A Subjective Model of Temporal Preferences[*]

Haluk Ergin[†]   Todd Sarver[‡]

September 30, 2009

Abstract

We study preferences for timing of resolution of objective uncertainty in a simple menu choice model with two stages of information arrival. We characterize two general classes of utility representations called hidden action representations. The representations can be interpreted as if an unobservable action is taken by the individual (or by the malevolent nature, depending on whether the preference is for early or late resolution of uncertainty) between the two periods. We illustrate that our general representations allow for a richer class of preferences for timing of resolution of uncertainty than was possible in Kreps and Porteus (1978), and provide a unified framework for studying a variety of well-known preferences in the literature. We show that subjective versions of the class of multi-prior preferences (Gilboa and Schmeidler (1989)) and variational preferences (Maccheroni, Marinacci, and Rustichini (2006)) overlap with the class of hidden action preferences exhibiting a preference for late resolution of uncertainty. The costly contemplation model (Ergin and Sarver (2009)) is characterized as a special case of the class of hidden action preferences exhibiting a preference for early resolution of uncertainty. A generalization of the Kreps and Porteus (1978) model that allows for preference for flexibility in the sense of Dekel, Lipman, and Rustichini (2001) is also characterized.

1 Introduction

This paper considers several new classes of dynamic preferences, providing representations for preferences for both early and late resolution of uncertainty. These preferences are examined in a simple menu-choice model with two-stage objective uncertainty. We axiomatize two hidden action representations corresponding to preferences for early and late resolution of uncertainty, and we show that, under a minimality condition, these representations are uniquely identified from the preference. In other words, preferences for early and late resolution of uncertainty can be interpreted as arising from the presence of an unobserved (hidden) action that can be taken between the resolution of the first and second period objective uncertainty.[1]

It is well known that an individual may prefer to have uncertainty resolve at an earlier date in order to be able to condition her future actions on the realization of this uncertainty. For example, an individual may prefer to have uncertainty about her future income resolve earlier so that she can smooth her consumption across time. Suppose an individual has the possibility of receiving a promotion with a substantial salary increase several years into the future. If she is able to learn the outcome of that promotion decision now, then even if she will not actually receive the increased income until a later date, she may choose to increase her current consumption by temporarily decreasing her savings or increasing her debt. On the other hand, if she is not told the outcome of the promotion decision, then by increasing her consumption now, she risks having larger debt and hence suboptimally low consumption in the future. In this example, changing the timing of the resolution of uncertainty benefits the individual by increasing her ability to condition her choices on the outcome of that uncertainty.

Kreps and Porteus (1978) considered a broader class of dynamic preferences that allow for a preference for early (or late) resolution of uncertainty even when the individual's ability to condition her (observed) actions on the outcome of this uncertainty is unchanged. For example, suppose the individual described above has no current savings and is unable to take on debt. Then, if she learns the outcome of the promotion decision now, she still is unable to increase her current consumption. Even in this case, the preferences considered by Kreps and Porteus (1978) allow the individual to have a strict preference for that uncertainty to resolve earlier (or later). This additional time-preference for the resolution of uncertainty has proved quite useful in applications, in particular, in macroeconomic models of asset pricing and business cycles.

The setting for our model is a simple two-stage version of the model considered by

---

[1]Similar observations were made by Kreps and Porteus (1979) and Machina (1984). We discuss how their models fit into our general class of preferences in Section 4.3.

Kreps and Porteus (1978). However, we allow for more general axioms, which permits us to model a richer set of preferences exhibiting a preference for early or late resolution of uncertainty. In particular, we relax the Strategic Rationality Axiom of Kreps and Porteus (1978) (see Axiom 7) to allow for a preference for flexibility as in Kreps (1979) and Dekel, Lipman, and Rustichini (2001, henceforth DLR). We are able to represent this general class of preferences for early and late resolution of uncertainty as if there is an unobserved (hidden) action that can be taken between the resolution of the first and second period objective uncertainty. In the case of a preference for early resolution of uncertainty, this hidden action can be thought of as an action chosen by the individual. Thus, the individual prefers to have objective uncertainty resolve in the first period so that she can choose this action optimally. In the case of a preference for late resolution of objective uncertainty, this hidden action could be thought of as an action chosen by the (malevolent) nature. In this case, the individual prefers to have objective uncertainty resolve in the second period, after this action has been selected by nature, so as to mitigate natures ability to harm her.

This paper not only provides representations for a more general class of preferences for early and late resolution of uncertainty, but also provides new ways to understand and interpret these temporal preferences. Our hidden action model is general enough to encompass the subjective versions of a number of well-known representations in the literature: the multiple priors model of Gilboa and Schmeidler (1989), the variational preferences model of Maccheroni, Marinacci, and Rustichini (2006), the costly contemplation model of Ergin and Sarver (2009), and a version of the temporal preferences model of Kreps and Porteus (1978) extended to allow for subjective uncertainty as in DLR (2001). We identify the preference for temporal resolution of uncertainty implied by each of these representations as well as their additional behavioral implications over the hidden action model. The general framework in this paper provides a unification of these well-known representations and provides simple axiomatizations.

