Temptation with Uncertain Normative Preferences

John Stovall

No 1036

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

Temptation with Uncertain Normative Preferences*

John E. Stovall†
University of Warwick

16 December 2013

Abstract

We model a decision maker who anticipates being affected by temptation but is also uncertain about what is normatively best. Our model is an extended version of Gul and Pesendorfer's (2001) where there are three time periods: in the ex-ante period the agent chooses a set of menus, in the interim period she chooses a menu from this set, and in the final period she chooses from the menu. We posit axioms from the ex-ante perspective. Our main axiom on preference states that the agent prefers to have the option to commit in the interim period. Our representation is a generalization of Dekel et al.'s (2009) and identifies the agent's multiple normative preferences and multiple temptations. We also characterize the uncertain normative preference analogue to the representation in Stovall (2010). Finally, we characterize the special case where normative preference is not uncertain. This special case allows us to uniquely identify the representations of Dekel et al. (2009) and Stovall (2010).

1 Introduction

We model a decision maker who anticipates being affected by temptation but is also uncertain about what is normatively best.

Consider an agent who wants to make a healthy choice for dinner but is afraid she will be tempted to choose an unhealthy choice. However, she is uncertain about what is healthiest because of conflicting information from health studies she has read.

Or consider a parent who must make some choices about her young child's future. She wants to provide him with enriching activities that will help develop some untapped talent (e.g. sports, music lessons, art classes), but she doesn't know what he will enjoy or be good at. In addition the parent wants to provide some discipline because she is afraid the child will be tempted to pursue other (less worthwhile) activities.

*I thank Peter Hammond, Takashi Hayashi, Jawwad Noor, and seminar participants at the University of Warwick for comments.

†Email: J.Stovall@warwick.ac.uk

What behavior would characterize an agent like those described above? In the standard approach where the agent has a preference over menus (i.e. sets of alternatives), there is some tension between the two phenomena. Uncertainty about future preferences induces a preference for expanding the menu. However the possibility of future temptation induces a preference for restricting the menu. Though such preferences are not mutually exclusive in this setting, there are problems identifying the different normative and temptation preferences. We discuss this in more detail in Section 2.

In order to separate and identify the effect on preferences of these two phenomena, we consider the expanded domain of preference over sets of menus, which we call neighborhoods. We think of a neighborhood as representing a choice problem over three periods. In the ex-ante period the agent chooses a neighborhood, in the interim period she chooses a menu from the neighborhood, and in the final period she chooses from the menu. This domain has been used by Kopylov (2009b) to generalize the seminal work of Gul and Pesendorfer (2001) and by Kopylov and Noor (2010) to model self-deception.

Our first main axiom on preference, Option to Commit, states that the agent prefers to have the option to commit in the interim period. Thus the agent prefers to defer commitment until just before temptation hits. Our second main axiom, Interim Negative Set Betweenness, roughly states that if the agent has a strict preference for the option to commit to a menu, then she thinks the menu is best in some (subjective) state. We provide representations which are analogues to those in Dekel, Lipman, and Rustichini (2009) (henceforth DLR) and Stovall (2010), but where the normative preference is uncertain.

The addition of uncertain normative preference to these models should be important to applications. For example, Amador et al. (2006) study a consumption-savings model where the agent values both commitment and flexibility. One of their models is in fact an uncertain normative version of Stovall's representation where the agent receives a taste shock to his normative preference and is also uncertain about the strength of temptation to consume rather than save. Thus we provide here an axiomatic foundation to preferences used in their model.

Additionally, we consider the special case where the normative preference is not uncertain. This allows us to uniquely identify all components of DLR's and Stovall's original representations, something which is not possible in the standard domain.

Recent work by Ahn and Sarver (2013) suggests an alternative approach to identifying the representations of DLR and Stovall. Ahn and Sarver consider a two-period model where both ex ante preference over menus and ex post (random) choice from the menu is observed. They ask what joint conditions on ex ante preference and ex post choice imply that the anticipated choice from a menu is the same as the actual choice from the menu. One implication of their result is that with both sets of data (ex ante preference over menus and ex post choice), one is able to uniquely identify the agent's subjective beliefs and state-dependent utilities. Though they do not consider ex ante preferences affected by temptation, Ahn and Sarver's approach suggests that ex post choice data may be useful in identifying the representations of DLR and Stovall. While this may be possible, the present work shows that identification is possible using just ex ante preference.

Related to this discussion is recent work by Dekel and Lipman (2012). They discuss a representation which they call a random Strotz representation. One thing they show is that every preference which has Stovall's representation also has a random Strotz representation. Additionally, they show that these representations imply different choice from a menu. Thus these two representations cannot be differentiated by ex ante preference but they can be by ex post choice. Similar results would apply here.

In the next section we discuss the model and the reasons for the expanded domain in more detail. The main axioms and results are presented in Section 3. Section 4 considers the case where normative preference is constant and the identification of the representation which it allows. Proofs are collected in the appendix.

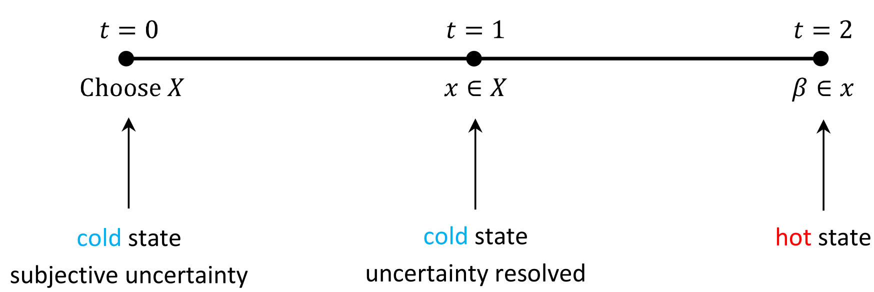
2 Model

Let Δ denote the set of probability distributions over a finite set, and call $\beta \in \Delta$ an alternative. Let $\mathcal{M}$ denote the set of closed, non-empty subsets of Δ, and call $x \in \mathcal{M}$ a menu. Let $\mathcal{N}$ denote the set of closed, non-empty subsets of $\mathcal{M}$, and call $X \in \mathcal{N}$ a neighborhood. Throughout, we will use $\alpha, \beta, \ldots$ to denote elements of Δ, $x, y, \ldots$ to denote elements of $\mathcal{M}$, and $X, Y, \ldots$ to denote elements of $\mathcal{N}$.

We think of a neighborhood as representing a choice problem over three time periods: the ex ante period where she chooses a neighborhood X, the interim period where she chooses a menu $x \in X$, and the ex post period where she chooses an alternative $\beta \in x$.

Our primitive is a binary relation $\succeq$ over $\mathcal{N}$ that represents the agent's ex ante preferences. We do not model choice in the interim or ex post periods explicitly. However, the agent's ex ante preferences will obviously be affected by her (subjective) expectations of her future preferences and temptations.

The time line we envision is the following: When choosing a neighborhood in the ex ante period, the agent is in a "cold" state, meaning she is not affected by temptation. She expects to be in a cold state in the interim as well. But in the ex post period, she expects to be in a "hot" state. She is also uncertain about what her normative preferences and temptations will be, but expects that uncertainty to be resolved in the interim period.



As most of the literature posits preference over the domain of menus $\mathcal{M}$, we now explain why this is not adequate to model our agent. We begin with some definitions. We say that $U : \mathcal{M} \to \mathbb{R}$ is a *Stovall temptation (T^S) representation* if

$$U(x) = \sum_{i=1}^{I} q_i \left\{ \max_{\beta \in x} [u(\beta) + v_i(\beta)] - \max_{\beta \in x} v_i(\beta) \right\}$$

where $q_i > 0$ for all i, $\sum_I q_i = 1$, and u and each v_i are expected-utility (EU) functions. The interpretation of this representation is that u is the agent's normative preference (it represents her preference over singleton menus, which are commitments), each v_i is a temptation, and q_i is the probability the agent assigns to temptation i being realized later. A version of this representation in which the normative preference varied across states would capture the idea that the agent is uncertain about her normative preferences.

We say that $U : \mathcal{M} \to \mathbb{R}$ is a *DLR temptation (T^{DLR}) representation* if

$$U(x) = \sum_{i=1}^{I} q_i \left\{ \max_{\beta \in x} \left[u(\beta) + \sum_{j=1}^{J_i} v_{ij}(\beta) \right] - \sum_{j=1}^{J_i} \max_{\beta \in x} v_{ij}(\beta) \right\}$$

where $q_i > 0$ for all i, $\sum_I q_i = 1$, and u and each v_{ij} are EU functions. Here each state i has multiple temptations which might affect the agent. Note that the T^S representation is a special case of the T^{DLR} representation.

We say that $U : \mathcal{M} \to \mathbb{R}$ is a *finite additive EU representation* if

$$U(x) = \sum_{k=1}^{K} \max_{\beta \in x} w_k(\beta) - \sum_{j=1}^{J} \max_{\beta \in x} v_j(\beta) \tag{1}$$

where each w_k and each v_j is an EU function.[1] Observe that both the T^S and T^{DLR} representations are special cases of (1).

