Temptation, Guilt, and Finite Subjective State Spaces*

Igor Kopylov

Institute for Mathematical Behavioral Sciences, and

Department of Economics, University of California, Irvine

Abstract

Building on the model of Dekel, Lipman and Rustichini (2001), we characterize an additive utility representation with a unique *finite* subjective state space. We describe the numbers of positive and negative subjective states explicitly in terms of preference. As a special case, Gul–Pesendorfer's (2001) model of temptation and self-control is derived from weaker assumptions on behavior (we relax Set Betweenness). The main novel application is a model of temptation and *guilt* that identifies separate psychological costs of these emotions. In this model the preference for commitment is strengthened by the decision maker's desire to reduce the cost of guilt.

1 Introduction

Kreps [14] derives a subjective state space from preference over *menus*. More precisely, he characterizes a utility representation

$$U(A) = \sum_{s \in S} \max_{x \in A} u_s(x), \tag{1}$$

where (i) each menu A is a set of consumption alternatives that become feasible ex post, and (ii) the finite state space S and the functions $u_s(\cdot)$ are derived from the ex ante ranking $\succeq$ of menus A. According to this model, the decision maker behaves *as if* she is uncertain about her future taste $\succeq_s$ represented by the state-dependent utility $u_s(\cdot)$. Unfortunately, the rankings $\succeq_s$ are not determined uniquely by preference over menus. The indeterminacy of the subjective state space is a problem not only for the interpretation, but also for potential applications of the model. In particular, this indeterminacy precludes a formal analysis of common knowledge,

*This version is preliminary and incomplete.

common priors, and subjective attitudes towards uncertainty associated with the space S.

Kreps's model has another potential limitation: it requires that *flexibility* is always valuable in the sense that each menu A is weakly preferred to any smaller menu $B \subset A$. Gul and Pesendorfer [10] (henceforth GP) argue that this condition is problematic when people are affected by temptations. For example, when planning a lunch in the morning, the decision maker may prefer a vegetarian dish (x) to a hamburger (y), but also anticipate that the latter will be tempting if the menu $\{x, y\}$ is available at the lunchtime. This temptation can be sufficiently strong for her to succumb and choose a hamburger ex post, which is an inferior decision from the ex ante perspective. Even if the decision maker has enough self-control to choose x from the menu $\{x, y\}$, the very presence of the temptation y can be psychologically unpleasant. Therefore, she should prefer *commitment* $\{x\} \succ \{x, y\}$.

Motivated by these concerns about the interpretation and the behavioral content of Kreps's model, Dekel, Lipman and Rustichini [2] (henceforth DLR) characterize the following additive utility representation[1]

$$U(A) = \int_S \max_{x \in A} u_s(x) d\mu(s). \tag{2}$$

Here, (i) the elements x in menus A are lotteries—objective distributions over final consumption, (ii) the functions $u_s(\cdot)$ have the expected utility form, and (iii) μ is a signed measure that has full support in S. Moreover, the class of rankings $\succeq_s$ that are represented by the functions $u_s(\cdot)$ is unique. In this sense, the subjective state space is determined uniquely by preference over menus. Clearly, the representation (2) accommodates preference for commitment because μ can be negative over some states in S.

DLR's model delivers no structure for the set S and the functions $u_s(\cdot)$, except the expected utility form for the latter. In contrast, applications may be very specific about the structure of S. For example, GP's model of temptation asserts that the state space S contains only two elements—one positive and one negative. The corresponding utility representation has the form

$$U(A) = \max_{x \in A}(u(x) + v(x)) - \max_{y \in A} v(y) = \max_{x \in A}[u(x) - \max_{y \in A}(v(y) - v(x))]. \tag{3}$$

Here, $u(\cdot)$ represents the commitment ranking of alternatives $x \in X$, and $v(\cdot)$ reflects the expected temptation costs. More precisely, the decision maker anticipates ex ante that if she selects x from a menu A ex post, then her temptation cost will be $\max_{y \in A}(v(y) - v(x))$. In order to balance her commitment value and temptation costs, she plans to choose x that maximizes $\max_{x \in A} u(x) + v(x)$. In

[1] DLR consider a non-additive representation as well. Dekel, Lipman, Rustichini, and Sarver [4] provide a corrigendum for DLR's results.

other applications, which we discuss later, the state space S has more than two elements, but is still finite.

The main objective of this paper is to characterize the representation (2) with a finite number of components, and to describe this number explicitly in terms of preference. Formally, we derive a utility function

$$U(A) = \sum_{i=1}^{P(\succeq)} \max_{x \in A} w_i(x) - \sum_{j=1}^{N(\succeq)} \max_{x \in A} v_j(x) \tag{4}$$

where the functions w_i and v_j are continuous, linear, and essentially unique. The numbers $P(\succeq)$ and $N(\succeq)$ are determined roughly by the maximal flexibility, and respectively by the maximal commitment, that the decision maker may desire. For example, the utility representation (3) corresponds to the case $P(\succeq) = N(\succeq) = 1$. This equality follows immediately from GP's axioms. We show that it can be obtained even if one of these axioms (Set Betweenness) is replaced by a weaker condition.

As an additional motivation for our results, we extend GP's model and characterize the following utility representation with $S = 3$:

$$\begin{aligned} U(A) &= \max_{x \in A}[(1+G)u(x) + v(x)] - \max_{y \in A} v(y) - G \max_{z \in A} u(z), \\ &= \max_{x \in A}[u(x) - \max_{y \in A}(v(y) - v(x)) - G \max_{z \in A}(u(z) - u(x))], \end{aligned} \tag{5}$$

Here $\max_{y \in A}(v(y) - v(x))$ and $G \max_{z \in A}(u(z) - u(x))$ are interpreted as the costs of temptation and guilt respective that are required for selecting x from the menu A. This representation accommodates a strong preference for commitment that may result if temptation is followed by guilt.

Newcomb (1885) writes "The fact that the benevolent gentleman may wish there were no beggars, and may be very sorry to see them, does not change the economic effect of his readiness to give them money ". In the words of Becker (1997): "People do not want to encounter beggars, even though they may contribute handsomely after an encounter." In this example, the decision maker has a normative preference to contribute to beggars (x), but still avoids meeting them, thus committing to give nothing (y). Therefore, her preference

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

violates the axioms in GP's model.

The representation (5) is a special case of "temptation-driven" preferences studied by DLR [3]. Yet our approach has different primitives, axioms, and interpretation. The structure that we impose on the components of (5) allows to distinguish guilt from temptation.

Moreover, we relax the primitives that are used by both DLR and GP, and take menus in an abstract convex metric space rather than in a simplex of probability distributions. In particular, menus may consist of Anscombe-Aumann acts. Such preferences have been recently used to study non-Bayesian updating (Epstein [6]) and time-varying pessimism (Epstein and Kopylov [7]). Despite the more general framework, our construction of the subjective state space is arguably more explicit and self-contained than in DLR. Each function w_i and v_j in the representation (4) is derived via an explicit formula that links it to the underlying choice behavior.

2 Main Representation Result

Consider an abstract version of the decision framework used by DLR[2, 3] and GP[10]. Let X be a convex subset of a linear space.[2] Endow X with a metric d such that (i) X is compact, (ii) the mixture operation $\alpha x + (1-\alpha)y$ that maps $[0,1] \times X \times X$ into X is *continuous*, (iii) the metric d is *quasi-convex*, that is, for all $\alpha \in [0,1]$ and $x, y, x', y' \in X$,

$$d(\alpha x + (1-\alpha)x', \alpha y + (1-\alpha)y') \leq d(x,y) \vee d(x',y'). \tag{6}$$

For example, these conditions hold if

(1) X is a compact convex subset of a normed linear space;

(2) $X = \Delta(Z)$ is the class of all probability distributions on a compact metric set Z, and d is Prohorov's [19] metric of weak convergence; this setting is used by GP and—for finite Z only—by DLR;

(3) $X = \mathcal{H} = (\Delta(Z))^{\Omega}$ is the class of all Anscombe–Aumann acts f that map a finite state space Ω into $\Delta(Z)$, and d is a standard product of metrics in $\Delta(Z)$—this framework is used by Epstein [6] to model non-Bayesian updating;

(4) $X = \mathcal{H}_f = \{\alpha f + (1-\alpha)c \,:\, \alpha \in [0,1], c \in \Delta(Z)\} \subset \mathcal{H}$ is the class of all mixtures of a given act $f \in \mathcal{H}$ and constant acts $c \in \Delta(Z)$—this domain is essential for the model of time-varying pessimism in Epstein–Kopylov [7].

Formally, X may consist of deterministic commodity bundles rather than lotteries or Anscombe–Aumann acts. However, the non-probabilistic specification may be problematic for the separability condition (i.e. Independence Axiom) that is imposed on choice behavior later.

Let $\mathcal{A}$ be the set of all *menus*—non-empty compact subsets $A \subset X$. Following Kreps [14], interpret each menu $A \in \mathcal{A}$ as a physical action that, if taken ex ante, makes the set of alternatives $A \subset X$ feasible ex post. The decision maker's weak preference over menus is given as a binary relation $\succeq$ on $\mathcal{A}$. The indifference $\sim$ and strict preference $\succ$ are respectively the symmetric and asymmetric parts of

[2]Even more generally, one can define the convexity structure in X axiomatically and then construct an embedding of X into some linear space. See Stone [21] or Mongin [16].

the relation $\succeq$. Note that the decision maker's ex post choice in a menu A is not taken as a primitive of the model, but her anticipation of this choice is essential for interpreting her ex ante behavior.

Following DLR and GP, adopt the following axioms.

Axiom 1 (Order). *$\succeq$ is complete and transitive.*

Axiom 2 (Continuity). *For all menus $A \in \mathcal{A}$, the sets $\{B \in \mathcal{A} : B \succeq A\}$ and $\{B \in \mathcal{A} : B \preceq A\}$ are closed in the Hausdorff metric topology.*[3]

Order and Continuity are standard postulates of rationality.