2   The Model

Let Z be a finite set of alternatives, and let $\triangle(Z)$ denote the set of all probability distributions on Z, endowed with the Euclidean metric d (generic elements $p, q, r \in \triangle(Z)$). Let $\mathcal{A}$ denote the set of all closed subsets of $\triangle(Z)$, endowed with the Hausdorff metric:

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

Elements of $\mathcal{A}$ are called menus (generic menus $A, B, C \in \mathcal{A}$). Let $\triangle(\mathcal{A})$ denote the set of all Borel probability measures on $\mathcal{A}$, endowed with the weak* topology (generic elements $P, Q, R \in \triangle(\mathcal{A})$).[2] The primitive of the model is a binary relation $\succsim$ on $\triangle(\mathcal{A})$, representing the individual's preferences over lotteries over menus.

We interpret $\succsim$ as corresponding to the individual's choices in the first period of a two-period decision problem. In the beginning of period 1, the individual chooses a lottery P over menus. Later in period 1, the uncertainty associated with the lottery P is resolved and P returns a menu A. In the unmodeled period 2, the individual chooses a lottery p out of A. Later in period 2, the lottery p resolves and returns an alternative z. We will refer to the uncertainty associated with the resolution of P as the *first-stage uncertainty*, and will refer to the uncertainty associated with the resolution of p as the *second-stage uncertainty*. Although the period 2 choice is unmodeled, it will be important for the interpretation of the representations.

Our model is a special case of Kreps and Porteus (1978) with only two periods and no consumption in period 1.[3] A lottery $P \in \triangle(\mathcal{A})$ over menus is a *temporal lottery* (Kreps and Porteus (1978)) if P returns a singleton menu with probability one. An individual facing a temporal lottery makes no choice in period 2, between the resolution of first and second stages of the uncertainty. Note that the set of temporal lotteries can be naturally associated with $\triangle(\triangle(Z))$.

For any $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$, the convex combination of these two menus is defined by $\alpha A + (1 - \alpha)B \equiv \{\alpha p + (1 - \alpha)q : p \in A \text{ and } q \in B\}$. We let $\mathrm{co}(A)$ denote the convex hull of the menu A and let $\delta_A \in \triangle(\mathcal{A})$ denote the degenerate lottery that puts probability 1 on the menu A. Then, $\alpha\delta_A + (1 - \alpha)\delta_B$ denotes the lottery that puts probability α on the menu A and probability $1 - \alpha$ on the menu B. Finally, for any continuous function $V : \mathcal{A} \to \mathbb{R}$ and $P \in \triangle(\mathcal{A})$, we let $\mathbb{E}_P[V]$ denote the expected value of V under the lottery P, i.e., $\mathbb{E}_P[V] = \int_{\mathcal{A}} V(A) \, P(dA)$.

3   General Representations

We will impose the following set of axioms in all the representation results in the paper. Therefore, it will be convenient to refer to them altogether as Axiom 1.

---

[2]Given a metric space X, the weak* topology on the set of all finite signed Borel measures on X is the topology where a net of signed measures $\{\mu_d\}_{d \in D}$ converges to a signed measure μ if and only if $\int_X f \, \mu_d(dx) \to \int_X f \, \mu(dx)$ for every bounded continuous function $f : X \to \mathbb{R}$.

[3]The same model is also used in Epstein and Seo (2007) and in Section 4 of Epstein, Marinacci and Seo (2007).

Axiom 1

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

2. (Continuity): *The upper and lower contour sets,* $\{P \in \triangle(\mathcal{A}) : P \succsim Q\}$ *and* $\{P \in \triangle(\mathcal{A}) : P \precsim Q\}$, *are closed in the weak* topology.*

3. (First-Stage Independence): *For any* $P, Q, R \in \triangle(\mathcal{A})$ *and* $\alpha \in (0,1)$,

$$P \succ Q \quad \Rightarrow \quad \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)R.$$

4. (L–Continuity): *There exist* $A^*, A_* \in \mathcal{A}$ *and* $M \geq 0$ *such that for every* $A, B \in \mathcal{A}$ *and* $\alpha \in [0,1]$ *with* $\alpha \geq M d_h(A, B)$,

$$(1-\alpha)\delta_A + \alpha \delta_{A^*} \succsim (1-\alpha)\delta_B + \alpha \delta_{A_*}.$$

5. (Indifference to Randomization (IR)): *For every* $A \in \mathcal{A}$, $\delta_A \sim \delta_{co(A)}$.

Axioms 1.1 and 1.2 are standard. Axiom 1.3 is the von Neumann-Morgenstern independence axiom imposed with respect to the first-stage uncertainty. Axioms 1.1–1.3 ensure that there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Given Axioms 1.1–1.3, Axiom 1.4 is a technical condition implying the Lipschitz continuity of V.[4] Axiom 1.5 was introduced in DLR (2001). It is justified if the individual choosing from the menu A in period 2 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 individual.

The next two axioms from Kreps and Porteus (1978) will be key in our representation results.

