

A Unique Costly Contemplation Representation*

Haluk Ergin[†] Todd Sarver[‡]

Current Draft:
December 4, 2006

PRELIMINARY AND INCOMPLETE

Abstract

We extend the costly contemplation model to preferences over sets of lotteries, assuming that the state-dependent utilities are von Neumann-Morgenstern. The contemplation costs are uniquely pinned down in a reduced form representation, where the decision-maker selects a subjective measure over expected utility functions instead of a subjective signal over a subjective state space. We show that in this richer setup, costly contemplation is characterized by Aversion to Contingent Planning, Indifference to Randomization, Independence of Degenerate Decisions, and Strong Continuity.

*We thank Faruk Gul, Bart Lipman, Massimo Marinacci, and numerous seminar participants for helpful comments.

[†]Correspondence address: MIT, Department of Economics, E52-383A, 50 Memorial Drive, Cambridge, MA 02142. Email: hergin@mit.edu.

[‡]Correspondence address: Northwestern University, Department of Economics, 2001 Sheridan Road, Evanston, IL 60208. Email: tsarver@northwestern.edu.

1 Introduction

1.1 Brief Overview

We will take a Decision Maker (DM) to one of two restaurants. The first one is a seafood restaurant that serves a tuna (t) and a salmon (s) dish, which we denote by $A = \{t, s\}$. The second one is a steak restaurant that serves a filet mignon (f) and a ribeye (r) dish, which we denote by $B = \{f, r\}$. We will flip a coin to determine which restaurant to go to. If it comes up heads then we will buy the DM the meal of her choice in A , if it comes up tails then we will buy her the meal of her choice in B .

We consider presenting the DM one of the two following decision problems:

Decision Problem 1

We ask the DM to make a complete contingent plan listing what she would choose conditional on each outcome of the coin flip.

Decision Problem 2

We first flip the coin and let the DM know its outcome. She then selects the dish of her choice from the restaurant determined by the coin flip.

Decision problem 1 corresponds to a choice out of $A \times B = \{(t, f), (t, r), (s, f), (s, r)\}$, where for instance (s, f) is the plan where the DM indicates that she will have the salmon dish from the seafood restaurant if the coin comes up heads and she will have the filet mignon from the steak restaurant if the coin comes up tails. Note that each choice of a contingent plan eventually yields a lottery over meals. For example if the DM chooses (s, f) then she will face the lottery $\frac{1}{2}s + \frac{1}{2}f$ that yields a salmon and a filet mignon dish, each with one-half probability. Hence decision problem 1 is identical to a choice out of the set of lotteries $\{\frac{1}{2}t + \frac{1}{2}f, \frac{1}{2}t + \frac{1}{2}r, \frac{1}{2}s + \frac{1}{2}f, \frac{1}{2}s + \frac{1}{2}r\}$.

It is conceivable that the DM prefers facing the second decision problem rather than the first one. In this case we say that her preferences (over decision problems) exhibit *Aversion to Contingent Planning (ACP)*. One explanation of ACP is that the DM finds it psychologically costly to figure out her tastes over meals. Because of this cost, she would rather not contemplate on an inconsequential decision: In our restaurant example, she would rather not contemplate about her choice out of A , were she to know that the coin came up tails and her actual choice set is B . In particular she prefers to learn which choice set (A or B) is the relevant one, before contemplating on her choice.

Our main results are a representation and a uniqueness theorem for preferences over sets of lotteries. We interpret that the preference arises from a choice situation where, initially the DM chooses from among sets of lotteries (menus, options sets, or decision problems) and subsequently chooses a lottery from that set. The only primitive of the model is the preference over sets of lotteries which corresponds to the DM's choice behavior in the first period, we do not explicitly model the second period choice out of the sets. The key axiom in our analysis is ACP and our representation is a reduced form of the costly contemplation representation introduced in Ergin (2003). We begin by a brief overview of Ergin's and our results before we present them more formally in the next section.

The primitive of Ergin's model is a preference over sets of alternatives. He shows that if this preference is monotone, in the sense that each option set is weakly preferred to its subsets, then the DM behaves as if she optimally contemplates her mood before making a choice out of her option set. Ergin models contemplation as a subjective information acquisition problem, where the DM optimally acquires a costly subjective signal over a subjective state space. He interprets these subjective signals as contemplation strategies. The subjective state space, signals, and costs are all parameters of the representation but not a part of the model. However, as we illustrate in the next section these parameters are hardly pinned down from the preference over sets.

We extend the costly contemplation model to preferences over lotteries assuming that the state dependent utilities are von Neumann-Morgenstern. We model a contemplation strategy as a subjective measure over expected utility functions instead of a subjective signal over a subjective state space. The extension of the domain of preferences and the reduction of the parameters of the representation make it possible to uniquely identify contemplation costs from the preference. We show that in this extended model, ACP, indifference to randomization, along with continuity and independence of degenerate decisions properties characterize costly contemplation. We also prove that the measures in our representation are positive if and only if the preference is monotone.

1.2 Background and detailed results

The costly contemplation representation in Ergin (2003) is the following:

$$V(A) = \max_{\pi \in \Pi} \left[\sum_{E \in \pi} \max_{z \in A} \sum_{\omega \in E} U(z, \omega) - c(\pi) \right] \quad (1)$$

The interpretation of the above formula is as follows. The DM has a finite subjective state space Ω representing her tastes over alternatives. She does not know the realization of the state $\omega \in \Omega$ but has a uniform prior on Ω . Her tastes over alternatives in Z are represented by a state dependent utility function and $U: Z \times \Omega \rightarrow \mathbb{R}$. Before making a choice out of a set $A \subset Z$, the DM may engage in contemplation. A *contemplation strategy* is modeled as a signal about the subjective state which corresponds to a partition π of Ω . If the DM carries out the contemplation strategy π , she incurs a psychological cost of contemplation $c(\pi)$, learns which event of the partition π the actual state lies in, and picks an alternative that yields the highest expected utility conditional on each event $E \in \pi$. The set of partitions of Ω is denoted by Π . Faced with the option set A , the DM chooses an optimal level of contemplation by maximizing the value minus the cost of contemplation. This yields $V(A)$ in (1) as the ex-ante value of the option set A .

The appeal of the above formula (1) comes from its similarity to an optimal information acquisition formula. It expresses the costly contemplation problem as a problem with which we are more familiar as economists. The difference from a standard information acquisition problem is that, in the costly contemplation formula the parameters (Ω, U, c) are not directly observable but need to be derived from the DM's preference. Ergin shows that a preference \succsim over sets of alternatives is monotone ($A \subset B \Rightarrow B \succsim A$) if and only if there exist parameters (Ω, U, c) such that \succsim is represented by the ex-ante utility function V in (1). Unfortunately, these parameters are not pinned down from the preference \succsim . To illustrate this, first consider the following example:

Example 1 Let $Z = \{z_1, z_2, z_3\}$. All of the following cost functions lead to the same preference over sets of alternatives:

U	ω_1	ω_2	ω_3		c	c'
z_1	5	0	0	$\{\{\omega_1, \omega_2, \omega_3\}\}$	0	0
z_2	0	5	0	$\{\{\omega_i\}, \{\omega_j, \omega_k\}\}$	3	4
z_3	0	0	2	$\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$	4	5

given by $\{z_1, z_2, z_3\} \succ \{z_1, z_2\} \succ \{z_1, z_3\} \sim \{z_2, z_3\} \sim \{z_1\} \sim \{z_2\} \succ \{z_3\}$.

The above example should not come as a big surprise. Because of the finiteness of Ergin's model, there is only a finite number preferences over sets, hence the preference data is not enough to pin down one among a continuum possibility of costs. This suggests that the non-uniqueness problem might be resolved by increasing the domain of the preferences to sets of lotteries. Let $\Delta(Z)$ stand for the set of lotteries over Z , and

let A now denote a compact set of lotteries. The natural generalization of the costly contemplation formula in (1) to sets of lotteries is:

$$V(A) = \max_{\pi \in \Pi} \left[\sum_{E \in \pi} \max_{p \in A} \sum_{\omega \in E} U(p, \omega) - c(\pi) \right] \quad (2)$$

where the state-dependent utilities $U(\cdot, \omega): \Delta(Z) \rightarrow \mathbb{R}$ are assumed to be von Neumann-Morgenstern to guarantee additional structure in the extended model. Assuming that the preference over lotteries has a representation as in (2), it is indeed possible to distinguish between the alternative cost functions in Example 1 from the preference. For instance $V(\{z_1, \frac{1}{2}z_2 + \frac{1}{2}z_3\}) = 5.5 > 5 = V(\{z_1\})$ when the cost function is c , whereas $V'(\{z_1, \frac{1}{2}z_2 + \frac{1}{2}z_3\}) = 5 = V'(\{z_1\})$ when the cost function is c' . However, as we show next, extending the model to sets of lotteries is still not enough to guarantee uniqueness of the parameters (Ω, U, c) .

Example 2 Let $Z = \{z_1, z_2\}$ and consider the following two specifications of state spaces $\omega = \{\omega_1, \omega_2, \omega_3\}$, $\hat{\Omega} = \{\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3\}$, and the corresponding partition costs:

$\Omega:$	<table style="border-collapse: collapse; border: none;"> <tr><td style="border-right: 1px solid black; padding: 5px;">U</td><td style="padding: 5px;">ω_1</td><td style="border-right: 1px solid black; padding: 5px;">ω_2</td><td style="padding: 5px;">ω_3</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">z_1</td><td style="padding: 5px;">1</td><td style="border-right: 1px solid black; padding: 5px;">-2</td><td style="padding: 5px;">2</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">z_2</td><td style="padding: 5px;">-1</td><td style="border-right: 1px solid black; padding: 5px;">2</td><td style="padding: 5px;">-2</td></tr> </table>	U	ω_1	ω_2	ω_3	z_1	1	-2	2	z_2	-1	2	-2	<table style="border-collapse: collapse; border: none;"> <tr><td style="border-right: 1px solid black; padding: 5px;">$c(\{\{\omega_1, \omega_2, \omega_3\}\})$</td><td style="padding: 5px;">0</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">$c(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$</td><td style="padding: 5px;">1</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">$c(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$</td><td style="padding: 5px;">2</td></tr> </table>	$c(\{\{\omega_1, \omega_2, \omega_3\}\})$	0	$c(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$	1	$c(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$	2
U	ω_1	ω_2	ω_3																	
z_1	1	-2	2																	
z_2	-1	2	-2																	
$c(\{\{\omega_1, \omega_2, \omega_3\}\})$	0																			
$c(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$	1																			
$c(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$	2																			
$\hat{\Omega}:$	<table style="border-collapse: collapse; border: none;"> <tr><td style="border-right: 1px solid black; padding: 5px;">\hat{U}</td><td style="padding: 5px;">$\hat{\omega}_1$</td><td style="border-right: 1px solid black; padding: 5px;">$\hat{\omega}_2$</td><td style="padding: 5px;">$\hat{\omega}_3$</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">z_1</td><td style="padding: 5px;">-1</td><td style="border-right: 1px solid black; padding: 5px;">-1</td><td style="padding: 5px;">3</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">z_2</td><td style="padding: 5px;">1</td><td style="border-right: 1px solid black; padding: 5px;">1</td><td style="padding: 5px;">-3</td></tr> </table>	\hat{U}	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\omega}_3$	z_1	-1	-1	3	z_2	1	1	-3	<table style="border-collapse: collapse; border: none;"> <tr><td style="border-right: 1px solid black; padding: 5px;">$\hat{c}(\{\{\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3\}\})$</td><td style="padding: 5px;">0</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2, \hat{\omega}_3\}\})$</td><td style="padding: 5px;">1</td></tr> <tr><td style="border-right: 1px solid black; padding: 5px;">$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2\}, \{\hat{\omega}_3\}\})$</td><td style="padding: 5px;">2</td></tr> </table>	$\hat{c}(\{\{\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3\}\})$	0	$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2, \hat{\omega}_3\}\})$	1	$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2\}, \{\hat{\omega}_3\}\})$	2
\hat{U}	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\omega}_3$																	
z_1	-1	-1	3																	
z_2	1	1	-3																	
$\hat{c}(\{\{\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3\}\})$	0																			
$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2, \hat{\omega}_3\}\})$	1																			
$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2\}, \{\hat{\omega}_3\}\})$	2																			

In each of the two cases above, assume that the cost of all other partitions is 2. Then for any compact set of lotteries A , the parameters (Ω, U, c) and $(\hat{\Omega}, \hat{U}, \hat{c})$ yield the same ex-ante utility value $V(A)$ in (2).

Let U_E denote the expected utility function conditional on an event E , defined by $U_E(p) = \sum_{\omega \in E} U(p, \omega)$. Each partition π induces a collection of such conditional expected utility functions $(U_E)_{E \in \pi}$. When the DM undertakes the contemplation strategy π , she perceives that as a result of her contemplation, she will end up with an ex-post utility function in $(U_E)_{E \in \pi}$. Moreover, it is enough for her to know the cost $c(\pi)$ and the ex-post utility functions $(U_E)_{E \in \pi}$ associated with each contemplation strategy π , to evaluate the ex-ante value $V(A)$ in (2). In particular, it is not possible to behaviorally distinguish between two sets of parameters (Ω, U, c) and $(\hat{\Omega}, \hat{U}, \hat{c})$ that induce the same collections of ex-post conditional utility functions at the same costs.

To be concrete, consider the following table which lists the ex-post utility functions corresponding to the three partitions of Ω in Example 2:

π	$(U_E)_{E \in \pi}$
$\{\{\omega_1, \omega_2, \omega_3\}\}$	$(1, -1)$
$\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$	$(1, -1), (0, 0)$
$\{\{\omega_1, \omega_3\}, \{\omega_2\}\}$	$(3, -3), (-2, 2)$
$\{\{\omega_1, \omega_2\}, \{\omega_3\}\}$	$(-1, 1), (2, -2)$
$\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$	$(1, -1), (-2, 2), (2, -2)$

where we denote an expected utility function $u: \Delta(Z) \rightarrow \mathbb{R}$ by the vector $(u(z_1), u(z_2))$.

Let \mathcal{U} denote the non-constant expected utility functions on $\Delta(Z)$, normalized up to positive affine transformations. In Example 2, we can take $\mathcal{U} = \{(1, -1), (-1, 1)\}$.¹ Each ex-post utility function U_E must be a non-negative affine transformation of some $u \in \mathcal{U}$. In particular, each partition π induces a measure μ_π on \mathcal{U} , where the weight $\mu_\pi(u)$ of $u \in \mathcal{U}$ is the sum of the non-negative affine transformation coefficients given by:

$$\{\alpha_E \mid U_E = \alpha_E u + \beta_E, \alpha_E \geq 0, \text{ and } E \in \pi\}.$$

The measures induced by the above partitions are given by:

π	$\mu_\pi(1, -1)$	$\mu_\pi(-1, 1)$
$\{\{\omega_1, \omega_2, \omega_3\}\}$	1	0
$\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$	1	0
$\{\{\omega_1, \omega_2\}, \{\omega_3\}\}$	2	1
$\{\{\omega_1, \omega_3\}, \{\omega_2\}\}$	3	2
$\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$	3	2

We can use these induced measures to rewrite (2) in the alternative form:

$$V(A) = \max_{\mu \in \mathcal{M}} \left[\sum_{u \in \mathcal{U}} \mu(u) \max_{p \in A} u(p) - c(\mu) \right] \quad (3)$$

where the set of partitions Π are replaced by the set of measures $\mathcal{M} = \{\mu_\pi \mid \pi \in \Pi\}$ over \mathcal{U} and the cost of a measure is defined by $c(\mu) = \min\{c(\pi) \mid \mu = \mu_\pi\}$. In this formulation, a contemplation strategy is expressed as measure over expected utility functions instead of a subjective signal π over the subjective state space Ω .

¹We chose this normalization for the readability of the example, the normalization that we use later for our results is different than the one above.

Note that the integral of u with respect to an induced measure gives the ex-ante utility function U_Ω . In the current example $U_\Omega = (1, -1)$. Therefore the measures in \mathcal{M} also satisfy the following *consistency condition*:

$$\forall \mu, \nu \in \mathcal{M} : \quad \sum_{u \in \mathcal{U}} \mu(u) u = \sum_{u \in \mathcal{U}} \nu(u) u. \quad (4)$$

Even though a DM's realized tastes *ex-post* contemplation can be very different from her tastes *ex-ante* contemplation, the condition above requires that the contemplation process should not affect the DM's tendencies on the average.

1.3 Outline and Related Literature

Kreps (1979) was the first to study preferences over menus and to associate such preferences with a subjective state space representation. Dekel, Lipman, and Rustichini (2001, henceforth DLR) extend Kreps' analysis to the current setting of preferences over menus of lotteries. They use the additional structure of this domain to obtain an essentially unique subjective state space. Our extension of the analysis of Ergin (2003) can be thought of analogously to the extension of Kreps (1979) undertaken in DLR (2001).

One of the representations considered by DLR (2001), the additive EU representation, requires a version of the independence axiom. In the next section, we argue that the independence axiom requires that contemplation be costless, and hence the formal statement of aversion to contingent planning will be a relaxation of independence.² Specific violations of the independence axiom in the setting of preferences over menus of lotteries is still a largely unexplored area of research, and the only other paper we know of is Epstein and Marinacci (2005). They study an agent who has an incomplete, or coarse, conception of the future. This coarse conception entails a degree of pessimism on the part of the agent, and their resulting representations are intuitively similar to the maxmin representation of Gilboa and Schmeidler (1989). Aside from the obvious difference that our representation is maxmax, our consistency condition (see Equation (4)) is not appropriate for a model of ambiguity or coarseness.

In terms of mathematical technique, our work is most similar to that of Maccheroni, Marinacci, and Rustichini (2004). Like them, we use classic duality results from convex analysis to establish our representation theorem. However, their interests and the setting of their model are quite different from the present work. Using the Anscombe-Aumann

²We discuss the relationship between our axioms and representation and those of DLR (2001) in more detail in Section 4.

setting, they represent ambiguity aversion using a generalization of the maxmin representation of Gilboa and Schmeidler (1989). In their representation, each prior of the agent is associated with a cost, and hence this cost function captures the agent’s ambiguity attitude. Thus our representation appears similar to theirs in that both associate a cost function with measures. Nonetheless, as discussed in the previous paragraph, our representation cannot easily be interpreted as a model of ambiguity. Furthermore, the setting of our model is quite different, which requires us to develop some additional mathematical results (see Appendix A).

The remainder of the paper is organized as follows. We describe our model of contemplation in greater detail in Section 2. We introduce and motivate our axioms in Section 2.1, and we give a more precise statement of our representation in Section 2.2. Our main results, the existence and uniqueness theorems for our representation, are presented in Section 3. Section 4 contains some concluding remarks. All proofs are relegated to the appendix.

2 Axioms and Representation

2.1 Axioms

Let Z be a finite set of alternatives, and let $\Delta(Z)$ denote the set of all probability distributions on Z , endowed with the Euclidean metric d .³ Let \mathcal{A} denote the set of all closed subsets of $\Delta(Z)$, endowed with the Hausdorff metric, which is defined by

$$d_h(A, B) = \max \left\{ \max_{p \in A} \min_{q \in B} d(p, q), \max_{q \in B} \min_{p \in A} d(p, q) \right\}.$$

Elements of \mathcal{A} are called menus or option sets. The primitive of our model is a binary relation \succsim on \mathcal{A} , representing the DM’s preferences over menus. We maintain the interpretation that, after committing to a particular menu A , the DM chooses a lottery out of A in an unmodeled second stage.

For any $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$, define the convex combination of these two menus by $\alpha A + (1 - \alpha)B \equiv \{\alpha p + (1 - \alpha)q : p \in A \text{ and } q \in B\}$. Let $co(A)$ denote the convex hull of the set A .

We impose the following order and continuity axioms.

³Since Z is finite, the topology generated by d is equivalent to the vague topology on $\Delta(Z)$.

Axiom 1 (Weak Order): \succsim is complete and transitive.

