A Representation Theorem for Utility Maximization

Mark Dean

G4840 - Behavioral Economics
A Representation Theorem

- When dealing with models that have latent (or unobservable) variables we will want to find a representation theorem.

- This consists of three things:
  - A data set
  - A model
  - A set of conditions on the data which are necessary and sufficient for it to be consistent with the model.

- Means testing these conditions is the same as testing the model itself.
We are now going to develop a representation theorem for the model of utility maximization. We want to do so properly, so we are going to have to use some notation. Don’t worry - we are just formalizing the ideas from before!
• The data we are going to use are the choices people make

• Notation:
  - $X$: Set of objects you might get to choose from
  - $2^X$: The power set of $X$ (i.e. all the subsets of $X$)
  - $\emptyset$: The empty set

• Our data is going to take the form of a choice correspondence which tells us what the person chose from each subset of $X$

**Definition**
A choice correspondence $C$ is a mapping $C : 2^X / \emptyset \rightarrow 2^X / \emptyset$ such that $C(A) \subset A$ for all $A \in 2^X / \emptyset$. 
• Don’t panic! This is just a way of recording what we described previously
• For example, if we offered someone the choice of Jaffa Cakes and Kit Kats, and they chose Jaffa Cakes, we would write

\[ C(\{\text{kitkat}, \text{jaffacakes}\}) = \{\text{jaffacakes}\} \]

• \( C \) is just a record of the choices made from all possible choice sets
  • i.e. all sets in \( 2^X \) apart from the empty set \( \emptyset \)
• We insist that the DM chooses something that was actually in the data set
  • i.e. \( C(A) \subset A \)
• Note that there is something a bit weird going on
• We allow for people to choose more than one option!
• i.e. we allow for data of the form

\[ C(\{kitkat, jaffacakes, lays\}) = \{jaffacakes, kitkat\} \]

• Which we interpret as something like “the decision maker would be happy with either jaffa cakes or lays from this choice set”
• This is very useful, but a bit dubious
  • We will come back to it later
The model we want to test is that of utility maximization

i.e. there exists a utility function $u : X \rightarrow \mathbb{R}$

Such that the things that are chosen are those which maximize utility

For every $A$

$$C(A) = \text{arg max}_{x \in A} u(x)$$

If this is true, we say that $u$ rationalizes $C$

If $C$ can be rationalized by some $u$ then we say it has a utility representation
We want to know when data is consistent with utility maximization

- i.e. it has a utility representation

So we would like to find a set of conditions on $C$ such that it has a utility representation if and only if these conditions are satisfied

- Testing these conditions is then the same as testing the model of utility maximization
You may remember a condition called the Weak Axiom of Revealed Preference from Intermediate Micro.

We will break WARP down into two parts:

**Axiom \( \alpha \) (AKA Independence of Irrelevant Alternatives)** If \( x \in B \subseteq A \) and \( x \in C(A) \), then \( x \in C(B) \)

**Axiom \( \beta \)** If \( x, y \in C(A) \), \( A \subseteq B \) and \( y \in C(B) \) then \( x \in C(B) \)

Notice we can test these conditions!

If we have data, we can see if they are satisfied.
These conditions form the basis of our first representation theorem.

**Theorem**

*A Choice Correspondence has a utility representation if and only if it satisfies axioms \( \alpha \) and \( \beta \)*

- **if**: if \( \alpha \) and \( \beta \) are satisfied then a utility representation exists
- **only if**: if a utility representation exists then \( \alpha \) and \( \beta \) are satisfied
• Because it is useful (and good for you) we are going to prove this (!)

• In order to do so, we are going to have to introduce another model based on **preferences**
  • Again, should be familiar from Intermediate Micro

• Reminder: A preference relation $\succeq$ is a way of comparing alternatives
  • If $x$ is ‘as good as’ $y$ we write $x \succeq y$
  • We write $x \succ y$ if $x \succeq y$ but not $y \succeq x$
  • We write $x \sim y$ if $\succeq y$ and $\succeq x$
• We demand that preferences have certain properties:
  
  • Completeness: for every $x$ and $y$ in $X$ either $x \succeq y$ or $y \succeq x$ (or both)
  • Transitivity: if $x \succeq y$ and $y \succeq z$ then $x \succeq z$
  • Reflexive: $x \succeq x$

• We say that preference relation $\succeq$ represents a choice function $C$ if, for every $A$

$$C(A) = \{ x \in A | x \succeq y \text{ for all } y \in A \}$$

• i.e. the things that are chosen are those that are preferred to everything else in the choice set
• Preferences are all well and good, but we were interested in the model of utility maximization!