Axiom 2 (Preference for Early Resolution of Uncertainty (PERU)) *For any* $A, B \in \mathcal{A}$ *and* $\alpha \in (0,1)$,
$$\alpha \delta_A + (1-\alpha)\delta_B \succsim \delta_{\alpha A + (1-\alpha)B}.$$

Axiom 3 (Preference for Late Resolution of Uncertainty (PLRU)) *For any* $A, B \in \mathcal{A}$ *and* $\alpha \in (0,1)$,
$$\delta_{\alpha A + (1-\alpha)B} \succsim \alpha \delta_A + (1-\alpha)\delta_B.$$

$$\alpha\delta_A + (1-\alpha)\delta_B \qquad\qquad\qquad \delta_{\alpha A+(1-\alpha)B}$$
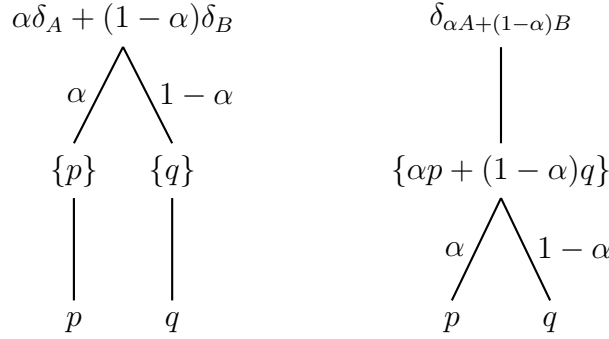

Figure 1: Timing of Resolution of Uncertainty with Temporal Lotteries

To understand Axioms 2 and 3, suppose first that $A = \{p\}$ and $B = \{q\}$ for some $p, q \in \triangle(Z)$. In this case, period 2 choice out of a menu is degenerate, and the two equations above are statements about the preference between the temporal lotteries $\alpha\delta_{\{p\}} + (1-\alpha)\delta_{\{q\}}$ and $\delta_{\alpha\{p\}+(1-\alpha)\{q\}}$. The temporal lottery $\alpha\delta_{\{p\}} + (1-\alpha)\delta_{\{q\}}$ corresponds to the first tree in Figure 1, in which the uncertainty regarding whether lottery p or q is selected resolves in period 1. The temporal lottery $\delta_{\alpha\{p\}+(1-\alpha)\{q\}}$ corresponds to the second tree in Figure 1, in which the same uncertainty resolves in period 2.[5] PERU requires a weak preference the first temporal lottery, whereas PLRU requires a weak preference for the second temporal lottery.

In the general case where A and B need not be singletons, the interpretations of PERU and PLRU are more subtle, since $\delta_{\alpha A+(1-\alpha)B}$ and $\alpha\delta_A + (1-\alpha)\delta_B$ involve non-degenerate period 2 choices. Note that in $\alpha\delta_A + (1-\alpha)\delta_B$, the uncertainty regarding whether the menu A or B is selected resolves in period 1, before the individual makes her choice out of the selected menu. On the other hand, in $\delta_{\alpha A+(1-\alpha)B}$ no uncertainty is resolved in period 1. The period 2 choice of a lottery $\alpha p + (1-\alpha)q$ from the convex combination menu $\alpha A+(1-\alpha)B$ is identical to a pair of choices $p \in A$ and $q \in B$, where after the individual chooses (p, q), p is selected with probability α and q is selected with probability $1 - \alpha$. Therefore, the period 2 choice out of the menu $\alpha A + (1-\alpha)B$ can be interpreted as a complete contingent plan out of the menus A and B.

The key distinction between the two lotteries over menus is that in $\delta_{\alpha A+(1-\alpha)B}$ the period 2 choice is made prior to the resolution of the uncertainty regarding whether the

---

[4]In models with preferences over menus over lotteries, analogous L–continuity axioms can be found in Dekel, Lipman, Rustichini, and Sarver (2007, henceforth DLRS), Sarver (2008), and Ergin and Sarver (2009).

[5]In both temporal lotteries, the remaining uncertainty, i.e., the outcome of p conditional on p being selected and the outcome of q conditional on q being selected, is resolved in period 2.

choice from A or the choice from B will be implemented, whereas in $\alpha\delta_A + (1-\alpha)\delta_B$ the same uncertainty is resolved in period 1 before the individual makes a choice out of the selected menu. Therefore, PERU can be interpreted as the individual's preference to make a choice out of the menu after learning which menu is selected, whereas PLRU is the individual's preference to make her choice out of the menus before learning which menu is selected.

We will also characterize the special case of our representation where the individual has a weak preference for larger menus.

Axiom 4 (Monotonicity) *For any $A, B \in \mathcal{A}$, $A \subset B$ implies $\delta_B \succsim \delta_A$.*

Since expected-utility functions on $\triangle(Z)$ are equivalent to vectors in $\mathbb{R}^Z$, we will use the notation $u(p)$ and $u \cdot p$ interchangeably for any expected utility function $u \in \mathbb{R}^Z$. We define the set of *normalized (non-constant) expected-utility functions* on $\triangle(Z)$ to be

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

We are ready to introduce our general representations:[6]

Definition 1 A *Maximum [Minimum] Hidden Action (max-HA [min-HA]) representation* is a pair $(\mathcal{M}, c)$ consisting of a compact set of finite signed Borel measures $\mathcal{M}$ on $\mathcal{U}$ and a lower semi-continuous function $c : \mathcal{M} \to \mathbb{R}$ such that:

1. $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (1) [(2)]:

$$V(A) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) - c(\mu) \right) \tag{1}$$

$$V(A) = \min_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) + c(\mu) \right). \tag{2}$$

2. The set $\mathcal{M}$ is *minimal*: For any compact proper subset $\mathcal{M}'$ of $\mathcal{M}$, the function V' obtained by replacing $\mathcal{M}$ with $\mathcal{M}'$ in Equation (1) [(2)] is different from V.