Menus can be mixed in a natural way.[4] For any $A, B \in \mathcal{A}$ and $\alpha \in [0,1]$, let

$$\alpha A + (1-\alpha)B = \{\alpha x + (1-\alpha)y \ : \ x \in A, \ y \in B\}.$$

Axiom 3 (Independence). *For all $\alpha \in [0,1]$ and menus $A, B, C \in \mathcal{A}$,*

$$A \succ B \quad \Rightarrow \quad \alpha A + (1-\alpha)C \succ \alpha B + (1-\alpha)C.$$

To motivate this axiom, adapt the arguments of DLR and GP. Assume that for any menus A and C, the decision maker is indifferent between the menu $\alpha A + (1-\alpha)C$ and a hypothetical lottery $\alpha \circ A + (1-\alpha) \circ C$ that is resolved *before* the ex post stage and yields the menus A or C with probabilities α and $1-\alpha$ respectively. (Here the preference $\succeq$ is extended to lotteries over menus.) Then the ranking $A \succ B$ should imply $\alpha \circ A + (1-\alpha) \circ C \succ \alpha \circ B + (1-\alpha) \circ C$ because the possibility of getting the menu C with probability $1-\alpha$ should not affect the comparison of A and B. Thus Independence follows.

The appeal of this argument depends on the specification of X. If X consists of deterministic consumption bundles, then the indifference between the mixture $\alpha A + (1-\alpha)C$ and the lottery $\alpha \circ A + (1-\alpha) \circ C$ seems arbitrary. (One encounters a similar problem when motivating Independence in Herstein–Milnor's abstract mixture space setting.) On the other hand, if X consists of lotteries or Anscombe-Aumann acts, then the indifference between $\alpha A + (1-\alpha)C$ and $\alpha \circ A + (1-\alpha) \circ C$ can be interpreted in terms of the timing of the resolution of the objective uncertainty. In these frameworks, one can interpret the menu $\alpha A + (1-\alpha)C$ as a lottery which yields the menus A and C with probabilities α and $1-\alpha$ respectively but is resolved *after* the ex post stage, that is, after the contingent choices are made in both menus A and C. Thus Independence seems reasonable when elements and mixtures in X have probabilistic nature.[5]

---

[3] The definition and some properties of the Hausdorff metric are given in the appendix.

[4] Yet the space of menus $\mathcal{A}$ is not a mixture space in the sense of Herstein–Milnor [12] because the property $\alpha A + (1-\alpha)A = A$ can be violated.

[5] Still there are situations where the timing of the resolution of objective uncertainty should matter for choice behavior. For example, this timing can be important if (i) contemplation is costly for the agent (see Ergin–Sarver [9]), (ii) the decision maker's belief depends on the menu at hand (Epstein–Kopylov [7]), or (iii) the decision maker has a coarse (i.e. incomplete) perception of future contingencies (see Epstein–Marinacci–Seo [8]).

To arrive at the preference $\succeq$, the decision maker may first rank elements in X, and then evaluate each menu A via the best alternative $x \in A$ in this menu. This two-step procedure motivates the principle of *strategic rationality* formulated by Kreps [14]: for all menus A and B:

$$A \succeq B \quad \Rightarrow \quad A \sim A \cup B.$$

This condition, together with Axioms 1–3, is sufficient to represent $\succeq$ by a utility function U that satisfies $U(A) = \max_{x \in A} U(\{x\})$.

However, the principle of strategic rationality can be violated by both a strict preference for flexibility, such as $A \cup B \succ A \succeq B$, and a preference for commitment, such as $A \succ A \cup B$. Moreover, these two patterns of behavior may co-exist.

Given any family of menus $A_1, \ldots, A_n$, say that A_k is *positive*, *negative*, or *neutral* in this family if $\cup_i A_i \succ \cup_{i \neq k} A_i$, $\cup_i A_i \prec \cup_{i \neq k} A_i$, or $\cup_i A_i \sim \cup_{i \neq k} A_i$ respectively.

Axiom 4 (Finiteness). *For any sufficiently large n and any family of menus $A_1, \ldots, A_n \in \mathcal{A}$, there is a neutral member A_k in this family.*

This axiom requires roughly that the decision maker has a bounded desire for flexibility and for commitment. In applications, Finiteness is usually derived from other assumptions rather than imposed directly. For example, it is implied by Set Betweenness in GP's model of temptation, which we discuss later.

Finiteness is necessary for a broad class of models with a finite subjective state space. Indeed, suppose that $\succeq$ can be represented by a utility function

$$U(A) = V(\sup_{x \in A} u_1(x), \ldots, \sup_{x \in A} u_S(x)) \tag{7}$$

with arbitrary $S \geq 1$ and arbitrary functions $u_1, \ldots, u_S : X \to \mathbb{R}$ and $V : \mathbb{R}^S \to \mathbb{R}$; these functions need not be additive or even continuous. Take any $n > S$ and $A_1, \ldots, A_n \in \mathcal{A}$. For any $j \leq S$, take $i(j) \in \{1, \ldots, n\}$ such that

$$\sup_{x \in A_{i(j)}} u_j(x) = \sup_{x \in \cup_i A_i} u_j(x).$$

As $n > S$, there is $k \leq n$ such that $k \neq i(j)$ for all $j \leq S$. Then

$$U(\cup_{i \neq k} A_i) = U(\cup_i A_i),$$

that is, A_k is neither positive nor negative in the family $A_1, \ldots, A_n$. Thus, Finiteness follows from the representation (7).

Let $P(A_1, \ldots, A_n)$ be the number of positive, and $N(A_1, \ldots, A_n)$ the number of negative menus in a family $A_1, \ldots, A_n$. Finiteness implies that there exist finite maxima

$$P(\succeq) = \max P(A_1, \ldots, A_n)$$
$$N(\succeq) = \max N(A_1, \ldots, A_n)$$
$$S(\succeq) = \max[P(A_1, \ldots, A_n) + N(A_1, \ldots, A_n)]$$

across all $n \geq 1$ and menus $A_1, \ldots, A_n \in \mathcal{A}$. To show this claim, take any menus $A_1, \ldots, A_n \in \mathcal{A}$ and let $S = P(A_1, \ldots, A_n) + N(A_1, \ldots, A_n)$. Wlog each of the $m \leq n$ menus $A_1, \ldots, A_S$ is either positive or negative in the family $A_1, \ldots, A_n$. For any $k \in \{1, \ldots, S\}$, let

$$A_k^* = A_k \cup A_{m+1} \cup \cdots \cup A_n. \tag{8}$$

Then there are no neutral menus in the family $A_1^*, \ldots, A_S^*$ because $\cup_{i=1}^n A_i = \cup_{j=1}^S A_j^*$ and $\cup_{i \neq k} A_i = \cup_{j \neq k} A_j^*$. By Finiteness, S cannot be arbitrarily large and hence, achieves a finite maximum $S(\succeq)$. A fortiori, $P(\succeq)$ and $N(\succeq)$ are also finite.

Say that a function $u : X \to \mathbb{R}$ is *linear* if $u(\alpha x+(1-\alpha)y) = \alpha u(x)+(1-\alpha)u(y)$ for all $\alpha \in [0,1]$ and $x, y \in X$. Say that a list of linear functions $u_1, \ldots, u_S$ is *redundant* if it contains a constant function, or if $u_i(\cdot) = \alpha u_j(\cdot) + \beta$ for some $j \neq i$, $\alpha > 0$ and $\beta \in \mathbb{R}$. In the latter case, the pair of functions u_i and u_j represents the same ranking of X.

The following is our main representation result.

Theorem 2.1. *The preference $\succeq$ satisfies Axioms 1–4 iff $\succeq$ is represented by*[6]

$$U(A) = \sum_{i=1}^{S} \gamma_i \max_{x \in A} u_i(x), \tag{9}$$

where $S \geq 0$, $\gamma_1, \ldots, \gamma_S \in \{-1, 1\}$, and $u_1, \ldots, u_S : X \to \mathbb{R}$ is a non-redundant list of continuous linear functions.

The representation (9) is unique up to a permutation of the indices i's, and a positive linear transformation of the functions u_i's.[7]

Moreover, $S(\succeq) = S$, $P(\succeq) = \sum_{i:\gamma_i=1} \gamma_i$, and $N(\succeq) = \sum_{i:\gamma_i=-1} -\gamma_i$.

The representation (9) is an extension of the *finite additive expected utility representation*, which DLR [3] define and characterize (Theorem 6) for menus in a finite dimensional simplex of probability distributions.

Theorem 2.1 extends DLR's results in several directions. First it relaxes the primitives by taking menus in an abstract convex metric space. In particular, it accommodates domains used by GP[10, 11], Epstein [6], and Epstein–Kopylov [7]. Of course, all applications of Theorem 2.1 that we formulate later are also valid for each of these domains.

Second, there is a difference in axioms. Our version of Finiteness complies with the general non-linear representation (7), while DLR's counterpart holds only for the additive model (4) and can be violated by (7). In this respect, our version is

[6] By convention, a sum over an empty index set is zero.

[7] More precisely, if $\succeq$ is represented by (9) with another list of components $\gamma_1', \ldots, \gamma_S'$ and $u_1', \ldots, u_S'$, then $\gamma_{\pi(i)}' = \gamma_i$ and $u_{\pi(i)}'(\cdot) = \alpha u_i(\cdot) + \beta_i$ for some $\alpha > 0$, $\beta_i \in \mathbb{R}$, and a permutation π of the set $\{1, \ldots, S\}$.

weaker than the one used by DLR and is potentially more useful for non-linear extensions of the model.[8]

Third, our construction of the representation (9) is arguably more direct than in DLR. Roughly it proceeds in three steps.

Step 1. Order, Continuity, and Independence imply by Herstein–Milnor's [12] Theorem that there is a linear and continuous utility representation $U_c(\cdot)$ on the mixture space of convex menus. These axioms also imply indifference between any menu A and the closed convex hull of this menu (see Lemma A.3 in the appendix). Thus the utility function $U_c(\cdot)$ can be extended to a linear representation $U(\cdot)$ on the domain of all menus.