Axiom 2 (Strong Continuity):

1. (Continuity): For all $A \in \mathcal{A}$, the sets $\{B \in \mathcal{A} : B \succsim A\}$ and $\{B \in \mathcal{A} : B \precsim A\}$ are closed.
2. (L-Continuity): There exist $p^*, p_* \in \Delta(Z)$ and $M > 0$ such that for every $A, B \in \mathcal{A}$ and $\alpha \in (0, 1)$ with $d_h(A, B) < \alpha/M$,

$$(1 - \alpha)A + \alpha\{p^*\} \succ (1 - \alpha)B + \alpha\{p_*\}.$$

The weak order axiom is entirely standard, as is the first part of the strong continuity axiom. The added assumption of L-continuity is used to obtain Lipschitz continuity of our representation in much the same way that the continuity axiom is used to obtain continuity.⁴ To interpret L-continuity, first note that $\{p^*\} \succ \{p_*\}$.⁵ For any $A, B \in \mathcal{A}$, continuity implies that there exists $\alpha \in (0, 1)$ such that $(1 - \alpha)A + \alpha\{p^*\} \succ (1 - \alpha)B + \alpha\{p_*\}$. L-continuity implies that such a preference holds for any $\alpha > Md_h(A, B)$, so as A and B get closer, the minimum required weight on p^* and p_* converges to 0 at a smooth rate. The constant M can be thought of as the sensitivity of this minimum α to the distance between A and B .

The next axiom is introduced in DLR (2001).

Axiom 3 (Indifference to Randomization (IR)): For every $A \in \mathcal{A}$, $A \sim co(A)$.

IR is justified if the DM choosing from the menu A can also randomly select an alternative from the menu, for example, by flipping a coin. In that case, the menus A and $co(A)$ offer the same set of options, and hence they are identical from the perspective of the DM.

The next axiom captures an important aspect of our model of costly contemplation.

Axiom 4 (Aversion to Contingent Planning (ACP)): If $A \succsim B$ and $\alpha \in [0, 1]$, then

$$A \succsim \alpha A + (1 - \alpha)B.$$

⁴Similar L-continuity axioms are used in DLRS (2005) and Sarver (2005). Although not immediately obvious, there is also some connection between our L-continuity axiom and the properness condition proposed by Mas-Colell (1986).

⁵Let $\alpha = 1/2$. Applying L-continuity with $A = B = \{p^*\}$ implies $\{p^*\} \succ \{\frac{1}{2}p^* + \frac{1}{2}p_*\}$, and applying L-continuity with $A = B = \{p_*\}$ implies $\{\frac{1}{2}p^* + \frac{1}{2}p_*\} \succ \{p_*\}$.

To interpret ACP, suppose we were to extend the DM's preferences to lotteries over menus. Let $\alpha \circ A \oplus (1 - \alpha) \circ B$ denote the lottery that yields the menu A with probability α and the menu B with probability $1 - \alpha$. The situation in which the DM makes a contingent choice, p if A and q if B , prior to the resolution of this lottery is equivalent to choosing the alternative $\alpha p + (1 - \alpha)q \in \alpha A + (1 - \alpha)B$.⁶ As discussed in the introduction, if contemplation is costly for the DM, then she will prefer that this lottery is resolved prior to her choosing an alternative. Hence,

$$\alpha \circ A \oplus (1 - \alpha) \circ B \succsim \alpha A + (1 - \alpha)B. \quad (5)$$

If in addition this extended preference satisfies stochastic dominance in the sense that $A \succsim B$ implies $A \succsim \alpha \circ A \oplus (1 - \alpha) \circ B$, then Equation (5) implies ACP.

Suppose the DM is asked to make a contingent plan, and she is told that she will be choosing from the menu A with probability α and from the menu $\{p\}$ with probability $1 - \alpha$. We refer to a choice from the singleton menu $\{p\}$ as a degenerate decision. When faced with a degenerate decision, there is no benefit to the DM from contemplating. Therefore, if the probability α that her contingent choice from A will be implemented decreases, then her benefit from contemplation decreases. Hence we should expect the DM to choose a less costly level of contemplation as α decreases. However, if α is fixed, then replacing the degenerate decision $\{p\}$ with another degenerate decision $\{q\}$ should not affect the DM's optimal level of contemplation. The following axiom states that if α is fixed, then there is a type of substitutability of degenerate decisions.

Axiom 5 (Independence of Degenerate Decisions (IDD)): *For any $A, B \in \mathcal{A}$, $\alpha \in [0, 1]$, and $p, q \in \Delta(Z)$,*

$$\alpha A + (1 - \alpha)\{p\} \succsim \alpha B + (1 - \alpha)\{p\} \implies \alpha A + (1 - \alpha)\{q\} \succsim \alpha B + (1 - \alpha)\{q\}.$$

Suppose the DM prefers making a contingent plan from $\alpha A + (1 - \alpha)\{p\}$ to making one from $\alpha B + (1 - \alpha)\{p\}$. As argued above, substituting q for p in both cases will not affect DM's optimal level of contemplation, and hence she will make the same contingent choices from A and B , respectively. IDD states that substituting the degenerate decision $\{q\}$ in the place of $\{p\}$ in both cases will not affect her preferences over these contingent plans. IDD allows for the possibility that the DM contemplates to obtain information about her ex-post utility, but it rules out the possibility that the DM changes her belief by becoming more optimistic about the utility she will obtain from a given lottery.

⁶We see here that any contingent choice from A and B can be represented as an alternative in $\alpha A + (1 - \alpha)B$. A partial converse to this statement is also true and can be found in Appendix D.

Finally, we will also consider the monotonicity axiom of Kreps (1979) in conjunction with our other axioms to obtain a refinement of our representation.

Axiom 6 (Monotonicity): *If $A \subset B$, then $B \succsim A$.*

2.2 Representation

The elements of our representation are a state space S , a Borel measurable state-dependent utility function $U : \Delta(Z) \times S \rightarrow \mathbb{R}$, a set of finite signed Borel measures \mathcal{M} on S , and a cost function $c : \mathcal{M} \rightarrow \mathbb{R}$. Note that any $\mu \in \mathcal{M}$ is a signed measure, so it can take both positive and negative values. We consider a representation $V : \mathcal{A} \rightarrow \mathbb{R}$ defined by

$$V(A) = \max_{\mu \in \mathcal{M}} \left[\int_S \max_{p \in A} U(p, s) \mu(ds) - c(\mu) \right]. \quad (6)$$

We want to think of the different measures as representing different levels of information or contemplation. However, the representation defined in Equation (6) is too general to always fit into this interpretation. We therefore impose the following restriction on the measures in our representation:

Definition 1 Given (S, \mathcal{M}, U, c) , the set of measures \mathcal{M} is said to be *consistent* if for each $\mu, \mu' \in \mathcal{M}$ and $p \in \Delta(Z)$,

$$\int_S U(p, s) \mu(ds) = \int_S U(p, s) \mu'(ds).$$

It is easy to see that this restriction is necessary for our contemplation interpretation. Suppose μ and μ' represent different levels of contemplation, and suppose $p \in \Delta(Z)$. Since the individual has only one choice when faced with the singleton menu $\{p\}$, she cannot change her choice based on her information. Therefore, the only effect of the individual's contemplation decision on her utility is the effect it has on her contemplation cost c . Thus the first term in the brackets in Equation (6) must be the same for both μ and μ' , which implies the set of measures \mathcal{M} must be consistent.

We also restrict attention to representations for which each $U(\cdot, s)$ is an expected-utility function. That is, for each $s \in S$, there exists $u \in \mathbb{R}^Z$ such that $U(p, s) = u \cdot p$ for all $p \in \Delta(Z)$. Define the set of *normalized (non-constant) expected-utility functions* on $\Delta(Z)$ to be

$$\mathcal{U} = \left\{ u \in \mathbb{R}^Z : \sum_{z \in Z} u_z = 0, \sum_{z \in Z} u_z^2 = 1 \right\}.$$

Since each $U(\cdot, s)$ is assumed to be an affine function, for each s there exists $u \in \mathcal{U}$ and constants $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$ such that for all $p \in \Delta(Z)$,

$$U(p, s) = \alpha(u \cdot p) + \beta.$$

Therefore, by appropriate normalization of the measures and the cost function, we can take $S = \mathcal{U}$ and $U(p, u) = u \cdot p$. That is, it is without loss of generality to assume that the set of states is actually the set of expected-utility functions. While our representation theorems do not require such a restriction, our uniqueness results will be easier to understand if we write the representation in this canonical form. Intuitively, if we do not normalize our expected-utility functions in some way, then there will be an extra “degree of freedom” that must be accounted for in the uniqueness results. Given this normalization, the elements of our representation are simply a pair (\mathcal{M}, c) , and the representation in Equation (6) can be written as

$$V(A) = \max_{\mu \in \mathcal{M}} \left[\int_{\mathcal{U}} \max_{p \in A} (u \cdot p) \mu(du) - c(\mu) \right]. \quad (7)$$

Another important issue related to the assumption that $S = \mathcal{U}$ is whether or not we are imposing a state space on the representation. We argue that the critical restriction here is that each $U(\cdot, s)$ is an expected-utility function, not that $S = \mathcal{U}$. The set \mathcal{U} makes available to the individual all possible expected-utility preferences and therefore does not restrict the state space. It is true that in some cases the set \mathcal{U} may contain “more” states than are necessary to represent an individual’s preferences. However, the support of the measures in the representation will reveal which of the expected-utility functions in \mathcal{U} are necessary. Thus, instead of considering minimality of the state space, it is sufficient to consider minimality of the set of measures in the representation, which leads us to our next definition:⁷

Definition 2 Given a compact set of measures \mathcal{M} and a cost function c , suppose V defined as in Equation (7) represents \succsim . The set \mathcal{M} is said to be *minimal* if for any compact proper subset \mathcal{M}' of \mathcal{M} , the function V' obtained by replacing \mathcal{M} with \mathcal{M}' in Equation (7) no longer represents \succsim .

We are now ready to formally define our representation:

⁷Note that we endow the set of all finite signed Borel measures on \mathcal{U} with the weak* topology, that is, the topology where a net $\{\mu_d\}_{d \in D}$ converges to μ if and only if $\int_{\mathcal{U}} f \mu_d(du) \rightarrow \int_{\mathcal{U}} f \mu(du)$ for every continuous function $f : \mathcal{U} \rightarrow \mathbb{R}$.

Definition 3 A *Reduced Form Costly Contemplation (RFCC) representation* is a compact set of finite signed Borel measures \mathcal{M} and a lower semi-continuous function $c : \mathcal{M} \rightarrow \mathbb{R}$ such that

1. $V : \mathcal{A} \rightarrow \mathbb{R}$ defined by Equation (7) represents \succsim .
2. \mathcal{M} is both consistent and minimal.
3. There exist $p, q \in \Delta(Z)$ such that $V(\{p\}) > V(\{q\})$.

The first two requirements of this definition have been explained above. The third condition is simply a technical requirement relating to the strong continuity axiom. If we take p^* and p_* as in the definition of L-continuity, then $\{p^*\} \succ \{p_*\}$, which gives rise to this third condition.

3 Main Results

The following is our main representation theorem:

Theorem 1 *A. The preference \succsim has a RFCC representation if and only if it satisfies weak order, strong continuity, IR, ACP, and IDD.*

B. The preference \succsim has a RFCC representation in which each $\mu \in \mathcal{M}$ is positive if and only if it satisfies weak order, strong continuity, monotonicity, ACP, and IDD.

The following result shows that a RFCC representation is essentially unique:

Theorem 2 *If (\mathcal{M}, c) and (\mathcal{M}', c') are two RFCC representations for \succsim , then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $\mathcal{M}' = \alpha\mathcal{M}$ and $c'(\alpha\mu) = \alpha c(\mu) + \beta$ for all $\mu \in \mathcal{M}$.*

4 Conclusion

We now briefly discuss the relation of our results to the additive EU representation of DLR (2001). The crucial axiom in their additive representation is the counterpart of the standard independence axiom adapted to sets.

Axiom 7 (Independence): *For any $A, B, C \in \mathcal{A}$ and $\alpha \in (0, 1)$, if $A \succ B$ then $\alpha A + (1 - \alpha)C \succ \alpha B + (1 - \alpha)C$.*

Under weak order and continuity, independence implies a form of indifference to contingent planning which can be stated as follows.

Axiom 8 (Indifference to Contingent Planning (ICP)): *For any $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$, if $A \sim B$ then $A \sim \alpha A + (1 - \alpha)B$.*

Since RFCC preferences may show a strict aversion to contingent planning ($A \sim B$ and $A \succ \alpha A + (1 - \alpha)B$), they in general need not satisfy independence. It is also not very difficult to see that under weak order and continuity, independence implies ACP, IR, and IDD. Therefore the axioms of the additive EU representation (weak order, strong continuity, and independence) are strictly more powerful than the axioms in our paper.⁸ In particular the additive EU representation of DLR (2001) is a special case of our RFCC representation.⁹

Definition 4 An *additive EU representation* is a nonempty closed subset S of \mathcal{U} and a finite signed Borel measure μ with full support on S such that \succsim is represented by

$$V(A) = \int_S \max_{p \in A} (u \cdot p) \mu(du). \quad (8)$$

The representations are formally related as follows. An additive EU representation is an RFCC representation with $\mathcal{M} = \{\mu\}$. Conversely, an RFCC representation with $\mathcal{M} = \{\mu\}$ is an additive EU representation with $S = \text{supp}(\mu)$. We now state the corresponding representation and uniqueness results from DLR (2001).

Theorem 3 (DLR 2001, DLRS 2005) *A. A preference \succsim satisfies weak order, strong continuity, and independence if and only if it has an additive EU representation.*

B. A preference \succsim satisfies weak order, strong continuity, independence, and monotonicity if and only if it has an additive EU representation with a positive measure.

Theorem 4 (DLR 2001) *If \succsim has two additive EU representations (S, μ) and (S', μ') , then $S = S'$ and there exists $\alpha > 0$ such that $\mu = \alpha \mu'$.*

⁸The continuity axiom in the original DLR (2001) paper is our Axiom 2.1. As discussed in the correction of DLR (2001) by Dekel, Lipman, Rustichini, and Sarver (2005, henceforth DLRS), Axiom 2.1 needs to be strengthened for the original DLR (2001) representation result to be valid. The stronger form of continuity stated in DLRS (2005) has a different form than our Axiom 2, but the two axioms are easily seen to be equivalent under weak order and independence.

⁹However, in their definition of the additive EU representation, DLR (2001) do not impose our normalization that $S \subset \mathcal{U}$.

As in our model, in the DLR (2001) model the subsequent choice out of A does not have a straightforward interpretation with signed measures. Different potential interpretations are possible under additional assumptions on the preference. These include anticipation of temptation (Gul and Pesendorfer, 2001; Dekel, Lipman, and Rustichini, 2005) or ex-post regret (Sarver, 2005). Also, although a preference with a RFCC representation does not satisfy independence, it does satisfy IR. Hence, it falls within the class of non-additive representations considered in DLR (2001). We conjecture that the unique DLR (2001) subjective state space corresponding to the RFCC representation (\mathcal{M}, c) is $\overline{\cup_{\mu \in \mathcal{M}} \text{supp}(\mu)}$.

In the introduction we discussed how a costly contemplation representation (Ω, U, c) induces a RFCC representation (\mathcal{M}, c) . It is possible to show a converse to this: For every finite monotone RFCC representation (\mathcal{M}, c) , there is a costly contemplation representation (Ω, U, c) that induces (\mathcal{M}, c) .¹⁰ It is also possible to show a counterpart of this result without monotonicity if we add a state-dependent temptation term to the costly contemplation representation. However the subsequent choice out of A has an unsuccessful interpretation under the latter, since the same RFCC representation may be expressed in multiple ways corresponding to different second period choices out of A .

We conclude by making an observation about the infinite regress issue. The infinite regress problem of bounded rationality can be informally explained as follows (see, e.g., Conlisk, 1996): Consider an abstract decision problem D . The standard rational economic agent is typically assumed to solve the problem D optimally without any constraints, no matter how difficult the problem might be. One may be tempted to make the model more “realistic” by explicitly taking account of the costs of solving it. This leads to a new optimization problem $F(D)$, the problem that incorporates into D the costs of solving D . However, typically $F(D)$ itself is a more difficult problem than D . So if one would like to have an even more “realistic” model, why not include the cost of solving $F(D)$ explicitly? The latter leads to the new decision problem $F^2(D) = F(F(D))$. This argument can be iterated ad infinitum. The fact that it is not clear at which level $F^n(D)$ one should stop, and how to stop if one stops at any level, corresponds to the infinite regress problem.

The representation result in this paper may be seen as giving an *as if* solution to the infinite regress problem. To the extent that one finds ACP a convincing behavioral aspect of bounded rationality arising from contemplation costs, there is no loss of generality from restricting attention to $F^1(D)$, the case where the decision maker optimally solves the problem of optimal contemplation subject to costs.

¹⁰A formal proof of this claim will be available in a subsequent version of this paper.

Appendix

A Mathematical Preliminaries

In this section we establish some general mathematical results that will be used to prove our representation and uniqueness theorems. Our main results will center around a classic duality relationship from convex analysis. After presenting some intermediate results, we describe this duality in Section A.2.

Suppose X is a real Banach space. We now introduce the standard definition of the subdifferential of a function.

Definition 5 Suppose $C \subset X$ is convex and $f : C \rightarrow \mathbb{R}$. For $x \in C$, the *subdifferential* of f at x is defined to be

$$\partial f(x) = \{x^* \in X^* : \langle y - x, x^* \rangle \leq f(y) - f(x) \text{ for all } y \in C\}.$$

The subdifferential is important in the approximation of a convex function by affine functions. In fact, it is straightforward to show that $x^* \in \partial f(x)$ if and only if the affine function $h(y) \equiv f(x) + \langle y - x, x^* \rangle$ satisfies $h \leq f$ and $h(x) = f(x)$. It should also be noted that when X is infinite-dimensional it is possible to have $\partial f(x) = \emptyset$ for some $x \in C$, even if f is convex. However, the following results show that under certain continuity assumptions on f , the subdifferential is always nonempty. For a convex subset C of X , a function $f : C \rightarrow \mathbb{R}$ is said to be *Lipschitz continuous* if there is some real number K such that for every $x, y \in C$, $|f(x) - f(y)| \leq K\|x - y\|$. The number K is called a *Lipschitz constant* of f .

Lemma 1 *Suppose C is a convex subset of a Banach space X . If $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous and convex, then $\partial f(x) \neq \emptyset$ for all $x \in C$. In particular, if $K \geq 0$ is a Lipschitz constant of f , then for all $x \in C$ there exists $x^* \in \partial f(x)$ with $\|x^*\| \leq K$.*

Proof: We begin by introducing the standard definition of the epigraph of a function $f : C \rightarrow \mathbb{R}$:

$$\text{epi}(f) = \{(x, t) \in C \times \mathbb{R} : t \geq f(x)\}.$$

Note that $\text{epi}(f) \subset X \times \mathbb{R}$ is a convex set because f is convex with a convex domain C . Now, define

$$H = \{(x, t) \in X \times \mathbb{R} : t < -K\|x\|\}.$$

It is easily seen that H is nonempty and convex. Also, since $\|\cdot\|$ is necessarily continuous, H is open (in the product topology).

Let $x \in C$ be arbitrary. Let $H(x)$ be the translate of H so that its vertex is $(x, f(x))$; that is, $H(x) = (x, f(x)) + H$. We claim that $\text{epi}(f) \cap H(x) = \emptyset$. To see this, note first that

$$\begin{aligned} H(x) &= \{(x + y, f(x) + t) \in X \times \mathbb{R} : t < -K\|y\|\} \\ &= \{(y, t) \in X \times \mathbb{R} : t < f(x) - K\|y - x\|\}. \end{aligned}$$

Now, suppose $(y, t) \in \text{epi}(f)$, so that $t \geq f(y)$. By Lipschitz continuity, we have $f(y) \geq f(x) - K\|y - x\|$. Therefore, $t \geq f(x) - K\|y - x\|$, which implies $(y, t) \notin H(x)$.