The pair $(\mathcal{M}, c)$ is an *HA representation* if it is a max-HA or a min-HA representation. An HA representation $(\mathcal{M}, c)$ is *monotone* if all measures in $\mathcal{M}$ are positive.

---

[6]Note that we endow the set of all finite signed Borel measures on $\mathcal{U}$ with the weak* topology (see footnote 2).

Before we interpret the representations, we will first argue that after appropriately renormalizing the set of ex post utility functions, one can reinterpret the integral term in Equations (1) and (2) as an expectation. More specifically, suppose that μ is a nonnegative measure with $\lambda = \mu(\mathcal{U}) > 0$. Consider the probability measure π on $\mathcal{V} = \lambda \mathcal{U}$ which (heuristically) puts $\mu(u)/\lambda$ weight on each $v = \lambda u \in \mathcal{V}$. Then, by a simple change of variables, the above integral can be rewritten as the expectation $\int_{\mathcal{V}} \max_{p \in A} v(p) \pi(dv)$ which can be interpreted as follows. The individual anticipates that her ex post utility functions will be distributed according to π. Conditional on each realization of her ex post utility $v \in \mathcal{V}$, she will chose a lottery in A that maximizes v. She aggregates across different possible realizations of v by taking expectation with respect to π.[7] Therefore, each measure μ can be interpreted as a reduced-form representation of the individual's subjective uncertainty about her ex post (period 2) utility function over $\triangle(Z)$. We conjecture that analogous interpretations are possible if μ is a signed measure, by introducing regret and temptation to the current framework.[8]

We next interpret Equation (1). In period 1, the individual anticipates that after the first-stage uncertainty is resolved but before she makes her choice in period 2, she will be able to select an action μ from a set $\mathcal{M}$. Each action μ affects the distribution of the individual's ex post utility functions over $\triangle(Z)$, at cost $c(\mu)$. As argued in the above paragraph, the integral in Equation (1) can be interpreted as a reduced-form representation for the value of the action μ when the individual chooses from menu A, which is linear in the menu A. For each menu A, the individual maximizes the value minus cost of her action.

The interpretation of Equation (2) is dual. In period 1, the individual anticipates that after the first-stage uncertainty is resolved but before she makes her choice in period 2, the (malevolent) nature will select an action μ from a set $\mathcal{M}$. The individual anticipates the nature to chose an action which minimizes the value to the individual plus a cost term. The function c can be interpreted as being related to the pessimism attitude of the individual. For constant c, she expects the nature to chose an action that outright minimizes her utility from a menu. Different cost functions put different restrictions on the individual's perception of the malevolent nature's objective.[9]

---

[7]It is noteworthy that any probability measure π' such that μ and π' are absolutely continuous with respect to each other can be used in the above argument. The reason is that, heuristically, in the HA-representations, the weight $\mu(u)$ on the ex post utility function u captures both the cardinality of the ex post utility and its probability, and these two effects cannot be separated from behavior as in models with state dependent utility. See Kreps (1988) for an elaborate discussion of the state-dependence issue.

[8]See Sarver (2008), Gul and Pesendorfer (2001), and DLR (2009).

[9]While we do not suggest that there literally exists a malevolent nature, it is a useful way to interpret a pessimistic or ambiguity averse attitude on the part of the decision-maker. This interpretation of the cost function is due to Maccheroni, Marinacci, and Rustichini (2006), whose results we discuss in more

In the above representations, both the set of available actions and and their costs are subjective in that they are part of the representation. Therefore $\mathcal{M}$ and c are not directly observable to the modeler and need to be identified from the individual's preferences. Note that in both Equations (1) and (2), it is possible to enlarge the set of actions by adding a new action μ to the set $\mathcal{M}$ at a prohibitively high cost $c(\mu)$ without affecting the equations. Therefore, in order to identify $(\mathcal{M}, c)$ from the preference, we also impose an appropriate minimality condition on the set $\mathcal{M}$.

We postpone more concrete interpretations of the set of actions and costs to the discussion of the applications of HA-representations in the following section. We are now ready to state our general representation result.

Theorem 1 *A. The preference $\succsim$ has a max-HA [min-HA] representation if and only if it satisfies Axiom 1 and PERU [PLRU].*

B. The preference $\succsim$ has a monotone max-HA [min-HA] representation if and only if it satisfies Axiom 1, PERU [PLRU], and monotonicity.[10]

The special case of HA representations satisfying indifference to timing of resolution of uncertainty (i.e., both PERU and PLRU) are those where $\mathcal{M}$ is a singleton. In that case, the constant cost can be dropped out from Equations (1) and (2), leading to an analogue of DLR (2001)'s additive representation where the individual is risk neutral with respect to the first-stage uncertainty.