Step 2. Let $S = S(\succeq)$. Use (8) to construct a family of menus $A_1^*, \ldots, A_S^*$ such that each A_i^* is either positive or negative in this family. Let $A^* = \cup_i A_i^*$ and for any j, $A_{-j}^* = \cup_{i \neq j} A_i^*$. Wlog let $U(A^*) = 0$.

Step 3. Let $\gamma_i = 1$ or $\gamma_i = -1$ if A_i^* is positive or, respectively, negative in the family $A_1^*, \ldots, A_S^*$. Take a small $\alpha > 0$. For any $y \in X$, let

$$u_i(y) = \frac{U(A_{-i}^* \cup (\alpha\{y\} + (1-\alpha)A_i^*))}{\alpha\gamma_i}. \tag{10}$$

This formula is based on the following intuition. Suppose that there is a representation (9) such that for any $i \in \{1, \ldots, S\}$, all elements that maximize $u_i(\cdot)$ on A^* belong to A_i^*. Then a slight linear perturbation of A_i^* in the union $A^* = \cup_i A_i^*$ by the element y should affect the overall utility through changing the only component $\max_{x \in A^*} u_i(x)$. Thus one can compute $u_i(y)$ by observing the corresponding change in utility and adjusting it by α and γ_i.

This construction is self-contained and do not invoke the theory of topological linear spaces but instead, provide an explicit formula for each function $u_i(\cdot)$.

Yet the main novelty of Theorem 2.1 is the assertion that the representation (9) has $S(\succeq)$ components, among which $P(\succeq)$ are *positive* and $N(\succeq)$ are *negative.* This distinction is based on the signs of the corresponding factors γ_i with $\sum_{i:\gamma_i=1} \gamma_i$ being the number of positive signs, and $\sum_{i:\gamma_i=-1} -\gamma_i$ the number of negative ones. It follows that $S(\succeq) = P(\succeq) + N(\succeq)$, and the utility $U(\cdot)$ can be rewritten as

$$U(A) = \sum_{i=1}^{P(\succeq)} \max_{x \in A} w_i(x) - \sum_{j=1}^{N(\succeq)} \max_{x \in A} v_j(x) \tag{11}$$

with a non-redundant list $w_1, \ldots, w_{P(\succeq)}, v_1, \ldots, v_{N(\succeq)}$ of continuous and linear functions.

[8] It is an open question though what these extensions might be.

By restricting $S(\succeq)$, $P(\succeq)$, and $N(\succeq)$, one can control the composition of the equivalent representations (9) and (11). For example, one can assume that in addition to Axioms 1–4, $\succeq$ satisfies Kreps' *monotonicity*: for all menus $A, B \in \mathcal{A}$

$$A \subset B \quad \Rightarrow \quad B \succeq A.$$

Then negative menus never exist, $S = S(\succeq) = P(\succeq)$, and the representation (9) takes the form

$$U(A) = \sum_{i=1}^{S} \max_{x \in A} u_i(x).$$

The decision maker as portrayed by this representation is uncertain ex ante about which of the utility functions $u_1, \ldots, u_S$ she will have ex post, and evaluates any menu A by putting an equal weight (or probability) on the maximum of each of these utilities in A. This interpretation has the standard degree of freedom because the positive weights of the components $\max_{x\in A} u_i(x)$ can vary freely as long as the linear functions $u_i(\cdot)$ are rescaled accordingly.

Next, Theorem 2.1 delivers the main representation result in GP's model of temptation. The central case of their model makes the following assumptions about the preference $\succeq$ in addition to Axioms 1–3.

Axiom 5 (Set Betweenness). *For all menus $A, B \in \mathcal{A}$,*

$$A \succeq B \quad \Rightarrow \quad A \succeq A \cup B \succeq B.$$

Axiom 6 (Self-Control). *$A \succ A \cup B \succ B$ for some menus $A, B \in \mathcal{A}$.*

Set Betweenness implies that $P(\succeq) \leq 1$ and $N(\succeq) \leq 1$. (The easy proof by contradiction is left to the reader.) Conversely, Self-Control requires that $P(\succeq) \geq 1$ and $N(\succeq) \geq 1$. Thus, by Theorem 2.1, the preference $\succeq$ satisfies Axioms 1–3, Set Betweenness, and Self-Control iff it can be represented by

$$U(A) = \max_{x \in A} w(x) - \max_{x \in A} v(x) = \max_{x \in A}(u(x) + v(x)) - \max_{x \in A} v(x), \qquad (12)$$

where the functions u and v are linear, continuous, and not redundant. These functions represent distinct *commitment* and *temptation* rankings on X.

In fact, Theorem 2.1 implies a tighter characterization for the representation (12), where Set Betweenness is replaced by the following weaker condition.

Axiom 7 (Two-States). *For any menus $A_1, A_2, A_3 \in \mathcal{A}$, either $A_1 \cup A_2 \cup A_3 \sim A_1 \cup A_2$, or $A_1 \cup A_2 \cup A_3 \sim A_1 \cup A_3$, or $A_1 \cup A_2 \cup A_3 \sim A_2 \cup A_3$.*

To motivate this axiom, assume that the decision maker's evaluation of any menu B is determined by her anticipated ex post choice $x \in B$ and greatest temptation $y \in B$ in this menu. Let $B = A_1 \cup A_2 \cup A_3$. If needed, relabel the menus A_1, A_2, A_3 so that $x, y \in A_1 \cup A_2$. Then the indifference $B \sim A_1 \cup A_2$ should hold as long as the decision maker's expected choice $x \in B$ and temptation $y \in B$ are unchanged when some of the other alternatives in B become unavailable.[9]

This motivation, unlike GP's rationale for Set Betweenness, does not require that temptations should be harmful. Accordingly, Two-States is strictly weaker than Set Betweenness, providing that $\succeq$ is complete and transitive. Indeed, take any menus $A_1, A_2, A_3 \in \mathcal{A}$. If needed, relabel them so that $A_1 \cup A_3 \succeq A_1 \cup A_2 \succeq A_2 \cup A_3$. Then by Set Betweenness,

$$A_1 \cup A_3 \succeq A_1 \cup A_2 \cup A_3 \succeq A_1 \cup A_2 \succeq A_1 \cup A_2 \cup A_3 \succeq A_2 \cup A_3,$$

that is, $A_1 \cup A_2 \cup A_3 \sim A_1 \cup A_2$. On the other hand, the representation (11) with two positive or alternatively with two negative components implies Two-States, but violates Set Betweenness.

Suppose that $\succeq$ satisfies Axioms 1–3, Two-States, and Self-Control. By Self-Control, $P(\succeq) \geq 1$ and $N(\succeq) \geq 1$. By Two-States, $S(\succeq) \leq 2$. (We omit the easy proof.) By Theorem 2.1, $P(\succeq) + N(\succeq) = S(\succeq)$, and $\succeq$ has the representation (11). Thus, $P(\succeq) = 1$, $N(\succeq) = 1$, and we have the following

Corollary 2.2. *$\succeq$ satisfies Axioms 1–3, Two-States, and Self-Control iff it can be represented by* (12).

This corollary obtains GP's main utility representation from weaker assumptions on primitives and behavior. Moreover, it can be easily extended to accommodate behavioral phenomena other than temptation. We describe some extensions of this sort in the following sections.

3 Modelling Temptation and Guilt

Given any alternatives $x, y \in X$ such that $\{x\} \succ \{y\}$, GP's model distinguishes three possible cases for temptation and self-control in the menu $\{x, y\}$.

(i) y does not tempt x, that is, $v(x) \geq v(y)$. Then the decision maker ranks

$$\{x\} \sim \{x, y\} \succ \{y\}. \tag{13}$$

(ii) y tempts x, but the positive cost of self-control $v(y) - v(x) > 0$ is smaller than the commitment utility difference $u(x) - u(y)$. Then the decision maker plans to resist the temptation y and ranks

$$\{x\} \succ \{x, y\} \succ \{y\}. \tag{14}$$

[9]This rationale essentially requires the Nash–Chernoff condition (or Sen's property α) for ex post choice and for temptation; GP use the same condition, though implicitly, to motivate Set Betweenness. Noor [17] argues that the Nash–Chernoff condition can be problematic in the presence of temptation, and relaxes GP's model accordingly.

(iii) y tempts x, and the cost of self-control $v(y) - v(x)$ is greater or equal than the commitment utility difference $u(x) - u(y)$. Then the decision maker plans to succumb to the temptation y and ranks

$$\{x\} \succ \{x,y\} \sim \{y\}. \tag{15}$$

Suppose that the decision maker expects that she will feel *guilty* after succumbing to the temptation y in the menu $\{x,y\}$, and that the guilt will be costly for her. Then the above rankings (13) and (14) are not affected, but (15) should be replaced by the strong preference for commitment[10]

$$\{x\} \succ \{y\} \succ \{x,y\}. \tag{16}$$

The psychological cost of guilt may affect the value of the menu $\{x,y\}$ in two ways. First, this cost is paid directly if the decision maker still plans to succumb to the temptation y in this menu. Alternatively, the potential punishment by guilt can make her change the ex post behavior and select x in the menu $\{x,y\}$ even though the anticipated cost of self-control $v(y) - v(x)$ is greater than the commitment utility difference $u(x) - u(y)$. Then the decision maker still has the strict preference $\{y\} \succ \{x,y\}$ because it is better for her to commit to the normatively inferior y ex ante than to exercise the intense self-control and select the normatively best x in the menu $\{x,y\}$ ex post. Note that the ranking (16) alone does not reveal the anticipated ex post choice.

One can challenge the above analysis by arguing that any person who is deterred by guilt ex post should be also deterred by guilt ex ante. Indeed, it may appear that the decision maker should feel just as guilty after committing to the future temptation $\{y\}$ ex ante than after choosing y in the menu $\{x,y\}$ ex post. In the words of Elster [5, p.65], "to want to be immoral is to be immoral". Yet in many settings there is no intuitive equivalence between the ex ante and ex post costs of guilt. For example, people tend to feel more guilty after meeting a beggar and not giving him any money than after avoiding this beggar in the first place. Guilt may also increase over time for recruits in firms and organizations, such as the military, which spend resources to instill guilt on its members and promote loyalty and team spirit. These situations are studied by Kandeal and Lazear [13]. Thus the preference (16) can be intuitive as long as the decision maker believes that she would feel *less* guilty after committing to the future temptation $\{y\}$ ex ante rather than after choosing y in the menu $\{x,y\}$ ex post.