Since $H(x)$ is open and nonempty, it has an interior point. We have also shown that $H(x)$ and $\text{epi}(f)$ are disjoint convex sets. Therefore, a version of the Separating Hyperplane Theorem implies there exists a nonzero continuous linear functional $(x^*, \lambda) \in X^* \times \mathbb{R}$ that separates $H(x)$ and $\text{epi}(f)$.¹¹ That is, there exists a scalar δ such that

$$\langle y, x^* \rangle + \lambda t \leq \delta \quad \text{if } (y, t) \in \text{epi}(f) \quad (9)$$

and

$$\langle y, x^* \rangle + \lambda t \geq \delta \quad \text{if } (y, t) \in H(x). \quad (10)$$

Clearly, we cannot have $\lambda > 0$. Also, if $\lambda = 0$, then Equation (10) implies $x^* = 0$. This would contradict (x^*, λ) being a nonzero functional. Therefore, $\lambda < 0$. Without loss of generality, we can take $\lambda = -1$, for otherwise we could renormalize (x^*, λ) by dividing by $|\lambda|$.

Since $(x, f(x)) \in \text{epi}(f)$, we have $\langle x, x^* \rangle - f(x) \leq \delta$. For all $t > 0$, we have $(x, f(x) - t) \in H(x)$, which implies $\langle x, x^* \rangle - f(x) + t \geq \delta$. Therefore, $\langle x, x^* \rangle - f(x) = \delta$, and thus for all $y \in C$,

$$\langle y, x^* \rangle - f(y) \leq \delta = \langle x, x^* \rangle - f(x).$$

Equivalently, we can write $f(y) - f(x) \geq \langle y - x, x^* \rangle$. Thus, $x^* \in \partial f(x)$.

It remains only to show that $\|x^*\| \leq K$. Suppose to the contrary. Then, there exists $y \in X$ such that $\langle y, x^* \rangle < -K\|y\|$, and hence there also exists $\varepsilon > 0$ such that $\langle y, x^* \rangle + \varepsilon < -K\|y\|$. Therefore,

$$\langle y + x, x^* \rangle - f(x) + K\|y\| + \varepsilon < \langle x, x^* \rangle - f(x) = \delta,$$

which, by Equation (10), implies $(y + x, f(x) - K\|y\| - \varepsilon) \notin H(x)$. However, this contradicts the definition of $H(x)$. Thus $\|x^*\| \leq K$. ■

The following simple lemma will also be useful.

Lemma 2 *Let $K \geq 0$ and let $\{x_d\}_{d \in D} \subset X$ and $\{x_d^*\}_{d \in D} \subset X^*$ be nets such that (i) $\|x_d^*\| \leq K$ for all $d \in D$, and (ii) $x_d \rightarrow x$ and $x_d^* \xrightarrow{w^*} x^*$ for some $x \in X$ and $x^* \in X^*$. Then, $\langle x_d, x_d^* \rangle \rightarrow \langle x, x^* \rangle$.*

Proof: We have

$$\begin{aligned} |\langle x_d, x_d^* \rangle - \langle x, x^* \rangle| &\leq |\langle x_d - x, x_d^* \rangle| + |\langle x, x_d^* - x^* \rangle| \\ &\leq \|x_d - x\| \|x_d^*\| + |\langle x, x_d^* - x^* \rangle| \\ &\leq \|x_d - x\| K + |\langle x, x_d^* - x^* \rangle| \rightarrow 0, \end{aligned}$$

¹¹See Aliprantis and Border (1999, Theorem 5.50) or Luenberger (1969, p133).

so that $\langle x_d, x_d^* \rangle \rightarrow \langle x, x^* \rangle$. ■

In the next Lemma, we assume that X is a Banach lattice.¹² Let $X_+ = \{x \in X : x \geq 0\}$ denote the *positive cone* of X . A function $f : C \rightarrow \mathbb{R}$ on a subset C of X is *monotone* if $f(x) \geq f(y)$ whenever $x, y \in C$ are such that $x \geq y$. A continuous linear functional $x^* \in X^*$ is *positive* if $\langle x, x^* \rangle \geq 0$ for all $x \in X_+$.

Lemma 3 *Suppose C is a convex subset of a Banach lattice X such that for any $x, x' \in C$, $x \vee x' \in C$. If $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous, convex, and monotone and if $K \geq 0$ is a Lipschitz constant of f , then for all $x \in C$ there exists a positive $x^* \in \partial f(x)$ with $\|x^*\| \leq K$.*

Proof: Let $\text{epi}(f)$, H , and $H(x)$ be as defined in the proof of Lemma 1. Remember that $\text{epi}(f)$ and $H(x)$ are non-empty and convex, $H(x)$ is open, and $\text{epi}(f) \cap H(x) = \emptyset$ for all $x \in C$. Define

$$I(x) = H(x) + X_+ \times \{0\}.$$

Then $I(x) \subset X \times \mathbb{R}$ is convex as the sum of two convex sets, and it has non-empty interior since it contains the nonempty open set $H(x)$.

Let $x \in C$ be arbitrary. We claim that $\text{epi}(f) \cap I(x) = \emptyset$. Suppose for a contradiction that $(x', s) \in \text{epi}(f) \cap I(x)$. Then $x' \in C$, and there exist $y \in X$, $z \in X_+$ such that $x' = x + y + z$, and $s - f(x) < -K\|y\|$. Let $\bar{x} = x \vee x' \in C$ and $\bar{y} = \bar{x} - x'$. Note that $|\bar{y}| = \bar{y} = (x - x')^+$ and $-y = x - x' + z \geq x - x'$, hence

$$|y| = |-y| \geq (-y)^+ \geq (x - x')^+ = |\bar{y}|.$$

Since X is a Banach lattice, the above inequality implies that $\|y\| \geq \|\bar{y}\|$. Monotonicity of f implies that $f(\bar{x}) \geq f(x)$. We therefore have $x' = \bar{x} + \bar{y}$ and $s - f(\bar{x}) \leq s - f(x) < -K\|y\| \leq -K\|\bar{y}\|$. Hence $(x', s) \in H(\bar{x})$, a contradiction to $\text{epi}(f) \cap H(\bar{x}) = \emptyset$.

We showed that $I(x)$ and $\text{epi}(f)$ are disjoint convex sets and $I(x)$ has nonempty interior. Therefore, the same version of the Separating Hyperplane Theorem used in the proof of Lemma 1 implies that there exists a nonzero continuous linear functional $(x^*, \lambda) \in X^* \times \mathbb{R}$ that separates $I(x)$ and $\text{epi}(f)$. That is, there exists a scalar δ such that

$$\langle y, x^* \rangle + \lambda t \leq \delta \quad \text{if } (y, t) \in \text{epi}(f) \tag{11}$$

and

$$\langle y, x^* \rangle + \lambda t \geq \delta \quad \text{if } (y, t) \in I(x). \tag{12}$$

Note that Equation (11) is the same as Equation (9), and Equation (12) implies Equation (10). Therefore by the exact same arguments as in the proof of Lemma 1, we can without loss of generality let $\lambda = -1$, and conclude that $\delta = \langle x, x^* \rangle - f(x)$, $x^* \in \partial f(x)$, and $\|x^*\| \leq K$.

¹²See Aliprantis and Border (1999, page 302) for a definition of Banach lattices.

It only remains to show that x^* is positive. Let $y \in X_+$. Then for any $\varepsilon > 0$, $(x + y, f(x) - \varepsilon) \in I(x)$. By equation (12)

$$\langle x + y, x^* \rangle - f(x) + \varepsilon \geq \delta = \langle x, x^* \rangle - f(x),$$

hence $\langle y, x^* \rangle \geq -\varepsilon$. Since the latter holds for all $\varepsilon > 0$ and $y \in X_+$, we have that $\langle y, x^* \rangle \geq 0$ for all $y \in X_+$. Therefore x^* is positive. \blacksquare

A.1 Variation of the Mazur Density Theorem

The Mazur density theorem is a classic result from convex analysis. It states that if X is a separable Banach space and $f : C \rightarrow \mathbb{R}$ is a continuous convex function defined on a convex open subset C of X , then the set of points x where $\partial f(x)$ is a singleton is a dense G_δ set in C .¹³ The notation G_δ indicates that a set is the countable intersection of open sets.

We wish to obtain a variation of this theorem by relaxing the assumption that C has a nonempty interior. However, it can be shown that the conclusion of the theorem does not hold for arbitrary convex sets. We will therefore require that the affine hull of C , defined below, is dense in X .

Definition 6 The *affine hull* of a set $C \subset X$, denoted $\text{aff}(C)$, is defined to be the smallest affine subspace of X that contains C .

That is, the affine hull of C is defined by $x + \text{span}(C - C)$ for any fixed $x \in C$. If C is convex, then it is straightforward to show that

$$\text{aff}(C) = \{\lambda x + (1 - \lambda)y : x, y \in C \text{ and } \lambda \in \mathbb{R}\}. \quad (13)$$

Intuitively, if we were to draw a line through any two points of C , then that entire line would necessarily be included in any affine subspace that contains C .

We are now ready to state our variation of Mazur's theorem. Essentially, we are able to relax the assumption that C has a nonempty interior and instead assume that $\text{aff}(C)$ is dense in X if we also replace the continuity assumption with the more restrictive assumption of Lipschitz continuity.

Proposition 1 *Suppose X is a separable Banach space and C is a closed and convex subset of X containing the origin, and suppose $\text{aff}(C)$ is dense in X . If $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous and convex, then the set of points x where $\partial f(x)$ is a singleton is a dense G_δ (in the relative topology) set in C .*

¹³See Phelps (1993, Theorem 1.20). An equivalent characterization in terms of closed convex sets and smooth points can be found in Holmes (1975, p171).

Proof: This proof is a variation of the proof of Mazur's theorem found in Phelps (1993). Since any subset of a separable Banach space is separable, $\text{aff}(C)$ is separable. Let $\{x_n\} \subset \text{aff}(C)$ be a sequence which is dense in $\text{aff}(C)$, and hence, by the density of $\text{aff}(C)$ in X , also dense in X . Let K be a Lipschitz constant of f . For each $m, n \in \mathbb{N}$, let $A_{m,n}$ denote the set of $x \in C$ for which there exist $x^*, y^* \in \partial f(x)$ such that $\|x^*\|, \|y^*\| \leq 2K$ and

$$\langle x_n, x^* - y^* \rangle \geq \frac{1}{m}.$$

We claim that if $\partial f(x)$ is not a singleton for $x \in C$, then $x \in A_{m,n}$ for some $m, n \in \mathbb{N}$. By Lemma 1, for all $x \in C$, $\partial f(x) \neq \emptyset$. Therefore, if $\partial f(x)$ is not a singleton, then there exist $x^*, y^* \in \partial f(x)$ such that $x^* \neq y^*$. This does not tell us anything about the norm of x^* or y^* , but by Lemma 1, there exists $z^* \in \partial f(x)$ such that $\|z^*\| \leq K$. Either $z^* \neq x^*$ or $z^* \neq y^*$, so it is without loss of generality that we assume the former. It is straightforward to verify that the subdifferential is convex. Therefore, for all $\lambda \in (0, 1)$, $\lambda x^* + (1 - \lambda)z^* \in \partial f(x)$, and

$$\|\lambda x^* + (1 - \lambda)z^*\| \leq \|z^*\| + \lambda\|x^* - z^*\| \leq 2K$$

for λ sufficiently small. For some such λ , let $w^* = \lambda x^* + (1 - \lambda)z^*$. Then, $w^* \neq z^*$ and $\|w^*\| \leq 2K$. Since $w^* \neq z^*$, there exists $y \in X$ such that $\langle y, w^* - z^* \rangle > 0$. By the continuity of $w^* - z^*$, there exists a neighborhood N of y such that for all $z \in N$, $\langle z, w^* - z^* \rangle > 0$. Since $\{x_n\}$ is dense in X , there exists $n \in \mathbb{N}$ such that $x_n \in N$. Thus $\langle x_n, w^* - z^* \rangle > 0$, and hence there exists $m \in \mathbb{N}$ such that $\langle x_n, w^* - z^* \rangle > \frac{1}{m}$. Therefore, $x \in A_{m,n}$.

We have just shown that the set of $x \in C$ for which $\partial f(x)$ is a singleton is $\bigcap_{m,n} (C \setminus A_{m,n})$. It remains only show that for each $m, n \in \mathbb{N}$, $C \setminus A_{m,n}$ is open (in the relative topology) and dense in C . Then, we can appeal to the Baire category theorem.

We first show that each $A_{m,n}$ is relatively closed. If $A_{m,n} = \emptyset$, then $A_{m,n}$ is obviously closed, so suppose otherwise. Consider any sequence $\{z_k\} \subset A_{m,n}$ such that $z_k \rightarrow z$ for some $z \in C$. We will show that $z \in A_{m,n}$. For each k , choose $x_k^*, y_k^* \in \partial f(z_k)$ such that $\|x_k^*\|, \|y_k^*\| \leq 2K$ and $\langle x_n, x_k^* - y_k^* \rangle \geq \frac{1}{m}$. Since X is separable, the closed unit ball of X^* endowed with the weak* topology is metrizable and compact, which implies any sequence in this ball has a weak*-convergent subsequence.¹⁴ Therefore, the closed ball of radius $2K$ around the origin of X^* has this same property. Thus, without loss of generality, we can assume there exist $x^*, y^* \in X^*$ with $\|x^*\|, \|y^*\| \leq 2K$ such that $x_k^* \xrightarrow{w^*} x^*$ and $y_k^* \xrightarrow{w^*} y^*$. Therefore, for any $y \in C$, we have

$$\langle y - z, x^* \rangle = \lim_{k \rightarrow \infty} \langle y - z_k, x_k^* \rangle \leq \lim_{k \rightarrow \infty} [f(y) - f(z_k)] = f(y) - f(z).$$

¹⁴ For metrizability, see Aliprantis and Border (1999, Theorem 6.34). Compactness follows from Alaoglu's theorem; see Aliprantis and Border (1999, Theorem 6.25). Note that compactness only guarantees that every net has a convergent subnet, but compactness and metrizability together imply that every sequence has a convergent subsequence.

The first equality follows from Lemma 2, the inequality from the definition of the subdifferential, and the last equality from the continuity of f . Therefore, $x^* \in \partial f(z)$. A similar argument shows $y^* \in \partial f(z)$. Finally, since

$$\langle x_n, x^* - y^* \rangle = \lim_{k \rightarrow \infty} \langle x_n, x_k^* - y_k^* \rangle \geq \frac{1}{m},$$

we have $z \in A_{m,n}$, and hence $A_{m,n}$ is relatively closed.

We now need to show that $C \setminus A_{m,n}$ is dense in C for each $m, n \in \mathbb{N}$. Consider arbitrary $m, n \in \mathbb{N}$ and $z \in C$. We will find a sequence $\{z_k\} \subset C \setminus A_{m,n}$ such that $z_k \rightarrow z$. Since C contains the origin, $\text{aff}(C)$ is a subspace of X . Hence, $z + x_n \in \text{aff}(C)$, so Equation (13) implies there exist $x, y \in C$ and $\lambda \in \mathbb{R}$ such that $\lambda x + (1 - \lambda)y = z + x_n$. Let us first suppose $\lambda > 1$; we will consider the other cases shortly. Note that $\lambda > 1$ implies $0 < \frac{\lambda-1}{\lambda} < 1$. Consider any sequence $\{a_k\} \subset (0, \frac{\lambda-1}{\lambda})$ such that $a_k \rightarrow 0$. Define a sequence $\{y_k\} \subset C$ by $y_k = a_k y + (1 - a_k)z$, and note that $y_k \rightarrow z$. We claim that for each $k \in \mathbb{N}$, $y_k + \frac{a_k}{\lambda-1}x_n \in C$. To see this, note the following:

$$\begin{aligned} y_k + \frac{a_k}{\lambda-1}x_n &= a_k y + (1 - a_k)z + \frac{a_k}{\lambda-1}(x_n + z - z) \\ &= a_k y + (1 - a_k)z + \frac{a_k}{\lambda-1}(\lambda x + (1 - \lambda)y - z) \\ &= (1 - a_k)z + \frac{a_k \lambda}{\lambda-1}x - \frac{a_k}{\lambda-1}z \\ &= \left(1 - \frac{a_k \lambda}{\lambda-1}\right)z + \frac{a_k \lambda}{\lambda-1}x \end{aligned}$$

Since $0 < a_k < \frac{\lambda-1}{\lambda}$, we have $0 < \frac{a_k \lambda}{\lambda-1} < 1$. Thus $y_k + \frac{a_k}{\lambda-1}x_n$ is a convex combination of z and x , so it is an element of C .

Consider any $k \in \mathbb{N}$. Because C is convex, we have $y_k + tx_n \in C$ for all $t \in (0, \frac{a_k \lambda}{\lambda-1})$. Define a function $g : (0, \frac{a_k \lambda}{\lambda-1}) \rightarrow \mathbb{R}$ by $g(t) = f(y_k + tx_n)$, and note that g is convex. It is a standard result that a convex function on an open interval in \mathbb{R} is differentiable for all but (at most) countably many points of this interval.¹⁵ Let t_k be any $t \in (0, \frac{a_k \lambda}{\lambda-1})$ at which $g'(t)$ exists, and let $z_k = y_k + t_k x_n$. If $x^* \in \partial f(z_k)$, then it is straightforward to verify that the linear mapping $t \mapsto t \langle x_n, x^* \rangle$ is a subdifferential to g at t_k . Since g is differentiable at t_k , it can only have one element in its subdifferential at that point. Therefore, for any $x^*, y^* \in \partial f(z_k)$, we have $\langle x_n, x^* \rangle = \langle x_n, y^* \rangle$, and hence $z_k \in C \setminus A_{m,n}$. Finally, note that since $0 < t_k < \frac{a_k \lambda}{\lambda-1}$ and $a_k \rightarrow 0$, we have $t_k \rightarrow 0$. Therefore, $z_k = y_k + t_k x_n \rightarrow z$.

We did restrict attention above the case of $\lambda > 1$. However, if $\lambda < 0$, then let $\lambda' = 1 - \lambda > 1$, $x' = y$, $y' = x$, and the analysis is the same as above. If $\lambda \in [0, 1]$, then note that $z + x_n \in C$. Similar to in the preceding paragraph, for any $k \in \mathbb{N}$, define a function $g : (0, \frac{1}{k}) \rightarrow \mathbb{R}$ by $g(t) = f(z + tx_n)$. Let t_k be any $t \in (0, \frac{1}{k})$ at which $g'(t)$ exists, and let $z_k = z + t_k x_n$. Then, as argued above, $z_k \in C \setminus A_{m,n}$ for all $k \in \mathbb{N}$ and $z_k \rightarrow z$.

We have now proved that for each $m, n \in \mathbb{N}$, $C \setminus A_{m,n}$ is open (in the relative topology) and dense in C . Since C is a closed subset of a Banach space, it is a Baire space, which

¹⁵See Phelps (1993, Theorem 1.16).

implies every countable intersection of (relatively) open dense subsets of C is also dense.¹⁶ This completes the proof. ■

A.2 Fenchel-Moreau Duality

Let X continue to denote a real Banach space. We now introduce the definition of the conjugate of a function.

Definition 7 Suppose $C \subset X$ is convex and $f : C \rightarrow \mathbb{R}$. The *conjugate* (or *Fenchel conjugate*) of f is the function $f^* : X^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in C} [\langle x, x^* \rangle - f(x)].$$

There is an important duality between f and f^* :¹⁷

Lemma 4 Suppose $C \subset X$ is convex and $f : C \rightarrow \mathbb{R}$. Then,

1. f^* is lower semicontinuous in the weak* topology.
2. $f(x) \geq \langle x, x^* \rangle - f^*(x^*)$ for all $x \in C$ and $x^* \in X^*$.
3. $f(x) = \langle x, x^* \rangle - f^*(x^*)$ if and only if $x^* \in \partial f(x)$.

Proof: (1): For any $x \in C$, the mapping $x^* \mapsto \langle x, x^* \rangle - f(x)$ is continuous in the weak* topology. Therefore, for all $\alpha \in \mathbb{R}$, $\{x^* \in X^* : \langle x, x^* \rangle - f(x) \leq \alpha\}$ is weak* closed. Hence,

$$\{x^* \in X^* : f^*(x^*) \leq \alpha\} = \bigcap_{x \in C} \{x^* \in X^* : \langle x, x^* \rangle - f(x) \leq \alpha\}$$

is closed for all $\alpha \in \mathbb{R}$. Thus f^* is lower semicontinuous.

