3 Lecture 3: Choices from Budget Sets

Up to now, we have been rather demanding about the data that we need in order to test our models. We have made two important assumptions: that we observe choices from all possible choice sets, and that we observe choice correspondences (i.e. we see all the options that a decision maker would be 'happy with'). In many cases, we may not be so lucky with our data.

Unfortunately, without these two properties, conditions α and β are no longer necessary or sufficient to guarantee a utility representation. Consider the following example of an incomplete data set.

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 $z \succ x$[3]. 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).

What about if 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 $C : 2^X/\varnothing \to X$ be a choice function. Under what conditions can we find a complete preference relation $\succeq$ on X such that

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

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

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 $C : 2^X/\varnothing \to X$ be a choice function. Under what conditions can we find a linear order (i.e. a preference relation that does not allow indifference) $\succ$ 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 $\{x \in A | x \succ y \ \forall \ y \in A\}$ contains only one item . 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).

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.

We will now consider how to get around both of these problems in a particular setting: choice from budget sets. By this we mean that we observe the choices of a person who has a certain amount of money to spend, and has to choose what amount of various different goods to buy. This should be a very familiar setup from ECON 1110. Also, as we shall see, there is plenty of economic data that comes in this form.

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.

We want to think of the choices that a person makes when faced with different budget sets - i.e. how much they will buy of each good when the price of good i is p_i and income is $\dot{I}$.

Choice sets are determined by a vector of prices $p \in \mathbb{R}^n_+$, giving a choice set

$$\left\{ x \in \mathbb{R}^n_+ | \sum_{i=1}^{n} p_i x_i \leq I \right\}$$

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.

The next thing we want to do is to introduce the concept of *revealed preference.* In general, we say that an object a is revealed (directly) preferred to an object b if a is chosen when b is available. In the context of budget sets, this means that a bundle x^j is chosen when it would have been cheaper to buy x^k

Definition 5 *A commodity bundle x^j is revealed directly preferred to a bundle x^k if*

$$p^j x^k \leq p^j x^j$$

in which case we write $x^j R^D x^k$

x^j is revealed preferred to x^k if we can find a sequence of bundles $x^1, ..., x^n$ such that

$$x^j R^D x^1 R^D x^2 ... R^D x^n R^D x^k$$

in which case we write $x^j R x^k$

It turns out that the concept of revealed preference will help us out with the first problem (that we do not observe choices from all choice sets) but not the second: it should be fairly easy to see that we can still rationalize choices from budget sets by assuming that the chooser is indifferent between all bundles. The problem is that choosing x^j over x^k tells us that x^j is *as good as* x^k but not necessarily that x^j *is better than* x^k i.e. we don't have any way of spotting 'strict revealed preference'. In order to get round this problem we need to introduce a new, relatively innocuous assumption: that people have preferences that are **locally non-satiated.**

Definition 6 *A preference relation $\succeq$ on a commodity space $\mathbb{R}^n_+$ is locally non-satiated*[4] *if, for any $x \in \mathbb{R}^n_{++}$, $\varepsilon > 0$ there exists some $y \in B_\varepsilon(y)$ such that $y \succ x$*

In other words, for any bundle x there is another bundle close to x such that is strictly preferred to it. One example of preferences that are locally non-satiated are 'more is better' - so a bundle x^j that contains more of every commodity than x^k is preferred to x^k

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 $p^j x^k < p^j x^j$. Now, 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 $x^j \succ x^k$. Why? Non satiation tells us that we can always find something close to x^k which is preferred to x^k. Moreover, because x^k is strictly cheaper than x^j we can find a bundle that is strictly preferred to x^k and is cheaper than x^j (call such a bundle x^l). We know that x^j is weakly preferred to x^l, which is strictly preferred to x^k. So, by transitivity x^j is strictly preferred to x^k

[4] 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

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

Definition 7 *A commodity bundle x^j is revealed strictly preferred to a bundle x^k if*

$$p^j x^k < p^j x^j$$

in which case we write $x^j S x^k$

Using these definitions, we can come up with a condition which (it turns out) is necessary and sufficient for utility maximization: the Generalized Axiom of Revealed Preference.

