \in X^n$ such that*

$$xR^D x_1 R^D x_2 R^D ... R^D x_n R^D y$$

in which case we write xRy. We say that x is revealed strictly preferred to y if, for some $\{x, y\} \subset A$, $x \in C(A)$ and $y \notin C(A)$, in which case we write xSy

Is the principle of revealed preference a sensible one? Let us take the two concepts in turn. First, the principle of weak revealed preference: is it sensible to say that if x is chosen when y is available, then it cannot be the case that y is preferred to x? I can certainly think of cases when this is not a sensible assumption. For example, say that I am in a large wine shop, and I choose a bottle to buy. Is this definitely my favorite wine in the shop? Almost certainly not, as I have not searched through the entire wine shop. There may be a better wine out there that I have simply not come across (we will deal with models that allow for this possibility in later lectures).

What about the principle of strict revealed preference? If anything, this is less convincing, as it relies on the assumption that we observe a 'choice correspondence', which we never do in the real world. All we ever get to see is the single object that a person actually chose. It is perfectly possible that, in fact, they were indifferent between several alternatives, and selected one of those from which they are indifferent, which would violate the principle of strict revealed preference.

Thus, in my opinion, it is best not to treat the principle of revealed preference as tautological - if you chose x over y then you **must** prefer x to y. Instead, it is an implication of a model - the model being that a DM's choices are equivalent to the maximal elements in that set according to their preference relation. From this observation we will be able to derive testable implications from this model.

With that caveat aside, we now return to our problem. Just to be clear, our question is as follows:

Problem 3 *Let X be a finite set, $\mathcal{X} \subset 2^X \varnothing$ be an arbitrary subset of the power set of X and C be a choice correspondence on X. Under what conditions is there a complete preference relation $\succeq$ on X that rationalizes C?*

The answer is the Generalized Axiom of Revealed Preference

Definition 3 *A choice correspondence C satisfies the Generalized Axiom of Revealed Preference (GARP) if, for any x,$y \in X$ such that xRy it is not the case that ySx*

It turns out that GARP is necessary and sufficient for choices to be represented by a complete transitive preference relation.

Theorem 4 *Let X be some non-empty set, and C a choice function on $\mathcal{A} \subseteq 2^X/\varnothing$. C satisfies GARP if and only if there exists a complete preference relation $\succeq$ that rationalizes C*

Proof. *First, note that GARP implies directly that S is the asymmetric part of R. Second, note that R is the transitive closure of R^D. Thus by Proposition 1 of the Order Theory notes there exisits a complete preference relation $\succeq$ such that xRy implies $x \succeq y$ and xSy implies $x \succ y$. Thus*

$$\begin{aligned} x &\in C(A) \\ &\Rightarrow xRy \ \forall \ y \in A \\ &\Rightarrow x \succeq y \ \forall \ y \in A \end{aligned}$$

and

$$\begin{aligned} x &\notin C(A) \\ &\Rightarrow \exists y \in A \ s.t. \ ySx \\ &\Rightarrow y \succ x \\ &\Rightarrow x \not\succeq y \end{aligned}$$

so $C(A) = \{x \in X | x \succeq y \forall y \in A\}$ ■

Proof. *To show that the representation implies GARP, note that if choices are made in order to maximize some complete preorder* $\succeq$*, then* xR^Dy *implies that* $x \succeq y$ *and* xSy *implies* $x \succ y$*, so* $xRySx$ *implies that there exists a chain* $\{x_1, x_2, ...x_n\}$ *such that*

$$x \succeq x_1 ... \succeq x_n \succeq y \succ x$$

a contradiction. ■

Note that the theorem also does not require the finiteness of X

What do we think of the weak cycles condition? From an aesthetic point of view, it is certainly not as beautiful as conditions α and β - it all seems a bit mechanical and brute force - in a sense the axioms seem to be stating the obvious. However, the flip side of this is that the GARP condition is very easy to test, as we will discuss in the next lecture.

A few additional points to note

- Note that, if rather that observing a choice *correspondence* we observe a choice *function*, then the $R = S$. The GARP condition reduces to the requirement that R is acyclic
- We might naturally want an equivalent relaxation of theorem 2. Let $\succeq$ be an arbitrary binary relation. Under what circumstances can we find a utility function such that $x \succeq y$ implies $u(x) \geq u(y)$ and $x \succ y$ implies $u(x) > u(y)$. In fact, if we put together all the bits that we have so far proved, we know the answer this question.