(2): For any $x \in C$ and $x^* \in X^*$, we have

$$f^*(x^*) = \sup_{x' \in C} [\langle x', x^* \rangle - f(x')] \geq \langle x, x^* \rangle - f(x),$$

and therefore $f(x) \geq \langle x, x^* \rangle - f^*(x^*)$.

(3): By the definition of the subdifferential, $x^* \in \partial f(x)$ if and only if

$$\langle y, x^* \rangle - f(y) \leq \langle x, x^* \rangle - f(x). \tag{14}$$

¹⁶See Theorems 3.34 and 3.35 of Aliprantis and Border (1999).

¹⁷For more on this relationship, see Ekeland and Turnbull (1983) or Holmes (1975). A finite-dimensional treatment can be found in Rockafellar (1970).

for all $y \in C$. By the definition of the conjugate, Equation (14) holds if and only if $f^*(x^*) = \langle x, x^* \rangle - f(x)$, which is equivalent to $f(x) = \langle x, x^* \rangle - f^*(x^*)$. \blacksquare

For the remainder of this section, assume that $C \subset X$ is convex and $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous and convex. Then, Lemma 1 implies that $\partial f(x) \neq \emptyset$ for all $x \in C$. Therefore, by parts 2 and 3 of Lemma 4, we have

$$f(x) = \max_{x^* \in X^*} [\langle x, x^* \rangle - f^*(x^*)] \quad (15)$$

for all $x \in C$. We have just proved a slight variation of the classic Fenchel-Moreau theorem.¹⁸

We now show that under the assumptions of Proposition 1, there is a minimal compact subset of X^* for which Equation (15) holds. Let C_f denote the set of all $x \in C$ for which the subdifferential of f at x is a singleton:

$$C_f = \{x \in C : \partial f(x) \text{ is a singleton}\}. \quad (16)$$

Let \mathcal{N}_f denote the set of functionals contained in the subdifferential of f at some $x \in C_f$:

$$\mathcal{N}_f = \{x^* \in X^* : x^* \in \partial f(x), x \in C_f\}. \quad (17)$$

Finally, let \mathcal{M}_f denote the closure of \mathcal{N}_f in the weak* topology:

$$\mathcal{M}_f = \overline{\mathcal{N}_f}. \quad (18)$$

Proposition 2 *Suppose X , C , and f satisfy the assumptions of Proposition 1. That is, suppose (i) X is a separable Banach space, (ii) C is a closed and convex subset of X containing the origin such that $\text{aff}(C)$ is dense in X , and (iii) $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous and convex. Then, \mathcal{M}_f is weak* compact, and for any weak* compact $\mathcal{M} \subset X^*$,*

$$\mathcal{M}_f \subset \mathcal{M} \iff f(x) = \max_{x^* \in \mathcal{M}} [\langle x, x^* \rangle - f^*(x^*)] \quad \forall x \in C.$$

Proof: If $K \geq 0$ is a Lipschitz constant of f , then Lemma 1 implies that for all $x \in C$ there exists $x^* \in \partial f(x)$ with $\|x^*\| \leq K$. Therefore, if $\partial f(x) = \{x^*\}$, then $\|x^*\| \leq K$. Thus, we have $\|x^*\| \leq K$ for all $x^* \in \mathcal{N}_f$, and hence also for all $x^* \in \mathcal{M}_f$. Since \mathcal{M}_f is a weak* closed and norm bounded set in X^* , it is weak* compact by Alaoglu's Theorem (see Aliprantis and Border, 1999, Theorem 6.25).

(\implies): Let $x \in C$ be arbitrary. By Proposition 1, C_f is dense in C , so there exists a net $\{x_d\}_{d \in D} \subset C_f$ such that $x_d \rightarrow x$. For all $d \in D$, take $x_d^* \in \partial f(x_d)$, and we have $\{x_d^*\}_{d \in D} \subset \mathcal{M}_f$ by the definition of \mathcal{M}_f . Since \mathcal{M}_f is weak* compact, every net in \mathcal{M}_f has a convergent

¹⁸The standard version of this theorem states that if $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous and convex, then $f(x) = f^{**}(x) \equiv \sup_{x^* \in X^*} [\langle x, x^* \rangle - f^*(x^*)]$. See, e.g., Proposition 1 in Ekeland and Turnbull (1983, p97).

subnet. Without loss of generality, suppose the net itself converges, so that $x_d^* \xrightarrow{w^*} x^*$ for some $x^* \in \mathcal{M}_f$. By Lemma 2, the definition of the subdifferential, and the continuity of f , for any $y \in C$,

$$\langle y - x, x^* \rangle = \lim_d \langle y - x_d, x_d^* \rangle \leq \lim_d [f(y) - f(x_d)] = f(y) - f(x),$$

which implies $x^* \in \partial f(x)$. Since $x \in C$ was arbitrary, we conclude that for all $x \in C$, there exists $x^* \in \mathcal{M}_f \subset \mathcal{M}$ such that $x^* \in \partial f(x)$. Then, by parts 2 and 3 of Lemma 4, we conclude that for all $x \in C$,

$$f(x) = \max_{x^* \in \mathcal{M}} [\langle x, x^* \rangle - f^*(x^*)].$$

(\Leftarrow): First note that the maximum taken over measures in \mathcal{M} is well-defined. The mapping $x^* \mapsto \langle x, x^* \rangle$ is weak* continuous, and f^* is weak* lower semicontinuous by part 1 of Lemma 4. Therefore, $x^* \mapsto \langle x, x^* \rangle - f^*(x^*)$ is weak* upper semicontinuous and hence attains a maximum on any weak* compact set.

Fix any $x \in C_f$. By the above, there exists $x^* \in \mathcal{M}$ such that $f(x) = \langle x, x^* \rangle - f^*(x^*)$, which implies $x^* \in \partial f(x)$ by part 3 of Lemma 4. However, $x \in C_f$ implies $\partial f(x) = \{x^*\}$, and hence $\partial f(x) \subset \mathcal{M}$. Since $x \in C_f$ was arbitrary, we have $\mathcal{N}_f \subset \mathcal{M}$. Because \mathcal{M} is weak* closed, we have $\mathcal{M}_f = \overline{\mathcal{N}_f} \subset \mathcal{M}$. \blacksquare

B Proof of Theorem 1

The necessity of the axioms in Theorem 1 is straightforward and left to the reader. For the sufficiency direction, let $\mathcal{A}^c \subset \mathcal{A}$ denote the collection of all convex menus. In both parts A and B of Theorem 1, \succsim satisfies IR. In part A, IR is directly assumed whereas in part B it is implied by weak order, continuity, monotonicity, and ACP (see Lemma 5). Therefore for all $A \in \mathcal{A}$, $A \sim co(A) \in \mathcal{A}^c$. Note that for any $u \in \mathcal{U}$, we have

$$\max_{p \in A} u \cdot p = \max_{p \in co(A)} u \cdot p.$$

Thus if we establish the representations in Theorem 1 for convex menus and then apply the same functional form to all of \mathcal{A} , then by IR the resulting function represents \succsim on \mathcal{A} . Note also that \mathcal{A} is a compact metric space since $\Delta(Z)$ is a compact metric space (see, e.g., Munkres (2000, p279)). It is a standard exercise to show that \mathcal{A}^c is a closed subset of \mathcal{A} , and hence \mathcal{A}^c is also compact.

We make some preliminary observations regarding our axioms in Section B.1. We then construct a function V with certain desirable properties in Section B.2. Finally, in Section B.3, we apply the duality results from Appendix A to complete the representation theorem.

B.1 Preliminary Observations

In this section we establish a number of simple implications of the axioms introduced in the text. These results will be useful in subsequent sections.

Lemma 5 *If \succsim satisfies weak order, ACP, monotonicity, and continuity, then it also satisfies IR.*

Proof: Let $A \in \mathcal{A}$. Monotonicity implies that $co(A) \succsim A$, hence we only need to prove that $A \succsim co(A)$. Let us inductively define a sequence of sets via $A_0 = A$ and $A_k = \frac{1}{2}A_{k-1} + \frac{1}{2}A_{k-1}$ for $k \geq 1$. ACP implies that $A_{k-1} \succsim A_k$, therefore by transitivity $A \succsim A_k$ for any k . It is straightforward to verify that $d_h(A_k, co(A)) \rightarrow 0$, so we have $A \succsim co(A)$ by continuity. ■

For proving our representation theorem, it will be useful to derive an alternative formulation of our IDD axiom. Before introducing this new axiom, we define the set of *translations* to be

$$\Theta \equiv \left\{ \theta \in \mathbb{R}^Z : \sum_{z \in Z} \theta_z = 0 \right\}. \quad (19)$$

Any $\theta \in \Theta$ can be thought of as a signed measure on Z such that $\theta(Z) = 0$. For $A \in \mathcal{A}$ and $\theta \in \Theta$, define $A + \theta \equiv \{p + \theta : p \in A\}$. Intuitively, adding θ to A in this sense simply “shifts” A . Also, note that for any $p, q \in \Delta(Z)$, we have $p - q \in \Theta$. We now give a formulation of IDD in terms of translations.

Axiom 9 (Translation Invariance (TI)): *For any $A, B \in \mathcal{A}$ and $\theta \in \Theta$ such that $A + \theta, B + \theta \in \mathcal{A}$,*

$$A \succsim B \implies A + \theta \succsim B + \theta.^{19}$$

Lemma 6 *\succsim satisfies IDD if and only if it satisfies TI.*

Proof: To see that TI implies IDD, assume that $A, B \in \mathcal{A}$, $p, q \in \Delta(Z)$ are such that $\lambda A + (1 - \lambda)\{q\} \succsim \lambda B + (1 - \lambda)\{q\}$. Let $A' = \lambda A + (1 - \lambda)\{q\}$, $B' = \lambda B + (1 - \lambda)\{q\}$ and $\theta = (1 - \lambda)(p - q)$. Note that $\theta \in \Theta$, $A' + \theta = \lambda A + (1 - \lambda)\{p\} \in \mathcal{A}$, and $B' + \theta = \lambda B + (1 - \lambda)\{p\} \in \mathcal{A}$. Hence by TI, $\lambda A + (1 - \lambda)\{p\} \succsim \lambda B + (1 - \lambda)\{p\}$.

To see that IDD implies TI, assume that $A, B \in \mathcal{A}$ and $\theta \in \Theta$ are such that $A + \theta, B + \theta \in \mathcal{A}$ and $A \succsim B$. If $\theta = 0$, the conclusion of TI holds trivially, so assume that $\theta \neq 0$. Let $Z^- = \{z \in Z : \theta_z < 0\}$. Define $\theta^+, \theta^- \in \mathbb{R}^Z$ by $\theta_z^+ = \max\{0, \theta_z\}$ and $\theta_z^- = \max\{0, -\theta_z\}$ for any $z \in Z$. Then let $\kappa \equiv \sum_{z \in Z} \theta_z^+ = \sum_{z \in Z} \theta_z^- > 0$.

We will first show that for any $r \in A \cup B$,

$$0 \leq r_z - \theta_z^- \leq 1 - \kappa \quad \text{for all } z \in Z. \quad (20)$$

¹⁹Note that TI implies its converse, for suppose $A + \theta \succsim B + \theta$. Then, by TI, $A = (A + \theta) + (-\theta) \succsim (B + \theta) + (-\theta) = B$.

Note that for any $z \in Z^-$, $r_z - \theta_z^- = r_z + \theta_z \geq 0$ since $r + \theta \in \Delta(Z)$. Note also that if $z \notin Z^-$ then $r_z - \theta_z^- = r_z \geq 0$ since $\theta_z^- = 0$. So for any $z \in Z$,

$$0 \leq r_z - \theta_z^- \leq \left(1 - \sum_{z' \in Z^- \setminus \{z\}} r_{z'}\right) - \theta_z^- \leq \left(1 - \sum_{z' \in Z^- \setminus \{z\}} \theta_{z'}^-\right) - \theta_z^- = 1 - \kappa,$$

establishing Equation (20). Therefore, since $\theta \neq 0$, we have $0 < \kappa \leq 1$. Then, $p \equiv \frac{1}{\kappa}\theta^+$, $q \equiv \frac{1}{\kappa}\theta^-$ are in $\Delta(Z)$, and $\theta = \kappa(p - q)$. There are two cases to consider:

First consider the case of $\kappa < 1$. Define subsets A' and B' of \mathbb{R}^Z by

$$A' \equiv \{r' \in \mathbb{R}^Z : r' = \frac{1}{1-\kappa}(r - \theta^-) \text{ for some } r \in A\},$$

$$B' \equiv \{r' \in \mathbb{R}^Z : r' = \frac{1}{1-\kappa}(r - \theta^-) \text{ for some } r \in B\}.$$

By Equation (20) and the definition of κ , we have that $A', B' \in \mathcal{A}$ and

$$(1 - \kappa)A' + \kappa\{q\} = A \succsim B = (1 - \kappa)B' + \kappa\{q\}. \quad (21)$$

Next consider the $\kappa = 1$ case. By Equation (20) we have $r = \theta^- = q$ for any $r \in A \cup B$. Therefore $A = B = \{q\}$, and hence Equation (21) holds for any choice of $A', B' \in \mathcal{A}$.

Since Equation (21) holds in each of the two cases above, we conclude by IDD that

$$A + \theta = (1 - \kappa)A' + \kappa\{p\} \succsim (1 - \kappa)B' + \kappa\{p\} = B + \theta.$$

Therefore TI is satisfied. ■

In light of Lemma 6, we will use IDD and TI interchangeably. We now present one useful consequence of translation invariance.

Lemma 7 *Suppose \succsim satisfies weak order, continuity, and TI. If $A \in \mathcal{A}$, $\theta \in \Theta$, and $\alpha \in (0, 1)$ are such that $A + \theta \in \mathcal{A}$, then*

$$A \succsim A + \theta \iff A \succsim A + \alpha\theta \iff A + \alpha\theta \succsim A + \theta. \quad (22)$$

Proof: We will make a simple induction argument. Suppose

$$A + \frac{m-1}{n}\theta \succsim A + \frac{m}{n}\theta$$

for some $m, n \in \mathbb{N}$ with $m < n$. Then adding $\frac{1}{n}\theta$ to each side of the above and applying TI yields

$$A + \frac{m}{n}\theta \succsim A + \frac{m+1}{n}\theta.$$

Now suppose that $A \succsim A + \frac{1}{n}\theta$. Then, using induction and the transitivity of \succsim , we obtain the following:

$$A \succsim A + \frac{1}{n}\theta \succsim \cdots \succsim A + (1 - \frac{1}{n})\theta \succsim A + \theta. \quad (23)$$

A similar line of reasoning shows that if $A \prec A + \frac{1}{n}\theta$, then we obtain the following:

$$A \prec A + \frac{1}{n}\theta \prec \cdots \prec A + (1 - \frac{1}{n})\theta \prec A + \theta. \quad (24)$$

In sum, Equations (23) and (24) imply that for any $m, n \in \mathbb{N}$, $1 \leq m < n$, we have

$$A \succsim A + \frac{1}{n}\theta \iff A \succsim A + \theta \iff A \succsim A + \frac{m}{n}\theta \iff A + \frac{m}{n}\theta \succsim A + \theta.$$

This establishes Equation (22) for $\alpha \in (0, 1) \cap \mathbb{Q}$. The continuity of \succsim implies that Equation (22) holds for all $\alpha \in (0, 1)$. \blacksquare

Although we do not assume that independence holds on \mathcal{A} , our other axioms imply that independence does hold for singleton menus.

Axiom 10 (Singleton Independence): *For all $p, q, r \in \Delta(Z)$ and $\lambda \in (0, 1)$,*

$$\{p\} \succsim \{q\} \iff \lambda\{p\} + (1 - \lambda)\{r\} \succsim \lambda\{q\} + (1 - \lambda)\{r\}.$$

Lemma 8 *If \succsim satisfies weak order, continuity, and TI, then it also satisfies singleton independence.*

Proof: Let $\theta = q - p$ and $\theta' = (1 - \lambda)(r - p)$. Then

$$\begin{aligned} \{p\} \succsim \{q\} = \{p\} + \theta &\iff \{p\} \succsim \{p\} + \lambda\theta = (1 - \lambda)\{p\} + \lambda\{q\} \\ &\iff \lambda\{p\} + (1 - \lambda)\{r\} \succsim \lambda\{q\} + (1 - \lambda)\{r\}, \end{aligned}$$

where the first equivalence follows from Lemma 7, and the second equivalence follows from TI, $\{p\} + \theta' = \lambda\{p\} + (1 - \lambda)\{r\}$, and $(1 - \lambda)\{p\} + \lambda\{q\} + \theta' = \lambda\{q\} + (1 - \lambda)\{r\}$. Therefore singleton independence is satisfied. \blacksquare

Before proceeding, we define the following important subset of \mathcal{A}^c :

$$\mathcal{A}^\circ \equiv \{A \in \mathcal{A}^c : \forall \theta \in \Theta \exists \alpha > 0 \text{ such that } A + \alpha\theta \in \mathcal{A}^c\}. \quad (25)$$

Thus \mathcal{A}° contains menus that can be translated at least a ‘‘little bit’’ in the direction of any vector in Θ . It is easily verified that \mathcal{A}° is convex. In addition, the following result gives an alternative characterization of \mathcal{A}° along with some other important properties.

Lemma 9 *The set \mathcal{A}° has the following properties:*

1. $\mathcal{A}^\circ = \{A \in \mathcal{A}^c : \exists \varepsilon > 0 \text{ such that } \forall p \in A, \forall z \in Z, p_z \geq \varepsilon\}$.
2. Suppose $p \in \Delta(Z)$ is such that $p_z > 0$ for all $z \in Z$. Then for any $A \in \mathcal{A}^c$ and $\lambda \in [0, 1)$, $\lambda A + (1 - \lambda)\{p\} \in \mathcal{A}^\circ$.
3. \mathcal{A}° is dense in \mathcal{A}^c .

Proof: (1): Let $\hat{\mathcal{A}}^\circ \equiv \{A \in \mathcal{A}^c : \exists \varepsilon > 0 \text{ such that } \forall p \in A, \forall z \in Z, p_z \geq \varepsilon\}$. To see that $\hat{\mathcal{A}}^\circ \subset \mathcal{A}^\circ$, take any $A \in \hat{\mathcal{A}}^\circ$ and $\theta \in \Theta$. Let $\varepsilon > 0$ be such that $p_z \geq \varepsilon$ for all $p \in A$ and $z \in Z$. Choose $\alpha > 0$ sufficiently small to ensure that $\alpha \cdot \max_{z \in Z} |\theta_z| \leq \varepsilon$. Then $p_z + \alpha \theta_z \geq p_z - \varepsilon \geq 0$ for all $p \in A$ and $z \in Z$, so $A + \alpha \theta \in \mathcal{A}^c$. Thus $A \in \mathcal{A}^\circ$.

To see that $\mathcal{A}^\circ \subset \hat{\mathcal{A}}^\circ$, take any $A \in \mathcal{A}^\circ$. Fix any $z \in Z$, and take any $\theta \in \Theta$ such that $\theta_z = -1$. Then let $\alpha_z > 0$ be such that $A + \alpha_z \theta \in \mathcal{A}^c$, so for any $p \in A$, $p_z + \alpha_z \theta_z = p_z - \alpha_z \geq 0$. We obtain such an $\alpha_z > 0$ for every $z \in Z$, so let $\varepsilon \equiv \min_{z \in Z} \alpha_z > 0$. Then for any $p \in A$ and $z \in Z$, $p_z \geq \alpha_z \geq \varepsilon$, so $A \in \hat{\mathcal{A}}^\circ$.

(2): Let $\varepsilon \equiv (1 - \lambda)(\min_{z \in Z} p_z) > 0$. Then for any $q \in A$ and $z \in Z$, $\lambda q_z + (1 - \lambda)p_z \geq \varepsilon$. Thus $\lambda A + (1 - \lambda)\{p\} \in \mathcal{A}^\circ$ by part 1.