Definition 8 *Let $\{x^1, ...x^l\}$ be a set of chosen commodity bundles at prices $\left\{p^1, ..., p^l\right\}$. We say that this data satisfies the generalized axiom of revealed preference (GARP) if $x^j R x^k$ means that it is not the case that $x^k S x^j$*

GARP is used in the following famous theorem by Sidney Afriat [1967]

Theorem 5 (Afriat) *Let $\{x^1, ...x^l\}$ be a set of chosen commodity bundles at prices $\left\{p^1, ..., p^l\right\}$. 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*
3. *There exists a continuous, concave, piecewise linear, strictly monotonic utility function u that rationalizes the data*

3.1 Testing for Utility Maximization in Practice

Here is an unpleasant fact about 'real life' data. GARP is almost *always* violated. In any actual data set, be it from the laboratory or from the 'real world', individuals will almost certainly fail the relevant axiom. Remember, one mistaken choice, one slip up, and the whole data set will fail GARP. This is problematic, as this is not a very interesting result: if we are going to classify everyone as irrational, then do we throw out all the machinery of economics, possibly due to a very small number of rogue choices? This seems too strong. Therefore, it would be nice to have some measure of how close a particular data set is from satisfying rationality. In this section, we are going to present some tools that will allow us to do just that.

3.1.1 HM-Index

One measure of rationality that does not have the robustness problem described above was proposed by Houtman and Maks [1985]: the size of the largest subset of choice observations that satisfy acyclicality (henceforth the HM index).

Imagine a choice experiment in which subject A exhibits the following behavior:

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

If we assume that choice is synonymous with (strict) revealed preference, then these data are not consistent with acyclicality, as z is revealed preferred to y, while y is revealed preferred to z, which is in turn revealed preferred to x. However, if one were to remove the observation $C_A(\{y, z\}) = \{y\}$, then the resulting system would be consistent with acyclicality.

Now imagine that subject B exhibits the following behavior:

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

This data set is also not consistent with acyclicality. However, in this case one would have to remove *two* observations before subject B's choices are consistent with acyclicality. In this sense, subject B could be described as *less rational* than subject A. This, in essence, is the meaning of the HM index. We usually describe the Houtman Maks index as the largest fraction of choices that are consistent with rationality. So the HM index of subject A would be 0.8 and subject B would be 0.6

Formally, if we have a set of observed choices X, where $x \in X$ generates a set of revealed preference relations $\succeq_x$, the HM index is the largest fraction of X such that the resulting revealed preference relations are acyclic. The HM index has some advantages (it can be applied to any type of choice) but also some disadvantages - one of which is that it is very computationally intensive to compute

3.1.2 The Afriat Measure

One problem with the HM index is that it looks only at the *number* of violations that need to be removed, not the *severity* of these violations. If a consumer's preference cycles only involve objects choices that are very close to indifference, or involve only small cost differences, then we may not find those violations very damning to the concept of utility maximization.

This shortcoming is easiest to illustrate in the case in which the observed choices are over bundles of commodities from different budget sets. Consider the following choice behavior for hypothetical consumers A and B from budget sets in a commodity space that contains two goods (x and y):

- Budget set 1 : income is 10, price of good x is 2, price of good y is 2
  - A buys 1 unit of good x and 4 units of good y
  - B buys 2 units of good x and 3 units of good y
- Budget set 2 : income is 10, price of good x is 3, price of good y is 1
  - A buys 3 unit of good x and 1 unit of good y
  - B buys 3 unit of good x and 1 units of good y

Figure 1 illustrates the choices of these two consumers.

