Temptation and Self Control

Mark Dean

Lecture Notes for Spring 2015 Behavioral Economics - Brown University

1 Lecture 1

1.1 Choosing Over Menus

I tried to convince you in the introductory lecture that one of the best ways to turn temptation and self control into observable phenomena was by thinking about cases in which people chose to restrict their own future choices. In order to model this, we need to make a change to the set-up that we have been using. Up until now, we have thought only about choices (or preferences) over objects x living in some set X. Now we need to think about choices (or preferences)[1] over *menus* of these objects. As a simple example, let's think of a set of objects $\{s, f, b\}$ which stands for salad, fish, and burger. Up till now we have been dealing with preferences over these objects. For example, we might write $s \succ f$, meaning that our DM prefers salad to fish, or if given a choice, they would choose salad over fish. Now, instead, we will be writing (for example) $\{s, f\} \succ \{s, b\}$, meaning that a *menu* that contains salad and fish is preferred to a menu that contains salad and burger. Intuitively, we are interpreting this the following way: tomorrow at lunchtime I am going to ask you what you would like for lunch. When I offer you the choice, would you like me to let you choose between the salad or fish, or between the salad and the burger?[2] Alternatively, you can think of this as meaning that the decision maker prefers to go to a restaurant that serves either

[1] From now on, I am going to couch everything in terms of preferences. But this shouldn't worry you. You know how to move between choices and preferences, yes?

[2] This is a common decision theory trick: when we want to model a new phenomena, we become more specific about the nature of the objects that are being chosen. You will see this when we talk about choice under risk and uncertainty later on.

salad or fish to one that serves either the salad or a burger. Before we go on, you should make sure you understand what this new data set is telling us, because if you don't, nothing that follows is going to make much sense.

How would a 'standard' decision maker behave in this setting? Presumably, in order to compare two menus, A and B, they would look at the item they most preferred in A (let's call it x_A) and the item that they most preferred in B (let's call it x_B).If x_A is better than x_B then A would be preferred to B Conversely, if x_B is preferred to x_A, then B would be preferred to A. So, in the fish/salad/burger example above, if fish is preferred to burger, and both are preferred to salad, then we would have $\{s, f\} \succ \{s, b\}$. If the burger is preferred to fish, and both are preferred to salad, then we would have $\{s, b\} \succ \{s, f\}$. If salad is preferred to burger and fish, we would have $\{s, f\} \sim \{s, b\}$. The key insight of this is the standard person will always prefer a bigger choice set to a smaller choice set (in the subset sense): so if $A \subset B$, the $B \succeq A$. Put another way, adding objects to a set cannot make it worse. Why? Well, either the new item is better than those already in there, in which case it makes the set better, or it isn't in which case it would be ignored.

For someone who has temptation and self control problems, this might not be true. Consider, for example, someone who is struggling on a diet. They would really like to lose weight, and so really want to eat the salad. However, they also know that, when it comes to lunchtime, they will be hungry, and that if they are offered the choice of the burger then they will find it difficult to resist. Such a person might actually prefer *not* to have the burger on the menu, and so prefer

$$\{s, f\} \succ \{s, f, b\}$$

This is what we will call a *preference for commitment.* Why might someone exhibit a preference for commitment? There are two possibilities. One is that, if the burger is on the menu then they may cave and eat it. The other is that if the burger is on the menu then they may be able to resist it, but it will be annoying for them to do so. We will now introduce a model that captures both of these effects

1.2 The Gul-Pesendorfer Model of Temptation and Self Control

In 2004, Faruk Gul and Wolfgang Pesendorfer introduced the following model of choice over menus

Definition 1 *Let X be some set of alternatives, and $Z = 2^X/\emptyset$. We say that a set of preferences* *on* $2^Z/\emptyset$[3] *have a temptation/self control representation if there exists a* $u : X \to \mathbb{R}$ *and* $v : X \to \mathbb{R}$ *such that*

$$U(A) = \max_{x \in A}(u(x) + v(x)) - \max_{y \in A} v(y)$$

represents the preferences on $2^Z/\emptyset$

The interpretation of this model as follows. u represents the agents long run, or 'commitment' utility over objects. In the absence of temptation, they would make choices in order to maximize u. However, the agent also suffers from temptation, represented by the function v. Temptation has two different effects. First, it affects the choices that the DM makes at the second stage - when choosing *from* the menu, the DM will choose the object that maximizes $u(x) + v(x)$, not just $u(x)$. Second, having a tempting object in the set makes it worse in the sense that, if the decision maker does not choose the most tempting object in the set, then they pay a cost. This is what the $-\max_{y \in A} v(y)$ bit of the expression does.