(3): It is easily verified that for any $A \in \mathcal{A}^c$, $(1 - 1/n)A + (1/n)\{p\} \rightarrow A$ as $n \rightarrow \infty$. Hence \mathcal{A}° is dense in \mathcal{A}^c by part 2. \blacksquare

Take p^* and p_* from the L-Continuity axiom, and let $\theta^* \equiv p^* - p_*$. We will utilize θ^* a great deal in the construction of our representation, and the following is an important property of θ^* .

Lemma 10 Suppose \succsim satisfies weak order, strong continuity, and TI, and take $\theta^* = p^* - p_*$. Let $A, B \in \mathcal{A}^\circ$ and $\alpha, \beta \in \mathbb{R}$ be such that $A \sim B$ and $A + \alpha\theta^*, B + \beta\theta^* \in \mathcal{A}^c$. Then,

$$A + \alpha\theta^* \succsim B + \beta\theta^* \iff \alpha \geq \beta. \quad (26)$$

Proof: We will first show that for any $A \in \mathcal{A}^\circ$, there exists $\gamma > 0$ such that $A + \gamma\theta^* \in \mathcal{A}^c$ and $A + \gamma\theta^* \succ A$. To see this, fix any $A \in \mathcal{A}^\circ$. It follows from part 1 of Lemma 9 that there exist $A' \in \mathcal{A}^c$ and $\gamma \in (0, 1)$ such that $A = (1 - \gamma)A' + \gamma\{p_*\}$.²⁰ By L-continuity we have

$$A + \gamma\theta^* = (1 - \gamma)A' + \gamma\{p^*\} \succ (1 - \gamma)A' + \gamma\{p_*\} = A.$$

Therefore, for any $A \in \mathcal{A}^\circ$ and $\alpha > 0$ such that $A + \alpha\theta^* \in \mathcal{A}^c$, take $\gamma > 0$ such that $A + \gamma\theta^* \in \mathcal{A}^c$ and $A + \gamma\theta^* \succ A$. Applying Lemma 7 to A and $\theta = \max\{\gamma, \alpha\}\theta^*$, we have $A + \alpha\theta^* \succ A$. A similar argument shows that if $A \in \mathcal{A}^\circ$ and $\alpha < 0$ are such that $A + \alpha\theta^* \in \mathcal{A}^c$, then $A \succ A + \alpha\theta^*$.

Now, let $A, B \in \mathcal{A}^\circ$ and $\alpha, \beta \in \mathbb{R}$ be such that $A \sim B$ and $A + \alpha\theta^*, B + \beta\theta^* \in \mathcal{A}^c$. We prove the equivalence from Equation (26) by considering three cases:

²⁰Take $\varepsilon > 0$ as in Lemma 9, and let $\gamma \equiv \varepsilon$ and $A' \equiv \{q \in \mathbb{R}^Z : q = \frac{1}{1-\gamma}(p - \gamma p_*) \text{ for some } p \in A\}$. It follows that $A' \in \mathcal{A}^c$ and $A = (1 - \gamma)A' + \gamma\{p_*\}$.

If $\alpha = \beta$ then $\alpha\theta^* = \beta\theta^*$. Hence by TI, $A + \alpha\theta^* \sim B + \beta\theta^*$.

If $\alpha > \beta$, there are three sub-cases to consider. First consider $\alpha > \beta \geq 0$, which implies $0 < \alpha - \beta \leq \alpha$ and hence $A + (\alpha - \beta)\theta^* \in \mathcal{A}^c$. From the above arguments $A + (\alpha - \beta)\theta^* \succ A \sim B$, so by TI, $A + \alpha\theta^* = [A + (\alpha - \beta)\theta^*] + \beta\theta^* \succ B + \beta\theta^*$. Similarly, if $0 \geq \alpha > \beta$, then $\beta \leq \beta - \alpha < 0$ and hence $B + (\beta - \alpha)\theta^* \in \mathcal{A}^c$. From the above arguments $A \sim B \succ B + (\beta - \alpha)\theta^*$, which implies by TI that $A + \alpha\theta^* \succ [B + (\beta - \alpha)\theta^*] + \alpha\theta^* = B + \beta\theta^*$. Finally, $\alpha > 0 > \beta$ implies $A + \alpha\theta^* \succ A \sim B \succ B + \beta\theta^*$.

If $\beta > \alpha$, then by symmetric arguments $B + \beta\theta^* \succ A + \alpha\theta^*$. ■

B.2 Construction of V

Recall that for any metric space (X, d) , $f : X \rightarrow \mathbb{R}$ is *Lipschitz continuous* if there is some real number K such that for every $x, y \in X$, $|f(x) - f(y)| \leq Kd(x, y)$. The number K is called a *Lipschitz constant* of f . We will construct a function $V : \mathcal{A}^c \rightarrow \mathbb{R}$ that represents \succsim on \mathcal{A}^c and has certain desirable properties. We next define the notion of translation-linearity in order to present the main result of this section. Recall that the set of translations, denoted by Θ , is defined in Equation (19).

Definition 8 Suppose that $V : \mathcal{A}^c \rightarrow \mathbb{R}$. Then V is *translation-linear* if there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}^c$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}^c$, we have $V(A + \theta) = V(A) + v \cdot \theta$.

Proposition 3 *If the preference \succsim satisfies weak order, strong continuity, ACP, and IDD, then there exists a function $V : \mathcal{A}^c \rightarrow \mathbb{R}$ with the following properties:*

1. For any $A, B \in \mathcal{A}^c$, $A \succsim B \iff V(A) \geq V(B)$.
2. V is Lipschitz continuous, convex, and translation-linear.
3. There exist $p, q \in \Delta(Z)$ such that $V(\{p\}) > V(\{q\})$.

Moreover, if V and V' are two functions that satisfy 1–3, then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V + \beta$.

The remainder of this section is devoted to the proof of Proposition 3. First note that by taking the p^* and p_* from the L-Continuity axiom, it follows that $\{p^*\} \succ \{p_*\}$. Thus part 3 of this proposition follows from part 1.

Let $\mathcal{S} \equiv \{\{p\} : p \in \Delta(Z)\}$ be the set all of singleton sets in \mathcal{A}^c . Given the assumptions of Proposition 3 and the results of Lemma 8, \succsim satisfies the von Neumann-Morgenstern axioms on \mathcal{S} . Therefore, there exists $v \in \mathbb{R}^Z$ such that for all $p, q \in \Delta(Z)$, $\{p\} \succsim \{q\}$ if and only if $v \cdot p \geq v \cdot q$. We will abuse notation and also treat v as a function $v : \mathcal{S} \rightarrow \mathbb{R}$ naturally defined by $v(\{p\}) = v \cdot p$. Note that v is translation-linear since $v(\{p\} + \theta) = v(\{p\}) + v \cdot \theta$ whenever $p \in \Delta(Z)$, $\theta \in \Theta$, and $p + \theta \in \Delta(Z)$.

We want to extend v to a function V on \mathcal{A}^c that represents \succsim and is translation-linear. The outline of the construction of the desired extension is the following: We first restrict attention to menus in \mathcal{A}° , as defined in Equation (25). This restriction will allow us to make extensive use of the translation invariance (TI) property defined in the previous section. We will construct a sequence of subsets of \mathcal{A}° , starting with $\mathcal{A}^\circ \cap \mathcal{S}$, such that each set is contained in its successor set. We will then extend v sequentially to each of these domains, while still representing \succsim and preserving translation-linearity (with respect to the vector v). The domain will grow to eventually contain all of the sets in \mathcal{A}° , and we show how to extend to all of \mathcal{A}^c by continuity. We prove that the resulting function is translation-linear, Lipschitz continuous, and convex.

As above, take p^* and p_* from the L-Continuity axiom, and let $\theta^* \equiv p^* - p_*$. Define a sequence $\mathcal{A}_0, \mathcal{A}'_0, \mathcal{A}_1, \mathcal{A}'_1, \dots$ of subsets of \mathcal{A}° inductively as follows: Let $\mathcal{A}_0 \equiv \mathcal{A}^\circ \cap \mathcal{S}$. By part 1 of Lemma 9, we have that $\mathcal{A}_0 = \{\{p\} : p \in \Delta(Z) \text{ and } \forall z \in Z, p_z > 0\}$. Define \mathcal{A}'_i for all $i \geq 0$ by

$$\mathcal{A}'_i \equiv \{A \in \mathcal{A}^\circ : A \sim B \text{ for some } B \in \mathcal{A}_i\},$$

and define \mathcal{A}_i for all $i \geq 1$ by

$$\mathcal{A}_i \equiv \{A \in \mathcal{A}^\circ : A = B + \alpha\theta^* \text{ for some } \alpha \in \mathbb{R}, B \in \mathcal{A}'_{i-1}\}.$$

Intuitively, we first extend \mathcal{A}_0 by including all $A \in \mathcal{A}^\circ$ that are viewed with indifference to some $B \in \mathcal{A}_0$. Then we extend to all translations by multiples of θ^* . We repeat the process, alternating between extension by indifference and extension by translation. Note that $\mathcal{A}_0 \subset \mathcal{A}'_0 \subset \mathcal{A}_1 \subset \mathcal{A}'_1 \subset \dots$.

We also define a sequence of functions, $V_0, V'_0, V_1, V'_1, \dots$, from these domains. That is, for all $i \geq 0$, $V_i : \mathcal{A}_i \rightarrow \mathbb{R}$ and $V'_i : \mathcal{A}'_i \rightarrow \mathbb{R}$. Define these functions recursively as follows:

1. Let $V_0 \equiv v|_{\mathcal{A}_0}$.
2. For $i \geq 0$, if $A \in \mathcal{A}'_i$, then $A \sim B$ for some $B \in \mathcal{A}_i$, so define V'_i by $V'_i(A) \equiv V_i(B)$.
3. For $i \geq 1$, if $A \in \mathcal{A}_i$, then $A = B + \alpha\theta^*$ for some $\alpha \in \mathbb{R}$ and $B \in \mathcal{A}'_{i-1}$, so define V_i by $V_i(A) \equiv V'_{i-1}(B) + \alpha(v \cdot \theta^*)$.

In a series of lemmas, we will show that these are well-defined functions which represent \succsim on their domains and are translation-linear.

First, we present some important properties of \mathcal{A}_i and \mathcal{A}'_i that will be used to prove Lemmas 12 and 13.

Lemma 11 *For any $i \geq 0$:*

1. *If $A \in \mathcal{A}_i$ and $\theta \in \Theta$, then there exists $\bar{\alpha} > 0$ such that:*

$$A + \alpha\theta \in \mathcal{A}_i, \quad \forall \alpha \in [0, \bar{\alpha}].^{21} \tag{27}$$

²¹As the proof of this lemma will illustrate, the same property holds for \mathcal{A}'_i .

2. For all $A, B \in \mathcal{A}'_i$ and $C \in \mathcal{A}^\circ$, $A \succsim C \succsim B$ implies $C \in \mathcal{A}'_i$.

Proof: (1): First, it follows immediately from part 1 of Lemma 9 that for any $A \in \mathcal{A}^\circ$ and $\theta \in \Theta$, there exists $\bar{\alpha} > 0$ such that

$$A + \alpha\theta \in \mathcal{A}^\circ, \quad \forall \alpha \in [0, \bar{\alpha}]. \quad (28)$$

We now prove by induction. To verify the property on $\mathcal{A}_0 = \mathcal{A}^\circ \cap \mathcal{S}$, take any $A \in \mathcal{A}_0$ and recall that $A = \{p\}$ for some $p \in \Delta(Z)$. Then take $\bar{\alpha} > 0$ such that Equation (28) holds. Then for all $\alpha \in [0, \bar{\alpha}]$, since $p + \alpha\theta \in \Delta(Z)$, we have $A + \alpha\theta \in \mathcal{A}_0$.

We now prove if the property holds for \mathcal{A}_i , then it also holds for \mathcal{A}_{i+1} . Take any $A \in \mathcal{A}_{i+1}$ and $\theta \in \Theta$. Then $A \sim B$ for some $B \in \mathcal{A}'_i$, and hence $B = C + \beta\theta^*$ for some $C \in \mathcal{A}_i$ and $\beta \in \mathbb{R}$. Choose $\bar{\alpha} > 0$ to be the minimum of that required to satisfy Equation (28) for A and B and to satisfy Equation (27) for C . Fix any $\alpha \in [0, \bar{\alpha}]$. Then $C + \alpha\theta \in \mathcal{A}_i$, and hence $C + \alpha\theta + \beta\theta^* = B + \alpha\theta \in \mathcal{A}'_i$. By TI, $A \sim B$ implies $A + \alpha\theta \sim B + \alpha\theta$, which implies $A + \alpha\theta \in \mathcal{A}_{i+1}$.

(2): We again prove by induction. To prove the result for \mathcal{A}'_0 , suppose $A, B \in \mathcal{A}'_0$ and $A \succsim C \succsim B$ for some $C \in \mathcal{A}^\circ$. Since $A, B \in \mathcal{A}'_0$, there exist $\{p\}, \{q\} \in \mathcal{A}_0$ such that $\{p\} \sim A \succsim C \succsim B \sim \{q\}$. Continuity implies there exists a $\lambda \in [0, 1]$ such that $\{\lambda p + (1 - \lambda)q\} \sim C$. By the convexity of $\mathcal{A}_0 = \mathcal{A}^\circ \cap \mathcal{S}$ and the definition of \mathcal{A}'_0 , this implies that $C \in \mathcal{A}'_0$.

We now show that if \mathcal{A}'_i satisfies the desired condition, then \mathcal{A}'_{i+1} does also. Suppose $A, B \in \mathcal{A}'_{i+1}$ and $A \succsim C \succsim B$ for some $C \in \mathcal{A}^\circ$. If there exist $A', B' \in \mathcal{A}'_i$ such that $A' \succsim C \succsim B'$, then $C \in \mathcal{A}'_i \subset \mathcal{A}'_{i+1}$ by the induction assumption. Thus without loss of generality, suppose $C \succ A'$ for all $A' \in \mathcal{A}'_i$. Since $A \in \mathcal{A}'_{i+1}$, there exists a $A' \in \mathcal{A}_{i+1}$ such that $A' \sim A \succsim C$. Since $A' \in \mathcal{A}_{i+1}$, there exists a $A'' \in \mathcal{A}'_i$ and $\alpha \in \mathbb{R}$ such that $A' = A'' + \alpha\theta^*$. Since $A'' \in \mathcal{A}'_i$ implies $C \succ A''$, this implies $A'' + \alpha\theta^* \succ C \succ A''$, and therefore $\alpha > 0$ by Lemma 10. By continuity, there exists a $\alpha' \in [0, \alpha]$ such that $A'' + \alpha'\theta^* \sim C$. But $A'' + \alpha'\theta^* \in \mathcal{A}_{i+1}$, so it must be that $C \in \mathcal{A}'_{i+1}$. ■

The following lemmas allow us to prove the desired properties of each V_i and V'_i by induction.

Lemma 12 For all $i \geq 0$, if V_i is well-defined, translation-linear, and represents \succsim on \mathcal{A}_i , then V'_i is also well-defined, translation-linear, and represents \succsim on \mathcal{A}'_i .

Proof: (Well-defined): Suppose $A \in \mathcal{A}'_i$ and $B, B' \in \mathcal{A}_i$ are such that $A \sim B$ and $A \sim B'$. Since V_i represents \succsim on \mathcal{A}_i and \succsim is transitive, $V_i(B) = V_i(B')$, and hence $V'_i(A)$ is uniquely defined.

(Represents \succsim): If $A, A' \in \mathcal{A}'_i$ then there exist $B, B' \in \mathcal{A}_i$ such that $A \sim B$ and $A' \sim B'$. Therefore, $V'_i(A) = V_i(B) \geq V_i(B') = V'_i(A')$ if and only if $B \succsim B'$ if and only if $A \succsim A'$, so V'_i represents \succsim on \mathcal{A}'_i .

(*Translation-linear*): Throughout we will use the fact if $\theta \in \Theta$ and $A, A + \theta \in \mathcal{A}'_i$, then $A + \alpha\theta \in \mathcal{A}'_i$ for all $\alpha \in [0, 1]$. This follows by part 2 of Lemma 11 because by Lemma 7, either $A + \theta \succsim A + \alpha\theta \succsim A$ or $A \succsim A + \alpha\theta \succsim A + \theta$.

We first show that V'_i satisfies the following local version of translation-linearity: For all $A \in \mathcal{A}'_i$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}'_i$, there exist $\bar{\alpha} > 0$ such that for all $\alpha \in [0, \bar{\alpha}]$,

$$V'_i(A + \alpha\theta) = V'_i(A) + \alpha(v \cdot \theta).$$

To see this property holds, suppose $\theta \in \Theta$ and $A, A + \theta \in \mathcal{A}'_i$. By the definition of \mathcal{A}'_i there exists $B \in \mathcal{A}_i$ such that $A \sim B$. By part 1 of Lemma 11, there exists $\bar{\alpha} \in (0, 1]$ such that $B + \alpha\theta \in \mathcal{A}_i$ for all $\alpha \in [0, \bar{\alpha}]$. Fix any $\alpha \in [0, \bar{\alpha}]$, and $A \sim B$ implies $A + \alpha\theta \sim B + \alpha\theta$ by TI. Therefore, using the translation-linearity of V_i on \mathcal{A}_i ,

$$V'_i(A + \alpha\theta) = V_i(B + \alpha\theta) = V_i(B) + \alpha(v \cdot \theta) = V'_i(A) + \alpha(v \cdot \theta).$$

We now show that this local version of translation-linearity implies translation-linearity. Fix any $A \in \mathcal{A}'_i$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}'_i$, and let

$$\alpha^* \equiv \sup\{\bar{\alpha} \in [0, 1] : V'_i(A + \alpha\theta) = V'_i(A) + \alpha(v \cdot \theta) \forall \alpha \in [0, \bar{\alpha}]\}.$$

Note that $V'_i(A + \alpha^*\theta) = V'_i(A) + \alpha^*(v \cdot \theta)$. If $\alpha^* = 0$, this is obvious. If $\alpha^* > 0$, then local translation-linearity applied to $A' = A + \alpha^*\theta \in \mathcal{A}'_i$ and $\theta' = -\alpha^*\theta$ implies there exists $\bar{\alpha} > 0$ such that $V'_i(A + \alpha^*\theta - \bar{\alpha}\theta) = V'_i(A + \alpha^*\theta) - \bar{\alpha}(v \cdot \theta)$. Therefore,

$$\begin{aligned} V'_i(A + \alpha^*\theta) &= V'_i(A + (\alpha^* - \bar{\alpha})\theta) + \bar{\alpha}(v \cdot \theta) \\ &= V'_i(A) + (\alpha^* - \bar{\alpha})(v \cdot \theta) + \bar{\alpha}(v \cdot \theta) \\ &= V'_i(A) + \alpha^*(v \cdot \theta), \end{aligned}$$

where the second equality follows by the definition of α^* since $0 < \alpha^* - \bar{\alpha} < \alpha^*$. It remains only to show that $\alpha^* = 1$. If not, then local translation-linearity applied to $A' = A + \alpha^*\theta \in \mathcal{A}'_i$ and $\theta' = (1 - \alpha^*)\theta$ implies there exists $\bar{\alpha} > 0$ such that for all $\alpha \in [0, \bar{\alpha}]$,

$$\begin{aligned} V'_i(A + \alpha^*\theta + \alpha\theta) &= V'_i(A + \alpha^*\theta) + \alpha(v \cdot \theta) \\ &= V'_i(A) + (\alpha^* + \alpha)(v \cdot \theta). \end{aligned}$$

This would imply $\alpha^* \geq \alpha^* + \bar{\alpha}$, a contradiction. Thus $\alpha^* = 1$. ■

Lemma 13 *For all $i \geq 1$, if V'_{i-1} is well-defined, translation-linear, and represents \succsim on \mathcal{A}'_{i-1} , then V_i is also well-defined, translation-linear, and represents \succsim on \mathcal{A}_i .*

Proof: (*Well-defined*): Suppose $A \in \mathcal{A}_i$ and $A = B + \alpha\theta^* = B' + \alpha'\theta^*$ for $B, B' \in \mathcal{A}'_{i-1}$ and $\alpha, \alpha' \in \mathbb{R}$. Then $B = B' + (\alpha' - \alpha)\theta^*$, so the translation-linearity of V'_{i-1} implies $V'_{i-1}(B) =$

$V'_{i-1}(B') + (\alpha' - \alpha)(v \cdot \theta^*)$. Therefore, $V'_{i-1}(B) + \alpha(v \cdot \theta^*) = V'_{i-1}(B') + \alpha'(v \cdot \theta^*)$, and hence $V_i(A)$ is uniquely defined.