• How can we relate the two?

• We say that a utility function $u$ represents preferences $\succeq$ if

\[
\begin{align*}
  u(x) &\geq u(y) \text{ if and only if } \\
  x &\succeq y
\end{align*}
\]
So if we can find

- A preference relation which represents choices
- A utility function which represents preferences

we are done!

Preferences represents choices means

$$C(A) = \{x \in A | x \succeq y \text{ for all } y \in A\}$$

Utility represents preferences means

$$u(x) \geq u(y) \iff x \succeq y$$

So

$$C(A) = \{x \in A | u(x) \geq u(y) \text{ for all } y \in A\} = \arg \max_{x \in A} u(x)$$
Thus, in order to prove that axioms $\alpha$ and $\beta$ are equivalent to utility maximization we will do the following:

1. Show that if the data satisfies $\alpha$ and $\beta$ then we can find a complete, transitive, reflexive preference relation $\succeq$ which represents the data.

2. Show that if the preferences are complete, transitive and reflexive then we can find a utility function $u$ which represents them.

3. Show that if the data has a utility representation then it must satisfy $\alpha$ and $\beta$.

- We will do 1 and 2 in class. You can do 3 for homework.
Our job is to show that, if choices satisfy $\alpha$ and $\beta$ then we can find a preference relation $\succeq$ which is

- Complete, transitive and reflexive
- Represents choices

**Theorem**

*A Choice Correspondence can be represented by a complete, transitive, reflexive preference relation if satisfies axioms $\alpha$ and $\beta*
How should we proceed?

1. Choose a candidate binary relation $\succ$
2. Show that it is complete, transitive and reflexive
3. Show that it represents choice
• If we observed choices, what do we think might tell us that $x$ is preferred to $y$?
• How about if $x$ is chosen when the only option is $y$?
• Let’s try that!
• We will **define** $\triangleright$ as saying
  
  $$x \triangleright y \text{ if } x \in C(x, y)$$

• Okay, great, we have defined $\triangleright$
• But we need it to have the right properties
Is $\supset$ complete?
Yes!
For any set $\{x, y\}$ either $x$ or $y$ must be chosen (or both)
In the former case $x \supset y$
In the latter $y \supset x$
• Is $\sqsupseteq$ reflexive?

• Yes! (though we have been a bit cheeky)

• Let $x = y$, so then $C(x, x) = C(x) = x$

• Implies $x \sqsupseteq x$
Is $\triangleright$ transitive?

Yes! (though this requires a little proving)

Assume not, then

$$x \triangleright y, \quad y \triangleright z$$

but not $x \triangleright z$

We need to show that this cannot happen

i.e. it violates $\alpha$ or $\beta$

These are conditions on the data, so what do we need to do?

Understand what this means for the data
• Translating to the data
  • $x \gg y$ means that $x \in C(x, y)$
  • $y \gg z$ means that $y \in C(y, z)$
  • not $x \gg z$ means that $x \notin C(x, z)$

• Claim: such data cannot be consistent with $\alpha$ and $\beta$

• Why not?
• What would the person choose from \{x, y, z\}
  
• x?
  
  • No! Violation of \(\alpha\) as x not chosen from \{x, z\}

• y?
  
  • No! This would imply (by \(\alpha\)) that \(y \in C(x, y)\)
  • By \(\beta\) this means that \(x \in C(x, y, z)\)
  • Already shown that this can’t happen

• z?
  
  • No! This would imply (by \(\alpha\)) that \(z \in C(y, z)\)
  • By \(\beta\) this means that \(y \in C(x, y, z)\)
  • Already shown that this can’t happen
Transitivity

- If we have $x \triangleright y$, $y \triangleright z$ but not $x \triangleright z$ then the data cannot satisfy $\alpha$ and $\beta$
- Thus if $\alpha$ and $\beta$ are satisfied, we know that $\triangleright$ must be transitive!
- Thus, we can conclude that, if $\alpha$ and $\beta$ are satisfied $\triangleright$ must have all three right properties!
Finally, we need to show that $\triangleright$ represents choices - i.e.

$$C(A) = \{x \in A | x \triangleright y \text{ for all } y \in A\}$$

How do we do this?