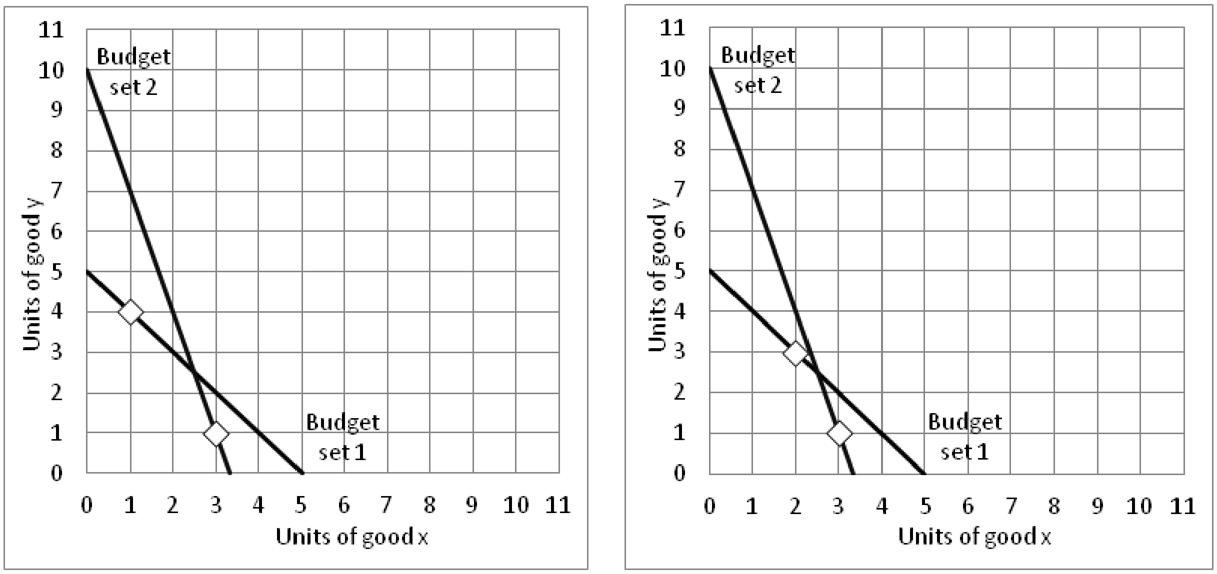

Figure 1

Both of these consumers violate acyclicality, as in both cases the bundle bought in budget set 2 was available in budget set 1, and vice versa. However, the 'cost' of the acyclicality violation for consumer A is higher than for consumer B. For A, the bundle chosen from budget set 2 was available at a cost of \$8 from budget set 1, while the bundle chosen in budget set 1 was available for \$7. For consumer B, the bundle chosen from set 1 was available at a cost of \$9 in set 2, while the bundle chosen in set 2 was available for \$8 in set 1. One could therefore think of the minimum 'cost' the acyclicality violation for A is \$2, while for B it is only \$1. Yet both consumers would have the same HM index.

One measure that tries to get at the cost of deviations from rationality was developed by Afriat [1972] for the case of choice from budget sets. The measure relies on the concept of being revealed preferred at an efficiency level

Definition 9 *We say that x is revealed preferred to y at efficiency level e if $ep^x x > p^x y$.*

Note that efficiency level 1 is the same as standard revealed preference, while for $e = 0$ the revealed preference relation is empty. Afriat's measure of rationality is the efficiency level e such that the resulting revealed preference relation is acyclic (the previous remark says that this has to be true for some $0 \leq e \leq 1$). The Afriat measure for consumer A is 0.8, while for consumer B it is 0.9

3.2 Power Measures

One issue with any of the rationality measures described above is that it is hard to interpret what a particular value tells us about the underlying data. For example, consider a data set in which we observe choices from two disjoint choice sets. In this case all our measures will give perfect rationality scores for *any* observed pattern of choice. In other words, such a data set offers no meaningful test of rationality. One way to address this shortcoming is to compare the values of our chosen index to the distribution of values we would see under some alternative 'null hypothesis' for behavior. Such a comparison allows one to determine whether observed behavior shows more, less or similar levels of rationality than the null hypothesis.

One popular benchmark is to compare index values to those that one would expect to see under random choice – in each choice set individuals have an equal chance of choosing any object in the choice set.[5] Although random choice is a relatively weak null hypothesis, it is applicable to almost any choice setting.

Once we have generated a benchmark, the next question is how to compare the experimental data to this benchmark. For a joint test of all subjects, one can compare the *distribution* of the index scores in the data with the *distribution* of index scores generated under the null hypothesis using some nonparametric measure of the difference between distributions (such as the Kolmogorov-Smirnoff test). In the case of a single observation, one can simply read off the percentile of the simulated data in which that observation falls. Another intuitive measure is to subtract the average simulated score from an actual score.

4 Suggested Readings

Kreps "Notes on the Theory of Choice" Chapters 1-3

Rubinstein "Lecture Notes in Microeconomic Theory" Chapters 1-3 (Available Online: Note that this is a graduate textbook, and goes beyond what you need to know. Don't worry if you find it a little bit tricky - you can ignore the stuff on continuity and the like)

Varian "Revealed Preference". This was originally Chapter 6 in a book entitled "Samuelsonian

[5] Or, in the case of budget sets, an equal chance of choosing any object on the budget line.

Economics in the 21st Century" but it is available for free online if you google "Varian Revealed Preference"