Assume that both temptation and guilt have non-trivial effects on behavior, and rankings such as (13), (14) and (16) can be observed.

Axiom 8 (Self-Control-With-Guilt). *There are menus $A, B, A', B' \in \mathcal{A}$ such that $A \cup B \succ B$ and $A' \succeq B' \succ A' \cup B'$.*

[10] The ranking $\{x\} \succ \{x,y\} \sim \{y\}$ remains intuitive only in the borderline situation when $v(y) - v(x) = u(x) - u(y)$.

Unlike Set Betweenness, the Two-States axiom is not violated directly by the strong preference for commitment (16). Yet it can be also problematic in the presence of guilt. For example, consider a menu $\{x, x', y\}$, where x is a vegetarian dish, x' is a chicken salad, and y is a hamburger. Suppose that the decision maker confronted with the menu

$$A = \{x, x', y\} = \{x, x'\} \cup \{x', y\} \cup \{x, y\},$$

anticipates that she will choose x' and be tempted by y ex post, and then feel guilty about not choosing the normatively best alternative x. Providing that both temptation and guilt are costly, the menus $\{x, x'\}$ and $\{x', y\}$ should be strictly preferred to A. On the other hand, it is also intuitive that the decision maker should strictly prefer A to $\{x, y\}$ because the former menu provides the better ex post choice x' in the presence of the same temptation y and the same normatively best opportunity x.

In order to accommodate both temptation and guilt, we relax Two-States as follows. For any $z \in X$, let

$$\mathcal{A}_z = \{A \in \mathcal{A} : z \in A \text{ and } z \succeq y \text{ for all } y \in A.\}$$

Axiom 9 (Two-Plus-One-States). *For any $z \in X$ and menus $A_1, A_2, A_3 \in \mathcal{A}_z$, either $A_1 \cup A_2 \cup A_3 \sim A_1 \cup A_2$, or $A_1 \cup A_2 \cup A_3 \sim A_1 \cup A_3$, or $A_1 \cup A_2 \cup A_3 \sim A_2 \cup A_3$.*

Clearly, this axiom is weaker than Two-States because it applies only to menus in a subclass $\mathcal{A}_z \subset \mathcal{A}$. To motivate Two-Plus-One-States, assume that the decision maker's evaluation of any menu B is determined by her anticipated ex post choice $x \in B$, greatest temptation $y \in B$, and normatively best alternative $z \in B$. Following GP, identify the decision maker's normative preference on X with her commitment ranking.[11] Suppose that $B = A_1 \cup A_2 \cup A_3$ where $A_1, A_2, A_3 \in \mathcal{A}_z$. If needed, relabel the menus A_1, A_2, A_3 so that $x, y \in A_1 \cup A_2$. Note that $z \in A_1 \cup A_2$ because $A_1, A_2 \in \mathcal{A}_z$. Therefore, the indifference $B \sim A_1 \cup A_2$ should hold as long as the ex post choice $x \in B$, temptation $y \in B$, and the normatively best alternative $z \in B$ are unchanged when some of the other alternatives in B become unavailable. This motivation does not put any restrictions on how the decision maker combines x, y and z to evaluate the menu B. It requires only that the availability of *other* alternatives in B is irrelevant for her evaluation.

Say that functions $u, v : X \to \mathbb{R}$ are *independent* if the elements $u, v, \vec{1}$ of the space $\mathbb{R}^X$ are linear independent. (Here $\vec{1}$ is the constant unity function on X.) Clearly if u and v are independent, then they are not redundant, but the converse is not necessarily true.[12]

---

[11] This identity is reasonable if at the ex ante stage, the decision maker has a "cool" mind, and her commitment ranking is not affected by future temptations or other visceral factors. Noor [18] relaxes this assumption and formulates a model where future consumption can be tempting.

[12] The stronger condition of independence prohibits situations where temptation exists but is always resisted and hence, produces no guilt. This happens when $v = -\alpha u$ for $\alpha \in (0, 1)$.

Corollary 3.1. *$\succeq$ satisfies Axioms 1–3, Two-Plus-One-States, and Self-Control-With-Guilt iff it can be represented by*

$$U(A) = \max_{x\in A}[(1+G)u(x) + v(x)] - \max_{y\in A} v(y) - G\max_{z\in A} u(z), \tag{17}$$

where $G > 0$, and the functions u and v are linear, continuous, and independent. Moreover, G is unique, and u, v are unique up to a positive linear transformation.

To interpret the utility representation (17), rewrite it as follows:

$$U(A) = \max_{x\in A}[u(x) - \max_{y\in A}(v(y) - v(x)) - G\max_{z\in A}(u(z) - u(x))].$$

Then $\max_{y\in A}(v(y)-v(x))$ and $G\max_{z\in A}(u(z)-u(x))$ are respectively the costs of temptation and guilt associated with selecting x from the menu A. Accordingly, the anticipated ex post choice $x \in A$ achieves a compromise between the normative values on one side and the costs of temptation and guilt on the other by solving the maximization problem

$$\max_{x\in A}[(1+G)u(x) + v(x)].$$

Note that for large $G > 0$ the solution to this problem is close to maximizing the normative values. In other words, strong guilt is a good disciplining device in the presence of temptations.

Based on this intuition, we characterize the behavioral meaning of the parameter $G > 0$. Suppose that two preferences $\succeq$ and $\succeq^*$ have representations (17) that have the same commitment and temptation rankings $u(\cdot) = u^*(\cdot)$ and $v(\cdot) = v^*(\cdot)$, but may have different parameters $G \geq G^*$. For instance, this situation is characterized by the equivalence

$$\{x\} \succ \{x,y\} \succ \{y\} \quad \Leftrightarrow \quad \{x\} \succ^* \{x,y\} \succ^* \{y\}. \tag{18}$$

Say that $\succeq$ is more disciplined than $\succeq^*$ if for any menu A and elements $x, y \in A$,

$$A \succ A\setminus\{x\} \text{ and } A \succ^* A\setminus\{y\} \quad \Rightarrow \quad \{x\} \succeq \{y\}.$$

Corollary 3.2. *Suppose that $\succeq$ and $\succeq^*$ have representations (17) and satisfy the equivalence (18). Then $\succeq$ is more disciplined than $\succeq^*$ iff $G \geq G^*$.*

This result asserts roughly that the parameter G reflects the extent to which the decision maker expects to abide by the normative values ex post.

A Proof of the Main Representation Result

The proof of Theorem 2.1 requires some preliminary facts about the algebraic and topological structure of the spaces X and $\mathcal{A}$.

Lemma A.1. *Let $S \geq 2$. A list of linear functions $u_1, \ldots, u_S : X \to \mathbb{R}$ is not redundant iff there are $x_1^*, \ldots, x_S^* \in X$ such that $u_i(x_i^*) > u_i(x_j^*)$ for all $i \neq j$.*

Proof. Let i, j, k, l vary in $\{1, \ldots, S\}$. Suppose that the list $u_1, \ldots, u_S : X \to \mathbb{R}$ is not redundant. Fix any $k \neq l$. The functions u_k and u_l represent different rankings of X. Then the inequalities $u_k(x) \geq u_k(y)$ and $u_l(x) \geq u_l(y)$ are not equivalent for some $x, y \in X$. Wlog $u_k(x) \geq u_k(y)$ and $u_l(y) > u_l(x)$. Morevoer, u_k and u_l are non-constant and linear. Take $x', y' \in X$ and $\varepsilon > 0$ such that $u_k(x') > u_k(y')$ and $u_l(\varepsilon y' + (1-\varepsilon)y) > u_l(\varepsilon x' + (1-\varepsilon)x)$. Let $x_{kl} = \varepsilon x' + (1-\varepsilon)x$ and $x_{lk} = \varepsilon y' + (1-\varepsilon)y$. Then $u_k(x_{kl}) > u_k(x_{lk})$ and $u_l(x_{lk}) > u_l(x_{kl})$. Thus there is a matrix $\|x_{kl}\| \in X^{S \times S}$ such that $u_k(x_{kl}) > u_k(x_{lk})$ for all $k \neq l$. (The diagonal of this matrix is arbitrary.)

For any i, let $x_i^* = \frac{1}{S^2} \sum_{k,l} x_{kl}^i$, where $x_{kl}^i = x_{kl}$ if $u_i(x_{kl}) \geq u_i(x_{lk})$ and $x_{kl}^i = x_{lk}$ otherwise. Fix any i, j such that $i \neq j$. Then for all k, l,

$$u_i(x_{kl}^i) = [u_i(x_{kl}) \vee u_i(x_{lk})] \geq u_i(x_{kl}^j).$$

Moreover, if $k = i$ and $l = j$, then $u_i(x_{kl}^i) = u_i(x_{ij}) > u_i(x_{ji}) = u_i(x_{kl}^j)$. Thus,

$$u_i(x_i^*) = \tfrac{1}{S^2} \sum_{k,l} u_i(x_{kl}^i) > \tfrac{1}{S^2} \sum_{k,l} u_i(x_{kl}^j) = u_i(x_j^*).$$

On the other hand, suppose that the inequalities $u_i(x_i^*) > u_i(x_j^*)$ and $u_j(x_j^*) > u_j(x_i^*)$ hold for any $i \neq j$. Then the functions u_i and u_j are not constant and represent different rankings on X. Therefore, the list $u_1, \ldots, u_S : X \to \mathbb{R}$ is not redundant. □

For any $x \in X$, $A \in \mathcal{A}$, and $\delta \geq 0$, let $d(x, A) = \min_{y \in A} d(x, y)$ and

$$N_\delta(A) = \{x \in X : d(x, A) \leq \delta\}.$$

Note that $N_\delta(A) \subset X$ is a compact set, that is, a menu.