We now show that any finite additive EU representation can be written as an uncertain normative version of the T^S representation. Start with a representation of the form (1). For every k, choose arbitrary $a_{k1}, a_{k2}, \ldots, a_{kJ}$ such that $a_{kj} \geq 0$ for every j and $\sum_J a_{kj} = 1$. Similarly, for every j, choose arbitrary $b_{1j}, b_{2j}, \ldots, b_{Kj}$ such that $b_{kj} \geq 0$ for every k and $\sum_K b_{kj} = 1$. Also, set $I \equiv KJ$ and let $\iota : K \times J \to I$ be any bijection. Finally, for every i, set $u_i \equiv a_{kj} w_k - b_{kj} v_j$ and $\hat{v}_i \equiv b_{kj} v_j$ where $i = \iota(k, j)$. Then we can rewrite (1) as

$$U(x) = \sum_{i=1}^{I} \left\{ \max_{\beta \in x} [u_i(\beta) + \hat{v}_i(\beta)] - \max_{\beta \in x} \hat{v}_i(\beta) \right\}$$

which is an uncertain normative version of the T^S representation with equal probabilities on each state i.

DLR showed a similar result for an uncertain normative version of the T^{DLR} representation. Hence there is no behavioral distinction (in the domain $\mathcal{M}$) between

[1] See DLR for a characterization of the finite additive EU representation. Also, Dekel et al. (2001, 2007) study a broader class of preferences of which this finite additive version is a special case.

an uncertain normative version of the T^S representation, an uncertain normative version of the T^{DLR} representation, and the finite additive EU representation. In addition, as is evident from the construction above, the u_i's are not identified and thus cannot be interpreted as representing the agent's various normative preferences.

As the results in the next section show, we are able to (essentially) uniquely identify the agent's various normative preferences and temptations when we expand the choice domain to $\mathcal{N}$.

3 Main Results

We begin with a set of axioms which are (arguably) independent of the issues of temptation: versions of Dekel et al.'s (2001) key axioms appropriately modified for this domain.[2]

DLR Axioms. $\succeq$ *satisfies Order, Continuity, Independence, and Finiteness.*

See the appendix for complete definitions for our setting. Our next axiom is a modification of the monotonicity axiom introduced by Kreps (1979).

Ex-ante Monotonicity. *If $X \subset Y$, then $Y \succeq X$.*

When an agent is uncertain what her future preferences will be, then she will desire flexibility by preferring larger choice sets. However note that Ex-ante Monotonicity only imposes this preference for flexibility on neighborhoods and not on the menus which make up the neighborhoods. Thus the agent values flexibility only between the ex-ante and interim periods. Since in our model the uncertainty is resolved in the interim period, flexibility per se is not valued after that.

We now introduce our first main axiom concerning temptation. It says that the agent values commitment between the interim and ex-post period.

Option to Commit. *For every x, y, and X, $\{x, y\} \cup X \succeq \{x \cup y\} \cup X$.*

Because of Ex-ante Monotonicity, the agent does not value commitment in the ex-ante period. However Option to Commit says that she does want to be able to commit in the interim period. Most importantly, Option to Commit does not put too much restriction on preferences to preclude uncertainty about normative preference. Consider the following example:

Example 1 *Consider the diet example from the introduction. The agent wants to choose the healthiest meal to eat. However, even though she is indifferent right now between committing to steak and committing to pasta (i.e. $\{\{s\}\} \sim \{\{p\}\}$), this is because she does not know which one will be healthier for her. If low fat diets are healthier, then she will want to choose pasta. However if high protein diets are healthier, then she will want to choose steak. In addition, she is afraid that no matter*

[2] Order, Continuity, and Independence are discussed in Dekel et al. (2001) for the domain $\mathcal{M}$, and in Kopylov (2009b) for the domain $\mathcal{N}$. See Noor and Takeoka (2010, 2011) for arguments against Independence in a temptation setting. Finiteness is discussed in DLR and Kopylov (2009a).

which one is healthier, she will be tempted to choose the other. She also knows that a study will be published tomorrow morning concerning what diet is healthier. Thus she has the preference $\{\{s\},\{p\}\} \succ \{\{s,p\}\}$ *because the former neighborhood gives her the option to commit after finding out which diet is healthier but before she enters the restaurant and is tempted, while the latter neighborhood does not allow her to commit to the (unknown) healthier option and instead guarantees she will experience temptation.*

Thus Option to Commit is consistent with an agent who thinks that an option might be normatively best in one state but tempting in another.

Our first representation takes the following form:[3]

Definition 1 *An* uncertain normative DLR temptation (UNT^{DLR}) representation *of* $\succeq$ *is a tuple* $\left\langle I, \{\langle u_i, J_i, \{v_{ij}\}_{J_i}\rangle\}_I\right\rangle$ *where*

1. $I \in \mathbb{N}_0$,
2. *for every* $i = 1, \ldots, I$, $\langle u_i, J_i, \{v_{ij}\}_{J_i}\rangle$ *is a tuple where*
   - *i.* u_i *is an EU function,*
   - *ii.* $J_i \in \mathbb{N}_0$, *and*
   - *iii. for every* $j = 1, \ldots, J_i$, v_{ij} *is an EU function,*

such that the function

$$\mathcal{U}(X) = \sum_{i=1}^{I} \max_{x \in X} \left\{ \max_{\beta \in x} \left[u_i(\beta) + \sum_{j=1}^{J_i} v_{ij}(\beta) \right] - \sum_{j=1}^{J_i} \max_{\beta \in x} v_{ij}(\beta) \right\}$$

represents $\succeq$.

The interpretation of this representation follows that given earlier: There are I subjective states. In state i, the normative preference is u_i while the v_{ij}'s are the temptations. For a fixed menu $x \in X$, the agent chooses the alternative in x which maximizes $u_i + \sum v_{ij}$ but experiences the disutility $\sum_{J_i} \max_{\beta \in x} v_{ij}(\beta)$, which is the forgone utility from the most tempting alternatives (in state i). For each state i, she chooses a possibly different menu from X which maximizes state i's indirect utility and sums across all states to get the total utility for X.

Note that the UNT^{DLR} representation is the uncertain normative analogue to the T^{DLR} representation. One key difference is that the UNT^{DLR} representation does not have probabilities associated with the states i. This is because such probabilities can not be identified due to the fact that the normative preferences u_i vary across states. We will see later that such probabilities can be identified when normative utility is constant across states.

We now state the first representation result. See the appendix for the definition of a minimal representation.[4]

[3]Some notation: $\mathbb{N}_0$ denotes the natural numbers with 0. For $N \in \mathbb{N}_0$, we use $\{A_n\}_N$ to denote the finite indexed family $\{A_n\}_{n \in \{1,\ldots,N\}}$. If $N = 0$, then $\{1, \ldots, N\}$ is the empty set and statements like "for $n = 1, \ldots, N$, we have ..." are vacuous.

[4]Intuitively, a minimal representation is one that has had all possible redundancies removed.

Theorem 1 *$\succeq$ satisfies the DLR axioms, Ex-ante Monotonicity, and Option to Commit if and only if $\succeq$ has a minimal UNT^{DLR} representation.*

Moreover, the representation is essentially unique. I.e. if $\left\langle I^n, \{\langle u_i^n, J_i^n, \{v_{ij}^n\}_{J_i^n}\rangle\}_{I^n}\right\rangle$, $n = 1, 2$, are both minimal UNT^{DLR} representations of $\succeq$, then:

1. *$I^1 = I^2 (\equiv I)$;*
2. *there exist π a permutation on $\{1, \ldots, I\}$ and scalars $a > 0$ and $b_1, \ldots, b_I$ such that for every i*
   (a) *$J_i^1 = J_{\pi(i)}^2 (\equiv J_i)$, and*
   (b) *$u_i^1 = a u_{\pi(i)}^2 + b_i$; and*
3. *for every i, there exists μ_i a permutation on $\{1, \ldots, J_i\}$ and scalars $c_{i1}, \ldots, c_{iJ_i}$ such that $v_{ij}^1 = a v_{\pi(i)\mu_i(j)}^2 + c_{ij}$ for every $j = 1, \ldots, J_i$.*

One problem with the previous representation is that it allows preferences which are arguably not purely temptation driven. Consider the following example.[5]

Example 2 *Suppose*

$$\{\{\alpha\}\} \sim \{\{\beta\}\} \sim \{\{\alpha\}, \{\beta\}\} \succ \{\{\alpha, \beta\}\}.$$

Since $\{\{\alpha\}\} \sim \{\{\alpha\}, \{\beta\}\}$, this suggests that there is no state in which the agent thinks β is strictly normatively better than α. Similarly, $\{\{\beta\}\} \sim \{\{\alpha\}, \{\beta\}\}$ suggests that there is no state in which the agent thinks α is strictly normatively better than β. Hence the agent thinks α and β are normatively the same across all possible states. However the strict preference for the option to commit $\{\{\alpha\}, \{\beta\}\} \succ \{\{\alpha, \beta\}\}$ suggests that the agent expects one to tempt the other. How can one option tempt the other if they are normatively the same across all states?