  Theorem 5 *Let* X *be a finite, non-empty set, and* $\succeq$ *be a binary relation on* X *and* $\succ$ *and* $\sim$ *be the asymmetric and symmetric parts of* $\succeq$ *Then there exists a function* $u : X \to \mathbb{R}$ *such* *that*

$$x \succeq y \text{ implies } u(x) \geq u(y)$$
$$x \succ y \text{ implies } u(x) > u(y)$$

If and only if $\succ, \sim$ *satisfy GARP*

Proof. *First, note that, without loss of generality we can assume that* $\sim$ *is reflexive. If not, we can add the reflexive relations to* $\succeq$*, and this will change neither whether or not* $\succeq$ *satisfies GARP nor whether a particular function will represent* $\succeq$ *in the sense above.*

Let T *be the transitive closure of* $\succeq$*. GARP guarantees both that* T *is an extension of* $\succeq$*, and that it cannot be that* $xTy \succ x$*. Thus, by theorem 4, there exists a complete preorder that extends* T*, and therefore* $\succeq$*. By theorem 2 there exists a utility function that represents this complete preorder and therefore* $\succeq$*. That a utility representable binary relation satisfies GARP is trivial from the observation that* $x \succeq y$ *implies* $u(x) \geq u(y)$ *and* $x \succ y$ *implies* $u(x) > u(y)$ ■

A couple of things to note. Firstly, in this case we DO need X to be finite, as theorem 2 does not necessarily hold otherwise. Second, note that the utility representation we have here is worse that that of theorem 2. In that case, we could go both ways: we could construct the preference relation from the utility function, or visa versa. That is not true in the case of theorem 5. If all we know is the utility function, and we see that $u(x) \geq u(y)$, then it could be that xRy, but it could be that the two are unrelated. For similar reasons, we can no longer guarantee that u is unique up to a strictly positive monotone transformation.

3 What If We Do Not Observe A Choice Correspondence?

Up until now, we have assumed that we can observe a choice correspondence - every choice set maps to a subset. However, if you think about it for a minute, this should you make you feel uncomfortable. A choice *function* is understandable - it is the thing that I observe you choose from any given choice set. But what is this correspondence? It is not at all clear how to interpret it. There are some suggestions in the literature - for example, if we observe a DM making choices multiple times then we could call $C(A)$ the set of objects that we *ever* see chosen from A, but this seems to be unsatisfactory. If we are in a world where we observe multiple choices from people which change from time to time, then surely we would want to model this explicitly? Perhaps by thinking about the resulting distribution of choices? (we will come on to models that take this approach later).

So can we drop the assumption that we observe a choice correspondence, and instead observe a choice function? For example, we could ask the following question:

Question 1 Let $C : 2^X/\varnothing \rightarrow X$ be a choice function. Under what conditions can we find a complete preference relation $\succeq$ on X such that

$$C(A) \in \{x \in A | x \succeq y \; \forall \; y \in A\}$$

In other words, under what conditions can we find a preference relation such that people always choose *one* of the best available options.

Unfortunately, it should be pretty easy to see that we can always find such a complete preference relation - we can just allow for everything to be indifferent! Then any object that the DM picks is automatically one of the best. So this approach won't get us very far.

Another thing we could do is just rule out indifference by assumption. In other words, we could ask the following question.

Question 2 Let $C : 2^X/\varnothing \rightarrow X$ be a choice function. Under what conditions can we find a linear order $\succ$ on X such that

$$C(A) = \{x \in A | x \succ y \; \forall \; y \in A\}$$

In other words, under what conditions can we find a preference relation *which does not allow indifference* such that people always choose the best available option.

Here we have solved the problem by assuming it away: by demanding that $\succ$ is a linear order we can no longer explain behavior by allowing people to be indifferent between everything, because we have ruled out indifference - in fact, we know that $\{x \in A | x \succ y \; \forall \; y \in A\}$ is a singleton. In this case, it is simple to check that our previous theorems will go through: in the case of a complete choice function the necessary and sufficient requirement is property α (β is unnecessary) while in the case of incomplete data, the necessary and sufficient requirement is GARP (though note that this condition just becomes acyclicality in this case).

Is this a sensible approach? It certainly is not ideal: in general it seems possible that people really are indifferent between two alternatives. If I am choosing between screwdrivers, I really don't care if the handle is blue or red. And if I am indifferent between the two, then it seems harsh to declare me irrational because in some cases I choose the red handled one and in some cases the blue handled one.

While there is no real agreed way out of this problem for general choice sets, we can do better in the case of choices from budget sets. For this section, we will assume that the objects of choice have a particular structure - that they are *commodity bundles* - there are n commodities in the word, and the DM has to select a bundle of these commodities, so $x \in X$ is now

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

where x_i is the amount of good i that is in the bundle. Choice sets are determined by a vector of prices $p \in \mathbb{R}^n_+$, giving a choice set

$$\left\{x \in \mathbb{R}^n_+ | p.x < I\right\}$$

In this case our data will consist of observations of choices made from different price vectors indexed p^j. We will assume that income levels are not observed. We will denote by x^j the bundle chosen when price $p^j$is in effect.