Let's consider an example to make life easier: Let's say that , in our fish/burger/ salad example our objects have the following long run and temptation utilities

Object	u	v
Salad	4	0
Fish	2	1
Burger	1	4

So, according to the long run utilities, the salad is better than the fish which is better than the burger, while the burger is most tempting followed by the fish, then the salad.

First, let's compare the DM's preference between two menus - one where he will only be able to choose the salad $\{s\}$, and one where they can choose from either the burger or the salad $\{b, s\}$. Which of these menus does our DM prefer? Well, first, we have to see what they would choose in the second stage - i.e. what they will choose *from* the menu. According to the model they will choose the x that has the highest value of $u(x) + v(x)$. In the case of $\{s\}$, this has to be the salad. In the case of $\{b, s\}$, we compare $u(s) + v(s) = 4$ to $u(b) + v(b) = 5$, so the decision maker will choose the burger from $\{b, s\}$ at the second stage.

[3] Actually, for various important reasons, these preferences are actually defined on sets of lotteries over these objects. Understanding why this is is beyond the scope of this course, but you should go and have a look at the original Gul and Pesendorfer paper if you are interested.

So, what is their preferences over menus? We need to compare

$$\begin{aligned} U(\{s\}) &= \max_{x\in\{s\}}(u(x)+v(x)) - \max_{y\in\{s\}} v(y) \\ &= 4+0-0 \\ &= 4 \end{aligned}$$

to

$$\begin{aligned} U(\{s,b\}) &= \max_{x\in\{s,b\}}(u(x)+v(x)) - \max_{y\in\{s,b\}} v(y) \\ &= 1+4-4 \\ &= 1 \end{aligned}$$

So the DM will prefer only being offered the salad at the second stage to being given the choice between the salad and the burger. The reason is that, at the second stage, if they are given the choice between the burger and the salad they will choose the burger, despite the fact that the burger has lower commitment utility than the salad. Thus they would prefer not to have the option of choosing the burger at the second stage

Second, let's compare the DM's preference between another two menus - one where he will only be able to choose the salad $\{s\}$, and one where they can choose from either the fish or the salad $\{f,s\}$. Which of these menus does our DM prefer? Well, first, we have to see what they would choose in the second stage. According to the model they will choose the x that has the highest value of $u(x)+v(x)$. In the case of $\{s\}$, this has to be the salad. In the case of $\{f,s\}$, we compare $u(s)+v(s)=4$ to $u(f)+v(f)=3$, so the decision maker will choose the salad from $\{f,s\}$ at the second stage.

So, what is their preferences over menus? We need to compare

$$\begin{aligned} U(\{s\}) &= \max_{x\in\{s\}}(u(x)+v(x)) - \max_{y\in\{s\}} v(y) \\ &= 4+0-0 \\ &= 4 \end{aligned}$$

to

$$\begin{aligned} U(\{s,f\}) &= \max_{x\in\{s,f\}}(u(x)+v(x)) - \max_{y\in\{s,f\}} v(y) \\ &= 4+0-1 \\ &= 3 \end{aligned}$$

So the decision maker will prefer the menu $\{s\}$ to the menu $\{s,f\}$ because they want to avoid the temptation of the fish. Note, though that in this case the decision maker doesn't give in to the temptation: they still choose the salad at the second stage. However, it is annoying for them to overcome the temptation of the fish if it is on the menu. You can see this from the maths from the fact that the decision maker receives a negative if the do not choose the most tempting option in the set - i.e. if the $v(x)$ of the x that satisfies $\max_{x\in\{s,f\}}(u(x)+v(x))$ is less that the $v(y)$ of the y that satisfies $\max_{y\in\{s,f\}} v(y)$

What is the justification of calling u the 'commitment' and v the 'temptation' utility? One way to see this is as follows. First, let us think about u. Note that, for any menu containing one item z, we have that

$$\begin{aligned} U(\{z\}) &= \max_{x\in\{z\}}(u(x)+v(x)) - \max_{y\in\{z\}} v(y) \\ &= u(z)+v(z)-v(z) \\ &= u(z) \end{aligned}$$

Thus, if we ask our DM to rank choices over these 'singleton menus', they will do so using the utility function u. Thus, u represents the preferences of the decision maker over consumption at the second stage if there is no temptation. Thus, we call them commitment, or long run preferences

Why can we interpret v as temptation? Well, remember, we said that the sign of temptation is that a subject would prefer to have an object removed from a choice set. Notice that, from a set $\{x,y\}$, the DM is only going to prefer to have y removed (i.e. $\{x\} \succ \{x,y\}$) if $v(y) > v(x)$. Thus there is also a tight link between v and our intuitive concept of temptation.