The previous example is consistent with Option to Commit, but not our next axiom.[6,7]

Interim Negative Set Betweenness. *If $\{x, x \cup y\} \cup X \succ \{x \cup y\} \cup X$, then $\{x, y\} \cup X \succ \{y\} \cup X$.*

Interim Negative Set Betweenness says that if the agent strictly prefers to have the option to commit to x over $x \cup y$, then that must be because she thinks x might be better than y.

The next representation takes the following form:

[5] Stovall provides a similar example in the preference-over-menus domain.

[6] This axiom is based on one introduced by DLR: a preference $\succeq$ on $\mathcal{M}$ satisfies Negative Set Betweenness if $x \succeq y$ implies $x \cup y \succeq y$. Negative Set Betweenness can be thought of as one half of Gul and Pesendorfer's (2001) Set Betweenness axiom.

[7] The example is inconsistent with Interim Negative Set Betweenness, Ex-ante Monotonicity, and Transitivity. Note that Ex-ante Monotonicity implies $\{\{\alpha\}, \{\alpha, \beta\}\} \succeq \{\{\alpha\}\}$. Transitivity then implies $\{\{\alpha\}, \{\alpha, \beta\}\} \succ \{\{\alpha, \beta\}\}$. But then Interim Set Betweenness is violated since $\{\{\alpha\}, \{\beta\}\} \sim \{\{\beta\}\}$.

Definition 2 *An* uncertain normative Stovall temptation (UNT^S) representation *of* $\succeq$ *is a tuple* $\langle I, \{u_i\}_I, \{v_i\}_I \rangle$ *where*

1. $I \in \mathbb{N}_0$, *and*
2. *for every* $i = 1, \ldots, I$, u_i *and* v_i *are EU functions,*

such that the function

$$\mathcal{U}(X) = \sum_{i=1}^{I} \max_{x \in X} \left\{ \max_{\beta \in x} [u_i(\beta) + v_i(\beta)] - \max_{\beta \in x} v_i(\beta) \right\}$$

represents $\succeq$.

Note that this is the uncertain normative analogue to the T^S representation. Also note that it is a special case of the UNT^{DLR} representation but with each $J_i = 1$.

Theorem 2 $\succeq$ *satisfies the DLR axioms, Ex-ante Monotonicity, Option to Commit, and Interim Negative Set Betweenness if and only if* $\succeq$ *is represented by a minimal* UNT^S *representation.*

Moreover, the representation is essentially unique. I.e. if $\langle I^n, \{u_i^n\}_{I^n}, \{v_i^n\}_{I^n} \rangle$, $n = 1, 2$, *are both minimal* UNT^S *representations of* $\succeq$, *then:*

1. $I^1 = I^2 (\equiv I)$; *and*
2. *there exist* π *a permutation on* $\{1, \ldots, I\}$ *and scalars* $a > 0$ *and* $b_1, \ldots, b_I, c_1, \ldots, c_I$ *such that for every* i, $u_i^1 = a u_{\pi(i)}^2 + b_i$ *and* $v_i^1 = a v_{\pi(i)}^2 + c_i$.

4 Constant Normative Preference

We now focus on the special case when there is no uncertainty about normative preference. This will allow us to give alternative characterizations of the T^{DLR} and T^S representations, but in the choice domain $\mathcal{N}$. Since characterizations of these representations have already been given in the domain $\mathcal{M}$, one may wonder why this is needed. The reason is that important parts of these representations are not identified. Specifically, the subjective state space is not identified, which means that the way the various temptations are assigned to states and the probabilities associated with each state are not identified either. By expanding preferences to the domain $\mathcal{N}$, we are able to (essentially) uniquely identify all parts of these representations.

To see why identification is important, consider the following example.

Example 3 *Suppose there are three final outcomes, and let* $w_1 = (2, 2, -4)$, $w_2 = (1, 2, -3)$, $v_1 = (-1, 2, -1)$, *and* $v_2 = (-2, 2, 0)$ *be vectors representing EU functions over* Δ. *Suppose* $\succeq$ *is a preference over* $\mathcal{M}$ *and has a finite additive EU representation*

$$U(x) = \sum_{k=1}^{2} \max_{\beta \in x} w_k(\beta) - \sum_{j=1}^{2} \max_{\beta \in x} v_j(\beta).$$

Then $\succeq$ can be written as two different T^S representations:

$$U(x) = \frac{1}{2}\left\{\max_{\beta\in x}\left[u(\beta)+\hat{v}_1(\beta)\right] - \max_{\beta\in x}\hat{v}_1(\beta)\right\} + \frac{1}{2}\left\{\max_{\beta\in x}\left[u(\beta)+\hat{v}_2(\beta)\right] - \max_{\beta\in x}\hat{v}_2(\beta)\right\}$$

where $u = w_1 + w_2 - v_1 - v_2 = (6, 0, -6)$, $\hat{v}_1 = 2v_1$, and $\hat{v}_2 = 2v_2$; and

$$U(x) = \frac{1}{3}\left\{\max_{\beta\in x}\left[u(\beta)+\bar{v}_1(\beta)\right] - \max_{\beta\in x}\bar{v}_1(\beta)\right\} + \frac{2}{3}\left\{\max_{\beta\in x}\left[u(\beta)+\bar{v}_2(\beta)\right] - \max_{\beta\in x}\bar{v}_2(\beta)\right\}$$

where u is as above, $\bar{v}_1 = 3v_1$, and $\bar{v}_2 = \frac{3}{2}v_2$.

Recall that for the T^S representation, the interpretation is that $u+v_i$ represents the choice preference in state i, and q_i represents the probability state i is realized. Hence the first representation suggests that the maximizer of $u+\hat{v}_1 = 2w_1$ is chosen $1/2$ of the time, while the second representation suggests that the maximizer of $u+\bar{v}_2 = \frac{3}{2}w_1$ is chosen $2/3$ of the time. But since $u+\hat{v}_1$ and $u+\bar{v}_2$ are cardinally equivalent, they represent the same preference over Δ. This means that the two representations suggest different (random) choice from menus even though they represent the same preference over menus.

So we now consider an axiom which imposes normative preference to be constant across states. This means that normative utility can be normalized across states, and thus the subjective probabilities can be identified.

Constant Normative Preference. *For every α and β, either $\{\{\alpha\},\{\beta\}\} \sim \{\{\alpha\}\}$ or $\{\{\alpha\},\{\beta\}\} \sim \{\{\beta\}\}$ (or both).*

Consider the neighborhood $\{\{\alpha\},\{\beta\}\}$. Since both $\{\alpha\}$ and $\{\beta\}$ are singleton menus, temptation is not an issue for the agent. (She can commit in the interim to either α or β.) If the agent was uncertain whether α or β was normatively best, then both $\{\{\alpha\},\{\beta\}\} \succ \{\{\alpha\}\}$ and $\{\{\alpha\},\{\beta\}\} \succ \{\{\beta\}\}$ would be true. Constant Normative Preference rules out such cases.

Constant Normative Preference is obviously necessary for the following representations:

Definition 3 *A* constant normative DLR temptation (CNT^{DLR}) representation *of $\succeq$ is a tuple $\left\langle I, u, \{q_i\}_I, \{\langle J_i, \{v_{ij}\}_{J_i}\rangle\}_I\right\rangle$ where*

1. *$I \in \mathbb{N}_0$,*
2. *u is an EU function,*
3. *$q_i > 0$ for every $i = 1, \dots, I$ and $\sum_I q_i = 1$,*
4. *for every $i = 1, \dots, I$, $\langle J_i, \{v_{ij}\}_{J_i}\rangle$ is a tuple where*
   - *(a) $J_i \in \mathbb{N}_0$, and*
   - *(b) for every $j = 1, \dots, J_i$, v_{ij} is an EU function,*

such that the function

$$\mathcal{U}(X) = \sum_{i=1}^{I} q_i \max_{x \in X} \left\{ \max_{\beta \in x} \left[u(\beta) + \sum_{j=1}^{J_i} v_{ij}(\beta) \right] - \sum_{j=1}^{J_i} \max_{\beta \in x} v_{ij}(\beta) \right\}$$

represents $\succeq$*.*

Definition 4 *A* constant normative Stovall temptation (CNT^S) representation *of* $\succeq$ *is a tuple* $\langle I, u, \{q_i\}_I, \{v_i\}_I \rangle$ *where*

1. $I \in \mathbb{N}_0$*,*
2. u *is an EU function,*
3. $q_i > 0$ *for every* $i = 1, \ldots, I$ *and* $\sum q_i = 1$*, and*
4. *for every* $i = 1, \ldots, I$*,* v_i *is an EU function,*

such that the function

$$\mathcal{U}(X) = \sum_{i=1}^{I} q_i \max_{x \in X} \left\{ \max_{\beta \in x} [u(\beta) + v_i(\beta)] - \max_{\beta \in x} v_i(\beta) \right\}$$

represents $\succeq$*.*

With our other axioms, Constant Normative Preference is also sufficient for a CNT^S representation.