What does revealed preference mean in this context? Well, using the definition above, we say that bundle x^j is strictly revealed preferred over bundle x^k if x^j is chosen and x^k is not when both are available. If we see someone buy a bundle x^j at prices p^j, we know that the could have bought any bundle y which is cheaper that x^j at prices p^j. Thus we have

$$\begin{aligned} & x^j R x^k \\ \iff & p^j x^k \leq p^j x^j \end{aligned}$$

However, this definition has all the attendant problems above: either we have to rule out indifference by assumption, or we have to realize that we can potentially explain any pattern of choices. To get round this problem, we can introduce a new, relatively innocuous assumption: that people have preferences that are **locally non-satiated.**

Definition 4 *A preference relation $\succeq$ on a commodity space $\mathbb{R}^n_+$ is locally non-satiated if, for any $x \in \mathbb{R}^n_{++}$, $\varepsilon > 0$ there exists some $y \in B_\varepsilon(y)$ such that $y \succ x$*[2]

In other words, for any bundle x there is another bundle close to x such that is strictly preferred to it. Is this a sensible assumption? Well, strictly monotonic preferences are locally non-satiated, so if you believe that people in general like more stuff, then it may not be a bad assumption.

How does this help us? Well, it allows us to resurrect the concept of strict revealed preference, even allowing for the possibility of indifference, and even in the case of choice functions. Consider two bundles x^j and x^k such that $p^j x^k < p^j x^j$. My claim is that, if our DM is choosing in order to maximize a complete locally non-satiated preference relation (in the sense of question 1 above), then it must be the case that $x^j \succ x^k$

Lemma 1 *Let x^j and x^k be two commodity bundles such that $p^j x^k < p^j x^j$. Then, if the DM's choices can be rationalized by a complete locally non-satiated preference relation, then it must be the case that $x^j \succ x^k$*

[2] Quick real analysis diversion. The notation $B_\varepsilon(x)$ is the 'open epsilon ball around x.' In other words it is the set of all objects that are a distance less than ε away from x.

$$B_\varepsilon(x) = \{y \in \mathbb{R}^n | d(x, y) < \varepsilon\}$$

where d is some metric. As we are in $\mathbb{R}^n$ we can define the distance function

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Proof. *The key step of the proof is to show that, for some* $\varepsilon > 0$*, then it must be the case that* $p^j y < p^j x^j \ \forall \ y \in B_\varepsilon(x^k)$*. To see this, let*

$$\varepsilon = \frac{p^j x^j - p^j x^k}{\sum_i p_i^j}$$

Now consider the ball $B_\varepsilon(x^k)$*. First, note that, for every* $y \in B_\varepsilon(x^k)$*, it must be the case that* $|y_i - x_i^k| < \varepsilon$ *all* i*. If not, then for some* i*,* $|y_i - x_i^k| \geq \varepsilon$ *and so*

$$\begin{aligned} \sqrt{\sum_{i=1}^{n}(x_i^k - y_i)^2} &\geq \sqrt{(x_i^k - y_i)^2} \\ &= |y_i - x_i^k| \geq \varepsilon \end{aligned}$$

Thus, it must be the case that, for every $y \in B_\varepsilon(x^k)$*, it must be the case that*

$$\begin{aligned} p.y &< .\sum p_i(x_i^k + \varepsilon) \\ &= px^k + \varepsilon \sum_i p_i^j \\ &= px^k + \frac{p^j x^j - p^j x^k}{\sum_i p_i^j} \sum_i p_i^j \\ &= p^j x^j \end{aligned}$$

So, there is a ball $B_\varepsilon(x^k)$ *such that everything in that ball is affordable. By the local non-satiation property, this implies there is a* $y \in \mathbb{R}^n_{++}$ *such that* $y \succ x^k$ *and* $p.y < p^j x^j$*. Thus, if* $x^k \succeq x^j$*, it would be the case that* $y \succ x^j$ *for some feasible bundle, contradicting the assumption that these preferences can rationalize choice.* ■

This points us toward a solution to our problem: we need to adjust our definition of revealed preference. If we see x^j chosen at prices p^j we cannot say anything about bundles that **cost the same** as x^j. However, if we believe in non-satiation, then we can say something about bundles that are **cheaper** than x^j. We will write $x^j S^* y$ if $p^j x^k < p^j x^j$