We next give a brief intuition about Theorem 1.A. Axiom 1 guarantees the existence of a Lipschitz continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $V(co(A)) = V(A)$ and $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. In terms of this expected utility representation, it is easy to see that PERU corresponds to convexity of V and PLRU corresponds to concavity of V. The set $\mathcal{A}^c$ of convex menus can be mapped one-to-one to the set Σ of support functions, preserving the metric and the linear operations. Therefore, by using the property $V(co(A)) = V(A)$ and mimicking the construction in DLR (2001), V can be thought of as a function defined on the subset Σ of the Banach space $C(\mathcal{U})$ of continuous real-valued functions on $\mathcal{U}$. We then apply a variation of the classic duality principle that convex [concave] functions can be written as the supremum [infimum] of affine functions lying below [above] them.[11] Finally, we apply the Riesz representation theorem to write each such continuous affine function as an integral against a measure μ minus [plus]

---

detail in Section 4.1.

[10]In part B, IR can be dropped for the case of the max-HA representation because it is implied by weak order, continuity, first-stage independence, PERU, and monotonicity.

[11]See Rockafellar (1970), Phelps (1993), and Appendix A of the current paper for variations of this duality result.

a scalar $c(\mu)$. Theorem 1.B states that additionally imposing monotonicity guarantees that all measures in the HA representation are positive.

We show that the uniqueness of the HA representations follows from the affine uniqueness of V and a result about the uniqueness of the dual representation of a convex function in the theory of conjugate convex functions (see Theorem 11 in Appendix A). A similar application of the duality and uniqueness results can be found in Ergin and Sarver (2009).

Theorem 2 *If $(\mathcal{M}, c)$ and $(\mathcal{M}', c')$ are two max-HA [min-HA] 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 Applications

4.1 Multiple Priors and Variational Preferences

A preference for late resolution of uncertainty could arise if an individual would like to delay the resolution of objective lotteries for hedging reasons. In this section, we formalize this intuition by showing that the monotone min-HA model is equivalent to two representations that have natural interpretations in terms of ambiguity-aversion. The following multiple-priors representation allows for ambiguity regarding the distribution over ex post subjective states and is intuitively similar to the multiple-priors representation proposed by Gilboa and Schmeidler (1989) in the Anscombe-Aumann setting.

Definition 2 A *Subjective-State-Space Multiple-Priors (SSMP) representation* is a quadruple $((\Omega, \mathcal{F}), U, \Pi)$ where Ω is a state space endowed with the σ-algebra $\mathcal{F}$, $U : \Omega \to \mathbb{R}^Z$ is a Z–dimensional, $\mathcal{F}$–measurable, and bounded random vector, and Π is a set of probability measures on $(\Omega, \mathcal{F})$, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by

$$V(A) = \min_{\pi \in \Pi} \int_\Omega \max_{p \in A} U(\omega) \cdot p \ \pi(d\omega), \tag{3}$$

and the minimization in Equation (3) has a solution for every $A \in \mathcal{A}$.

The next representation is similar in spirit to the variational representation considered by Maccheroni, Marinacci, and Rustichini (2006) in the Anscombe-Aumann setting.

Definition 3 A *Subjective-State-Space Variational (SSV) representation* is a quintuple $((\Omega, \mathcal{F}), U, \Pi, c)$ where Ω is a state space endowed with the σ-algebra $\mathcal{F}$, $U : \Omega \to \mathbb{R}^Z$ is a Z-dimensional, $\mathcal{F}$-measurable, and bounded random vector, Π is a set of probability measures on $(\Omega, \mathcal{F})$, and $c : \Pi \to \mathbb{R}$ is a function, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by

$$V(A) = \min_{\pi \in \Pi} \left(\int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega) + c(\pi) \right), \tag{4}$$

and the minimization in Equation (4) has a solution for every $A \in \mathcal{A}$.[12]

The SSV representation generalizes the SSMP representation by allowing a "cost" $c(\pi)$ to be assigned to each measure π in the representation. In the Anscombe-Aumann setting, the class of variational preferences considered by Maccheroni, Marinacci, and Rustichini (2006) is strictly larger than the class of multiple-prior expected-utility preferences considered by Gilboa and Schmeidler (1989). However, we show that in the current setting, the SSMP and SSV representations are equivalent in the sense that the set of preferences that can be represented using an SSMP representation is precisely the set of preferences that can be represented using an SSV representation. The reason for this equivalence in the subjective versions of the representations is the state-dependence of the utility functions in the representations. The following Theorem formalizes this claim and, moreover, states that a preference $\succsim$ can be represented by one of these representations if and only if it has a monotone min-HA representation.