Theorem 3 $\succeq$ *satisfies the DLR axioms, Ex-ante Monotonicity, Option to Commit, Interim Negative Set Betweenness, and Constant Normative Preference if and only if* $\succeq$ *is represented by a minimal CNT^S representation.*

Moreover, the representation is essentially unique. I.e. if $\langle I^n, u^n, \{q_i^n\}_{I^n}, \{v_i^n\}_{I^n} \rangle$*,* $n = 1, 2$*, are both minimal CNT^S representations of* $\succeq$*, then:*

1. $I^1 = I^2 (\equiv I)$*;*
2. *there exists scalars* $a > 0$ *and* b *such that* $u^1 = au^2 + b$*; and*
3. *there exists* π *a permutation on* $\{1, \ldots, I\}$ *and scalars* $c_1, \ldots, c_I$ *such that for every* i*,* $q_i^1 = q_{\pi(i)}^2$ *and* $v_i^1 = av_{\pi(i)}^2 + c_i$*.*

However, adding Constant Normative Preference to the list of axioms in Theorem 1 is not sufficient to obtain a CNT^{DLR} representation. To see this, note that the representation

$$\mathcal{U}(X) = \sum_{i=1}^{\hat{I}} q_i \max_{x \in X} \left\{ \max_{\beta \in x} \left[u(\beta) + \sum_{j=1}^{J_i} v_{ij}(\beta) \right] - \sum_{j=1}^{J_i} \max_{\beta \in x} v_{ij}(\beta) \right\} + \sum_{i=\hat{I}+1}^{I} \max_{x \in X} \left\{ \max_{\beta \in x} \left[\sum_{j=1}^{J_i} v_{ij}(\beta) \right] - \sum_{j=1}^{J_i} \max_{\beta \in x} v_{ij}(\beta) \right\} \quad (2)$$

would satisfy Constant Normative Preference but it does not in general have a CNT^{DLR} representation.[8]

So consider the following strengthening of Constant Normative Preference.

Monotonicity of Commitments. *If* $\{\{\alpha\}\} \cup X \succ X$ *and* $\{\{\beta\}\} \succ \{\{\alpha\}\}$*, then* $\{\{\beta\}, \{\alpha\}\} \cup X \succ \{\{\alpha\}\} \cup X$.

Consider the statement $\{\{\alpha\}\} \cup X \succ X$. Since $\{\{\alpha\}\}$ represents commitment to the alternative α, this is saying that commitment to α improves the neighborhood X. If commitment to α improves the neighborhood X, then any commitment strictly better than α must improve the neighborhood $\{\{\alpha\}\} \cup X$. This is the content of the axiom.

It is not hard to show that Monotonicity of Commitments implies Constant Normative Preference.

Lemma 1 *If $\succeq$ satisfies Monotonicity of Commitments and Continuity, then $\succeq$ satisfies Constant Normative Preference.*

Proof. Suppose $\{\{\beta\}\} \succ \{\{\alpha\}\}$. Then if we also had $\{\{\alpha\}, \{\beta\}\} \succ \{\{\beta\}\}$, Monotonicity of Commitments would imply $\{\{\alpha\}, \{\beta\}\} \succ \{\{\alpha\}, \{\beta\}\}$ (taking $X = \{\{\beta\}\}$), a contradiction. Hence if $\{\{\beta\}\} \succ \{\{\alpha\}\}$, then we must have $\{\{\alpha\}, \{\beta\}\} \sim \{\{\beta\}\}$. Similarly, if $\{\{\alpha\}\} \succ \{\{\beta\}\}$, then we must have $\{\{\alpha\}, \{\beta\}\} \sim \{\{\alpha\}\}$. Continuity guarantees that if $\{\{\alpha\}\} \sim \{\{\beta\}\}$, then we must have $\{\{\alpha\}, \{\beta\}\} \sim \{\{\alpha\}\} \sim \{\{\beta\}\}$. Hence, no matter how $\{\{\alpha\}\}$ and $\{\{\beta\}\}$ are ranked, Constant Normative Preference holds. ■

Returning to the representation in equation (2), note that if $\hat{I} = 0$, then such a representation would trivially satisfy Monotonicity of Commitments as the commitment preference would be constant. One could write this representation as a CNT^{DLR} representation by simply defining u to be the constant function 0, but then the probabilities q_i would not be identified. To rule out such a case, we include the following non-triviality axiom.

Conditional Non-triviality. *If there exist X and Y such that $X \succ Y$, then there exist α and β such that $\{\{\alpha\}\} \succ \{\{\beta\}\}$.*

Note that we would not want to impose Conditional Non-triviality in the case of uncertain normative preference. Returning to Example 1, we had $\{\{s\}\} \sim \{\{p\}\}$ yet $\{\{s\}, \{p\}\} \succ \{\{s, p\}\}$ as the agent was uncertain which dish would be healthiest and which would tempt.

Theorem 4 *$\succeq$ satisfies the DLR axioms, Ex-ante Monotonicity, Option to Commit, Monotonicity of Commitments, and Conditional Non-triviality if and only if $\succeq$ is represented by a minimal CNT^{DLR} representation with a non-constant commitment preference.*

[8] Indeed this set of axioms characterizes this representation. (This result follows directly from Lemma 6.) Note also that this representation is the analogue to what DLR call a weak temptation representation.

Moreover, the representation is essentially unique. I.e. if $\left\langle I^n, u^n, \{q_i^n\}_{I^n}, \{\langle J_i^n, \{v_{ij}^n\}_{J_i^n}\rangle\}_{I^n}\right\rangle$, $n = 1, 2$, *are both minimal CNT*^{DLR} *representations with a non-constant commitment preference and both represent* $\succeq$*, then:*

1. $I^1 = I^2 (\equiv I)$;
2. *there exist scalars* $a > 0$ *and* b *such that* $u^1 = au^2 + b$;
3. *there exists* π *a permutation on* $\{1, \ldots, I\}$ *such that* $q_i^1 = q_{\pi(i)}^2$ *and* $J_i^1 = J_{\pi(i)}^2 (\equiv J_i)$ *for every* i*; and*
4. *for every* i*, there exist* μ_i *a permutation on* $\{1, \ldots, J_i\}$ *and scalars* $c_{i1}, \ldots, c_{iJ_i}$ *such that* $v_{ij}^1 = av_{\pi(i)\mu_i(j)}^2 + c_{ij}$ *for every* $j = 1, \ldots, J_i$.

One interesting aspect of this last theorem is that we are able to obtain DLR's representation without a technical axiom like their Approximate Improvements Are Chosen (AIC). The intuition behind AIC is complicated and relies on considering the closure of the set of improvements of a menu. (An improvement of a menu is simply an alternative that, when added to the menu, improves that menu.) Though our domain is certainly more complicated than the one used by DLR, the axioms imposed are much more intuitive than AIC.

Appendix

A Preliminaries

In what follows, we will abuse some notation: We use the cardinality of a set to also denote the set itself. Thus we will write $i \in I$ instead of $i \in \{1, \ldots, I\}$, etc. We will identify an EU function with its corresponding vector in Euclidean space consisting of utilities of pure outcomes, e.g. $u(\beta) = u \cdot \beta$. We will identify a finite additive EU representation U of the form given in equation (1) with the tuple $\langle K, J, \{w_k\}_K, \{v_j\}_J \rangle$, where K and J are finite, and each w_k and v_j is an EU function. If such is the case then we will alternatively write $U = \langle K, J, \{w_k\}_K, \{v_j\}_J \rangle$.

Throughout, we use $\mathbf{0}$ and $\mathbf{1}$ to represent the vectors of 0's and 1's respectively. Note that if the vector w satisfies $w \cdot \mathbf{1} = 0$ and if β is in the interior of Δ, then $\beta + \epsilon w$ is also a lottery for small enough ϵ.

We use the usual metric over Δ. We endow $\mathcal{M}$ with the Hausdorff topology and define the mixture of two menus $x, y \in \mathcal{M}$ as

$$\lambda x + (1 - \lambda) y \equiv \{\lambda \beta + (1 - \lambda) \beta' : \beta \in x, \beta' \in y\}$$

for $\lambda \in [0, 1]$.

Similarly, we endow $\mathcal{N}$ with the Hausdorff topology and define the mixture of two neighborhoods $X, Y \in \mathcal{N}$ as

$$\lambda X + (1 - \lambda) Y \equiv \{\lambda x + (1 - \lambda) y : x \in X, y \in Y\}$$

for $\lambda \in [0, 1]$.

The following are the standard von Neumann–Morgenstern axioms modified for this domain.

Order. *$\succeq$ is complete and transitive.*

Continuity. *The sets $\{X : X \succeq Y\}$ and $\{X : Y \succeq X\}$ are closed.*

Independence. *If $X \succ Y$, then for $Z \in \mathcal{N}$ and $\lambda \in (0, 1]$,*

$$\lambda X + (1 - \lambda) Z \succ \lambda Y + (1 - \lambda) Z.$$

Now we introduce our finiteness axiom. Before we state the axiom, we need some definitions.

Definition 5 *$Y \subset X$ is* critical for *X if if for all Y' where $Y \subset Y' \subset X$, we have $Y' \sim X$.*

Note that every neighborhood is critical for itself.

Definition 6 *y is* critical for *$x \in X$ if for all y' where $y \subset y' \subset x$, we have $(X \setminus \{x\}) \cup \{y'\} \sim X$.*

Note that every menu is critical for itself in any neighborhood.