Following from lemma 1 it is easy to show that the maximization of a non-satiated set of preferences implies GARP. In fact, the relation goes deeper than that, as described in the following celebrated result from Afriat:

Theorem 6 (Afriat) *Let* $\{x^1, ...x^l\}$ *be a set of chosen commodity bundles at prices* $\{p^1, ..., p^l\}$*. The following statements are equivalent:*

1. *The data set can be rationalized by a locally non-satiated set of preferences $\succeq$ that can be represented by a utility function*

2. *The data set satisfies GARP (i.e. xRy implies not yS^*x)*

3. *There exists positive $\left\{u^i, \lambda^i\right\}_{i=1}^{l}$ such that*

$$u^i \leq u^j + \lambda^j p^j (x^i - x^j) \;\forall\; i, j$$

4. *There exists a continuous, concave, piecewise linear, strictly monotonic utility function u that rationalizes the data*

We will not prove this result, as it is quite cumbersome.[3] However, it is worth noting two things:

1. GARP is equivalent to rationalizability by a locally non-satiated preference relation. Thus, if we are prepared to admit non-satiation, then utility maximization does have testable implications in this setting

2. The data set can be represented by a non-satiated utility function if and only if it can be a represented by a concave, continuous, piecewise linear, strictly monotonic utility function. Thus, for this finite data set, concavity, continuity, piecewise linearity and strict monotonicity do not have any testable implications beyond ensuring non-satiation. In the case of continuity and piecewise linearity this might not be so surprising, but the fact that we get concavity for free is a very interesting result.

[3] If you are interested, have a look at 'Two New Proofs of Afriat's Theorem' By Fostel, Scarf and Todd, available here http://ecommons.cornell.edu/bitstream/1813/9258/1/TR001381.pdf

4 What If X Is Not Finite

The final restriction that we want to look at is the finiteness of our set of alternatives X. While finiteness is, in many cases, a reasonable assumption, there are at least two reasons to be interested in the case in which we do not have it. First, there are some cases where it will not hold. For example, if we are going to extend this model to choosing lotteries, then even two prizes can generate an uncountably infinite number of options. Second, many of the reasons for being interested in utility representations require us to pretend that these functions are defined on uncountable spaces: if not, then there is no way for us to use tools such as differentiation.

4.1 Countability: A Reminder

Before proceeding - a quick reminder about the nature of numbers, and infinity.[4] The most basic numbers are the natural, or counting numbers.

Definition 5 *The natural, or counting numbers denoted by* $\mathbb{N}$ *are the set of numbers* $\{1, 2, 3, ...\}$

The nature of the natural numbers is defined formally by the Peano axioms, which you can read up on if you are interested, but for this course, your intuition about what they are will get you through.

The next most complicated set of numbers are the integers. These allow us to include negative numbers and zero

Definition 6 *The integers, denoted by* $\mathbb{Z}$ *are the set of numbers* $\{..., -3, -2, -1, 0, 1, 2, 3, ..\}$

These can basically be defined using the natural numbers.

Next most complicated are the rational numbers. These are any numbers that can be generated by dividing an integer by a natural number

Definition 7 *A rational number, denoted by* $\mathbb{Q}$ *is the set of numbers*

$$\mathbb{Q} = \left\{ \frac{a}{b} | a \in \mathbb{Z}, \ b \in \mathbb{N} \right\}$$

[4] We will not go into detail here. For more information, you can look at Ok chapters A-2 and B, or at these notes: http://www.econ.brown.edu/fac/Mark_Dean/Maths_Real_10.pdf

Clearly, once we have defined the natural and the integers, we know what the rational numbers are.

Unfortunately, the rational numbers are not enough for us to cope with all the concepts we may want to deal with. For example, numbers such as π, $\sqrt{2}$ and e are not rational numbers - they cannot be expressed as a rational divided by a natural number. In other words, the rational numbers have 'holes' in. The set of numbers that plugs these holes are what we call the real numbers, denoted by $\mathbb{R}$. Unfortunately, the construction of the real numbers is not something we can attempt formally here, but you can think informally of it including all the rational numbers and the limit of all sequences of rational numbers. What we usually think of as 'numbers' are really the real numbers. The real line includes things like π, $\sqrt{2}$ and e. We call such numbers (that are in the set of reals, but not rationals) irrational numbers.

We call a set *countably infinite* if there is a bijection between that set and the natural numbers - in other words, a function between that set and the real numbers which is one-to-one and onto. Any set that is countably infinite can therefore be indexed by the natural numbers. I.e. if X is a countably finite set, then we can refer to the elements of x as $x_1, x_2, x_{3,...}$. Clearly, the natural numbers are countable. Perhaps more surprisingly, both the integers and the rational numbers are countable - despite the fact that there seem to be more of the rational numbers, there are not 'importantly more'. However, the real numbers are *not* countable - there is no way we can enumerate the natural numbers as r_1, r_2, r_3. Here are some properties that we are going to use:

Remark 1 *Here are some properties of* $\mathbb{Q}$ *and* $\mathbb{R}$.