(*Translation-linear*): Suppose $\theta \in \Theta$ and $A, A + \theta \in \mathcal{A}_i$. Then there exist $B, B' \in \mathcal{A}'_{i-1}$ and $\alpha, \alpha' \in \mathbb{R}$ such that $A = B + \alpha\theta^*$ and $A + \theta = B' + \alpha'\theta^*$. Then $B' = B + (\alpha - \alpha')\theta^* + \theta$, so the translation-linearity of V'_{i-1} implies $V'_{i-1}(B') = V'_{i-1}(B) + v \cdot [(\alpha - \alpha')\theta^* + \theta]$. By the definition of V_i , we therefore have

$$V_i(A + \theta) = V'_{i-1}(B') + \alpha'(v \cdot \theta^*) = V'_{i-1}(B) + \alpha(v \cdot \theta^*) + v \cdot \theta = V_i(A) + v \cdot \theta.$$

(*Represents \succsim*): Suppose $A, A' \in \mathcal{A}_i$, so that $A = B + \alpha\theta^*$ and $A' = B' + \alpha'\theta^*$ for some $B, B' \in \mathcal{A}'_{i-1}$ and $\alpha, \alpha' \in \mathbb{R}$. There are several different cases to consider, and the interest of brevity we only work through one of them here: $A, A' \succsim B' \succsim B$. Thus $B + \alpha\theta^* \succsim B' \succsim B$, which implies $\alpha \geq 0$ by Lemma 10. Continuity implies there exists $\alpha'' \in [0, \alpha]$ such that $B + \alpha''\theta^* \sim B'$, which therefore implies $B + \alpha''\theta^* \in \mathcal{A}'_{i-1}$. Thus by Lemma 10 and the definition of V_i , we have $A \succsim A'$ if and only if $\alpha - \alpha'' \geq \alpha'$ if and only if

$$\begin{aligned} V_i(A) &= V'_{i-1}(B + \alpha''\theta^*) + (\alpha - \alpha'')(v \cdot \theta^*) \\ &= V'_{i-1}(B') + (\alpha - \alpha'')(v \cdot \theta^*) \\ &\geq V'_{i-1}(B') + \alpha'(v \cdot \theta^*) = V_i(A'). \end{aligned}$$

The other cases are similar.²² ■

Using induction and the results of Lemmas 12 and 13, we have proved that for all $i \geq 0$, $V_i : \mathcal{A}_i \rightarrow \mathbb{R}$ is well-defined, translation-linear, and represents \succsim on \mathcal{A}_i . We now define a function $\hat{V} : \bigcup_i \mathcal{A}_i \rightarrow \mathbb{R}$ by $\hat{V}(A) \equiv V_i(A)$ if $A \in \mathcal{A}_i$. This is well-defined because if $A \in \mathcal{A}_i$ and $A \in \mathcal{A}_j$, then without loss of generality suppose $j \geq i$, so $\mathcal{A}_i \subset \mathcal{A}_j$. Then $V_j(B) = V_i(B)$ for all $B \in \mathcal{A}_i$, and hence $V_j(A) = V_i(A)$. Note that \hat{V} represents \succsim on $\bigcup_i \mathcal{A}_i$ and is translation-linear.

By the following lemma, we have now established a translation-linear representation for \succsim on all of \mathcal{A}° .

Lemma 14 $\mathcal{A}^\circ = \bigcup_i \mathcal{A}_i$.

Proof: That $\bigcup_i \mathcal{A}_i \subset \mathcal{A}^\circ$ follows immediately from the definition of \mathcal{A}_i , so it remains only to prove that $\mathcal{A}^\circ \subset \bigcup_i \mathcal{A}_i$. Consider any set $A \in \mathcal{A}^\circ$. By part 1 of Lemma 9, there exists some $\alpha > 0$ such that $A + \alpha\theta^*, A - \alpha\theta^* \in \mathcal{A}^\circ$. Fix any $p \in A$, and we therefore have $\{p\} + \alpha\theta^*, \{p\} - \alpha\theta^* \in \mathcal{A}_0 \subset \mathcal{A}^\circ$. For every $\lambda \in [0, 1]$, define $A(\lambda) \equiv \lambda A + (1 - \lambda)\{p\}$. Note

²²The only substantively different cases are the variations of the following: $B' \succsim A, A' \succsim B$. However, in this case we can apply Lemma 11, which implies $A, A' \in \mathcal{A}'_{i-1}$, and hence the result is obtained by assumption.

that $A(\lambda) + \alpha\theta^*, A(\lambda) - \alpha\theta^* \in \mathcal{A}^\circ$, which follows from the convexity of \mathcal{A}° since

$$A(\lambda) + \alpha\theta^* = \lambda A + (1 - \lambda)\{p\} + \alpha\theta^* = \lambda(A + \alpha\theta^*) + (1 - \lambda)(\{p\} + \alpha\theta^*),$$

and similarly for $A(\lambda) - \alpha\theta^*$. By Lemma 10 for all $\lambda \in [0, 1]$, $A(\lambda) + \alpha\theta^* \succ A(\lambda) \succ A(\lambda) - \alpha\theta^*$. By continuity, for each λ there exists an open (relative to $[0, 1]$) interval $e(\lambda)$ such that $\lambda \in e(\lambda)$ and for all $\lambda' \in e(\lambda)$,

$$A(\lambda) + \alpha\theta^* \succ A(\lambda') \succ A(\lambda) - \alpha\theta^*.$$

Thus $\{e(\lambda) : \lambda \in [0, 1]\}$ is an open cover of $[0, 1]$. Since $[0, 1]$ is compact, there exists a finite subcover, $\{e(\lambda_1), \dots, e(\lambda_n)\}$. Assume the λ_i 's are ordered so that $e(\lambda_i) \cap e(\lambda_{i+1}) \neq \emptyset$, $\{p\} = A(0) \in e(\lambda_1)$, and $A = A(1) \in e(\lambda_n)$. That is, as i increases, $e(\lambda_i)$ moves “farther” from $\{p\}$ and “closer” to A . We can prove that $A(\lambda_1) \in \mathcal{A}_1$ by first observing that

$$A(\lambda_1) + \alpha\theta^* \succ A(0) = \{p\} \succ A(\lambda_1) - \alpha\theta^*,$$

which by continuity implies there exists $\alpha' \in (-\alpha, \alpha)$ such that $A(\lambda_1) + \alpha'\theta^* \sim \{p\}$. This implies $A(\lambda_1) + \alpha'\theta^* \in \mathcal{A}'_0$, which implies that $A(\lambda_1) \in \mathcal{A}_1$. We now show that $A(\lambda_i) \in \mathcal{A}_i$ implies $A(\lambda_{i+1}) \in \mathcal{A}_{i+1}$. If $A(\lambda_i) \in \mathcal{A}_i$, then we also have $A(\lambda_i) + \alpha'\theta^* \in \mathcal{A}_i$ for all $\alpha' \in (-\alpha, \alpha)$. Since $e(\lambda_i) \cap e(\lambda_{i+1}) \neq \emptyset$, choose any $\lambda \in e(\lambda_i) \cap e(\lambda_{i+1})$. Then,

$$\begin{aligned} A(\lambda_i) + \alpha\theta^* &\succ A(\lambda) \succ A(\lambda_i) - \alpha\theta^* \\ A(\lambda_{i+1}) + \alpha\theta^* &\succ A(\lambda) \succ A(\lambda_{i+1}) - \alpha\theta^* \end{aligned}$$

By continuity, there exist $\alpha', \alpha'' \in (-\alpha, \alpha)$ such that $A(\lambda_i) + \alpha'\theta^* \sim A(\lambda) \sim A(\lambda_{i+1}) + \alpha''\theta^*$, which implies $A(\lambda_{i+1}) + \alpha''\theta^* \in \mathcal{A}'_i$. Hence, $A(\lambda_{i+1}) \in \mathcal{A}_{i+1}$. By induction, we conclude that $A(\lambda_i) \in \mathcal{A}_i$ for $i = 1, \dots, n$, and also that $A \in \mathcal{A}'_n \subset \mathcal{A}_{n+1} \subset \bigcup_i \mathcal{A}_i$. \blacksquare

We have now proved that \hat{V} is translation-linear and represents \succsim on \mathcal{A}° . Before extending \hat{V} to \mathcal{A}^c , we first establish that \hat{V} is Lipschitz continuous and convex.

Lemma 15 \hat{V} is Lipschitz continuous.

Proof: For all $\delta \in (0, 1)$, define:

$$\mathcal{A}_\delta^\circ \equiv \{A \in \mathcal{A}^c : \forall p \in A, \forall z \in Z : p_z \geq \delta\}.$$

We next summarize some straightforward facts about \mathcal{A}_δ° whose proofs we omit:

1. \mathcal{A}_δ° is a convex subset of \mathcal{A}° .
2. For all $A \in \mathcal{A}_\delta^\circ$ and $\alpha \in (0, \delta)$ there exists a unique menu $A^\alpha \in \mathcal{A}^\circ$ such that $A = (1 - \alpha)A^\alpha + \alpha\{p_*\}$.²³

²³The menu A^α is given by $A^\alpha = \left\{q \in \mathbb{R}^Z : q = \frac{1}{1-\alpha}(p - \alpha p_*) \text{ for some } p \in A\right\}$.

3. For all $A, B \in \mathcal{A}_\delta^\circ$, $\alpha \in (0, \delta)$: $(1 - \alpha)d_h(A^\alpha, B^\alpha) = d_h(A, B)$.

4. For all $A \in \mathcal{A}_\delta^\circ$, $\alpha \in (0, \delta)$: $A + \alpha\theta^* \in \mathcal{A}^\circ$.

Let $K \equiv 2M(v \cdot \theta^*) > 0$ and $\delta \in (0, 1/2)$. We first show that:

$$A, B \in \mathcal{A}_\delta^\circ \ \& \ d_h(A, B) < \frac{\delta}{2M} \ \Rightarrow \ |\hat{V}(A) - \hat{V}(B)| \leq Kd_h(A, B). \quad (29)$$

Suppose that A, B are as in the left hand side of (29). Let $\alpha \in (2Md_h(A, B), \delta)$. Then

$$d_h(A^\alpha, B^\alpha) = \frac{1}{1 - \alpha}d_h(A, B) \leq 2d_h(A, B) < \alpha/M < 1/M,$$

where the weak inequality follows from $\alpha < \delta < 1/2$. Applying L-Continuity we have:

$$A + \alpha\theta^* = (1 - \alpha)A^\alpha + \alpha\{p^*\} \succ (1 - \alpha)B^\alpha + \alpha\{p_*\} = B.$$

Since \hat{V} represents \succsim and is translation-linear on \mathcal{A}° , we have $\hat{V}(A) + \alpha(v \cdot \theta^*) > \hat{V}(B)$, implying

$$\alpha(v \cdot \theta^*) > \hat{V}(B) - \hat{V}(A).$$

Since the above inequality holds for any $\alpha \in (2Md_h(A, B), \delta)$, we conclude that

$$\hat{V}(B) - \hat{V}(A) \leq 2Md_h(A, B)(v \cdot \theta^*) = Kd_h(A, B).$$

Interchanging the roles of A and B above we also have that $\hat{V}(A) - \hat{V}(B) \leq Kd_h(A, B)$, proving Equation (29).

Next, we use the argument in the proof of Lemma 8 in the supplementary appendix of DLRS (2005) to show that

$$A, B \in \mathcal{A}_\delta^\circ \ \Rightarrow \ |\hat{V}(A) - \hat{V}(B)| \leq Kd_h(A, B), \quad (30)$$

i.e., the requirement $d_h(A, B) < \frac{\delta}{2M}$ in Equation (29) is not necessary. To see this, take any sequence $0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \lambda_{n+1} = 1$ such that $(\lambda_{i+1} - \lambda_i)d_h(A, B) < \frac{\delta}{2M}$. Let $A_i = \lambda_i A + (1 - \lambda_i)B$. It is straightforward to verify that

$$d_h(A_{i+1}, A_i) = (\lambda_{i+1} - \lambda_i)d_h(A, B) < \frac{\delta}{2M}.$$

Combining this with the triangular inequality and Equation (29), we obtain

$$\begin{aligned} |\hat{V}(A) - \hat{V}(B)| &\leq \sum_{i=0}^n |\hat{V}(A_{i+1}) - \hat{V}(A_i)| \\ &\leq K \sum_{i=0}^n d_h(A_{i+1}, A_i) = K \sum_{i=0}^n (\lambda_{i+1} - \lambda_i)d_h(A, B) = Kd_h(A, B). \end{aligned}$$

To conclude the proof, note that by part 1 of Lemma 9, for any $A, B \in \mathcal{A}^\circ$, there exists a small enough $\delta \in (0, 1/2)$ such that $A, B \in \mathcal{A}_\delta^\circ$. Hence by Equation (30), \hat{V} is Lipschitz continuous on \mathcal{A}° with the Lipschitz constant K . \blacksquare

Lemma 16 \hat{V} is convex.

Proof: The argument given here is similar to that used in Lemma 20 of Maccheroni, Marinacci, and Rustichini (2004). We will show that every $A_0 \in \mathcal{A}^\circ$ has a convex and open neighborhood in \mathcal{A}° on which \hat{V} is convex. By a standard result from convex analysis, this implies that \hat{V} is convex on \mathcal{A}° .

Let $A_0 \in \mathcal{A}^\circ$. Define \mathcal{C} to be the collection of all closed and bounded non-empty convex subsets of $\{p \in \mathbb{R}^Z : \sum_{z \in Z} p_z = 1\}$, endowed with the Hausdorff metric topology. It follows from part 1 of Lemma 9 that there exists an $\varepsilon > 0$ such that $B_\varepsilon(A_0) \subset \mathcal{A}^\circ$, where we define

$$B_\varepsilon(A_0) \equiv \{A \in \mathcal{C} : d_h(A, A_0) < \varepsilon\}.$$

Note that $d_h(\cdot, \cdot)$ indicates the Hausdorff metric. For any $\theta \in \Theta$ and $A \in \mathcal{C}$, we have $A + \theta \in \mathcal{C}$ and $d_h(A, A + \theta) = \|\theta\|$, where $\|\cdot\|$ indicates the Euclidean norm. There exists $\theta \in \Theta$ such that $\|\theta\| < \varepsilon$ and $v \cdot \theta > 0$.²⁴ This implies that $A_0 + \theta \in B_\varepsilon(A_0)$ and $A_0 + \theta \succ A_0$. By continuity, there exists $\rho \in (0, \frac{1}{3})$ such that for all $A \in B_{\rho\varepsilon}(A_0)$, $|\hat{V}(A) - \hat{V}(A_0)| < \frac{1}{3}(v \cdot \theta)$. Therefore, if $A, B \in B_{\rho\varepsilon}(A_0)$, then

$$|\hat{V}(A) - \hat{V}(B)| \leq |\hat{V}(A) - \hat{V}(A_0)| + |\hat{V}(A_0) - \hat{V}(B)| < \frac{2}{3}(v \cdot \theta).$$

Let $\alpha \equiv \frac{\hat{V}(A) - \hat{V}(B)}{v \cdot \theta}$, which implies $|\alpha| < \frac{2}{3}$. Then, we have

$$\begin{aligned} d_h(A_0, B + \alpha\theta) &\leq d_h(A_0, B) + d_h(B, B + \alpha\theta) \\ &< \rho\varepsilon + \|\alpha\theta\| \\ &< \frac{1}{3}\varepsilon + \frac{2}{3}\varepsilon = \varepsilon, \end{aligned}$$

so $B + \alpha\theta \in B_\varepsilon(A_0) \subset \mathcal{A}^\circ$. Thus \hat{V} is defined at $B + \alpha\theta$. Note that $\alpha(v \cdot \theta) = \hat{V}(A) - \hat{V}(B)$, so that $\hat{V}(B + \alpha\theta) = \hat{V}(B) + \alpha(v \cdot \theta) = \hat{V}(A)$. Since \succsim satisfies ACP, for any $\lambda \in [0, 1]$,

$$\hat{V}(A) \geq \hat{V}(\lambda A + (1 - \lambda)(B + \alpha\theta)).$$

Therefore,

$$\begin{aligned} \hat{V}(A) &\geq \hat{V}(\lambda A + (1 - \lambda)B) + (1 - \lambda)\alpha(v \cdot \theta) \\ &= \hat{V}(\lambda A + (1 - \lambda)B) + (1 - \lambda)(\hat{V}(A) - \hat{V}(B)), \end{aligned}$$

²⁴For instance $\theta = \alpha\theta^*$ for any $\alpha \in (0, \varepsilon/\|\theta^*\|)$ where $\theta^* = p^* - p_*$.

so we have

$$\lambda \hat{V}(A) + (1 - \lambda) \hat{V}(B) \geq \hat{V}(\lambda A + (1 - \lambda)B).$$

Therefore, \hat{V} is convex on the convex and open neighborhood $B_{\rho\varepsilon}(A_0)$ of A_0 in \mathcal{A}° . \blacksquare

Since \mathcal{A}° is dense in \mathcal{A}^c (see Lemma 9), we can extend \hat{V} to \mathcal{A}^c by continuity. That is, define a function $V : \mathcal{A}^c \rightarrow \mathbb{R}$ as follows: For any $A \in \mathcal{A}^c$ there exists a sequence $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}^\circ$ such that $A_n \rightarrow A$, so define $V(A) \equiv \lim_{n \rightarrow \infty} \hat{V}(A_n)$. Since \hat{V} is Lipschitz continuous, the following lemma establishes that V is well-defined and also Lipschitz continuous. Furthermore, this extension V of \hat{V} represents \succsim on \mathcal{A}^c and preserves the translation-linearity and convexity of \hat{V} .

Lemma 17 *The function $V : \mathcal{A}^c \rightarrow \mathbb{R}$ is well-defined, and it satisfies properties 1–3 from Proposition 3.*

Proof: By Lemma 9, \mathcal{A}° is dense in \mathcal{A}^c . Since \mathcal{A}^c is a compact metric space, it is complete. Since \hat{V} is Lipschitz continuous it is uniformly continuous (see Aliprantis and Border (1999, page 76)). Therefore by Lemma 3.8 in Aliprantis and Border (1999, page 77), V is well-defined and it is the unique continuous extension of \hat{V} to \mathcal{A}^c . To see that V is Lipschitz continuous, let $K > 0$ be a Lipschitz constant for \hat{V} on \mathcal{A}° and let $A, B \in \mathcal{A}^c$. Take sequences $\{A_n\}_{n \in \mathbb{N}}, \{B_n\}_{n \in \mathbb{N}} \subset \mathcal{A}^\circ$ such that $A_n \rightarrow A$ and $B_n \rightarrow B$. Then,

$$|V(A) - V(B)| = \lim_{n \rightarrow \infty} |\hat{V}(A_n) - \hat{V}(B_n)| \leq \lim_{n \rightarrow \infty} K d_h(A_n, B_n) = K d_h(A, B).$$

Hence V is Lipschitz continuous with the same constant K .

To see that V is translation-linear, let $A, A + \theta \in \mathcal{A}^c$ for some $\theta \in \Theta$. Fix any $p \in \Delta(Z)$ such that $p_z > 0$ for all $z \in Z$. For all $n \in \mathbb{N}$, define $A_n \equiv (1 - 1/n)A + (1/n)\{p\}$ and $\theta_n \equiv (1 - 1/n)\theta$. By Lemma 9, for all $n \in \mathbb{N}$, $A_n \in \mathcal{A}^\circ$ and $A_n + \theta_n = (1 - 1/n)(A + \theta) + (1/n)\{p\} \in \mathcal{A}^\circ$. Moreover, $A_n \rightarrow A$ and $A_n + \theta_n \rightarrow A + \theta$ as $n \rightarrow \infty$. Therefore,

$$V(A + \theta) - V(A) = \lim_{n \rightarrow \infty} [\hat{V}(A_n + \theta_n) - \hat{V}(A_n)] = \lim_{n \rightarrow \infty} v \cdot \theta_n = v \cdot \theta.$$

Thus we see that V is translation-linear on all of \mathcal{A}^c . The proof that V is convex is straightforward and follows from a similar line of reasoning; it is therefore omitted.