Finiteness. *There exists $N \in \mathbb{N}$ such that for every neighborhood X: (i) there exists Y critical for X where $|Y| < N$, and (ii) for every menu $x \in X$, there exists y critical for $x \in X$ where $|y| < N$.*

Lemma 2 *Finiteness implies Kopylov's (2009a) Finiteness.*

Proof. By way of contradiction, suppose not. Then there exists $X_1, \ldots, X_N$ such that $\bigcup_{n \in N} X_n \equiv X_\sigma \not\sim X_{-m} \equiv \bigcup_{n \in N \setminus \{m\}} X_n$ for every $m \in N$. By Finiteness, there exists Y critical for X_σ such that $|Y| < N$. But this implies that there exists m^* such that $Y \subset X_{-m^*}$. (If not, then $|Y| \geq N$.) But since Y is critical for X_σ and since $X_{-m^*} \subset X_\sigma$, we have $X_{-m^*} \sim X_\sigma$, a contradiction. ■

The following representation will serve as a foundation for all subsequent representations. Let $\mathcal{U} : \mathcal{N} \to \mathbb{R}$ take the form

$$\mathcal{U}(X) = \sum_{i=1}^{I} \max_{x \in X} U_i(x), \tag{3}$$

where each U_i is a finite additive EU representation. Identify this representation with the tuple $\langle I, \{U_i\}_I \rangle$, where I is finite and $U_i = \langle K_i, J_i, \{w_{ik}\}_K, \{v_{ij}\}_J \rangle$ is a finite additive EU representation for each i.

Before characterizing this representation, the following definitions will be useful for the uniqueness results. Let $\mathcal{L}$ be the set of all continuous linear functions $f : L \to \mathbb{R}$, where L is a linear space. By the Mixture Space Theorem (Herstein and Milnor, 1953), $f, g \in \mathcal{L}$ represent the same ordering over L if and only if there exists $a > 0$ and $b \in \mathbb{R}$ such that $f = ag + b$. We say an indexed family of linear functions $\{f_i\}_I$ is *redundant* if there exists a constant function in this set, or if there exist $i, j \in I$ where $i \neq j$ such that f_i and f_j represent the same ordering over L. For any $f, g \in \mathcal{L}$ and $a > 0$, we write $f \bowtie_a g$ if there exists $b \in \mathbb{R}$ such that $f = ag + b$. More generally, for any two indexed families of linear functions with the same index set $\{f_i\}_I, \{g_i\}_I$, and for any $a > 0$, we write (abusing notation)

$$\{f_i\}_I \bowtie_a \{g_i\}_I$$

if there exists a permutation π over I such that $f_i \bowtie_a g_{\pi(i)}$ for every i.

We say a finite additive EU representation $\langle K, J, \{w_k\}_K, \{v_j\}_J \rangle$ is *minimal* if $\{w_k\}_K \cup \{v_j\}_J$ is not redundant.

The uniqueness result from Kopylov (2009a, Theorem 2.1) implies that two minimal finite additive EU representations $U^n = \left\langle K^n, J^n, \{w_k^n\}_{K^n}, \{v_j^n\}_{J^n} \right\rangle$, $n = 1, 2$, represent the same preference over $\mathcal{M}$ if and only if $K^1 = K^2$, $J^1 = J^2$, and there exists $a > 0$ such that $\{w_k^1\}_{K^1} \bowtie_a \{w_k^2\}_{K^2}$ and $\{v_j^1\}_{J^1} \bowtie_a \{v_j^2\}_{J^2}$. Thus these conditions hold if and only if $U^1 \bowtie_a U^2$.

We can now state the definition of a minimal representation that we use in all of our theorems.

Definition 7 *We say that a representation taking the form of (3) is* minimal *if* $\{U_i\}_I$ *is not redundant and each* U_i *is minimal.*

Theorem 5 $\succeq$ *satisfies Order, Continuity, Independence, Finiteness, and Ex-ante Monotonicity if and only if* $\succeq$ *has a representation in the form of (3) that is minimal.*

Moreover, the representation is essentially unique. I.e. if $\langle I^n, \{U_i^n\}_{I^n}\rangle$, $n = 1, 2$ *are both minimal and represent* $\succeq$*, then:*

1. $I^1 = I^2$*; and*

2. *there exists* $a > 0$ *such that* $\{U_i^1\}_{I^1} \bowtie_a \{U_i^2\}_{I^2}$*.*

Proof. Showing the axioms are necessary is a straightforward exercise. So turn to sufficiency. By Lemma 2, we apply the result from Kopylov (2009a, Theorem 2.1) to obtain the representation of $\succeq$:

$$\mathcal{U}(X) = \sum_{i=1}^{I} \max_{x\in X} U_i(x) - \sum_{m=1}^{M} \max_{x\in X} V_m(x)$$

where $I, M \geq 0$, each U_i and V_m is a continuous linear function from $\mathcal{M}$ to $\mathbb{R}$, and the set $\{U_1, \ldots, U_I, V_1, \ldots, V_M\}$ is not redundant. Furthermore, since $\succeq$ satisfies Ex-ante Monotonicity, the same result from Kopylov implies that $M = 0$.

For every $i \in I$, let $\succeq_i$ by the binary relation over $\mathcal{M}$ implied by U_i. Since $U_i : \mathcal{M} \to \mathbb{R}$ is a continuous linear function for every i, then $\succeq_i$ satisfies the analogues to Order, Continuity, and Independence.

We show now that for every $i \in I$, $\succeq_i$ satisfies Kopylov's (2009a) Finiteness for a preference relation over $\mathcal{M}$. So fix i^*. By way of contradiction, suppose not. Then there exists $x_1, \ldots, x_N$ such that $U_{i^*}(x_\sigma) \neq U_{i^*}(x_{-m})$ for every $m \in N$ (where $x_\sigma \equiv \cup_{n\in N} x_n$ and $x_{-m} \equiv \cup_{n\in N\setminus\{m\}} x_n$ for every m). By Kopylov (2009a, Lemma A.1), there exists $z_1, \ldots, z_I$ such that $U_j(z_j) > U_j(z_k)$ for every $j, k \in I$ where $j \neq k$. Hence (by continuity) there exists $\epsilon > 0$ such that

$$U_j((1-\epsilon)z_j + \epsilon x_\sigma) > U_j((1-\epsilon)z_k + \epsilon x) \tag{4}$$

for every $j, k \in I$ where $j \neq k$, and for every $x \in \{x_\sigma, x_{-1}, \ldots, x_{-N}\}$. Set $\bar{z}_j \equiv (1-\epsilon)z_j + \epsilon x_\sigma$ for every $j \in I$. Set $X \equiv \{\bar{z}_j\}_I$. By Finiteness, there exists y critical for $\bar{z}_i \in X$ such that $|y| < N$. But then there must exist $m^* \in N$ such that $y \subset (1-\epsilon)z_{i^*} + \epsilon x_{-m^*} \equiv \bar{x}_{m^*}$. (If not, then $|Y| \geq N$.) Note that $\bar{x}_{m^*} \subset \bar{z}_{i^*}$ since $x_{-m^*} \subset x_\sigma$. Since y is critical for $\bar{z}_{i^*} \in X$, this implies that $(X \setminus \{\bar{z}_{i^*}\}) \cup \{\bar{x}_{m^*}\} \sim X$. Equation (4) implies that $\bar{z}_j = \arg\max_{x\in X\cup\{\bar{x}_{m^*}\}} U_j(x)$ for every $j \neq i^*$. Hence it must be that $U_{i^*}(\bar{z}_{i^*}) = U_{i^*}(\bar{x}_{m^*})$. But the linearity of U_{i^*} implies

$$\begin{aligned}
U_{i^*}(\bar{z}_{i^*}) &= U_{i^*}(\bar{x}_{m^*}) \\
U_{i^*}((1-\epsilon)z_{i^*} + \epsilon x_\sigma) &= U_{i^*}((1-\epsilon)z_{i^*} + \epsilon x_{-m^*}) \\
(1-\epsilon)U_{i^*}(z_{i^*}) + \epsilon U_{i^*}(x_\sigma) &= (1-\epsilon)U_{i^*}(z_{i^*}) + \epsilon U_{i^*}(x_{-m^*}) \\
U_{i^*}(x_\sigma) &= U_{i^*}(x_{-m^*}).
\end{aligned}$$

But this contradicts $U_{i^*}(x_\sigma) \neq U_{i^*}(x_{-m^*})$.

We apply the result from Kopylov (2009a, Theorem 2.1) to $\succeq_i$ to get a finite additive EU representation for U_i. Moreover, U_i is minimal.