1. $\mathbb{Q}$ *is countable*
2. $\mathbb{R}$ *is uncountable*
3. *For every* $a, b \in \mathbb{R}$ *such that* $a < b$, *there exists a* $c \in \mathbb{Q}$ *such* $a < c < b$ *(i.e.* $\mathbb{Q}$ *is dense in* $\mathbb{R}$*)*

4.2 Utility Representations on Infinite X

For most of this section we will be interested in the circumstances under which preferences can be represented by utility functions, as this is where most of the action is in terms of going from finite to infinite X.

4.2.1 Countable X

First let us move from a finite X to a countable X. Luckily, this is an easy case - it turns out that theorem 2 goes through directly.

Theorem 7 *Let X be a* ***countable*** *set. A binary relation $\succeq$ on X has a utility representation if and only if it is a complete preference relation.*

Proof. The proof that a representable binary relation must be a complete preference relation is unchanged from the finite case.

As X is countable, we can enumerate the set as $X = \{x_1, x_2, ...\}$. The other direction relies on the following claim:

Claim 8 *For every $n \in \mathbb{N}$ there is a utility function $u_n : \{x_1, ...x_n\} \to \mathbb{R}$ such that*

1. *u_n represents preferences on $\{x_1, ...x_n\}$*
2. *For all x_j, $j < n$, $u_{n-1}(x_j) = u_n(x_j)$*

The proof of this claim is by induction, and proceeds in the same way as the proof of the existence of a utility function on finite X. There, we provided a procedure by which a utility function on a set of size $n-1$ was extended to represent preferences on a set of size n, keeping the utility function the same for all $n-1$ elements.

Given this claim, we can construct a utility function $u : X \to \mathbb{R}$ as

$$u(x_n) = u_n(x_n) \ \forall \ n \in \mathbb{N}$$

Then, for any x_n,x_m let $k = \max\{x_n, x_m\}$ and note that

$$\begin{aligned} x_n &\succeq x_m \\ \text{iff } u_k(x_n) &\geq u_k(x_m) \\ \text{iff } u(x_n) &\geq u(x_m) \end{aligned}$$

■

4.2.2 Uncountable X

so can we extend the claim to cases where X is uncountable. The answer is no, as the following example demonstrates

Example 2 (Lexecographic preferences) *Let $\succeq$ be a binary relation on $\mathbb{R} \times \{1, 2\}$ such that*

$$\begin{aligned} \{a, b\} &\succeq \{c, d\} \text{ if} \\ \text{(i) } a &> c \\ \text{or (ii) } a &= c \text{ and } b \geq d \end{aligned}$$

You should check that you agree that $\succeq$ is a complete preference relation. There is no utility function that rationalizes $\succeq$.

Proof. *We can prove this by contradiction. Let $u : \mathbb{R} \times \{1, 2\} \to \mathbb{R}$ be a utility function that rationalizes $\succeq$. Then, for every $a \in \mathbb{R}$ it must be the case that $u(a, 2) > u(a, 1)$. Moreover, it must be the case that, if $b > a$ then $u(b, 1) > u(a, 2)$. Thus every real number maps to an interval on the real line, and none of these intervals overlap. By the remark above, each of these intervals contains a different rational, thus we can define a one-to-one mapping from the reals to the rationals. This contradicts the fact that the rationals are countable and the reals are not.* ■

So we clearly need another condition. One standard condition that does give us enough power is the existence of a **countable order-dense subset** of X

Definition 8 *Let X be a non empty set, and $\succeq$ be a preorder on X. We say that a set $S \subset X$ is $\succeq -$dense if, for every $x, y \in X$ such that $x \succ y$, there exists a $s \in S$ such that $x \succ s \succ y$.*

So, for example, $\mathbb{Q}$ is a countable $\geq -$dense subset of $\mathbb{R}$.