For any menus $A, B \in \mathcal{A}$, define the Hausdorff metric

$$\mu(A, B) = \left(\max_{x \in B} d(x, A) \right) \vee \left(\max_{x \in A} d(x, B) \right).$$

Equivalently, $\mu(A, B) = \min\{\delta \geq 0 \,:\, B \subset N_\delta(A) \text{ and } A \subset N_\delta(B)\}$. The metric μ turns $\mathcal{A}$ into a compact metric space.

For any menu $A \in \mathcal{A}$, let $\overline{\text{co}}\,(A)$ be the closure of the convex hull of A. In other words, $\overline{\text{co}}\,(A)$ is the smallest convex menu that includes A.

Lemma A.2. *The metric μ and the operator $\overline{\mathrm{co}}(\cdot)$ satisfy the following properties:*

(i) for all $n \geq 1$ and menus $A_1, \ldots, A_n, B_1, \ldots, B_n \in \mathcal{A}$,

$$\mu\left(\cup_{i=1}^n A_i, \cup_{i=1}^n B_i\right) \leq \max_i \mu(A_i, B_i), \tag{19}$$

(ii) for all menus $A, A', B, B' \in \mathcal{A}$,

$$\mu\left(\tfrac{A+A'}{2}, \tfrac{B+B'}{2}\right) \leq \mu(A,B) \vee \mu(A', B'), \tag{20}$$

(iii) if a menu $A \in \mathcal{A}$ consists of n elements, then

$$\overline{\mathrm{co}}(A) = \mathrm{co}(A) = \tfrac{1}{n} A + \tfrac{n-1}{n} \overline{\mathrm{co}}(A), \tag{21}$$

(iv) $\overline{\mathrm{co}}(\cdot)$ is a continuous mapping on the space $\mathcal{A}$,

(v) for all $\alpha \in [0,1]$ and menus $A, A' \in \mathcal{A}$,

$$\overline{\mathrm{co}}(\alpha A + (1-\alpha)A') = \alpha \overline{\mathrm{co}}(A) + (1-\alpha)\overline{\mathrm{co}}(A'). \tag{22}$$

Proof. Take any $n \geq 1$ and menus A, A', B, B', A_i, B_i for $i = 1, \ldots, n$.

Show (i). Let $\delta = \max_i \mu(A_i, B_i)$. Then $B_i \subset N_\delta(A_i)$ for all i, and hence,

$$\cup_{i=1}^n B_i \subset \cup_{i=1}^n N_\delta(A_i) = N_\delta(\cup_{i=1}^n A_i).$$

Similarly, $\cup_{i=1}^n A_i \subset N_\delta(\cup_{i=1}^n B_i)$. Therefore, $\mu\left(\cup_{i=1}^n A_i, \cup_{i=1}^n B_i\right) \leq \delta$.

Turn to (ii). Let $\delta = \mu(A,B) \vee \mu(A', B')$. Then $B \subset N_\delta(A)$ and $B' \subset N_\delta(A')$. As the metric d is quasi-convex, then

$$\tfrac{B+B'}{2} \subset \tfrac{N_\delta(A)+N_\delta(A')}{2} \subset N_\delta\left(\tfrac{A+A'}{2}\right).$$

Similarly, $\frac{A+A'}{2} \subset N_\delta\left(\frac{B+B'}{2}\right)$. Therefore, $\mu\left(\frac{A+A'}{2}, \frac{B+B'}{2}\right) \leq \delta$.

Turn to (iii). Suppose that $A = \{x_1, \ldots, x_n\}$ consists of n elements. Then

$$\mathrm{co}(A) = \left\{ \sum_{i=1}^n \alpha_i x_i : \alpha_i \geq 0, \sum_{i=1}^n \alpha_i = 1 \right\}$$

is compact because the n-element mixture operation is continuous, and the $(n-1)$-dimensional simplex is compact. Therefore $\overline{\mathrm{co}}(A) = \mathrm{co}(A)$. Moreover, $\sum_{i=1}^n \alpha_i = 1$ implies that $\alpha_i \geq \frac{1}{n}$ for some i. Thus, $\mathrm{co}(A) = \frac{1}{n}A + \frac{n-1}{n}\mathrm{co}(A)$. Substitute $\overline{\mathrm{co}}(A) = \mathrm{co}(A)$ to obtain (21).

Turn to (iv). Let $\delta = \mu(A,B)$. Then $B \subset N_\delta(A) \subset N_\delta(\overline{\mathrm{co}}(A))$. The menu $N_\delta(\overline{\mathrm{co}}(A))$ is convex because d is quasi-convex. Then $\overline{\mathrm{co}}(B) \subset N_\delta(\overline{\mathrm{co}}(A))$. Similarly, $\overline{\mathrm{co}}(A) \subset N_\delta(\overline{\mathrm{co}}(B))$. Thus, $\mu(\overline{\mathrm{co}}(A), \overline{\mathrm{co}}(B)) \leq \delta = \mu(A,B)$, and $\overline{\mathrm{co}}(\cdot)$ is continuous.

Finally, if the menus A, A' are finite, then the equation (22) follows from (21). By continuity, (22) holds for all menus $A, A' \in \mathcal{A}$. □

A.1 Axioms $\Rightarrow$ Representation

Assume that $\succeq$ satisfies Order, Continuity, and Independence.

Lemma A.3. *For all menus $A \in \mathcal{A}$, $A \sim \overline{\mathrm{co}}\,(A)$.*

Proof. Take a menu $A \in \mathcal{A}$ such that $A \succ \overline{\mathrm{co}}\,(A)$. Let $B = \frac{A+A}{2}$. By Independence,

$$B \succ \tfrac{A+\overline{\mathrm{co}}(A)}{2} \succ \tfrac{\overline{\mathrm{co}}(A)+\overline{\mathrm{co}}(A)}{2} = \overline{\mathrm{co}}\,(B)\,.$$

Take a sequence of finite menus B_k that converges to B. Then $\overline{\mathrm{co}}\,(B_k)$ converges to $\overline{\mathrm{co}}\,(B)$. By Continuity, there is k^* such that $B_{k^*} \succ \frac{A+\overline{\mathrm{co}}(A)}{2} \succ \overline{\mathrm{co}}\,(B_{k^*})$. Let n be the size of the finite menu B_{k^*}. By (21),

$$\tfrac{1}{n}B_{k^*} + \tfrac{n-1}{n}\overline{\mathrm{co}}\,(B_{k^*}) = \overline{\mathrm{co}}\,(B_{k^*}) = \tfrac{1}{n}\overline{\mathrm{co}}\,(B_{k^*}) + \tfrac{n-1}{n}\overline{\mathrm{co}}\,(B_{k^*})\,.$$

Yet by Independence, $\frac{1}{n}B_{k^*} + \frac{n-1}{n}\overline{\mathrm{co}}\,(B_{k^*}) \succ \frac{1}{n}\overline{\mathrm{co}}\,(B_{k^*}) + \frac{n-1}{n}\overline{\mathrm{co}}\,(B_{k^*})$. This is a contradiction. Similarly, $A \prec \overline{\mathrm{co}}\,(A)$ is impossible. $\square$

Let $\mathcal{C} \subset \mathcal{A}$ be the set of all convex menus:

$$\mathcal{C} = \{C \in \mathcal{A} \,:\, C = \alpha C + (1-\alpha)C \quad \text{for all } \alpha \in [0,1]\}.$$

Then $\mathcal{C}$ is a *mixture space* as defined by Herstein–Milnor [12] (see also Kreps [15]). By the Herstein–Milnor Theorem, the preference $\succeq$ restricted to convex menus can be represented by a function $U_c : \mathcal{C} \to \mathbb{R}$ such that for all $\alpha \in [0,1]$ and $C, C' \in \mathcal{C}$,

$$U_c(\alpha C + (1-\alpha)C') = \alpha U_c(C) + (1-\alpha)U_c(C'). \tag{23}$$

Show that U_c is continuous. Fix any $\gamma \in \mathbb{R}$. By (23), either $U_c(C) > \gamma$ for all $C \in \mathcal{C}$, or $U_c(C) < \gamma$ for all $C \in \mathcal{C}$, or $U_c(C_\gamma) = \gamma$ for some $C_\gamma \in \mathcal{C}$. In each of these cases, the sets $\{C \in \mathcal{C} : U_c(C) \geq \gamma\}$ and $\{C \in \mathcal{C} : U_c(C) \leq \gamma\}$ are closed in $\mathcal{C}$. For instance, in the non-trivial case when $U_c(C_\gamma) = \gamma$, the set

$$\{C \in \mathcal{C} : U_c(C) \geq \gamma\} = \mathcal{C} \cap \{A \in \mathcal{A} : A \succeq C_\gamma\}$$

is closed by Continuity.

For all menus $A \in \mathcal{A}$, let

$$U(A) = U_c\left(\overline{\mathrm{co}}\,(A)\right). \tag{24}$$

By Lemma A.3, the function $U : \mathcal{A} \to \mathbb{R}$ represents $\succeq$ on the entire $\mathcal{A}$. The utility U is continuous because it is a composition of continuous mappings. Moreover, U satisfies

$$U(\alpha A + (1-\alpha)B) = \alpha U(A) + (1-\alpha)U(B) \tag{25}$$

for all $\alpha \in [0,1]$ and menus $A, B \in \mathcal{A}$ because

$$U_c(\overline{\text{co}}\,(\alpha A + (1-\alpha)B)) = U_c(\alpha\overline{\text{co}}\,(A) + (1-\alpha)\overline{\text{co}}\,(B)) = \alpha U_c(\overline{\text{co}}\,(A)) + (1-\alpha)U_c(\overline{\text{co}}\,(B)).$$

Assume further that $\succeq$ satisfies Finiteness. Then the finite maximum

$$S(\succeq) = \max_{n\geq 1, A_1,\dots,A_n \in \mathcal{A}} [P(A_1,\dots,A_n) + N(A_1,\dots,A_n)]$$

is well-defined.