Well, first note that we are trying to show that two sets are equal

- The set of things that are chosen
- The set of things that are best according to $\triangleright$

We do this by showing two things

1. That if $x$ is in $C(A)$ it must also be $x \triangleright y$ for all $y \in A$
2. That if $x \triangleright y$ for all $y \in A$ then $x$ is in $C(A)$
• Say that $x \in C(A)$
• For $\triangleright$ to represent choices it must be that $x \triangleright y$ for every $y \in A$
• Note that, if $y \in A$, $\{x, y\} \subset A$
• So by $\alpha$ if

\[
x \in C(A) \\
\Rightarrow x \in C(x, y)
\]

• And so, by definition

\[
x \triangleright y
\]
Things that are Preferred must be Chosen

- Say that $x \in A$ and $x \triangleright y$ for every $y \in A$
- Can it be that $x \notin C(A)$
- No! Take any $y \in C(A)$
- By $\alpha$, $y \in C(x, y)$
- As $x \triangleright y$ it must be the case that $x \in C(x, y)$
- So, by $\beta$, $x \in C(A)$
- Contradiction!
Q.E.D.
• Well, unfortunately we are not really done
• We wanted to test the model of **utility maximization**
• So far we have shown that $\alpha$ and $\beta$ are equivalent to preference maximization
• Need to show that preference maximization is the same as utility maximization

**Theorem**

*If a preference relation $\succeq$ is complete, transitive and reflexive then there exists a utility function $u : X \to \mathbb{R}$ which represents $\succeq$, i.e.*

$$ u(x) \geq u(y) \iff x \succeq y $$
I am going to sketch the proof because you might find it interesting.

However, I won’t ask you to reproduce this on an exam, so you can relax if you so wish.
Proof By Induction

- We are going to proceed using **proof by induction**
  - We want to show that our statement is true regardless of the size of $X$
  - We do this using induction on the size of the set
  - Let $n = |X|$, the size of the set

- Induction works in two stages
  - Show that the statement is true if $n = 1$
  - Show that, if it is true for $n$, it must also be true for any $n + 1$

- This allows us to conclude that it is true for $n$
  - It is true for $n = 1$
  - If it is true for $n = 1$ it is true for $n = 2$
  - If it is true for $n = 2$, it is true for $n = 3$....

- You have to be a bit careful with proof by induction
  - Or you can prove that all the horses in the world are the same color
So in this case we have to show that we can find a utility representation if $|X| = 1$

- Trivial

And show that if a utility representation exists for $|X| = n$, then it exists for $|X| = n + 1$

- Not trivial
• Take a set such that $|X| = n + 1$ and a complete, transitive reflexive preference relation $\succeq$

• Remove some $x^* \in X$

• Note that the new set $X/x^*$ has size $n$

• So, by the inductive assumption, there exists some $\nu : X/x^* \rightarrow \mathbb{R}$ such that

\[
\nu(x) \geq \nu(y) \iff x \succeq y
\]

• So now all we need to do is assign a utility number to $x^*$ which makes it work with $\nu$

• How would you do this?
• Four possibilities

1. $x^* \sim y$ for some $y \in X/x^*$
   - Set $\nu(x^*) = \nu(y)$

2. $x^* \succeq y$ for all $y \in X/x^*$
   - Set $\nu(x^*) = \max_{y \in X/x^*} \nu(y) + 1$

3. $x^* \preceq y$ for all $y \in X/x^*$
   - Set $\nu(x^*) = \min_{y \in X/x^*} \nu(y) - 1$

4. None of the above
What do we do in case 4?

We divide $X$ in two: those objects better than $x^*$ and those worse than $x^*$

$$X_* = \{ y \in X / x^* | x^* \preceq x \}$$

$$X^* = \{ y \in X / x^* | x \succeq x^* \}$$

Figure out the highest utility in $X_*$ and the lowest utility in $X^*$ and fit the utility of $x^*$ in between them

$$v(x^*) = \frac{1}{2} \min_{y \in X_*} v(y) + \frac{1}{2} \max_{y \in X_*} v(y)$$
• Note that everything in $X^*$ has higher utility than everything in $X^*$
  • Pick an $x \in X^*$ and $y \in X^*$
  • $x \succeq x^*$ and $x^* \succeq y$
  • Implies $x \succeq y$ (why?)
  • and so $v(x) \geq v(y)$
  • In fact, because we have ruled out indifference $v(x) > v(y)$

• This implies that

$$v(x) > v(x^*) > v(y)$$

• And so
  • The utility of everything better than $x^*$ is higher than $v(x^*)$
  • The utility of everything worse than $x^*$ is lower than $v(x^*)$
- Verify that \( \nu \) represents \( \succeq \) in all of the four cases
- That sounds exhausting
- You can look in the lecture notes if you so wish
Q.E.D.
The final step is to show that, if a choice correspondence has a utility representation then it satisfies $\alpha$ and $\beta$

This closes the loop and shows that all the statements are equivalent

- A choice correspondence satisfies $\alpha$ and $\beta$
- A choice correspondence has a preference relation
- A choice correspondence has a utility representation

Will leave you to do that for homework!
• We have now achieved our aim!
• We know how to test the model of utility maximization

**Theorem**

A *Choice Correspondence has a utility representation* if and only if it satisfies axioms $\alpha$ and $\beta$

• We just test $\alpha$ and $\beta$
• Before we move on to something more fun, I want to discuss two potential issues
  • How seriously should we take utility?
  • What happens if our data is not as good as we would like it to be?
How Unique is Our Utility Function?