In order to make the following proof flow, it is going to be worth making the following definitions:

Definition 9 *Let $\succeq$ be a preference relation on X. The weak and strict upper $\succeq -$contour sets of $x \in X$ are defined as*

$$\begin{aligned} U_{\succeq}(x) &= \{y \in X | y \succeq x\} \\ U_{\succ}(x) &= \{y \in X | y \succ x\} \end{aligned}$$

Similarly, the weak and strict lower $\succeq -$*contour sets of* $x \in X$ *are defined as*

$$L_{\succeq}(x) = \{y \in X | x \succeq y\}$$
$$L_{\succ}(x) = \{y \in X | x \succ y\}$$

While not particularly intuitive, the countable order-dense condition at least gives a starting point. We will therefore state and prove the necessary proposition here:

Proposition 1 *Let* X *be a non-empty set and* $\succeq$ *be a complete preference ordering on* X*. If* X *has a countable* $\succeq -$*dense subset, then* $\succeq$ *can be represented by a utility function* $u : X \to \mathbb{R}$

Proof. *The case where* $\succ$ *is empty is easy, so we concentrate on the non-trivial case.*

Let Y *be a countable* $\succeq -$*dense subset of* X*. By theorem 7, we know that there is a function* $w : Y \to \mathbb{R}$ *that represents* $\succeq$ *on* Y*. By theorem 4 of the previous notes, we can always rescale* w *to be between* 0 *and* 1 *as long as we do so with a strictly increasing transformation. Call such a rescaling* v.[5] *. Now for any* $x \in X$ *such that* $L_{\succ}(x)$ *is non-empty, define*

$$\alpha_x = \sup \{v(t) | t \in L_{\succ}(x) \cap Y\}$$

You should check that you understand that this well defined.

Now define the utility function $u : X \to \mathbb{R}$ *as*

$$\begin{aligned} u(x) &= 1 \text{ if } U_{\succ}(x) = \emptyset \\ &= 0 \text{ if } L_{\succ}(x) = \emptyset \\ &= \alpha_x \text{ otherwise} \end{aligned}$$

The rest of the proof is to show that u *actually represents* $\succeq$*. We will so one case - all the others are similar. The case we will do is the one on which* $x \succ z$ *and* $L_{\succ}(z) \neq \emptyset$ *and* $U_{\succ}(x) \neq \emptyset$*. In this case,* $u(x) = \alpha_x$ *and* $u(z) = \alpha_z$*. Now, by the order denseness of* Y*, we know that there exists* y_1 *and* y_2 *in* Y *such that*

$$x \succ y_1 \succ y_2 \succ z$$

Thus, $\alpha_x \geq v(y_1) > v(y_2) \geq \alpha_z$*, and so* $u(x) > u(z)$ ■

[5] For example, let $v(x) = \frac{1}{2}\left(\frac{w(x)}{1-|w(x)|} + 1\right)$

Note that, as written, this condition is sufficient, but clearly not necessary. For an alternative condition that is both necessary and sufficient is the following:

Definition 10 *A preordered set $(X, \succeq)$ is $\succeq$ −separable if there exists a countable set $Y \subset X$ such that, for every $x, z \in X$ such that $x \succ z$, there exists a $y \in Y$ such that $x \succeq y \succeq z$*

It turns out that $\succeq$ −separability is both necessary and sufficient for the existence of a utility representation

Theorem 9 *Let $\succeq$ be a complete preference relation on X. $\succeq$ admits a utility representation if and only if $(X, \succeq)$ is $\succeq$ −separable*

Proof. *Homework* ■

Okay, so this gives us some condition for a utility function on an uncountable set, but there doesn't seem to be a lot of intuition in the order density condition (note that, because such a condition cannot be tested, we are reduced to using our intuition about various conditions). Luckily, there is a more 'intuitive condition', that relates to the concept of the continuity of a preference relation.

4.2.3 A Diversion into Continuity

You should all be familiar with the concept of the continuity of a function. Intuitively, a continuous function is one that does not jump (you may have learned in high school that a continuous function is one that you can draw without taking your pen off the paper). We define a continuous function as follows:

Definition 11 *Let (X, d) and (Y, ρ) be 2 metric spaces. A function $f : X \to Y$ is continuous at a point $x \in X$ if, for every $\varepsilon > 0$, $\exists\ \delta > 0$ such that*

$$\begin{aligned} d(x, x') &< \delta \\ &\Rightarrow \rho(f(x), f(x')) < \varepsilon \end{aligned}$$

Equivalently, $f(B(x, \delta)) \subset B(f(x), \varepsilon)$. We say a function is continuous if it is continuous at every point in its domain.

We can also think of a preference relation as being continuous. There are many (equivalent) ways of describing a continuous preference relations, but I find the most intuitive is the following:

Definition 12 *Let* $\succeq$ *be a complete, transitive and reflexive preference relation on some metric space. We say that* $\succeq$ *is continuous if,for any* x*,* y *such that* $x \succ y$ *(i.e. not* $y \succeq x$*), there exists* r*,*s *such that* $x' \succ y'$ *for all* $x' \in B(x, r)$ *and* $y' \in B(y, s)$

Why does this mean that we can think of the preference relation as continuous? Well, what is means is that, if x is strictly preferred to y, then there is some neighborhood round x such that everything in that neighborhood is preferred to y. In other words, preferences do not jump in the sense that x is strictly preferred to y, but, however close you get to x there is some object x' such that $y \succeq x'$.