The uniqueness result follows as well from Kopylov (2009a, Theorem 2.1). ■

Lemma 3 *Let $\{U_i\}_I$ be a non-redundant indexed family of minimal finite additive EU representations, where $U_i = \langle K_i, J_i, \{w_{ik}\}_{K_i}, \{v_{ij}\}_{J_i}\rangle$ for every $i \in I$. For $i \in I$ and $m \in K_i \cup J_i$, let $u_{im} = w_{im}$ if $m \in K_i$ and $u_{im} = v_{im}$ if $m \in J_i$. Let $u_{im} \cdot \mathbf{1} = 0$ for every $i \in I$ and every $m \in K_i \cup J_i$.*

Then there exists $x_1, ..., x_I$ (in the interior of Δ) such that

1. *$U_i(x_i) > U_i(x_j)$ for every $i \neq j$, and*
2. *for any i, for any $m, n \in K_i \cup J_i$ where $m \neq n$, $\arg\max_{\beta \in x_i} u_{im}(\beta)$ is a singleton and $\arg\max_{\beta \in x_i} u_{im}(\beta) \neq \arg\max_{\beta \in x_i} u_{in}(\beta)$.*

Proof. Let S denote the set of all w_{ik}'s and v_{ij}'s normalized to have unit length, i.e.

$$S \equiv \left\{ s = \frac{u_{im}}{\sqrt{u_{im} \cdot u_{im}}} : i \in I \text{ and } m \in K_i \cup J_i \right\}.$$

Obviously S is finite. Also, for every $s \in S$, $s \cdot \mathbf{1} = 0$ and there exists $m \in K_i \cup J_i$ such that u_{im} and s represent the same ordering over Δ. Thus for every i, we can write

$$U_i(y) = \sum_{s \in S} \max_{\beta \in y} b_{is} s \cdot \beta$$

where $b_{is} > 0$ if there exists $k \in K_i$ such that w_{ik} and s represent the same ordering, $b_{is} < 0$ if there exists $j \in J_i$ such that v_{jk} and s represent the same ordering, and $b_{is} = 0$ otherwise. (Since U_i is minimal, exactly one of these holds for every s.)

Let x^* denote a sphere in the interior of Δ. For every $s \in S$, set $\beta_s \equiv \arg\max_{\beta \in x^*} s \cdot \beta$. Note that for $s \neq s'$, we have $\beta_s \neq \beta_{s'}$. Set $x \equiv \{\beta_s\}_S$. Hence $U_i(x) = \sum_{s \in S} b_{is} s \cdot \beta_s$. For $a \in \mathbb{R}^S$ and $\epsilon > 0$, set

$$\bar{x}(\epsilon, a) \equiv \{\beta_s + \epsilon a_s s\}_{s \in S}.$$

For fixed a, there exists ϵ_a small enough such that $\beta_s + \epsilon_a a_s s$ is in the interior of Δ and such that $\beta_s + \epsilon_a a_s s = \arg\max_{\beta \in \bar{x}(\epsilon_a, a)} s \cdot \beta$. For every i, set

$$a_i \equiv \left\{ \frac{b_{is}}{\sqrt{\sum_{s' \in S} b_{is'}^2}} \right\}_{s \in S}.$$

Set $\epsilon \equiv \min_i \epsilon_{a_i}$. For every i, set $x_i \equiv \bar{x}(\epsilon, a_i)$. Note that $U_i(x_i) = U_i(x) + \epsilon \sum_{s \in S} a_{is} b_{is}$. Hence $x_i = \arg\max_{i' \in I} U_i(x_{i'})$ since $\{U_i\}_I$ is not redundant and since a_i is the unique solution to the constrained maximization problem: $\max_{\bar{a}} \sum_{s \in S} \bar{a}_s b_{is}$ subject to $\sum_{s \in S} \bar{a}_s^2 = 1$. This proves the first part.

The second part follows from the fact that each U_i is minimal and that $\beta_s + \epsilon a_{is} s = \arg\max_{\beta \in x_i} s \cdot \beta$ for every s. ■

B Proofs

B.1 Proof for Theorem 1

First we show that Option to Commit is necessary. Let x, y, X be given. Fix i. Set $w_i \equiv u_i + \sum_{j \in J_i} v_j$. Observe that either $\max_{\beta \in x} w_i(\beta) = \max_{\beta \in x \cup y} w_i(\beta)$ or $\max_{\beta \in y} w_i(\beta) = \max_{\beta \in x \cup y} w_i(\beta)$. Note also that for every $j \in J_i$, we have $\max_{\beta \in x} v_j(\beta) \leq \max_{\beta \in x \cup y} v_j(\beta)$ and $\max_{\beta \in y} v_j(\beta) \leq \max_{\beta \in x \cup y} v_j(\beta)$. Hence either

$$\max_{\beta \in x} w_i(\beta) - \sum_{j \in J_i} \max_{\beta \in x} v_j(\beta) \geq \max_{\beta \in x \cup y} w_i(\beta) - \sum_{j \in J_i} \max_{\beta \in x \cup y} v_j(\beta).$$

or

$$\max_{\beta \in y} w_i(\beta) - \sum_{j \in J_i} \max_{\beta \in y} v_j(\beta) \geq \max_{\beta \in x \cup y} w_i(\beta) - \sum_{j \in J_i} \max_{\beta \in x \cup y} v_j(\beta).$$

Thus for every i,

$$\max_{x' \in \{x,y\} \cup X} \{\max_{\beta \in x'} [u_i(\beta) + \sum_{j \in J_i} v_j(\beta)] - \sum_{j \in J_i} \max_{\beta \in x'} v_j(\beta)\}$$
$$\geq \max_{x' \in \{x \cup y\} \cup X} \{\max_{\beta \in x'} [u_i(\beta) + \sum_{j \in J_i} v_j(\beta)] - \sum_{j \in J_i} \max_{\beta \in x'} v_j(\beta)\}.$$

Now we show that the axioms are sufficient. We will need the following lemma.

Lemma 4 *Let $\succeq$ have a representation in the form of (3) that is minimal. If $\succeq$ also satisfies Option to Commit, then $|K_i| \leq 1$ for every i.*

Proof. Since $\{U_i\}_I$ is not redundant, take $x_1, ..., x_I$ from Lemma 3 and set $X \equiv \{x_1, ..., x_I\}$. According to the uniqueness result of Theorem 5, we can assume without loss of generality that $w_{ik} \cdot \mathbf{1} = 0$ for every i and every $k \in K_i$. Fix i^* and by way of contradiction suppose $|K_{i^*}| > 1$. For any $k \in K_{i^*}$, set $\alpha_k \equiv \arg\max_{\beta \in x_{i^*}} w_{i^*k}(\beta)$. (Lemma 3 guarantees this max is a singleton.) For any $\epsilon > 0$, set $x_k^\epsilon \equiv x_{i^*} \cup \{\alpha_k + \epsilon w_{i^*k}\}$. Take $k, k' \in K_{i^*}$ such that $k \neq k'$. By Lemma 3, $U_i(x_i) > U_i(x_{i^*})$ for every $i \neq i^*$ and $\max_{\beta \in x_{i^*}} v_{i^*j}(\beta) > \max\{v_{i^*j}(\alpha_k), v_{i^*j}(\alpha_{k'})\}$ for every $j \in J_{i^*}$. Hence, there exists $\epsilon > 0$ such that the following hold for every $i \neq i^*$ and $j \in J_{i^*}$:

$$U_i(x_i) > \max\{U_i(x_k^\epsilon \cup x_{k'}^\epsilon), U_i(x_k^\epsilon), U_i(x_{k'}^\epsilon)\}$$

and

$$\max_{\beta \in x_{i^*}} v_{i^*j}(\beta) = \max_{\beta \in x_k^\epsilon} v_{i^*j}(\beta) = \max_{\beta \in x_{k'}^\epsilon} v_{i^*j}(\beta).$$

This implies

$$U_{i^*}(x_k^\epsilon \cup x_{k'}^\epsilon) > U_{i^*}(x_k^\epsilon), U_{i^*}(x_{k'}^\epsilon) > U_{i^*}(x_{i^*}).$$

Hence $\mathcal{U}(\{x_k^\epsilon \cup x_{k'}^\epsilon\} \cup X) > \mathcal{U}(\{x_k^\epsilon, x_{k'}^\epsilon\} \cup X)$, violating Option to Commit. ■

So by Theorem 5, $\succeq$ has a representation $\mathcal{U}$ in the form of (3) which is minimal, and by Lemma 4 $|K_i| \leq 1$ for every i. The UNT^{DLR} representation follows by setting $u_i = w_i - \sum_{J_i} v_j$ for every i, where $w_i = w_k$ for $k \in K_i$ if $|K_i| = 1$ and $w_i = \mathbf{0}$ if $|K_i| = 0$. The uniqueness result follows from Theorem 5.

B.2 Proof for Theorem 2

First we show that Interim Negative Set Betweenness is necessary. So let x, y, X satisfy $\{x, x \cup y\} \cup X \succ \{x \cup y\} \cup X$. Then there exists $i \in I$ such that

$$\max_{\beta \in x}[u_i(\beta) + v_i(\beta)] - \max_{\beta \in x} v_i(\beta) > \max_{\beta \in z}[u_i(\beta) + v_i(\beta)] - \max_{\beta \in z} v_i(\beta)$$

for every $z \in \{x \cup y\} \cup X$. Specifically, this holds when $z = x \cup y$. This implies

$$\max_{\beta \in x} v_i(\beta) < \max_{\beta \in x \cup y} v_i(\beta) = \max_{\beta \in y} v_i(\beta).$$

Since

$$\max_{\beta \in x \cup y}[u_i(\beta) + v_i(\beta)] \geq \max_{\beta \in y}[u_i(\beta) + v_i(\beta)]$$

we have

$$\max_{\beta \in x \cup y}[u_i(\beta) + v_i(\beta)] - \max_{\beta \in x \cup y} v_i(\beta) \geq \max_{\beta \in y}[u_i(\beta) + v_i(\beta)] - \max_{\beta \in y} v_i(\beta).$$

Hence

$$\max_{\beta \in x}[u_i(\beta) + v_i(\beta)] - \max_{\beta \in x} v_i(\beta) > \max_{\beta \in z}[u_i(\beta) + v_i(\beta)] - \max_{\beta \in z} v_i(\beta)$$

for every $z \in \{y\} \cup X$, which implies that $\{x, y\} \cup X \succ \{y\} \cup X$.