- We now know that if $\alpha$ and $\beta$ are satisfied, we can find some utility function that will explain choices.
- Is it the only one?
How Unique is Our Utility Function?

<table>
<thead>
<tr>
<th>Croft’s Choices</th>
<th>Available Snacks</th>
<th>Chosen Snack</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Jaffa Cakes, Kit Kat</td>
<td>Jaffa Cakes</td>
</tr>
<tr>
<td></td>
<td>Kit Kat, Lays</td>
<td>Kit Kat</td>
</tr>
<tr>
<td></td>
<td>Lays, Jaffa Cakes</td>
<td>Jaffa Cakes</td>
</tr>
<tr>
<td></td>
<td>Kit Kat, Jaffa Cakes, Lays</td>
<td>Jaffa Cakes</td>
</tr>
</tbody>
</table>

- These choices could be explained by $u(J) = 3$, $u(K) = 2$, $u(L) = 1$
- What about $u(J) = 100000$, $u(K) = -1$, $u(L) = -2$?
- Or $u(J) = 1$, $u(K) = 0.9999$, $u(L) = 0.998$?
How Unique is Our Utility Function?

- In fact, if a data set obeys $\alpha$ and $\beta$ there will be many utility functions which will rationalize the data.

**Theorem**

Let $u : X \to \mathbb{R}$ be a utility representation for a Choice Correspondence $C$. Then $v : X \to \mathbb{R}$ will also represent $C$ if and only if there is a strictly increasing function $T$ such that

$$v(x) = T(u(x)) \ \forall \ x \in X$$

- Strictly increasing function means that if you plug in a bigger number you get a bigger number out.
How Unique is Our Utility Function?

<table>
<thead>
<tr>
<th>Snack</th>
<th>$u$</th>
<th>$v$</th>
<th>$w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jaffa Cake</td>
<td>3</td>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>Kit Kat</td>
<td>2</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Lays</td>
<td>1</td>
<td>$-100$</td>
<td>3</td>
</tr>
</tbody>
</table>

- $v$ is a strictly increasing transform on $u$, and so represents the same choices.
- $w$ is not, and so doesn’t.
  - For example think of the choice set $\{k, l\}$
  - $u$ says they should choose kit cat
  - $w$ says they should choose lays
Why Does This Matter?

• It is important that we know how much the data can tell us about utility
  • Or other model objects we may come up with
• For example, our results tell us that there is a point in designing a test to tell whether people maximize utility
• But there is no point in designing a test to see whether the utility of Kit Kats is twice that of Lays
  • Assuming \( \alpha \) and \( \beta \) is satisfied, we can always find a utility function for which this is true
  • And another one for which this is false!
• We can use choices to help us determine that the utility of Kit Kats is higher than the utility of Lays
• But nothing in our data tells us how much higher is the utility of Kit Kats
• Imagine running an experiment to try and test $\alpha$ and $\beta$
• The data that we need is the choice correspondence

\[ C : 2^X / \emptyset \rightarrow 2^X / \emptyset \]

• How many choices would we have to observe?
• Lets say $|X| = 10$
  • Need to observe choices from every $A \in 2^X / \emptyset$
  • How big is the power set of $X$?
  • If $|X| = 10$ need to observe 1024 choices
  • If $|X| = 20$ need to observe 1048576 choices

• This is not going to work!
• So how about we forget about the requirement that we observe choices from all choice sets
• Are $\alpha$ and $\beta$ still enough to guarantee a utility representation?

\[
\begin{align*}
C(\{x, y\}) &= x \\
C(\{y, z\}) &= y \\
C(\{x, z\}) &= z
\end{align*}
\]

• If this is our only data then there is no violation of $\alpha$ or $\beta$
• But no utility representation exists!
• We need a different approach!
Revealed Preference

- We say that $x$ is **directly revealed preferred to** $y$ if, for some choice set $A$
  
  \[
  y \in A \\
  x \in C(A) 
  \]