We can also define the concept of upper and lower semi-continuity of preferences in the following way

Definition 13 *A preference relation* $\succeq$ *on a set* X *is upper semi-continuous if* $L_{\succ}(x)$ *is an open subset of* X *for every* x*, and is lower semi-continuous if* $U_{\succ}(x)$ *is an open subset of* X *for every* x*.*

You should convince yourself that a preference relation that is upper and lower continuous is continuous. How to think of upper- and lower semi continuity? Well the following example may help.

Example 3 *Consider the following preferences on* $(0, 1)$

$$\begin{aligned} x &\succ y \text{ if } x \geq 0.5 \text{ and } y < 0.5 \\ x &\sim y \text{ otherwise.} \end{aligned}$$

For any x*, the set* $L_{\succ}(x)$ *is open, as it is either the empty set or* $(0, 0.5)$*. However, the set* $U_{\succ}(x)$ *is not open, as, for* $x < 0.5$ *it is* $[0.5, 1)$*. These preferences are therefore upper semi-continuous but not lower semi-continuous.*

Now consider the preferences on $(0, 1)$

$$\begin{aligned} x &\succ y \text{ if } x > 0.5 \text{ and } y \leq 0.5 \\ x &\sim y \text{ otherwise.} \end{aligned}$$

In this case, $U_\succ(x)$ is open but $L_\succ(x)$ is not for every x. Thus, these preferences are lower but not upper semi- continuous. This illustrates the notion that an USC preference relation is one that may 'jump' down but does not jump up, while a LSC preference relation is one that may jump up, but not down.

We can analogously define the concept of an USC and LSC function

Definition 14 *A real valued function $f : X \to \mathbb{R}$ is upper semi-continuous if, for every $\alpha \in \mathbb{R}$, the set $\{y \in X | f(y) < \alpha\}$ is open. It is lower semi continuous if the set $\{y \in X | f(y) > \alpha\}$ is open.*

One more thing we are going to need is the concept of a separable metric space. A separable metric space is one that is 'not too big', in the sense that it contains a countable dense subset.

Definition 15 *Let X be a metric space. A set $Y \subset X$ is dense in X if $cl(Y) = X$. A set is separable if it has a countable dense subset.*

The intuition is that a separable metric space is not 'too big', as there is a countable space such that each element of the base space is arbitrarily close to an element on the countable set. The easiest example of a separable space is $\mathbb{R}$. Can you show this?

One result that we will use, but not prove is the following:

Theorem 10 *Let X be a separable metric space,Then there exists a countable class $\mathcal{O}$ of open subsets of X such that*

$$U = \cup \{O \in \mathcal{O} | O \subseteq U\}$$

for any open set $U \subset X$

4.2.4 Utility Representations with Continuity.

We will now discuss how the concept of continuity can help us guarantee the existence of a utility representation, and in fact can guarantee the existence of a *continuous* utility representation. We will do this by stating a set of lemmas. The first one we will prove, the later ones you will have to take on trust.

Theorem 11 (Rader's Utility Representation Theorem 1) *Let X be a separable metric space, and $\succeq$ be a complete preference relation on X. If $\succeq$ is upper (or lower) semi-continuous then it can be represented by a utility function*

Proof. *By theorem 10, we know that there is a countable class $\mathcal{O}$ of open subsets of X such that*

$$U = \cup \{O \in \mathcal{O} | O \subseteq U\}$$

for any open set $U \subset X$. Enumerate these sets $\mathcal{O} = \{O_1, O_2, ...\}$, and let

$$m(x) = \{i \in \mathbb{N} | O_i \subseteq L_{\succ}(x)\}$$

And define

$$u(x) = \sum_{i \in M(x)} \frac{1}{2^i}$$

Notice that

$$\begin{aligned} x &\succeq y \\ &\iff L_{\succ}(x) \supseteq L_{\succ}(y) \\ &\implies m(x) \supseteq m(y) \\ &\iff u(x) \geq u(y)^6 \end{aligned}$$

So the only thing we have to show is that $m(x) \supseteq m(y) \implies L_{\succ}(x) \supseteq L_{\succ}(y)$. To show this, take x, y such that $L_{\succ}(x) \supseteq L_{\succ}(y)$ is false. This implies that there exists a z such that $z \in L_{\succ}(y)$ but not $z \in L_{\succ}(x)$. As $L_{\succ}(y)$ is open, then there exists an open set that contains z and is a subset of $L_{\succ}(y)$. But this set cannot be a subset of $L_{\succ}(x)$, so $m(x) \supseteq m(y)$ is false, so we are done. ■