Theorem 3 *Let $V : \mathcal{A} \to \mathbb{R}$. Then, the following are equivalent:*

1. *There exists a monotone min-HA representation such that V is given by Equation (2).*

2. *There exists an SSMP representation such that V is given by Equation (3).*

3. *There exists an SSV representation such that V is given by Equation (4).*

---

[12]Note that for simplicity, we directly assume in the SSMP and SSV representations that the minimization in Equations (3) and (4) have a solution. One alternative approach that does not require this indirect assumption on the parameters would be to replace the minimums in Equations (3) and (4) with infima, in which case Theorem 3 would continue to hold. A second alternative is to impose topological assumptions on the parameters that would guarantee the existence of a minimum, for instance assuming that Ω is a metric space, $\mathcal{F}$ is the Borel σ-algebra on Ω, U is continuous, Π is weak*-compact, and c is lower semi-continuous.

The following immediate corollary provides the axiomatic foundation for the SSMP and SSV representations.

Corollary 1 *A preference $\succsim$ has a SSMP representation if and only if it has a SSV representation if and only if it satisfies Axiom 1, PLRU, and monotonicity.*

We next outline the intuition behind Theorem 3 for the case where in the min-HA, SSMP, and SSV representations, the sets $\mathcal{M}, \Pi, \Omega$ are finite, $\mathcal{F} = 2^\Omega$, and the measures in $\mathcal{M}$ have finite support. In this special case, the compactness, lower semi-continuity, measurability, and boundedness properties in the representations are automatically satisfied.

First consider $(1) \Rightarrow (3)$. Fix a monotone min-HA representation $(\mathcal{M}, c)$, and define V by Equation (2). Take any $\mu \in \mathcal{M}$, and define a measure π_μ on the set $\mu(\mathcal{U})\mathcal{U}$ by $\pi_\mu(v) = \frac{\mu(u)}{\mu(\mathcal{U})}$ for $v = \mu(\mathcal{U})u$, $u \in \mathcal{U}$. This transformation ensures that π_μ is a probability measure. Moreover, since $\pi_\mu(v)v = \mu(u)u$ for $v = \mu(\mathcal{U})u$, for any $A \in \mathcal{A}$,

$$\int_{\mu(\mathcal{U})\mathcal{U}} \max_{p \in A} v(p) \, \pi_\mu(dv) = \int_\mathcal{U} \max_{p \in A} u(p) \, \mu(du).$$

Let $\Omega = \bigcup_{\mu \in \mathcal{M}} \mu(\mathcal{U})\mathcal{U}$, and define $U : \Omega \to \mathbb{R}^Z$ by $U(\omega) = \omega$. Let $\Pi = \{\pi_\mu : \mu \in \mathcal{M}\}$ and $\tilde{c}(\pi_\mu) = c(\mu)$. Then, V can be expressed in the following SSV form:

$$V(A) = \min_{\pi \in \Pi} \left(\int_\Omega \max_{p \in A} U(\omega) \cdot p \, \pi(d\omega) + \tilde{c}(\pi) \right).$$

The idea of this construction is identical to the interpretation we gave for the integral term in the HA-representations: Since utility is state-dependent, the weight of the states in the linear aggregator cannot be uniquely pinned down, and integration against the positive measure μ can be reexpressed as integration against a probability measure π_μ after appropriately rescaling the state-dependent utility functions.

To see the intuition for $(3) \Rightarrow (2)$, consider an SSV representation $((\Omega, \mathcal{F}), U, \Pi, c)$, and define V by Equation (4). Let $\tilde{\Omega} = \Omega \times \Pi$ and $\tilde{\mathcal{F}} = 2^{\tilde{\Omega}}$. Let $\mathbf{1} \in \mathbb{R}^Z$ denote the vector whose coordinates are equal to 1, and define $\tilde{U} : \tilde{\Omega} \to \mathbb{R}^Z$ by $\tilde{U}(\omega, \pi) = U(\omega) + c(\pi)\mathbf{1}$ for any $\tilde{\omega} = (\omega, \pi) \in \tilde{\Omega}$. Take any probability measure $\pi \in \Pi$, and define a new measure ρ_π on $\tilde{\Omega}$ by $\rho_\pi(\omega, \pi) = \pi(\omega)$ and $\rho_\pi(\omega, \pi') = 0$ for any $\pi' \neq \pi$. It is immediate that ρ_π is

a probability measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$. Also, for any $A \in \mathcal{A}$,

$$\int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \; \rho_\pi(d\tilde{\omega}) = \int_{\Omega \times \Pi} \left[\max_{p \in A} U(\omega) \cdot p + c(\pi') \right] \rho_\pi(d\omega, d\pi')$$
$$= \int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega) + c(\pi).$$

Letting $\tilde{\Pi} = \{\rho_\pi : \pi \in \Pi\}$, we see that V can be expressed in the following SSMP form:

$$V(A) = \min_{\rho \in \tilde{\Pi}} \int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \; \rho(d\tilde{\omega}).$$

The idea behind this argument is also a consequence of state-dependence of utility which allows any constant to be absorbed in the integral term. Above, integration against the probability measure π plus the constant $c(\pi)$ is reexpressed as integration against the probability measure ρ_π whose support is a subset of states in $\Omega \times \{\pi\}$ where the utility functions are "shifted" by $c(\pi)\mathbf{1}$.