- We say that $x$ is **revealed preferred to** $y$ if we can find a set of alternatives $w_1, w_2, ..., w_n$ such that
  - $x$ is directly revealed preferred to $w_1$
  - $w_1$ is directly revealed preferred to $w_2$
  - ...
  - $w_{n-1}$ is directly revealed preferred to $w_n$
  - $w_n$ is directly revealed preferred to $y$

- We say $x$ is **strictly revealed preferred to** $y$ if, for some choice set $A$
  
  \[
  y \in A \textbf{ but not } y \in C(A) \\
  x \in C(A) 
  \]
The Generalized Axiom of Revealed Preference

- Note that we can observe revealed preference and strict revealed preference from the data
- With these definitions we can write an axiom to replace $\alpha$ and $\beta$

**Definition**
A choice correspondence $C$ satisfies the Generalized Axiom of Revealed Preference (GARP) if it is never the case that $x$ is revealed preferred to $y$, and $y$ is strictly revealed preferred to $x$

**Theorem**
A choice correspondence $C$ satisfies GARP if and only if it has a utility representation. This is true even if $C$ is incomplete (i.e. does not report choices from all choice sets)
Another weird thing about our data is that we assumed we could observe a choice *correspondence*

This is not an easy thing to do!

What about if we only get to observe a choice function?

How do we deal with indifference?
Choice Correspondence?

- One of the things we could do is assume that the decision maker chooses **one of** the best options

\[ C(A) \in \arg \max_{x \in A} u(x) \]

- Is this going to work?
- No!
- Any data set can be represented by this model
  - Why?
    - We can just assume that all alternatives have the same utility!
- Need some way of identifying when an alternative \( x \) is **better than** alternative \( y \)
  - i.e. some way to identify strict preference
One case in which we can do this is if our data comes from people choosing consumption bundles from budget sets.

- Should be familiar from intermediate micro.

The objects that the DM has to choose between are bundles of different commodities:

\[ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \]

- And they can choose any bundle which satisfies their budget constraint:

\[ \left\{ x \in \mathbb{R}_+^n \mid \sum_{i=1}^{n} p_i x_i \leq l \right\} \]
Budget constraint is $p_1 x_1 + p_2 x_2 = l$. 
• Claim: We can use choice from budget sets to identify strict preference
  • Even if we only see a single bundle chosen from each budget set
  • As long as we assume more is better

\[
x_n \geq y_n \text{ for all } n \text{ and } x_n \succ y_n \text{ for some } n
\]

implies that \( x \succ y \)

• i.e. preferences are strictly monotonic
Monotonicity

Everything in this quadrant better than x

X

X_1

X_2
Claim: if $p^x$ is the prices at which the bundle $x$ was chosen

\[ p^x x > p^x y \] implies $x \succ y$

Why?
Because $x$ was chosen, we know $x \preceq y$

Because $p^x x > p^x y$ we know that $y$ was inside the budget set when $x$ was chosen

Could it be that $y \preceq x$?
Because $y$ is inside the budget set, there is a $z$ which is better than $y$ \textbf{and} affordable when $x$ was chosen

- Implies that $x \preceq z$ and (by monotonicity) $z \succ y$
- By transitivity $x \succ y$
• When dealing with choice from budget sets we say
  
  • \( x \) is **directly revealed preferred to** \( y \) if \( p^x x \geq p^x y \)
  
  • \( x \) is **revealed preferred to** \( y \) if we can find a set of alternatives \( w_1, w_2, \ldots, w_n \) such that
    
    • \( x \) is directly revealed preferred to \( w_1 \)
    
    • \( w_1 \) is directly revealed preferred to \( w_2 \)
    
    • ...
    
    • \( w_{n-1} \) is directly revealed preferred to \( w_n \)
    
    • \( w_n \) is directly revealed preferred to \( y \)

  • \( x \) is **strictly revealed preferred to** \( y \) if \( p^x x > p^x y \)
Theorem (Afriat)

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

1. The data set can be rationalized by a strictly monotonic 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
We have completed our review of what is sometimes called 'revealed preference theory'.

- Phew

- Here are the takeaways
Testing models which have unobserved (latent) variables is tricky

- For example the model of utility maximization

The gold standard is a ‘representation theorem’

- Conditions on the data which are equivalent to testing the model
- Don’t have to make any specific assumptions about the nature of utility

In the case of utility maximization, we have such conditions

- $\alpha$ and $\beta$

These work if we can observe choice correspondences from every choice set

- Otherwise we need to use GARP

The utility numbers we find are not unique

- Only tell us ordering, not magnitudes