Show that the utility function U has the required form (9). If $S(\succeq) = 0$, then the preference $\succeq$ is degenerate because for all menus $A, B \in \mathcal{A}$, both A and B are neutral in the family A, B and hence, $A \sim A \cup B \sim B$.

Lemma A.4. *If $S(\succeq) = 1$, then there is $\gamma_1 \in \{-1, 1\}$ and a continuous linear function $u_1 : X \to \mathbb{R}$ such that $U(A) = \gamma_1 \max_{x\in A} u_1(x)$ for all menus $A \in \mathcal{A}$.*

Proof. Suppose that $S(\succeq) = 1$. Then for any menus $B, B' \in \mathcal{A}$, $U(B \cup B') = U(B)$ or $U(B \cup B') = U(B')$. It follows by induction with respect to n that any finite family of menus $B_1, \dots, B_n$ contains a member B_k such that $U\left(\cup_{i=1}^{n} B_i\right) = U(B_k)$.

With some abuse of notation, let $U(x) = U(\{x\})$ and $U(x,y) = U(\{x,y\})$ for all $x, y \in X$. Consider three possible cases.

Case 1. $U(x,y) = U(x) \vee U(y)$ for all $x, y \in X$. Let $\gamma_1 = 1$ and $u_1 \equiv U$ on X. Fix any finite menu A. Take $x^* \in A$ such that $U(x^*) = \max_{x\in A} U(x)$. The finite union $A = \cup_{x\in A}\{x^*, x\}$ has a component $\{x^*, y\}$ such that $U(A) = U(x^*, y)$. Then

$$U(A) = U(x^*, y) = U(x^*) \vee U(y) = U(x^*) = \max_{x\in A} U(x) = \gamma_1 \max_{x\in A} u_1(x).$$

By continuity, this equality extends from finite menus to the entire $\mathcal{A}$.

Case 2. $U(x,y) = U(x) \wedge U(y)$ for all $x, y \in X$. Let $\gamma_1 = -1$ and $u_1 \equiv -U$ on X. Analogously to the previous case, for all menus $A \in \mathcal{A}$,

$$U(A) = \min_{x\in A} U(x) = \gamma_1 \max_{x\in A} u_1(x).$$

Case 3. $U(x,y) \neq U(x) \vee U(y)$ and $U(x', y') \neq U(x') \wedge U(y')$ for some $x, y, x', y' \in X$. Wlog $U(x) \geq U(y)$ and $U(x') \geq U(y')$. Then $U(x) > U(y) = U(x,y)$ and $U(x') > U(y') = U(x', y')$.

For $\alpha \in [0,1]$, let $x_\alpha = \alpha x + (1-\alpha)x'$ and $y_\alpha = \alpha y + (1-\alpha)y'$. Let

$$I_x = \{\alpha \in [0,1] : U(x_\alpha) = U(x_\alpha, y_\alpha)\}$$
$$I_y = \{\alpha \in [0,1] : U(y_\alpha) = U(x_\alpha, y_\alpha)\}.$$

Then the sets I_x and I_y are (i) non-empty because $0 \in I_x$ and $1 \in I_y$, (ii) disjoint because $U(x_\alpha) > U(y_\alpha)$ by (25), (iii) closed because U is continuous, and (iv) cover the segment $[0,1]$ because for every $\alpha \in [0,1]$, $U(x_\alpha, y_\alpha) = U(x_\alpha)$ or $U(x_\alpha, y_\alpha) = U(y_\alpha)$. This is a contradiction because the segment $[0,1]$ is connected. □

Suppose that $S(\succeq) \geq 2$. Let $S = S(\succeq)$, and let indices i, j vary in $\{1, \ldots, S\}$. Given menus $A_1^*, \ldots, A_S^*$, write $A^* = \cup_i A_i^*$ and $A_{-j}^* = \cup_{i \neq j} A_i^*$.

We claim that there are convex menus $A_1^*, \ldots, A_S^*$ such that for all j,

$$U(A^*) \neq U\left(A_{-j}^*\right). \tag{26}$$

By definition of S, there is $n \geq S$ and a family of menus $A_1, \ldots, A_n \in \mathcal{A}$ such that

$$P(A_1, \ldots, A_n) + N(A_1, \ldots, A_n) = S,$$

and wlog each of the first S menus $A_1, \ldots, A_S$ in this family is either positive or negative. Let $B = \cup_{k=S+1}^n A_i$ and for all i, $A_i^* = \overline{\text{co}}\,(A_i \cup B)$. Then

$$U(A^*) = U(\overline{\text{co}}\,(A^*)) = U(\overline{\text{co}}\,(\cup_{k \leq n} A_k)) = U(\cup_{k \leq n} A_k).$$

Similarly,

$$U\left(A_{-j}^*\right) = U\left(\overline{\text{co}}\,\left(A_{-j}^*\right)\right) = U(\overline{\text{co}}\,(\cup_{k \leq n, k \neq j} A_k)) = U(\cup_{k \leq n, k \neq j} A_k).$$

Then the inequality (26) follows from the fact that A_j is either positive or negative in the family $A_1, \ldots, A_n$.

Wlog $U(A^*) = 0$. Take $\varepsilon > 0$ such that for all j,

$$|U(A_{-j}^*)| = |U(A_{-j}^*) - U(A^*)| > 4\varepsilon. \tag{27}$$

As $\mathcal{A}$ is compact, then U is uniformly continuous, and hence, there is $\delta > 0$ such that for all menus $B, B' \in \mathcal{A}$,

$$\mu(B, B') < \delta \quad \Rightarrow \quad |U(B) - U(B')| < \varepsilon. \tag{28}$$

Next we claim that there is $\alpha > 0$ such that for all menus $B, B', C \in \mathcal{A}$,

$$\mu(\alpha B + (1-\alpha)C, \alpha B' + (1-\alpha)C) < \delta. \tag{29}$$

Show this claim by contradiction: suppose that there are sequences $\alpha_n \in \mathbb{R}$ and $B_{\alpha_n}, B'_{\alpha_n}, C_{\alpha_n} \in \mathcal{A}$ such that $\alpha_n \to 0$ and

$$\mu(\alpha_n B_{\alpha_n} + (1-\alpha_n)C_{\alpha_n}, \alpha_n B'_{\alpha_n} + (1-\alpha_n)C_{\alpha_n}) \geq \delta.$$

As $\mathcal{A}$ is compact, one can take the sequences $B_{\alpha_n}, B'_{\alpha_n}, C_{\alpha_n}$ to converge in $\mathcal{A}$. Thus $\mu(\lim_n C_{\alpha_n}, \lim_n C_{\alpha_n}) \geq \delta$, which is a contradiction.

For any family of S menus $B_1, \ldots, B_S$, let

$$[B_1, \ldots, B_S] = \bigcup_i (\alpha B_i + (1-\alpha)A_i^*).$$

The ranking of such menus satisfies the following additivity property.

Lemma A.5. *For all menus $B_1, \dots, B_S$ and $B'_1, \dots, B'_S$,*

$$U\left(\frac{[B_1,\dots,B_S]+[B'_1,\dots,B'_S]}{2}\right) = U\left(\left[\frac{B_1+B'_1}{2}, \dots, \frac{B_S+B'_S}{2}\right]\right).$$

Proof. Fix any menus $B_1, \dots, B_S$ and $B'_1, \dots, B'_S$. For any i, let

$$A_i = \alpha\frac{B_i+B'_i}{2} + (1-\alpha)A^*_i.$$

Let $D = \bigcup_{i,j:i\neq j}\left(\alpha\frac{B_i+B'_j}{2} + (1-\alpha)\frac{A^*_i+A^*_j}{2}\right)$. Then

$$\frac{[B_1,\dots,B_S]+[B'_1,\dots,B'_S]}{2} = A_1 \cup \dots \cup A_S \cup D$$
$$\left[\frac{B_1+B'_1}{2}, \dots, \frac{B_S+B'_S}{2}\right] = A_1 \cup \dots \cup A_S.$$

Let $A = \cup_i A_i$ and $A_{-j} = \cup_{i\neq j} A_i$. We claim that for any j,

$$U(A \cup D) \neq U(A_{-j} \cup D). \tag{30}$$

Then each menu $A_1, \dots, A_S$ is either positive or negative in the family $A_1, \dots, A_S, D$. By definition of S, this family cannot contain more than S members that are not neutral. It follows that D is neutral, that is,

$$U(A_1 \cup \dots \cup A_S \cup D) = U(A_1 \cup \dots \cup A_S)$$

completing the proof of the lemma.