To see the intuition for $(2) \Rightarrow (1)$, consider an SSMP representation $((\Omega, \mathcal{F}), U, \Pi)$, and define V by Equation (3). By the definition of $\mathcal{U}$, since $U(\omega)$ is an expected-utility function for each $\omega \in \Omega$, there exist $u(\omega) \in \mathcal{U}$, $\alpha(\omega) \geq 0$, and $\beta(\omega) \in \mathbb{R}$ such that

$$U(\omega) = \alpha(\omega)u(\omega) \cdot p + \beta(\omega)\mathbf{1}.$$

For each $\pi \in \Pi$, define the measure μ_π on $\mathcal{U}$ by $\mu_\pi(u) = \sum_{\omega \in \Omega : u(\omega) = u} \alpha(\omega)\pi(\omega)$ for each $u \in \mathcal{U}$. Define the function $c : \Pi \to \mathbb{R}$ by $c(\pi) = \sum_{\omega \in \Omega} \beta(\omega)\pi(\omega)$. Then, for any $A \in \mathcal{A}$,

$$\sum_{\omega \in \Omega} \max_{p \in A} U(\omega) \cdot p \; \pi(\omega) = \sum_{\omega \in \Omega} \max_{p \in A} [u(\omega) \cdot p] \, \alpha(\omega)\pi(\omega) + \sum_{\omega \in \Omega} \beta(\omega)\pi(\omega)$$
$$= \int_{\mathcal{U}} \max_{p \in A} u(p)\mu_\pi(du) + c(\pi).$$

Let $\mathcal{M} = \{\mu_\pi : \pi \in \Pi\}$ and let $\tilde{c}(\mu) = \inf\{c(\pi) : \pi \in \Pi \text{ and } \mu_\pi = \mu\}$. Then, V can be expressed in the following min-HA form:

$$V(A) = \min_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) + \tilde{c}(\mu) \right). \tag{5}$$

It can also be shown that sequentially removing measures from $\mathcal{M}$ that are not strictly optimal in Equation (5) for some $A \in \mathcal{A}$ leads to a minimal set of measures $\mathcal{M}' \subset \mathcal{M}$.

Our SSMP representation bears some similarity to a representation considered by Epstein, Marinacci and Seo (2007, Theorem 1). One main distinction between our representation and theirs is that they impose a normalization on the state-dependent utility function in their representation. As the above arguments illustrate, the key to the equivalence proposed in Theorem 3 is the state-dependence of the utility functions in the SSMP and SSV representations. If a normalization as in Epstein, Marinacci and Seo (2007) were imposed on the utility functions in the SSV and SSMP representations, then the equivalence of these representations would no longer hold, as it would not be possible to "absorb" the cost function of the SSV representation into the utility function to obtain an SSMP representation. Moreover, although these representations would continue to be special cases of the monotone min-HA representation, it would not be possible to write every monotone min-HA representation as an SSV representation since it would not always be possible to "absorb" the magnitude of the measure into the utility function. Theorem 3 illustrates that imposing either a normalization on utility functions (as in the min-HA representation) or a normalization that measures be probabilities (as in the SSMP and SSV representations) is not restrictive; however, imposing both normalizations simultaneously would place a non-trivial additional restriction on the representations.

4.2   Costly Contemplation

A special case of the max-HA representation is one of subjective information acquisition/costly contemplation, where the measures in $\mathcal{M}$ can be interpreted as a reduced-form representation of the individual's information about her tastes and the integral term $\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du)$ can be interpreted as the ex ante value of information μ given menu A. In this application, we need to additionally impose the following axiom:

Axiom 5 (Reversibility of Degenerate Decisions (RDD)) *For any $A \in \mathcal{A}$, $p, q \in \triangle(Z)$, and $\alpha \in [0,1]$,*

$$\beta \delta_{\alpha A + (1-\alpha)\{p\}} + (1-\beta)\delta_{\{q\}} \sim \beta \delta_{\alpha A + (1-\alpha)\{q\}} + (1-\beta)\delta_{\{p\}}$$

where $\beta = 1/(2-\alpha)$.

We will call a choice out of a singleton menu as a degenerate decision. To interpret Axiom 5, consider Figure 2. The first tree represents $\beta \delta_{\alpha A + (1-\alpha)\{p\}} + (1-\beta)\delta_{\{q\}}$, where the individual makes a choice out of the menu $\alpha A + (1-\alpha)\{p\}$ with probability β, and makes a degenerate choice out of the menu $\{q\}$ with probability $1 - \beta$. A choice out of

$$\beta \quad 1-\beta \qquad\qquad \beta \quad 1-\beta$$

$$\alpha A + (1-\alpha)\{p\} \quad \{q\} \qquad \sim \qquad \alpha A + (1-\alpha)\{q\} \quad \{p\}$$

$$\alpha \quad 1-\alpha \qquad\qquad\qquad \alpha \quad 1-\alpha$$

$$\text{choice} \quad p \qquad q \qquad\qquad \text{choice} \quad q \qquad p$$
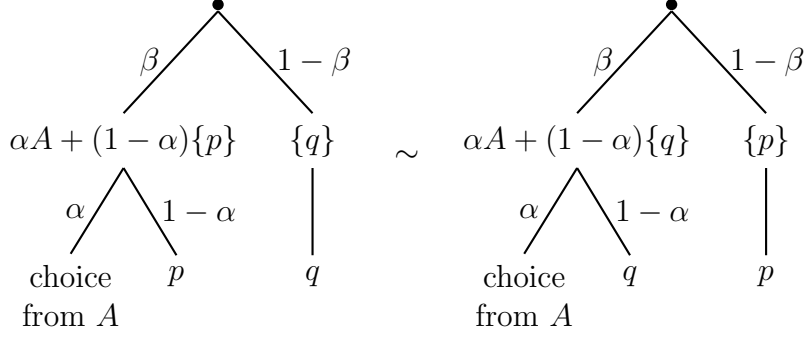$$\text{from } A \qquad\qquad\qquad\qquad \text{from } A$$