In order to show that V represents \succsim on \mathcal{A}^c , we prove $A \succ B \iff V(A) > V(B)$. Let $\{A_n\}_{n \in \mathbb{N}}, \{B_n\}_{n \in \mathbb{N}} \subset \mathcal{A}^\circ$ be such that $A_n \rightarrow A$ and $B_n \rightarrow B$ as $n \rightarrow \infty$.

To see “ \implies ”, suppose $A \succ B$. By the continuity of \succsim , $\{C \in \mathcal{A}^c : A \succ C \succ B\}$ is nonempty and open.²⁵ Since \mathcal{A}° is dense in \mathcal{A}^c , there exists $\bar{A} \in \mathcal{A}^\circ$ such that $A \succ \bar{A} \succ B$. Repeating

²⁵Note that the sets $\{\lambda \in [0, 1] : \lambda A + (1 - \lambda)B \succ B\}$ and $\{\lambda \in [0, 1] : A \succ \lambda A + (1 - \lambda)B\}$ are nonempty and open relative to $[0, 1]$ (by continuity of \succsim and continuity of convex combinations), and their union is $[0, 1]$. Since $[0, 1]$ is connected, their intersection must be non-empty. Hence the set $\{C \in \mathcal{A}^c : A \succ C \succ B\}$ is also nonempty.

the same argument for $\bar{A} \succ B$, there exists $\bar{B} \in \mathcal{A}^\circ$ such that $\bar{A} \succ \bar{B} \succ B$. By continuity, $\{C \in \mathcal{A}^c : C \succ \bar{A}\}$ is a neighborhood of A , so there exists $N \in \mathbb{N}$ such that $A_n \succ \bar{A}$ for all $n \geq N$. A similar argument implies there exists $N' \in \mathbb{N}$ such that $\bar{B} \succ B_n$ for all $n \geq N'$. Therefore,

$$V(A) = \lim_{n \rightarrow \infty} \hat{V}(A_n) \geq \hat{V}(\bar{A}) > \hat{V}(\bar{B}) \geq \lim_{n \rightarrow \infty} \hat{V}(B_n) = V(B).$$

To show “ \Leftarrow ”, we will apply a similar argument using the continuity of V . Suppose $V(A) > V(B)$. By continuity of V , $\{C \in \mathcal{A}^c : V(A) > V(C) > V(B)\}$ is nonempty and open. Since \mathcal{A}° is dense in \mathcal{A}^c , there exists $\bar{A} \in \mathcal{A}^\circ$ such that $V(A) > V(\bar{A}) > V(B)$. Repeating the same argument for $V(\bar{A}) > V(B)$, there exists $\bar{B} \in \mathcal{A}^\circ$ such that $V(\bar{A}) > V(\bar{B}) > V(B)$. By continuity, $\{C \in \mathcal{A}^c : V(C) > V(\bar{A})\}$ is a neighborhood of A , so there exists $N \in \mathbb{N}$ such that $\hat{V}(A_n) = V(A_n) > V(\bar{A}) = \hat{V}(\bar{A})$ for all $n \geq N$. A similar argument implies there exists $N' \in \mathbb{N}$ such that $\hat{V}(\bar{B}) > \hat{V}(B_n)$ for all $n \geq N'$. Therefore by continuity of \succsim ,

$$A = \lim_{n \rightarrow \infty} A_n \succsim \bar{A} \succ \bar{B} \succsim \lim_{n \rightarrow \infty} B_n = B.$$

Finally, since V represents \succsim on \mathcal{A}^c and $p^* \succ p_*$, we also have that $V(\{p^*\}) > V(\{p_*\})$. ■

The following lemma establishes uniqueness of the representation, completing the proof of Proposition 3.

Lemma 18 *Suppose \succsim satisfies weak order, strong continuity, ACP, and TI. If $V : \mathcal{A}^c \rightarrow \mathbb{R}$ and $V' : \mathcal{A}^c \rightarrow \mathbb{R}$ are two functions that satisfy 1–3 from Proposition 3, then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V + \beta$.*

Proof: Translation-linearity implies that V and V' are affine on singletons, and therefore the standard vNM uniqueness result implies $V'|_{\mathcal{S}} = \alpha V|_{\mathcal{S}} + \beta$ for some $\alpha > 0$, $\beta \in \mathbb{R}$. By translation-linearity and property 1 from Proposition 3, a simple induction argument shows that $V'|_{\mathcal{A}_i} = \alpha V|_{\mathcal{A}_i} + \beta$ for all i . Hence $V'|_{\mathcal{A}^\circ} = \alpha V|_{\mathcal{A}^\circ} + \beta$. Since \mathcal{A}° is dense in \mathcal{A}^c (see Lemma 9) and the functions V and V' are continuous on \mathcal{A}^c , we conclude that $V' = \alpha V + \beta$ on \mathcal{A}^c . ■

B.3 Application of Duality Results

In this section, we apply the duality results from Appendix A to the function V constructed in Section B.2 to obtain the desired RFCC representation. Thus in the remainder of this section assume that V satisfies 1–3 from Proposition 3. Note that if \succsim also satisfies monotonicity, then V is *monotone* in the sense that for all $A, B \in \mathcal{A}^c$ such that $A \subset B$, we have $V(A) \leq V(B)$. We explicitly assume monotonicity of V at the end of this section to prove the stronger representation of Theorem 1.B.

We follow a construction similar to the one in DLR (2001) to obtain from V a function W whose domain is the set of support functions. For any $A \in \mathcal{A}^c$, the support function $\sigma_A : \mathcal{U} \rightarrow \mathbb{R}$ of A is defined by $\sigma_A(u) = \max_{p \in A} u \cdot p$. For a more complete introduction to support functions, see Rockafellar (1970) or Schneider (1993). Let $C(\mathcal{U})$ denote the set of continuous real-valued functions on \mathcal{U} . When endowed with the supremum norm $\|\cdot\|_\infty$, $C(\mathcal{U})$ is a Banach space. Define an order \geq on $C(\mathcal{U})$ by $f \geq g$ if $f(u) \geq g(u)$ for all $u \in \mathcal{U}$. Let $\Sigma = \{\sigma_A \in C(\mathcal{U}) : A \in \mathcal{A}^c\}$. For any $\sigma \in \Sigma$, let

$$A_\sigma = \bigcap_{u \in \mathcal{U}} \left\{ p \in \Delta(Z) : u \cdot p = \sum_{z \in Z} u_z p_z \leq \sigma(u) \right\}.$$

Lemma 19 1. For all $A \in \mathcal{A}^c$ and $\sigma \in \Sigma$, $A_{(\sigma_A)} = A$ and $\sigma_{(A_\sigma)} = \sigma$. Hence σ is a bijection from \mathcal{A}^c to Σ .

2. For all $A, B \in \mathcal{A}^c$, $\sigma_{\lambda A + (1-\lambda)B} = \lambda \sigma_A + (1-\lambda) \sigma_B$.

3. For all $A, B \in \mathcal{A}^c$, $d_h(A, B) = \|\sigma_A - \sigma_B\|_\infty$.

Proof: These are standard results that can be found in Rockafellar (1970) or Schneider (1993). For instance in Schneider (1993), part 1 can be found on p39 (Theorem 1.7.1), part 2 can be found on p37, and part 3 can be found on p53 (Theorem 1.8.11). ■

Lemma 20 Σ is convex and compact, and $0 \in \Sigma$.

Proof: The set Σ is convex by the convexity of \mathcal{A}^c and part 2 of Lemma 19. As discussed above, the set \mathcal{A}^c is compact, and hence by parts 1 and 3 of Lemma 19, Σ is a compact subset of the Banach space $C(\mathcal{U})$. Also, if we take $q = (1/|Z|, \dots, 1/|Z|) \in \Delta(Z)$, then $q \cdot u = 0$ for all $u \in \mathcal{U}$. Thus $\sigma_{\{q\}} = 0$, and hence $0 \in \Sigma$. ■

Define the function $W : \Sigma \rightarrow \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$. Then, by part 1 of Lemma 19, $V(A) = W(\sigma_A)$ for all $A \in \mathcal{A}^c$. We say the function W is *monotone* if for all $\sigma, \sigma' \in \Sigma$ such that $\sigma \leq \sigma'$ we have $W(\sigma) \leq W(\sigma')$.

Lemma 21 W is convex and Lipschitz continuous with the same Lipschitz constant as V . If V is monotone, then W is monotone.

Proof: To see that W is convex, let $A, B \in \mathcal{A}^c$. Then,

$$\begin{aligned} W(\lambda \sigma_A + (1-\lambda) \sigma_B) &= W(\sigma_{\lambda A + (1-\lambda)B}) = V(\lambda A + (1-\lambda)B) \\ &\leq \lambda V(A) + (1-\lambda) V(B) = \lambda W(\sigma_A) + (1-\lambda) W(\sigma_B) \end{aligned}$$

by parts 1 and 2 of Lemma 19 and convexity of V . The function W is Lipschitz continuous with the same Lipschitz constant as V by parts 1 and 3 of Lemma 19. The function W inherits monotonicity from V because of the following fact which is easy to see from part 1 of Lemma 19: for all $A, B \in \mathcal{A}^c$, $A \subset B$ iff $\sigma_A \leq \sigma_B$. ■

We denote the set of continuous linear functionals on $C(\mathcal{U})$ (the dual space of $C(\mathcal{U})$) by $C(\mathcal{U})^*$. It is well-known that $C(\mathcal{U})^*$ is the set of finite signed Borel measures on \mathcal{U} , where the duality is given by:

$$\langle f, \mu \rangle = \int_{\mathcal{U}} f(u) \mu(du)$$

for any $f \in C(\mathcal{U})$ and $\mu \in C(\mathcal{U})^*$.²⁶

Define Σ_W , \mathcal{N}_W , and \mathcal{M}_W as in Equations (16), (17), and (18), respectively:

$$\begin{aligned} \Sigma_W &= \{\sigma \in \Sigma : \partial W(\sigma) \text{ is a singleton}\}, \\ \mathcal{N}_W &= \{\mu \in C(\mathcal{U})^* : \mu \in \partial W(\sigma), \sigma \in \Sigma_W\}, \\ \mathcal{M}_W &= \overline{\mathcal{N}_W}, \end{aligned}$$

where the closure is taken with respect to the weak* topology. We now apply Proposition 2 to the current setting.

Lemma 22 \mathcal{M}_W is weak* compact, and for any weak* compact $\mathcal{M} \subset C(\mathcal{U})^*$,

$$\mathcal{M}_W \subset \mathcal{M} \iff W(\sigma) = \max_{\mu \in \mathcal{M}} [\langle \sigma, \mu \rangle - W^*(\mu)] \quad \forall \sigma \in \Sigma.$$

Proof: We simply need to verify that $C(\mathcal{U})$, Σ , and W satisfy the assumptions of Proposition 2. Since \mathcal{U} is a compact metric space, $C(\mathcal{U})$ is separable.²⁷ By Lemma 20, Σ is a closed and convex subset of $C(\mathcal{U})$ containing the origin. Although the result is stated slightly differently, it is shown in Hörmander (1954) that $\text{aff}(\Sigma)$ is dense in $C(\mathcal{U})$. This result is also proved in DLR (2001). Finally, W is Lipschitz continuous and convex by Lemma 21. ■

One consequence of Lemma 22 is that for all $\sigma \in \Sigma$,

$$W(\sigma) = \max_{\mu \in \mathcal{M}_W} [\langle \sigma, \mu \rangle - W^*(\mu)].$$

Therefore, for all $A \in \mathcal{A}^c$,

$$V(A) = \max_{\mu \in \mathcal{M}_W} \left[\int_{\mathcal{U}} \max_{p \in A} (u \cdot p) \mu(du) - W^*(\mu) \right].$$

²⁶Since \mathcal{U} is a compact metric space, by the Riesz representation theorem (see Royden, 1988, p357) each continuous linear functional on $C(\mathcal{U})$ corresponds uniquely to a finite signed Baire measure on \mathcal{U} . Since \mathcal{U} is a locally compact separable metric space, the Baire sets and the Borel sets of \mathcal{U} coincide (see Royden, 1988, p332). Hence the set of Baire and Borel finite signed measures also coincide.

²⁷See Theorem 8.48 of Aliprantis and Border (1999).

The function W^* is lower semicontinuous by part 1 of Lemma 4, and \mathcal{M}_W is compact by Lemma 22. It remains only to show that \mathcal{M}_W is consistent and minimal and that monotonicity of W implies each $\mu \in \mathcal{M}_W$ is positive.

Since V is translation-linear, there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}^c$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}^c$, we have $V(A + \theta) = V(A) + v \cdot \theta$. The following result shows that a certain subset of \mathcal{M}_W must “agree” with v in a way that will imply the consistency of this subset. In what follows, let $q = (1/|Z|, \dots, 1/|Z|) \in \Delta(Z)$ and let $\mathcal{A}^\circ \subset \mathcal{A}^c$ be defined as in Equation (25).

Lemma 23 *If $A \in \mathcal{A}^\circ$ and $\mu \in \partial W(\sigma_A)$, then $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$ for all $p \in \Delta(Z)$.*

Proof: Fix any $A \in \mathcal{A}^\circ$ and $\mu \in \partial W(\sigma_A)$. We can apply the definition of the support function to $\theta \in \Theta$, so that $\sigma_{\{\theta\}}(u) = u \cdot \theta$ for $u \in \mathcal{U}$. It is easily verified that for any $A \in \mathcal{A}^c$ and $\theta \in \Theta$, $\sigma_{A+\theta} = \sigma_A + \sigma_{\{\theta\}}$.

We first prove that $\langle \sigma_{\{\theta\}}, \mu \rangle = v \cdot \theta$ for all $\theta \in \Theta$. Fix any $\theta \in \Theta$. Since $A \in \mathcal{A}^\circ$, there exists a $k > 0$ such that $A + k\theta, A - k\theta \in \mathcal{A}^c$. By the translation-linearity of V , we have

$$k(v \cdot \theta) = V(A + k\theta) - V(A) = W(\sigma_{A+k\theta}) - W(\sigma_A).$$

Since $\mu \in \partial W(\sigma_A)$, by part 3 of Lemma 4, $W(\sigma_A) = \langle \sigma_A, \mu \rangle - W^*(\mu)$. Also, by part 2 of the same lemma, $W(\sigma_{A+k\theta}) \geq \langle \sigma_{A+k\theta}, \mu \rangle - W^*(\mu)$. Therefore, we have

$$k(v \cdot \theta) \geq \langle \sigma_{A+k\theta}, \mu \rangle - \langle \sigma_A, \mu \rangle = \langle \sigma_{\{k\theta\}}, \mu \rangle = k \langle \sigma_{\{\theta\}}, \mu \rangle.$$

A similar argument can be used to show that

$$-k(v \cdot \theta) = W(\sigma_{A-k\theta}) - W(\sigma_A) \geq -k \langle \sigma_{\{\theta\}}, \mu \rangle.$$

Hence, we have $k(v \cdot \theta) = k \langle \sigma_{\{\theta\}}, \mu \rangle$, or equivalently, $v \cdot \theta = \langle \sigma_{\{\theta\}}, \mu \rangle$.

We now prove that $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$ for all $p \in \Delta(Z)$. Since $\sum_z u_z = 0$ for $u \in \mathcal{U}$, we have $u \cdot q = 0$ for all $u \in \mathcal{U}$. Clearly, this implies that $\sigma_{\{q\}} = 0$, so that $\langle \sigma_{\{q\}}, \mu \rangle = 0$. For any $p \in \Delta(Z)$, $p - q \in \Theta$, so the above results imply

$$\langle \sigma_{\{p\}}, \mu \rangle = \langle \sigma_{\{p-q\}}, \mu \rangle + \langle \sigma_{\{q\}}, \mu \rangle = \langle \sigma_{\{p-q\}}, \mu \rangle = v \cdot (p - q),$$

which completes the proof. ■

We showed in Section B.2 that if $q = (1/|Z|, \dots, 1/|Z|)$, then $\lambda A + (1 - \lambda)\{q\} \in \mathcal{A}^\circ$ for any $A \in \mathcal{A}^c$ and $\lambda \in (0, 1)$. Therefore, we can use Lemma 23 and the continuity of W to prove the consistency of \mathcal{M}_W .

Lemma 24 *If $\mu \in \mathcal{M}_W$, then $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$ for all $p \in \Delta(Z)$.*

Proof: Define $\mathcal{M} \subset \mathcal{M}_W$ by

$$\mathcal{M} \equiv \{\mu \in \mathcal{M}_W : \langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q) \text{ for all } p \in \Delta(Z)\}.$$

It is easily verified that \mathcal{M} is a closed subset of \mathcal{M}_W and is therefore compact. We want to show $\mathcal{M}_W \subset \mathcal{M}$, which would imply $\mathcal{M} = \mathcal{M}_W$. By Lemma 22, we only need to verify that $W(\sigma) = \max_{\mu \in \mathcal{M}} [\langle \sigma, \mu \rangle - W^*(\mu)]$ for all $\sigma \in \Sigma$.

Let $\sigma \in \Sigma$ be arbitrary. For all $\lambda \in (0, 1)$, we have $\lambda A_\sigma + (1 - \lambda)\{q\} \in \mathcal{A}^\circ$. Note that $\sigma_{\lambda A_\sigma + (1-\lambda)\{q\}} = \lambda \sigma_{A_\sigma} + (1 - \lambda)\sigma_{\{q\}} = \lambda \sigma$. Therefore, Lemma 23 implies that for all $\lambda \in (0, 1)$, $\mathcal{M}_W \cap \partial W(\lambda \sigma) \subset \mathcal{M}$. By Lemma 22, there exists $\mu \in \mathcal{M}_W$ such that $W(\lambda \sigma) = \langle \lambda \sigma, \mu \rangle - W^*(\mu)$, which implies $\mu \in \partial W(\lambda \sigma)$ by part 3 of Lemma 4. Thus $\mathcal{M}_W \cap \partial W(\lambda \sigma) \neq \emptyset$.

Take any net $\{\lambda_d\}_{d \in D}$ such that $\lambda_d \rightarrow 1$, and let $\sigma_d \equiv \lambda_d \sigma$, so that $\sigma_d \rightarrow \sigma$. From the above, for all $d \in D$ there exists $\mu_d \in \mathcal{M}_W \cap \partial W(\sigma_d) \subset \mathcal{M}$. Since \mathcal{M} is weak* compact, every net in \mathcal{M} has a convergent subnet. Without loss of generality, suppose the net itself converges, so that $\mu_d \xrightarrow{w^*} \mu$ for some $\mu \in \mathcal{M}$. By Lemma 2, the definition of the subdifferential, and the continuity of W , for any $\sigma' \in \Sigma$,

$$\langle \sigma' - \sigma, \mu \rangle = \lim_d \langle \sigma' - \sigma_d, \mu_d \rangle \leq \lim_d [W(\sigma') - W(\sigma_d)] = W(\sigma') - W(\sigma),$$

which implies $\mu \in \partial W(\sigma)$.²⁸ Hence, $W(\sigma) = \langle \sigma, \mu \rangle - W^*(\mu)$ by part 3 of Lemma 4. Since $\sigma \in \Sigma$ was arbitrary, this completes the proof. \blacksquare

The consistency of \mathcal{M}_W follows immediately from Lemma 24 since for any $\mu, \mu' \in \mathcal{M}_W$ and $p \in \Delta(Z)$, we have

$$\int_{\mathcal{U}} (u \cdot p) \mu(du) = \langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q) = \langle \sigma_{\{p\}}, \mu' \rangle = \int_{\mathcal{U}} (u \cdot p) \mu'(du).$$

Before proving the minimality of \mathcal{M}_W , we note the following useful result.