In fact, we can say more, and Radner did. It turns out that, if the above conditions are met, we can in fact represent the preferences with an upper semi continuous utility function

Theorem 12 (Rader's Utility Representation Theorem 2) *Let X be a separable metric space and $\succeq$ be a complete preference relation on X. If $\succeq$ is upper (lower) semi-continuous then it can be represented by an upper (lower) semi-continuous utility function*

Proof. *The theorem is provable for us, but it is tedious. If you are interested, have a look at Ok section D.5.2* ■

The daddy of theories such as this is Debreu's representation theorem. This extends Radner 2 to tell us that continuous preference relations can be mapped to continuous utility functions

Theorem 13 (Debreu) *Let X be a separable metric space, and $\succeq$ be a complete preference relation on X. If $\succeq$ is continuous, then it can be represented by a continuous utility function.*

Proof. *This cannot readily be proved with the tools that we have, but can be proved assuming a result called the 'Open Gap Lemma'. We will not do so, but if you are interested, look at Ok D.5.2*

■

So far, we have not said anything about the link between choice and preference on uncountable sets. To some extent there is not much to say - as we discussed before, properties α and β (or GARP) are enough to guarantee rationalizability. However, there are two issues that we have not looked into yet

1. Do all choices that result from maximizing a complete preference relation satisfy α and β?
2. Given that continuity of $\succeq$ is important for utility representations, what conditions on choice do we need to have them be rationalizable by a continuous preference relation?

The answer to the first question is clearly 'no', for the following simple reason: in general, for infinite X, the maximization of a preference relation may not even give rise to a choice function. To see this, just consider the set $X = [0, 1]$, and the preferences represented by $x \succeq y$ iff $x \geq y$. The set $\{x \in X | x \succeq y \ \forall \ y \in (0, 1)\}$ is clearly empty.

Luckily there is an approach which will deal with both these issues. This approach is two-fold

1. We will consider only compact choice sets on metric spaces
2. We will define a continuous choice correspondence.

To do this, we need to define a metric on choice sets: what we want is for choices from choice sets that are 'close' to each other to also be close to each other. In order to make this formal, we need some notion of closeness of sets. For this, we will use the Hausdorff metric:

Definition 16 (The Hausdorff metric) *Let (X, d) be a metric space, and $cb(X)$ be the set of all closed and bounded subsets of X. We will define the metric space $(cb(X), d^h)$, where d^h is the Hausdorff metric induced by d, and is defined as follows: For any $A, B \in cb(X)$, define $\Lambda(A, B)$ as* $\sup_{x \in A} d(x, B)$. *Now define*

$$d^H(A, B) = \max \{\Lambda(A, B), \Lambda(B, A)\}$$

Now we have defined a metric on closed and bounded sets, we can define the concept of a continuous choice correspondence.

Definition 17 *Let X be a compact metric space and Ω_X be the set of all closed subsets of X and $C : \Omega_X \to 2^X$ be a choice correspondence. If $S_m \to S$ for $S_m, S \in \Omega_X$, $x_m \in C(S_m)\ \forall\ m$ and $x_m \to x$, implies that $x \in C(S)$, then we say C is continuous.*

It turns out that continuity, plus α and β, is enough to give us our desired results

Theorem 14 *Let X be a compact metric space and Ω_X be the set of all closed subsets of X and $C : \Omega_X \to 2^X$ be a choice correspondence. C satisfies properties α , β and continuity if and only if there is a complete, continuous preference relation $\succeq$ on X that rationalizes C.*

Proof. *For brevity, we will only prove the fact that continuity of choice correspondences implies continuity of preferences. We know that it must be the case that*

$$x \succeq y \text{ iff } x \in C(\{x, y\})$$

You should also check, but the continuity of preferences is equivalent to the claim that, if $x_m \to x$, $y_m \to y$ such that $x_m \succeq y_m\ \forall\ m$, then it cannot be the case that $y \succ x$. Now note that

$$\Lambda(\{x, y\}, \{x_m, y_m\}) = \Lambda(\{x_m, y_m\}, \{x, y\}) \leq \max d(x_m, x), d(y_m, y)$$

so $\{x_m, y_m\} \to \{x, y\}$ in the Hausdorff metric. Also, if $x_m \succeq y_m\ \forall\ m$, then $x_m \in C(\{x_m, y_m\})$, then by continuity we have that $x \in C(\{x, y\})$, and so not $y \succ x$. ■