Now we show the axioms are sufficient. We will need the following lemma.

Lemma 5 *Let $\succeq$ have a representation in the form of (3) that is minimal. If $\succeq$ satisfies Interim Negative Set Betweenness, then $|J_i| \leq 1$ for every i.*

Proof. Since $\{U_i\}_I$ is not redundant, take $x_1, ..., x_I$ from Lemma 3. According to the uniqueness result of Theorem 5, we can assume without loss of generality that $v_{ij} \cdot \mathbf{1} = 0$ for every i and every $j \in J_i$. Fix i^* and by way of contradiction suppose $|J_{i^*}| > 1$. For any $j \in J_{i^*}$, set $\alpha_j \equiv \arg\max_{\beta \in x_{i^*}} v_{i^*j}(\beta)$. (Lemma 3 guarantees this max is a singleton.) For any $\epsilon > 0$, set $x_j^\epsilon \equiv x_{i^*} \cup \{\alpha_j + \epsilon v_{i^*j}\}$. Set $X_{-i^*} \equiv \{x_1, ..., x_I\} \setminus \{x_{i^*}\}$. Take $j, j' \in J_{i^*}$ such that $j \neq j'$. By Lemma 3, $U_i(x_i) > U_i(x_{i^*})$ and $U_{i^*}(x_{i^*}) > U_{i^*}(x_i)$ for every $i \neq i^*$, $\max_{\beta \in x_{i^*}} w_{i^*k}(\beta) > w_{i^*k}(\alpha_j)$ for every $k \in K_{i^*}$, $v_{i^*j'}(\alpha_{j'}) > v_{i^*j'}(\alpha_j)$, and $v_{i^*j}(\alpha_j) > v_{i^*j}(\alpha_{j'})$. Hence, there exists $\epsilon > 0$ such that the following hold for every $i \neq i^*$ and $k \in K_{i^*}$:

$$U_i(x_i) > \max\{U_i(x_j^\epsilon \cup x_{j'}^\epsilon), U_i(x_j^\epsilon), U_i(x_{j'}^\epsilon)\},$$

$$U_{i^*}(x_j^\epsilon \cup x_{j'}^\epsilon) > U_{i^*}(x_i),$$

$$\max_{\beta \in x_{i^*}} w_{i^*k}(\beta) = \max_{\beta \in x_j^\epsilon} w_{i^*k}(\beta) = \max_{\beta \in x_{j'}^\epsilon} w_{i^*k}(\beta),$$

$$v_{i^*j'}(\alpha_{j'}) > v_{i^*j'}(\alpha_j + \epsilon v_{i^*j}),$$

and

$$v_{i^*j}(\alpha_j) > v_{i^*j}(\alpha_{j'} + \epsilon v_{i^*j'}).$$

Hence $U_{i^*}(x_j^\epsilon) > U_{i^*}(x_j^\epsilon \cup x_{j'}^\epsilon)$ and $U_{i^*}(x_{j'}^\epsilon) > U_{i^*}(x_j^\epsilon \cup x_{j'}^\epsilon)$. Without loss of generality, assume $U_{i^*}(x_j^\epsilon) \geq U_{i^*}(x_{j'}^\epsilon)$. It is easy to verify then that $\mathcal{U}(\{x_{j'}^\epsilon, x_j^\epsilon \cup x_{j'}^\epsilon\} \cup X_{-i^*}) > \mathcal{U}(\{x_j^\epsilon \cup x_{j'}^\epsilon\} \cup X_{-i^*})$ and $\mathcal{U}(\{x_j^\epsilon, x_{j'}^\epsilon\} \cup X_{-i^*}) = \mathcal{U}(\{x_j^\epsilon\} \cup X_{-i^*})$, violating Interim Negative Set Betweenness. ■

So by Theorem 5, $\succeq$ has a representation $\mathcal{U}$ in the form of (3) which is minimal, and by Lemmas 4 and 5, $|K_i| \leq 1$ and $|J_i| \leq 1$ for every i. For every i where $|K_i| = 1$ set $w_i = w_k$ where $k \in K_i$, otherwise set $w_i = \mathbf{0}$. Similarly, if $|J_i| = 1$ then set $v_i = v_j$ for $j \in J_i$, otherwise set $v_i = \mathbf{0}$. The UNT^S representation follows by setting $u_i \equiv w_i - v_i$. The uniqueness result follows from Theorem 5.

B.3 Proof for Theorem 3

The necessity of Constant Normative Preference is obvious. The sufficiency part relies on the following lemma.

Lemma 6 *Let $\succeq$ have a representation in the form of (3) that is minimal. If $\succeq$ satisfies Constant Normative Preference, then there exists an EU function u such that for every $i \in I$, either $u_i \equiv \sum_{K_i} w_k - \sum_{J_i} v_j$ and u represent the same preference over Δ or u_i is constant.*

Proof. The proof is trivial if $|I| = 0$ or 1. So assume $|I| \geq 2$ and that there exist $i, i' \in I$ such that u_i and $u_{i'}$ are both non-constant and represent different preferences over Δ. Then there exist α and α' such that $u_i(\alpha) > u_i(\alpha')$ and $u_{i'}(\alpha') > u_{i'}(\alpha)$. But this implies $\{\{\alpha\}, \{\alpha'\}\} \succ \{\{\alpha\}\}$ and $\{\{\alpha\}, \{\alpha'\}\} \succ \{\{\alpha'\}\}$, violating Constant Normative Preference. ■

So let $\succeq$ satisfy the stated axioms. By Theorem 2, $\succeq$ has a minimal UNT^S representation

$$\mathcal{U}(X) = \sum_{i \in I} \max_{x \in X} \left\{ \max_{\beta \in x} [u_i(\beta) + v_i(\beta)] - \max_{\beta \in x} v_i(\beta) \right\}.$$

By Lemma 6, there exists u such that for every i, $u_i = q_i u + b_i$ for some $q_i \geq 0$ and $b_i \in \mathbb{R}$. By the previous uniqueness results, we can assume without loss of generality that $\sum_I q_i = 1$ and that $b_i = 0$ for every i. Also, the minimality of $\mathcal{U}$ implies that $q_i > 0$ for every i. Thus for every i, set $\hat{v}_i \equiv v_i / q_i$. This gives us the CNT^S representation

$$\mathcal{U}(X) = \sum_{i \in I} q_i \max_{x \in X} \left\{ \max_{\beta \in x} [u(\beta) + \hat{v}_i(\beta)] - \max_{\beta \in x} \hat{v}_i(\beta) \right\}.$$

The uniqueness follows from the previous results.

B.4 Proof for Theorem 4

First we show the necessity of the axioms. So let $\succeq$ have the CNT^{DLR} representation

$$\mathcal{U}(X)=\sum_{i\in I}q_i\max_{x\in X}\left\{\max_{\beta\in x}\left[u(\beta)+\sum_{j\in J_i}v_{ij}(\beta)\right]-\sum_{j\in J_i}\max_{\beta\in x}v_{ij}(\beta)\right\}$$

where u is non-constant.

For Conditional Non-triviality, if $X\succ Y$, then that implies that $I>0$. Since u is non-constant, then there exists α and β such that $u(\alpha)>u(\beta)$. But since $I>0$, this implies $U(\{\{\alpha\}\})>U(\{\{\beta\}\})$.

For Monotonicity of Commitments, suppose $\{\{\alpha\}\}\cup X\succ X$ and $\{\{\beta\}\}\succ\{\{\alpha\}\}$. Since $\{\{\alpha\}\}\cup X\succ X$, it must be that there exists i such that

$$u(\alpha)>\max_{x\in X}\left\{\max_{\beta\in x}\left[u(\beta)+\sum_{j\in J_i}v_{ij}(\beta)\right]-\sum_{j\in J_i}\max_{\beta\in x}v_{ij}(\beta)\right\}.$$

Since $\{\{\beta\}\}\succ\{\{\alpha\}\}$, we have $u(\beta)>u(\alpha)$. But then we must have

$$u(\beta)>\max_{x\in\{\{\alpha\}\}\cup X}\left\{\max_{\beta\in x}\left[u(\beta)+\sum_{j\in J_i}v_{ij}(\beta)\right]-\sum_{j\in J_i}\max_{\beta\in x}v_{ij}(\beta)\right\}.$$

Hence $\{\{\beta\},\{\alpha\}\}\cup X\succ\{\{\alpha\}\}\cup X$.