Show the claim (30). For all i, $A^*_i = \alpha A^*_i + (1-\alpha)A^*_i$ and hence, by (29),

$$\mu(A^*_i, \alpha B_i + (1-\alpha)A^*_i) < \delta.$$

Let $B = [B_1, \dots, B_S]$ and $B' = [B'_1, \dots, B'_S]$. By (19),

$$\mu(A^*, B) \leq \max_i \mu(A^*_i, \alpha B_i + (1-\alpha)A^*_i) < \delta$$
$$\mu(A^*, B') \leq \max_i \mu(A^*_i, \alpha B'_i + (1-\alpha)A^*_i) < \delta,$$

By the uniform continuity (28), $|U(B)| < \varepsilon$ and $|U(B')| < \varepsilon$. Then

$$|U(A \cup D)| = \left|U\left(\tfrac{B+B'}{2}\right)\right| < \varepsilon. \tag{31}$$

Fix any j. Let $B_{-j} = \cup_{i\neq j}(\alpha B_i + (1-\alpha)A^*_i)$ and $B'_{-j} = \cup_{i\neq j}(\alpha B'_i + (1-\alpha)A^*_i)$. By (19) and (20),

$$\mu(A^*_{-j}, B_{-j}) \leq \max_{i\neq j} \mu(A^*_i, \alpha B_i + (1-\alpha)A^*_i) < \delta$$
$$\mu(A^*_{-j}, B'_{-j}) \leq \max_{i\neq j} \mu(A^*_i, \alpha B'_i + (1-\alpha)A^*_i) < \delta$$
$$\mu\left(\frac{A^*+A^*_{-j}}{2}, \frac{B+B'_{-j}}{2}\right) \leq \mu(A^*, B) \vee \mu(A^*_{-j}, B'_{-j}) < \delta$$
$$\mu\left(\frac{A^*+A^*_{-j}}{2}, \frac{B'+B_{-j}}{2}\right) \leq \mu(A^*, B') \vee \mu(A^*_{-j}, B_{-j}) < \delta.$$

Note that $A_{-j} \cup D = \frac{B'+B_{-j}}{2} \cup \frac{B+B'_{-j}}{2}$. Therefore by (19),

$$\mu\left(\frac{A^*+A^*_{-j}}{2}, A_{-j} \cup D\right) \leq \mu\left(\frac{A^*+A^*_{-j}}{2}, \frac{B+B'_{-j}}{2}\right) \vee \left(\frac{A^*+A^*_{-j}}{2}, \frac{B'+B_{-j}}{2}\right) < \delta.$$

By (28), $\left|U\left(A_{-j} \cup D\right) - U\left(\frac{A^*+A^*_{-j}}{2}\right)\right| < \varepsilon$. By (27), $\left|U\left(\frac{A^*+A^*_{-j}}{2}\right)\right| > 2\varepsilon$. Then

$$\left|U\left(A_{-j} \cup D\right)\right| > \varepsilon.$$

□

By (31), $U\left(A \cup D\right) \neq U\left(A_{-j} \cup D\right)$ and the claim (30) is true.

For any i, define a function $W_i : \mathcal{A} \to \mathbb{R}$ over menus $B \in \mathcal{A}$ by

$$W_i(B) = \tfrac{1}{\alpha} U\left(A^*_{-i} \cup (\alpha B + (1-\alpha) A^*_i)\right).$$

Lemma A.6. *For any i, there is $\gamma_i \in \{-1, 1\}$ and a continuous linear function $u_i : X \to \mathbb{R}$ such that $W_i(A) = \gamma_i \max_{x \in A} u_i(x)$ for all menus $A \in \mathcal{A}$.*

Proof. The required components γ_i and u_i are derived analogously to γ_1 and u_1 in Lemma A.4 providing that W_i satisfies the following properties:

(i) the function $W_i : \mathcal{A} \to \mathbb{R}$ is continuous,

(ii) for all $x, x' \in X$, $W_i(\{\alpha x + (1-\alpha)x'\}) = \alpha W_i(\{x\}) + (1-\alpha) W_i(\{x'\})$,

(iii) for all $B, B' \in \mathcal{A}$, either $W_i(B \cup B') = W_i(B)$ or $W_i(B \cup B') = W_i(B')$.

For concreteness, we show these properties for W_S. The continuity of W_S is implied by the continuity of U and the mixture operation.

Turn to (ii). By Lemma A.5, for all $x, x' \in X$,

$$W_S\left(\left\{\tfrac{x+x'}{2}\right\}\right) = \tfrac{1}{\alpha} U\left(\left[A^*_1, \ldots, A^*_{S-1}, \left\{\tfrac{x+x'}{2}\right\}\right]\right) = \tfrac{1}{\alpha} \frac{U\left([A^*_1, \ldots, A^*_{S-1}, \{x\}]\right) + U\left([A^*_1, \ldots, A^*_{S-1}, \{x'\}]\right)}{2} = \frac{W_S(\{x\}) + W_S(\{x'\})}{2}.$$

By induction with respect to n, for all dyadic rationals $\gamma \in \left\{0, \frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n-1}{2^n}, 1\right\}$,

$$W_S(\{\gamma x + (1-\gamma)x'\}) = \gamma W_S(\{x\}) + (1-\gamma) W_S(\{x'\}).$$

By continuity, the function W_S is linear over singleton menus.

Turn to (iii). Fix any menus $B, B' \in \mathcal{A}$ such that $W_S(B \cup B') \neq W_S(B')$. Let $A_i = A^*_i$ for all $i < S$, $A_S = \alpha B + (1-\alpha) A^*_S$, and $D = \alpha B' + (1-\alpha) A^*_S$. Let $A = \cup_i A_i$ and $A_{-j} = \cup_{i \neq j} A_i$. By (29), $\mu(A, A^*) \leq \max_i \mu(A_i, A^*_i) < \delta$, and

$$\mu(A \cup D, A^*) = \mu(A \cup D, A^* \cup A^*_S) \leq \mu(A, A^*) \vee \mu(D, A^*_S) < \delta.$$

By (28),$|U(A \cup D)| < \varepsilon$. Take any $j < S$. Then

$$\mu(A_{-j} \cup D, A^*_{-j}) = \mu(A_{-j} \cup D, A^*_{-j} \cup A^*_S) \leq \mu(A_{-j}, A^*_{-j}) \vee \mu(D, A^*_S) < \delta.$$

By (28), $|U(A_{-j} \cup D) - U(A^*_{-j})| < \varepsilon$. By (27), $|U(A_{-j} \cup D)| > 3\varepsilon > |U(A \cup D)|$. Therefore,

$$U(A \cup D) \neq U(A_{-j} \cup D).$$

Moreover, $W_S(B \cup B') \neq W_S(B')$ implies that

$$U(A \cup D) = W_S(B \cup B') \neq W_S(B') = U(A_{-S} \cup D).$$

Then each menu $A_1, \ldots, A_S$ is either positive or negative in the family $A_1, \ldots, A_S, D$. By definition of S, this family does contain more than S menus that are not neutral. Therefore D is neutral and $U(A \cup D) = U(A)$. Thus $W_S(B \cup B') = W_S(B)$. □

By Lemma A.5, for all menus $B_1, \ldots, B_S$,

$$U([B_1, \ldots, B_S]) = U([B_1, \ldots, B_S]) + U(A^*) = U([B_1, \ldots, B_S]) + U([A^*_1, \ldots, A^*_S]) =$$
$$2U\left(\left[\tfrac{B_1 + A^*_1}{2}, \ldots, \tfrac{B_S + A^*_S}{2}\right]\right) =$$
$$U([A^*_1, B_2, \ldots, B_S]) + U([B_1, A^*_2, \ldots, A^*_S]) = U([A^*_1, B_2 \ldots, B_S]) + \alpha W_1(B_1).$$

Similarly, $U([A^*_1, B_2 \ldots, B_S]) = U([A^*_1, A^*_2, B_3, \ldots, B_S]) + \alpha W_2(B_2)$. Inductively,

$$U([B_1, \ldots, B_S]) = \sum_i \alpha W_i(B_i).$$

Thus, for all menus $A \in \mathcal{A}$,

$$\alpha U(A) = U(\alpha A + (1-\alpha)A^*) = U([A, \ldots, A]) = \alpha \sum_i W_i(A),$$

that is, $U(A) = \sum_i W_i(A) = \sum_i \gamma_i \max_{x \in A} u_i(x)$.

To show that the list of functions $u_1, \ldots, u_S$ is not redundant, we use Lemma A.1. For any i, take $x^*_i \in A^*_i$ such that $u_i(x^*_i) = \max_{x \in A^*_i} u_i(x)$. Fix any $i \neq j$. Suppose that $u_i(x^*_j) \geq u_i(x^*_i)$. Then

$$\max_{x \in A^*_{-i}} u_i(x) \geq \max_{x \in A^*_i} u_i(x)$$

and hence, $\max_{x \in A^*} u_i(x) = \max_{x \in A^*_{-i}} u_i(x)$. Therefore $W_i(A^*) = W_i(A^*_{-i})$. Yet direct calculation shows that $W_i(A^*) = U(A^*)$ because

$$\overline{\text{co}}(A^*) = \overline{\text{co}}\left(A^*_{-i} \cup (\alpha A^* + (1-\alpha)A^*_i)\right),$$

but $W_i(A^*_{-i}) = U(A^*_{-i})$ because

$$\overline{\text{co}}\left(\alpha A^*_{-i} + (1-\alpha)A^*)\right) = \overline{\text{co}}\left(A^*_{-i} \cup (\alpha A^*_{-i} + (1-\alpha)A^*_i)\right).$$

Thus $W_i(A^*) = 0 \neq W_i(A^*_{-i})$. This contradiction shows that $u_i(x^*_i) > u_i(x^*_j)$, and hence, the list $u_1, \ldots, u_S$ is not redundant by Lemma A.1.

A.2 Representation $\Rightarrow$ Axioms + Uniqueness

Suppose that $\succeq$ has the representation (9). Then $\succeq$ satisfies Order, Continuity, and Independence (we omit the standard proofs).

Turn to Finiteness. The representation (9) has S components, among which $S_+ = \sum_{i:\gamma_i=1} \gamma_i$ are positive and $S_- = \sum_{i:\gamma_i=-1} \gamma_i = S - S_+$ are negative. Let i, j vary in $\{1, \ldots, S\}$. For any i and any menu A, let $U_i(A) = \max_{x \in A} u_i(x)$.

Lemma A.7. *$P(\succeq) \leq S_+$, $N(\succeq) \leq S_-$, and $S(\succeq) \leq S$.*

Proof. Take any $n \geq 1$ and a family of menus $A_1, \ldots, A_n$. Let k, l vary in $\{1, \ldots, n\}$. Let $A = \cup_k A_k$ and for any l, $A_{-l} = \cup_{k \neq l} A_k$. If the menu A_l is positive, then

$$U(A) - U(A_{-l}) = \sum_i \gamma_i [U_i(A) - U_i(A_{-l})] > 0,$$

and hence, there is i_l such that $\gamma_{i_l} = 1$ and $U_{i_l}(A) > U_{i_l}(A_{-l})$. The latter inequality implies that $U_{i_l}(A_l) > U_{i_l}(A_k)$ for all $k \neq l$. Similarly, if $k \neq l$ and the menu A_k is also positive, then $U_{i_k}(A_k) > U_{i_k}(A_l)$, and hence, $i_l \neq i_k$. Therefore,

$$P(A_1, \ldots, A_n) = \sum_{k: A_k \text{ is positive}} \gamma_{i_k} \leq \sum_{i:\gamma_i=1} \gamma_i = S_+.$$

Analogously, $N(A_1, \ldots, A_n) \leq S_-$. Thus $P(\succeq) \leq S_+$, $N(\succeq) \leq S_-$, and hence, $S(\succeq) \leq S$. $\square$

Lemma A.7 implies that for any $n > S$ and any family $A_1, \ldots, A_n$,

$$P(A_1, \ldots, A_n) + N(A_1, \ldots, A_n) \leq S_+ + S_- = S < n,$$

and hence, there is a neutral menu A_k in this family. Finiteness follows.