Lemma 25 *For all $\mu \in C(\mathcal{U})^*$ there exists $\sigma \in \Sigma$ such that $W^*(\mu) = \langle \sigma, \mu \rangle - W(\sigma)$.*

Proof: Fix any $\mu \in C(\mathcal{U})^*$. Since W is continuous, the mapping $\sigma \mapsto \langle \sigma, \mu \rangle - W(\sigma)$ is continuous and hence attains a maximum on the compact set Σ . \blacksquare

We now prove the minimality of \mathcal{M}_W .

Lemma 26 *\mathcal{M}_W is minimal.*

Proof: Suppose $\mathcal{M}' \subsetneq \mathcal{M}_W$ is compact and $(\mathcal{M}', W^*|_{\mathcal{M}'})$ is a RFCC representation for λ . We will show that this is a contradiction.

Define $V' : \mathcal{A}^c \rightarrow \mathbb{R}$ as in Equation (7), and define $W' : \Sigma \rightarrow \mathbb{R}$ by $W'(\sigma) = V'(A_\sigma)$, so that

$$W'(\sigma) = \max_{\mu \in \mathcal{M}'} [\langle \sigma, \mu \rangle - W^*(\mu)]$$

²⁸Note that Lemma 2 requires that $\{\mu_d\}_{d \in D}$ be norm bounded, but this follows from the compactness of \mathcal{M} and Alaoglu's Theorem (see Aliprantis and Border, 1999, Theorem 6.25).

for all $\sigma \in \Sigma$. By the uniqueness part of Proposition 3, there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V - \beta$, which implies $W' = \alpha W - \beta$.

Take any $\bar{\mu} \in \operatorname{argmin}_{\mu \in \mathcal{M}} W^*(\mu)$. Such a $\bar{\mu}$ must exist by the compactness of \mathcal{M} and the lower semicontinuity of W^* . By Lemma 24, for any $p \in \Delta(Z)$,

$$W(\sigma_{\{p\}}) = \max_{\mu \in \mathcal{M}} [v \cdot (p - q) - W^*(\mu)] = v \cdot (p - q) - W^*(\bar{\mu}).$$

Likewise, by taking $\bar{\mu}' \in \operatorname{argmin}_{\mu \in \mathcal{M}'} W^*(\mu)$, we have that for any $p \in \Delta(Z)$,

$$W'(\sigma_{\{p\}}) = \max_{\mu \in \mathcal{M}'} [v \cdot (p - q) - W^*(\mu)] = v \cdot (p - q) - W^*(\bar{\mu}').$$

By singleton nontriviality, there exist $p, p' \in \Delta(Z)$ such that $\{p\} \succ \{p'\}$. Thus,

$$W(\sigma_{\{p\}}) - W(\sigma_{\{p'\}}) = v \cdot (p - p') = W'(\sigma_{\{p\}}) - W'(\sigma_{\{p'\}}) > 0,$$

which implies $\alpha = 1$.

Thus $W' = W - \beta$. Since $\mathcal{M}' \subsetneq \mathcal{M}_W$, Lemma 22 requires that there is some $\sigma \in \Sigma$ for which $W(\sigma) \neq W'(\sigma)$. We therefore have $\beta \neq 0$. However, take any $\mu' \in \mathcal{M}'$, and by Lemma 25 there exists $\sigma' \in \Sigma$ such that $W^*(\mu') = \langle \sigma', \mu' \rangle - W(\sigma')$, or equivalently, $W(\sigma') = \langle \sigma', \mu' \rangle - W^*(\mu')$. But then $W'(\sigma') = W(\sigma')$, which requires that $\beta = 0$, a contradiction. ■

We have now completed the proof of Theorem 1.A. The following lemma completes the proof of Theorem 1.B.

Lemma 27 *If W is monotone, then each $\mu \in \mathcal{M}_W$ is positive.*

Proof: $C(\mathcal{U})$ is a Banach lattice (Aliprantis and Border (1999, page 302)) and Σ has the property that if $\sigma, \sigma' \in \Sigma$ then $\sigma \vee \sigma' \in \Sigma$. Therefore by Lemma 3, any $\mu \in \mathcal{N}_W$ must be positive. Since the set of positive measures are weak* closed in $C(\mathcal{U})^*$, we conclude that each measure $\mu \in \mathcal{M}_W = \overline{\mathcal{N}_W}$ is also positive. ■

C Proof of Theorem 2

In the following, let (\mathcal{M}, c) be an RFCC representation of \succsim . Let V be as in (7) and define $W: \Sigma \rightarrow \mathbb{R}$ by $W(\sigma) \equiv V(A_\sigma)$. Then W is Lipschitz continuous, convex, and it satisfies

$$W(\sigma) = \max_{\mu \in \mathcal{M}} [\langle \sigma, \mu \rangle - c(\mu)] \tag{31}$$

for all $\sigma \in \Sigma$. Since V satisfies 1–3 from Proposition 3, we can use the results of Appendix B.3 as needed.

Lemma 28 *Let $K \geq 0$ and let $\{\mu_d\}_{d \in D}$ be a net in $C(\mathcal{U})^*$ such that (i) $\|\mu_d\| \leq K$ for all $d \in D$, and (ii) $\mu_d \xrightarrow{w^*} \hat{\mu}$ for some $\hat{\mu} \in C(\mathcal{U})^*$. Then $W^*(\mu_d) \rightarrow W^*(\hat{\mu})$.*

Proof: By Lemma 25, for each $d \in D$, there exists $\sigma_d \in \Sigma$ such that

$$W^*(\mu_d) = \langle \sigma_d, \mu_d \rangle - W(\sigma_d). \quad (32)$$

Since Σ is compact, there exists a subnet on which $\sigma_d \rightarrow \hat{\sigma}$ for some $\hat{\sigma} \in \Sigma$. Without loss of generality, let that subnet be the net itself. By Lemma 2, we then have

$$\langle \sigma_d, \mu_d \rangle \rightarrow \langle \hat{\sigma}, \hat{\mu} \rangle. \quad (33)$$

Let $\sigma \in \Sigma$, by the choice of σ_d we have

$$\langle \sigma_d, \mu_d \rangle - W(\sigma_d) \geq \langle \sigma, \mu_d \rangle - W(\sigma).$$

Taking limits above, we obtain

$$\langle \hat{\sigma}, \hat{\mu} \rangle - W(\hat{\sigma}) \geq \langle \sigma, \hat{\mu} \rangle - W(\sigma)$$

by (33) and continuity of W . Since the above inequality holds for any $\sigma \in \Sigma$, we have that

$$W^*(\hat{\mu}) = \langle \hat{\sigma}, \hat{\mu} \rangle - W(\hat{\sigma}). \quad (34)$$

By (33) and continuity of W , the limit of the right hand side in Equation (32) is the right hand side in Equation (34). Hence $W^*(\mu_d) \rightarrow W^*(\hat{\mu})$. \blacksquare

Let $\mu \in \mathcal{M}$. Then, by Equation (31), $W(\sigma) \geq \langle \sigma, \mu \rangle - c(\mu)$, and hence $c(\mu) \geq \langle \sigma, \mu \rangle - W(\sigma)$. Taking the supremum of the right hand side of the latter with respect to $\sigma \in \Sigma$ gives:

$$c(\mu) \geq W^*(\mu) \text{ for all } \mu \in \mathcal{M}. \quad (35)$$

Note also that if $\mu \in \mathcal{M}$ and $\sigma \in \Sigma$ then:

$$W(\sigma) = \langle \sigma, \mu \rangle - c(\mu) \implies \mu \in \partial W(\sigma). \quad (36)$$

To see (36), let $W(\sigma) = \langle \sigma, \mu \rangle - c(\mu)$. For all $\sigma' \in \Sigma$ we have $W(\sigma') \geq \langle \sigma', \mu \rangle - c(\mu)$. Hence $W(\sigma') - W(\sigma) \geq \langle \sigma' - \sigma, \mu \rangle$, which implies $\mu \in \partial W(\sigma)$. We also have

$$\mathcal{M}_W \subset \mathcal{M}. \quad (37)$$

To see (37), let $\mu \in \mathcal{N}_W$. Then there exists $\sigma \in \Sigma$ such that $\partial W(\sigma) = \{\mu\}$. By (36), any maximizer of (31) must be in $\partial W(\sigma) = \{\mu\}$, so μ must be the maximizer of (31), in particular $\mu \in \mathcal{M}$. Hence $\mathcal{N}_W \subset \mathcal{M}$. Since \mathcal{M} is closed, $\mathcal{M}_W = \overline{\mathcal{N}_W} \subset \mathcal{M}$.

Lemma 29 *If $\mu \in \mathcal{M}_W$ then $c(\mu) = W^*(\mu)$.*

Proof: First let $\mu \in \mathcal{N}_W$, so there exists $\sigma \in \Sigma$ such that $\partial W(\sigma) = \{\mu\}$. Let $\mu' \in \mathcal{M}$ be a maximizer of (31) for σ . By (36) $\mu' \in \partial W(\sigma)$, so $\mu = \mu'$. Hence μ maximizes (31) for σ , so

$$W(\sigma) = \langle \sigma, \mu \rangle - c(\mu),$$

implying that

$$c(\mu) = \langle \sigma, \mu \rangle - W(\sigma) \leq W^*(\mu).$$

Together with (35) and (37), the above inequality implies that $c(\mu) = W^*(\mu)$.

Now let $\mu \in \mathcal{M}_W$. Then there is a net $\{\mu_d\}_{d \in D}$ in \mathcal{N}_W converging to μ . For each $d \in D$, there is $\sigma_d \in \Sigma$ such that $\partial W(\sigma_d) = \{\mu_d\}$. By Lemma 1 there is an element of $\partial W(\sigma_d)$ with norm less than or equal to K , where $K \geq 0$ denotes a Lipschitz constant of W . We therefore have $\|\mu_d\| \leq K$. Then,

$$W^*(\mu) \leq c(\mu) \leq \liminf_d c(\mu_d) = \liminf_d W^*(\mu_d) = W^*(\mu),$$

where the first inequality follows from equations (35) and (37), the second inequality follows from lower semi continuity of c (see Theorem 2.39 in Aliprantis and Border, 1999, p43), the third equality follows from the above paragraph, and the final equality follows from Lemma 28. We conclude again that $c(\mu) = W^*(\mu)$. ■

As established in Appendix B.3, $(\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ is an RFCC representation of \succsim . By $\mathcal{M}_W \subset \mathcal{M}$ and Lemma 29, the minimality of (\mathcal{M}, c) implies that $\mathcal{M} = \mathcal{M}_W$ and $c = W^*|_{\mathcal{M}_W}$.

To conclude the uniqueness proof, let (\mathcal{M}, c) and (\mathcal{M}', c') be two RFCC representations of \succsim . Let $V, V', W,$ and W' be defined accordingly. By the uniqueness part of Proposition 3, there exist $\alpha, \beta \in \mathbb{R}$ with $\alpha > 0$ such that $V' = \alpha V - \beta$. This implies that $W' = \alpha W - \beta$. For any $\mu \in C(\mathcal{U})^*$ and $\sigma, \sigma' \in \Sigma$, note that:

$$W(\sigma') - W(\sigma) \geq \langle \sigma' - \sigma, \mu \rangle \iff W'(\sigma') - W'(\sigma) \geq \langle \sigma' - \sigma, \alpha\mu \rangle,$$

hence $\partial W'(\sigma) = \alpha \partial W(\sigma)$. In particular, $\Sigma_{W'} = \Sigma_W$ and $\mathcal{N}_{W'} = \alpha \mathcal{N}_W$. Taking closures we also have that $\mathcal{M}_{W'} = \alpha \mathcal{M}_W$. Since from our earlier arguments $\mathcal{M}' = \mathcal{M}_{W'}$ and $\mathcal{M} = \mathcal{M}_W$, we conclude that $\mathcal{M}' = \alpha \mathcal{M}$.

Finally, let $\mu \in \mathcal{M}$. Then,

$$c'(\alpha\mu) = \sup_{\sigma \in \Sigma} [\langle \sigma, \alpha\mu \rangle - W'(\sigma)] = \alpha \sup_{\sigma \in \Sigma} [\langle \sigma, \mu \rangle - W(\sigma)] + \beta = \alpha c(\mu) + \beta,$$

where the first and last equalities follow from our earlier findings that $c' = W'^*|_{\mathcal{M}_{W'}}$ and $c = W^*|_{\mathcal{M}_W}$. This concludes the proof of the theorem.

D Interpretation of ACP

For any $A, B \in \mathcal{A}$ and $\lambda \in (0, 1)$, the triple (λ, A, B) is *invertible* if for all $r \in \lambda A + (1 - \lambda)B$ there exist unique $p \in A$ and $q \in B$ for which $r = \lambda p + (1 - \lambda)q$.

Lemma 30 *Let $|Z| \geq 2$, $A, B \in \mathcal{A}$, $\lambda \in (0, 1)$, and $\varepsilon > 0$. Then there exist $A', B' \in \mathcal{A}$ such that $d_h(A, A') < \varepsilon$, $d_h(B, B') < \varepsilon$, and (λ, A', B') is invertible.*

Proof: Let $A, B \in \mathcal{A}$, $\lambda \in (0, 1)$, and $\varepsilon > 0$. For any $p \in \Delta(Z)$, let $|p|$ denote the Euclidean norm of p in \mathbb{R}^Z . Let $N_\varepsilon(p) = \{q \in \Delta(Z) : |p - q| < \varepsilon\}$ denote the open ε -ball around p relative to $\Delta(Z)$. Since $\{N_\varepsilon(p) : p \in A\}$ is an open covering of A and A is compact, there is a finite subset A' of A such that $\{N_\varepsilon(p) : p \in A'\}$ covers A . Note that by construction $d_h(A, A') < \varepsilon$.

Similarly there are finitely many lotteries $q_1, \dots, q_n \in B$ such that $N_\varepsilon(q_1), \dots, N_\varepsilon(q_n)$ cover B . We will construct the desired B' , by inductively selecting a $q'_i \in N_\varepsilon(q_i)$ and making sure at each step that $(\lambda, A', \{q'_1, \dots, q'_i\})$ is invertible. Let $q'_1 = q_1$, then clearly $(\lambda, A', \{q'_1\})$ is invertible. Suppose that $1 \leq i < n$, $(\lambda, A', \{q'_1, \dots, q'_i\})$ is invertible, and define the sets

$$C = \lambda A' + (1 - \lambda)\{q'_1, \dots, q'_i\} \quad \text{and} \quad D = \left\{ -\frac{\lambda}{1 - \lambda}p + \frac{1}{1 - \lambda}r \in \mathbb{R}^Z : p \in A', r \in C \right\}.$$

Since $|Z| \geq 2$, $N_\varepsilon(q_{i+1})$ is uncountable. Since D is finite there exists $q'_{i+1} \in N_\varepsilon(q_{i+1}) \setminus D$. We claim that $(\lambda, A', \{q'_1, \dots, q'_i, q'_{i+1}\})$ is invertible. To see this, it is enough to show that

$$C \cap (\lambda A' + (1 - \lambda)\{q'_{i+1}\}) = \emptyset,$$

since by the inductive assumption $(\lambda, A', \{q'_1, \dots, q'_i\})$ is invertible. Suppose for a contradiction that $r \in C \cap (\lambda A' + (1 - \lambda)\{q'_{i+1}\})$, then there exists $p \in A'$ such that $r = \lambda p + (1 - \lambda)q'_{i+1}$, which can be rewritten as $q'_{i+1} = -\frac{\lambda}{1 - \lambda}p + \frac{1}{1 - \lambda}r$, a contradiction to $q'_{i+1} \notin D$. Set $B' = \{q'_1, \dots, q'_n\}$, then $d_h(B, B') < \varepsilon$ and (λ, A', B') is invertible by induction. ■

Axiom 11 (Weak Aversion to Contingent Planning (WACP)): *For any $A, B \in \mathcal{A}$, if $A \succsim B$, $\lambda \in (0, 1)$, and (λ, A, B) is invertible then $A \succsim \lambda A + (1 - \lambda)B$.*

Lemma 31 *If \succsim satisfies weak order, WACP, and continuity, then it also satisfies ACP.*

Proof: If $|Z| = 1$, then \mathcal{A} is a singleton so ACP holds vacuously. Suppose that $|Z| \geq 2$, $A, B \in \mathcal{A}$, $A \succsim B$, and $\lambda \in (0, 1)$. By Lemma 30, for each integer n , there exist $A_n, B_n \in \mathcal{A}$ such that $d_h(A_n, A) < \frac{1}{n}$, $d_h(B_n, B) < \frac{1}{n}$, and (λ, A_n, B_n) is invertible. There is a subsequence n_k for which $A_{n_k} \succsim B_{n_k}$ for all k or $B_{n_k} \succsim A_{n_k}$ for all k . Suppose first that the former is true, then by Weak ACP, $A_{n_k} \succsim \lambda A_{n_k} + (1 - \lambda)B_{n_k}$ for all k . Continuity implies that in the limit $A \succsim \lambda A + (1 - \lambda)B$. Suppose finally that the subsequence is such that $B_{n_k} \succsim A_{n_k}$ for all k . By definition, (λ, A, B) is invertible implies that $(1 - \lambda, B, A)$ is invertible, hence

$B_{n_k} \succsim (1 - \lambda)B_{n_k} + \lambda A_{n_k}$ for all k . Continuity implies that in the limit $B \succsim (1 - \lambda)B + \lambda A$, which yields the desired conclusion since $A \succsim B$. ■

References

- Aliprantis, C., and K. Border (1999): *Infinite Dimensional Analysis*. Berlin, Germany: Springer-Verlag.
- Conlisk, J. (1996), “Why Bounded Rationality,” *Journal of Economic Literature*, 34, 669–700.
- Dekel, E., B. Lipman, and A. Rustichini (2001): “Representing Preferences with a Unique Subjective State Space,” *Econometrica*, 69, 891–934.
- Dekel, E., B. Lipman, and A. Rustichini (2005): “Temptation-Driven Preferences,” Mimeo, Boston University.
- Dekel, E., B. Lipman, A. Rustichini, and T. Sarver (2005): “Representing Preferences with a Unique Subjective State Space: Corrigendum,” Mimeo, Boston University.
- Ekeland, I., and T. Turnbull (1983): *Infinite-Dimensional Optimization and Convexity*. Chicago: The University of Chicago Press.
- Epstein, L. G., and M. Marinacci (2005): “Coarse Contingencies,” Mimeo.
- Ergin, H. (2003): “Costly Contemplation,” Mimeo, MIT.
- Gilboa, I., and D. Schmeidler (1989): “Maxmin Expected Utility with Non-Unique Priors,” *Journal of Mathematical Economics*, 18, 141–153.
- Gul, F., and W. Pesendorfer (2001): “Temptation and Self-Control,” *Econometrica*, 69, 1403–1435.
- Holmes, R. (1975): *Geometric Functional Analysis and Its Applications*. New York: Springer-Verlag.
- Hörmander, L. (1954): “Sur la Fonction d’Appui des Ensembles Convexes dans un Espace Localement Convexe,” *Arkiv för Matematik*, 3, 181–186.
- Kreps, D. (1979): “A Representation Theorem for Preference for Flexibility,” *Econometrica*, 47, 565–578.

- Luenberger, D. G. (1969): *Optimization by Vector Space Methods*. New York: John Wiley and Sons, Inc.
- Maccheroni, F., M. Marinacci, and A. Rustichini (2004): “Ambiguity Aversion, Malevolent Nature, and the Variational Representation of Preferences,” Mimeo.
- Mas-Colell, A. (1986): “The Price Equilibrium Existence Problem in Topological Vector Lattices,” *Econometrica*, 54, 1039–1054.
- Munkres, J. (2000): *Topology*. Upper Saddle River, NJ: Prentice Hall.
- Phelps, R. R. (1993): *Convex Functions, Monotone Operators, and Differentiability*. Berlin, Germany: Springer-Verlag.
- Rockafellar, R. T. (1970): *Convex Analysis*. Princeton, NJ: Princeton University Press.
- Royden, H. L. (1988): *Real Analysis*. Englewood Cliffs, NJ: Prentice Hall.
- Sarver, T. (2005), “Anticipating Regret: Why Fewer Options May Be Better,” Mimeo, Boston University.
- Schneider, R. (1993): *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge: Cambridge University Press.