Now we show the axioms are sufficient. By Theorem 1, $\succeq$ has a minimal UNT^{DLR} representation

$$\mathcal{U}(X)=\sum_{i\in I}\max_{x\in X}\left\{\max_{\beta\in x}\left[u_i(\beta)+\sum_{j\in J_i}v_{ij}(\beta)\right]-\sum_{j\in J_i}\max_{\beta\in x}v_{ij}(\beta)\right\}$$

For every i, set

$$V_i(x)\equiv\max_{\beta\in x}\left[u_i(\beta)+\sum_{j\in J_i}v_{ij}(\beta)\right]-\sum_{j\in J_i}\max_{\beta\in x}v_{ij}(\beta)$$

If $\succeq$ is constant, then minimality implies $I=0$ and there is nothing to prove. So assume $\succeq$ is not constant. Then by Conditional Non-triviality, there exists α' and α''' such that $\{\{\alpha'\}\}\succ\{\{\alpha'''\}\}$. By Continuity, we can assume there exists α'' such that $\{\{\alpha'\}\}\succ\{\{\alpha''\}\}\succ\{\{\alpha'''\}\}$ and that all are in the interior of Δ.

By Lemma 1, $\succeq$ satisfies Constant Normative Preference. So by Lemma 6, there exists u such that for every i, we have $u_i=q_iu+b_i$ for some $q_i\geq 0$ and $b_i\in\mathbb{R}$. By the previous uniqueness results, we can assume without loss of generality that $u(\alpha')=0$, $\sum_I q_i=1$, $b_i=0$ for every i, and $v_{ij}\cdot\mathbf{1}=0$ for every i and for every $j\in J_i$. Thus for any β, we have $\mathcal{U}(\{\{\beta\}\})=u(\beta)$, which implies u is not constant.

We now show that for every i, $q_i > 0$. Set $I^+ \equiv \{i \in I : q_i > 0\}$ and $I^0 \equiv \{i \in I : q_i = 0\}$. Fix $\epsilon \in (0, -u(\alpha''))$. Set $J \equiv \max_{i\in I} J_i$, $\bar{v} \equiv \max_{i\in I, j\in\bigcup_{i'\in I} J_{i'}} \sqrt{v_{ij} \cdot v_{ij}}$, and $a \equiv \frac{\epsilon}{\bar{v}J}$. Let x^* denote the sphere centered around α' with radius a. (If x^* is not in the interior of Δ, then choose a smaller a.) Thus any $\beta \in x^*$ can be written as $\beta = \alpha' + as$ where s is a vector such that $s \cdot \mathbf{1} = 0$ and $s \cdot s = 1$.

Since $\mathcal{U}$ is minimal, apply Lemma 3 to get $x_1, \ldots, x_I$. As is evident from the construction of $x_1, \ldots, x_I$ in Lemma 3, we can assume $\alpha' \in x_i \subset x^*$ for every i. Hence for every i and for every $j \in J_i$, we have

$$\begin{aligned}
\max_{\beta\in x_i} v_{ij}(\beta) &\leq \max_{\beta\in x^*} v_{ij}(\beta) \\
&= v_{ij} \cdot \left(\alpha' + a\frac{v_{ij}}{\sqrt{v_{ij} \cdot v_{ij}}}\right) \\
&= v_{ij} \cdot \alpha' + a\frac{v_{ij} \cdot v_{ij}}{\sqrt{v_{ij} \cdot v_{ij}}} \\
&= v_{ij}(\alpha') + a\sqrt{v_{ij} \cdot v_{ij}}.
\end{aligned}$$

This implies for every i

$$\begin{aligned}
\sum_{j\in J_i} \max_{\beta\in x_i} v_{ij}(\beta) &\leq \sum_{j\in J_i} \left(v_{ij}(\alpha') + a\sqrt{v_{ij} \cdot v_{ij}}\right) \\
&= \sum_{j\in J_i} v_{ij}(\alpha') + a\sum_{j\in J_i} \sqrt{v_{ij} \cdot v_{ij}} \\
&\leq \sum_{j\in J_i} v_{ij}(\alpha') + a\sum_{j\in J_i} \bar{v} \\
&\leq \sum_{j\in J_i} v_{ij}(\alpha') + a\bar{v}J \\
&= \sum_{j\in J_i} v_{ij}(\alpha') + \epsilon.
\end{aligned}$$

Since $\alpha' \in x_i$, we have for every i,

$$\begin{aligned}
V_i(x_i) &= \max_{\beta\in x_i} \left[q_i u(\beta) + \sum_{j\in J_i} v_{ij}(\beta)\right] - \sum_{j\in J_i} \max_{\beta\in x_i} v_{ij} \cdot \beta \\
&\geq \left[q_i u(\alpha') + \sum_{j\in J_i} v_{ij}(\alpha')\right] - \sum_{j\in J_i} \max_{\beta\in x_i} v_{ij} \cdot \beta \\
&\geq q_i u(\alpha') + \sum_{j\in J_i} v_{ij}(\alpha') - \sum_{j\in J_i} v_{ij}(\alpha') - \epsilon \\
&= -\epsilon \\
&> u(\alpha'').
\end{aligned}$$

Note that for every $i \in I^0$, we must have $J_i \geq 2$ (otherwise $\mathcal{U}$ would not be minimal). Recall by Lemma 3, $\arg\max_{\beta\in x_i} v_{ij}(\beta) \neq \arg\max_{\beta\in x_i} v_{ij'}(\beta)$ for every $j, j' \in J_i$ where

$j \neq j'$. Hence for every $i \in I^0$, we have

$$V_i(x_i) = \max_{\beta \in x_i} \left[\sum_{j \in J_i} v_{ij} \cdot \beta \right] - \sum_{j \in J_i} \max_{\beta \in x_i} v_{ij} \cdot \beta < 0$$

Note that for every i, we have $V_i(\{\beta\}) = q_i u(\beta)$. Hence for every $i \in I^+$ we have

$$V_i(x_i) > u(\alpha'') \geq q_i u(\alpha'') = V_i(\{\alpha''\})$$

and

$$V_i(x_i) > u(\alpha''') \geq q_i u(\alpha''') = V_i(\{\alpha'''\}),$$

while for every $i \in I^0$ we have

$$V_i(x_i) < 0 = V_i(\{\alpha''\}) = V_i(\{\alpha'''\}).$$

Let $X = \{x_1, \ldots, x_I\}$. Hence if I^0 is not empty, we must have $\{\{\alpha'''\}\} \cup X \succ X$ (since $V_i(\{\alpha'''\}) > \max_{x \in X} V_i(x)$ for every $i \in I^0$) and $\{\{\alpha'', \alpha'''\}\} \cup X \sim \{\{\alpha'''\}\} \cup X$ (since $V_i(\{\alpha''\}) \leq \max_{x \in X \cup \{\{\alpha'''\}\}} V_i(x)$ for every i). Yet $\{\{\alpha''\}\} \succ \{\{\alpha'''\}\}$, which violates Monotonicity of Commitments. Hence I^0 is empty, so $q_i > 0$ for every i.

For every i and for every $j \in J_i$, set $\hat{v}_{ij} \equiv v_{ij}/q_i$. Thus we can write $\mathcal{U}$ as

$$\mathcal{U}(X) = \sum_{i \in I} q_i \max_{x \in X} \left\{ \max_{\beta \in x} \left[u(\beta) + \sum_{j \in J_i} \hat{v}_{ij}(\beta) \right] - \sum_{j \in J_i} \max_{\beta \in x} \hat{v}_{ij}(\beta) \right\},$$

which is a CNT^{DLR} representation where u is not constant. The uniqueness result follows from the previous results.

References

Ahn, David S. and Todd Sarver (2013), "Preference for flexibility and random choice." *Econometrica*, 81, 341–361.

Amador, Manuel, Iván Werning, and George-Marios Angeletos (2006), "Commitment vs. flexibility." *Econometrica*, 74, 365–396.

Dekel, Eddie and Barton L. Lipman (2012), "Costly self-control and random self indulgence." *Econometrica*, 80, 1271–1302.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2001), "Representing preferences with a unique subjective state space." *Econometrica*, 69, 891–934.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2009), "Temptation-driven preferences." *Review of Economic Studies*, 76, 937–971.

Dekel, Eddie, Barton L. Lipman, Aldo Rustichini, and Todd Sarver (2007), "Representing preferences with a unique subjective state space: Corrigendum." *Econometrica*, 75, 591–600.

Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and self-control." *Econometrica*, 69, 1403–1435.

Herstein, Israel N and John Milnor (1953), "An axiomatic approach to measurable utility." *Econometrica*, 291–297.

Kopylov, Igor (2009a), "Finite additive utility representations for preferences over menus." *Journal of Economic Theory*, 144, 354–374.

Kopylov, Igor (2009b), "Temptations in general settings." *The B.E. Journal of Theoretical Economics*, 9.

Kopylov, Igor and Jawwad Noor (2010), "Self-deception and choice." Working paper, Boston University.

Kreps, David M. (1979), "A representation theorem for 'preference for flexibility'." *Econometrica*, 47, 565–578.

Noor, Jawwad and Norio Takeoka (2010), "Uphill self-control." *Theoretical Economics*, 5, 127–158.

Noor, Jawwad and Norio Takeoka (2011), "Menu-dependent self-control." Working paper, Boston University.

Stovall, John E. (2010), "Multiple temptations." *Econometrica*, 78, 349–376.