Show the uniqueness statement in Theorem 2.1. Suppose that $S = 1$ and u_1 is non-constant. Take $x', y' \in X$ such that $u_1(x') > u_1(y')$. If $\gamma_1 = 1$, then the menu $\{x'\}$ is positive in the family $\{x'\}$, $\{y'\}$. Therefore $P(\succeq) \geq 1$. By Lemma A.7, $P(\succeq) = 1 = S_+$ and $N(\succeq) = 0 = S_-$. Analogously, if $\gamma_1 = -1$, then $P(\succeq) = 0 = S_+$ and $N(\succeq) = 1 = S_-$. By Herstein–Milnor's theorem, $u_1(\cdot)$ is unique up to a positive linear transformation because $\gamma_1 u_1$ is a linear representation of the commitment ranking on X (i.e. ranking of singleton menus).

Hereafter, suppose that $S \geq 2$. By Lemma A.1, there are $x_1^*, \ldots, x_S^* \in X$ such that $u_i(x_i^*) > u_i(x_j^*)$ for all $i \neq j$.

Let $A^* = \cup_i \{x_i^*\}$. For any j, let $A_{-j}^* = \cup_{i \neq j} \{x_i^*\}$. Then for all i, j such that $i \neq j$, $U_i(A^*) = u_i(x_i^*) = U_i(A_{-j}^*)$ and $U_j(A^*) = u_j(x_j^*) > U_j(A_{-j}^*)$. Therefore,

$$U(A^*) - U(A_{-j}^*) = \sum_i \gamma_i [U_i(A^*) - U_i(A_{-j}^*)] = \gamma_j [U_j(A^*) - U_j(A_{-j}^*)]. \quad (32)$$

Then $A^* \succ A^*_{-j}$ if $\gamma_j = 1$ and $A^* \prec A^*_{-j}$ if $\gamma_j = -1$. Therefore the family $\{x^*_1\}, \ldots, \{x^*_S\}$ contains S_+ positive menus and S_- negative ones. Thus $P(\succeq) \geq S_+$, $N(\succeq) \geq S_-$, and $S(\succeq) \geq S$. By Lemma A.7, $P(\succeq) = S_+$, $N(\succeq) = S_-$, and $S(\succeq) = S$.

Next, suppose that $\succeq$ has another utility representation (9) with S components: for all $A \in \mathcal{A}$,

$$U'(A) = \sum_i \gamma'_i \max_{x \in A} u'_i(x) = \sum_i \gamma'_i U'_i(A),$$

where $\gamma'_i \in \{-1, 1\}$, u'_i is linear and continuous, and $U'_i(A) = \max_{x\in A} u'_i(x)$.

For any j, consider two cases.

Case 1. $\gamma_j = 1$. By (32), $A^* \succ A^*_{-j}$ and

$$U'(A^*) - U'(A^*_{-j}) = \sum_i \gamma'_i \left[U'_i(A^*) - U'_i(A^*_{-j})\right] > 0.$$

There is $\pi(j) \leq S$ such that $\gamma'_{\pi(j)} = \gamma_j = 1$ and $U'_{\pi(j)}(A^*) - U'_{\pi(j)}(A^*_{-j}) > 0$.

Case 2. $\gamma_j = -1$. By (32), $A^* \prec A^*_{-j}$ and

$$U'(A^*) - U'(A^*_{-j}) = \sum_i \gamma'_i \left[U'_i(A^*) - U'_i(A^*_{-j})\right] < 0.$$

There is $\pi(j) \leq S$ such that $\gamma'_{\pi(j)} = \gamma_j = -1$ and $U'_{\pi(j)}(A^*) - U'_{\pi(j)}(A^*_{-j}) > 0$.

In either case, $U'_{\pi(j)}(A^*) - U'_{\pi(j)}(A^*_{-j}) > 0$. Then for all $i \neq j$,

$$u'_{\pi(j)}(x^*_j) > u'_{\pi(j)}(x^*_i), \tag{33}$$

and hence, $\pi(i) \neq \pi(j)$. Thus, π is a permutation of the finite set S.

For any i, let $u''_i \equiv u'_{\pi(i)}$. Then for all menus $A \in \mathcal{A}$,

$$U'(A) = \sum_i \gamma'_i U'_i(A) = \sum_i \gamma'_{\pi(i)} U'_{\pi(i)}(A) = \sum_i \gamma_i U''_i(A),$$

where $U''_i(A) = \max_{x\in A} u''_i(x)$.

Lemma A.8. *There are $\alpha > 0$ and $\beta_i \in \mathbb{R}$ such that $u''_i \equiv \alpha u_i + \beta_i$ for all i.*

Proof. The Herstein-Milnor Theorem implies that there are $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $U'(\cdot) = \alpha U(\cdot) + \beta$ on the mixture space $\mathcal{C}$ of convex menus. By Lemma A.3, $U'(\cdot) = \alpha U(\cdot) + \beta$ on the entire $\mathcal{A}$.

If $i \neq j$, then $u_i(x^*_i) > u_i(x^*_j)$ by Lemma A.1, $u''_i(x^*_i) > u''_i(x^*_j)$ by (33), and hence there is a sufficiently small $\varepsilon > 0$ such that for all $x, y \in X$,

$$u_i(\varepsilon x + (1-\varepsilon)x^*_i) > u_j(\varepsilon y + (1-\varepsilon)x^*_j)$$
$$u''_i(\varepsilon x + (1-\varepsilon)x^*_i) > u''_j(\varepsilon y + (1-\varepsilon)x^*_j)).$$

For any j and $x \in X$, let $A^{\varepsilon}_{-j} = (\varepsilon x + (1-\varepsilon)x^*_j) \cup A^*_{-j}$. If $i \neq j$, then $U_i(A^*) = u_i(x_i) = U_i(A^{\varepsilon}_{-j})$ and $U''_i(A^*) = u''_i(x_i) = U''_i(A^{\varepsilon}_{-j})$. Then

$$U(A^*) - U(A^{\varepsilon}_{-j}) = \sum_i \gamma_i [U_i(A^*) - U_i(A^{\varepsilon}_{-j})] = \gamma_j \varepsilon (u_j(x^*_j) - u_j(x))$$
$$U'(A^*) - U'(A^{\varepsilon}_{-j}) = \sum_i \gamma_i [U''_i(A^*) - U''_i(A^{\varepsilon}_{-j})] = \gamma_j \varepsilon (u''_j(x^*_j) - u''_j(x)).$$

As $U'(A^*) - U'(A^{\varepsilon}_{-j}) = \alpha(U(A^*) - U(A^{\varepsilon}_{-j}))$, then

$$\gamma_j \varepsilon (u''_j(x^*_j) - u''_j(x)) = \alpha \gamma_j \varepsilon (u_j(x^*_j) - u_j(x))$$

and $u''_j(x) = \alpha u_j(x) + \beta_j$, where $\beta_j = u''_j(x^*_j) - \alpha u_j(x^*_j)$. □

References

[1] M. Amador, I. Werning, and G.-M. Angeletos. Commitment vs. flexibility. *Econometrica*, 72, 2006.

[2] E. Dekel, B. L. Lipman, and A. Rustichini. Representing preferences with a unique subjective state space. *Econometrica*, 69:891–934, 2001.

[3] E. Dekel, B. L. Lipman, and A. Rustichini. Temptation-driven preferences. Working Paper, 2006.

[4] E. Dekel, B. L. Lipman, A. Rustichini, and T. Sarver. Representing preferences with a unique subjective state space. *Econometrica*, 2007.

[5] J. Elster. Emotions and economic theory. *Journal of Economic Literature*, 36:47–74, 1998.

[6] L. Epstein. An axiomatic model of non-Bayesian updating. *Review of Economic Studies*, 73:413–436, 2006.

[7] L. Epstein and I. Kopylov. An axiomatic model of 'cold feet'. Working Paper, 2007.

[8] L. Epstein, M. Marinacci, and K. Seo. Coarse contingencies. Working Paper, 2007.

[9] H. Ergin and T. Sarver. A unique costly contemplation representation. Working Paper, 2005.

[10] F. Gul and W. Pesendorfer. Temptation and self-control. *Econometrica*, 69:1403–1435, 2001.

[11] F. Gul and W. Pesendorfer. Self-control and the theory of consumption. *Econometrica*, 72:119–158, 2004.

[12] I. Herstein and J. Milnor. An axiomatic approach to measurable utility. *Econometrica*, 21:291–297, 1953.

[13] E. Kandel and E. Lazear. Peer pressure and partnerships. *The Journal of Political Economy*, 100:801–817, 1992.

[14] D. Kreps. A representation theorem for 'preference for flexibiility'. *Econometrica*, 47:565–577, 1979.

[15] D. M. Kreps. *Notes on the Theory of Choice.* Westview Press, Boulder and London, 1988.

[16] P. Mongin. A note on mixture sets in decision theory. *Decisions in Economics and Finance*, 24:59–69, 2001.

[17] J. Noor. Menu-dependent self-control. Research Paper, 2006.

[18] J. Noor. Commitment and self-control. *Journal of Economic Theory*, 2007.

[19] Y. Prohorov. Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.*, 1:413–436, 1956.

[20] T. Sarver. Anticipating regret: Why fewer options may be better. Working Paper, 2005.

[21] M. Stone. Postulates for the barycentric coordinates. *Annali di Matematica*, XXIX:25–30, 1949.