If you were paying attention in the early lectures in for this course, you should be feeling a little uncomfortable at this stage. When we were talking about utility maximization, we thought long and hard about how to test a model that had an unobservable in it (in that case the utility).

Now we have introduced a model that has TWO unobservables in it, u and v. Surely we should be worried about what the observable implications of this model are, and whether it even has any!

Luckily, Gul and Pesendorfer have done precisely this: They derive axioms on preferences $\succ$ such that are necessary and sufficient of the model we have just described to exist. Unfortunately, the full derivation is too complex for this course (again, see the original paper if you are interested) However, we can discuss the key behavioral restriction of the paper, which is called 'set betweenness'. Remember we said that the standard model required that people always weakly preferred bigger sets to smaller sets (in the subset sense). Clearly, this is not necessarily true in the Gul Pesendorfer model. However, there are still some restrictions on these preferences specifically the following

Axiom 1 (Set Betweenness) *For any $A, B \in 2^Z/\emptyset$, such that $A \succeq B$*

$$A \succeq A \cup B \succeq B$$

In other words, the union of A and B has to be (weakly) better than B and (weakly) worse than A.

While we cannot discuss the sufficiency of this axiom, we can show why it is necessary (i.e. why, if the above model holds it must be the case that set betweenness holds.). We will show why it must be the case that $A \succeq A \cup B$. A similar argument can be used to show that $A \cup B \succeq B$.

First of all, what do we know? Well, by assumption, we know that $A \succeq B$. Let x_A be the item in A that maximizes $u(x) + v(x)$ in A, and y_A be the item that maximizes $v(y)$ in A. Define x_B and y_B similarly. Then, the fact that

$$A \succeq B$$

tells us that

$$u(x_A) + v(x_A) - v(y_A) \geq u(x_B) + v(x_B) - v(y_B)$$

This means that one of two things has to be true. Either $u(x_A) + v(x_A) \geq u(x_B) + v(x_B)$ (i.e. the chosen object in A has higher total utility than the chosen object in B), or $v(y_B) \geq v(y_A)$ (i.e. the most tempting object in B is worse that the most tempting in A) (of course it could also be true that both these things hold). Let's consider the two cases separately (we will show that it must be the case that $U(A \cup B) \leq U(A)$, you should check that you can prove that $U(A \cup B) \geq U(B)$

Case 1 $u(x_A) + v(x_A) \geq u(x_B) + v(x_B)$. If this is the case, then the chosen object in $A \cup B$ will be x_A, and the utility of the set will be

$$\begin{aligned} & U(A \cup B) \\ = \;& u(x_A) + v(x_A) - \max(v(y_A), v(y_B)) \\ \leq \;& u(x_A) + v(x_A) - v(y_A) \\ = \;& U(A) \end{aligned}$$

Case 2 $v(y_B) \geq v(y_A)$. As we already covered the case that $u(x_A)+v(x_A) \geq u(x_B)+v(x_B)$ above, lets assume that $u(x_A) + v(x_A) < u(x_B) + v(x_B)$. In this case, $u(x_B) + v(x_B)$ will be chosen from $A \cup B$, and so

$$\begin{aligned} & U(A \cup B) \\ = \;& u(x_B) + v(x_B) - v(y_B) \\ = \;& U(B) \\ \leq \;& U(A) \end{aligned}$$

Which gives us our conclusion.

Using this model, we can define when it is that a decision maker is tempted, and when they exhibit self control. In fact, we have already discussed temptation: we say that y is tempting relative to x if

$$\{x\} \succ \{x, y\}$$

In the context of the model, this means that $u(x) \geq u(y)$ and $v(x) < v(y)$

What about self control? We say that a DM exhibits self control if they are tempted by an alternative, but they do not give in to that temptation. How can we spot that in a DM. Well, consider the following set of preferences

$$\{x\} \succ \{x, y\} \succ \{y\}$$

What can we conclude from this? We know that y is tempting relative to x, by the fact that $\{x\} \succ \{x, y\}$. However, the fact that $\{x, y\} \succ \{y\}$ tells us that the DM must have overcome the temptation. Why? Assume not, and that the DM would actually choose y from $\{x, y\}$. Then the utility for this set would be

$$
\begin{aligned}
& U(\{x, y\}) \\
= \; & u(y) + v(y) - v(y) \\
= \; & u(y) \\
= \; & U(\{y\})
\end{aligned}
$$

So if the DM was going to give in, then they should be indifferent between $\{x, y\}$ and $\{y\}$. It is only the fact that they are going to overcome temptation that gives them a motive for strictly preferring $\{x, y\}$ to $\{y\}$.