Figure 2: Reversibility of Degenerate Decisions $(\beta = 1/(2-\alpha))$

the menu $\alpha A + (1-\alpha)\{p\}$ can be interpreted as a contingent plan, where initially in period 2 the individual determines a lottery out of A, and then her choice out of A is executed with probability α and the fixed lottery p is executed with the remaining $1-\alpha$ probability. Similarly, the second tree represents $\beta\delta_{\alpha A+(1-\alpha)\{q\}} + (1-\beta)\delta_{\{p\}}$ where the roles of p and q are reversed.

If one interprets the individual's behavior as one of costly contemplation/subjective information acquisition, then her optimal contemplation strategy might change as the probability α that her choice out of A is executed changes since her return to contemplation will be higher for higher values of α. However, since the probability that her choice out of A will be executed is the same in both $\alpha A + (1-\alpha)\{p\}$ and $\alpha A + (1-\alpha)\{q\}$, it is reasonable to expect that her contemplation strategy would be the same for both contingent planning problems, although she need not be indifferent between $\delta_{\alpha A+(1-\alpha)\{p\}}$ and $\delta_{\alpha A+(1-\alpha)\{q\}}$ depending on her preference between $\delta_{\{p\}}$ and $\delta_{\{q\}}$. The RDD axiom requires the individual to be indifferent between the two trees in Figure 2 when the probabilities of the paths leading to lotteries p and q are the same, i.e., when $\beta(1-\alpha) = 1-\beta$ or, equivalently, $\beta = 1/(2-\alpha)$.

Given a max-HA representation $(\mathcal{M}, c)$, we show that RDD is equivalent to the following consistency requirement on the set of measures $\mathcal{M}$:

Definition 4 A *Reduced-Form Costly Contemplation (RFCC) representation* is a max-HA representation $(\mathcal{M}, c)$ where the set $\mathcal{M}$ is *consistent*: For each $\mu, \nu \in \mathcal{M}$ and $p \in \triangle(Z)$,

$$\int_{\mathcal{U}} u(p)\, \mu(du) = \int_{\mathcal{U}} u(p)\, \nu(du).$$

An RFCC representation $(\mathcal{M}, c)$ is *monotone* if all the measures in $\mathcal{M}$ are positive.

14

We show in Ergin and Sarver (2009) that the consistency condition above is key for the interpretation of the max-HA representation as a subjective information acquisition problem. More specifically, V satisfies Equation (1) for some monotone RFCC representation $(\mathcal{M}, c)$ if and only if

$$V(A) = \max_{\mathcal{G} \in \mathbf{G}} \left(\mathbb{E} \left[\max_{p \in A} \mathbb{E}[U|\mathcal{G}] \cdot p \right] - c(\mathcal{G}) \right) \tag{6}$$

where $(\Omega, \mathcal{F}, P)$ is a probability space, $U : \Omega \to \mathbb{R}^Z$ is a random vector interpreted as the individual's state-dependent utility, $\mathbf{G}$ is a collection of sub-σ-algebras of $\mathcal{F}$ representing the set of signals that the individual can acquire, and $c(\mathcal{G})$ denotes the cost of subjective signal $\mathcal{G} \in \mathbf{G}$.[13]

As a special case of the max-HA representation, the RFCC representation satisfies PERU. However, it always satisfies indifference to timing of resolution of uncertainty when restricted to temporal lotteries, i.e., for all $p, q \in \triangle(Z)$ and $\alpha \in (0, 1)$:

$$\alpha \delta_{\{p\}} + (1 - \alpha) \delta_{\{q\}} \sim \delta_{\{\alpha p + (1-\alpha)q\}}.[14]$$

Therefore, an individual with RFCC preferences never has a strict PERU unless if she has non-degenerate choices in period 2.

We next present an RFCC representation theorem as an application of Theorem 1. Since RFCC representation is a special case of the max-HA representation, the uniqueness of the RFCC representation is immediately implied by the uniqueness of the max-HA representation.

Theorem 4 *A. The preference $\succsim$ has a RFCC representation if and only if it satisfies Axiom 1, PERU, and RDD.*

B. The preference $\succsim$ has a monotone RFCC representation if and only if it satisfies

---

[13]The costly contemplation representation in Equation (3) is similar to the functional form considered by Ergin (2003), whose primitive is a preference over menus taken from a finite set of alternatives.

[14]This property can also be established directly as a consequence of RDD and first-stage independence. Fix any $p, q \in \triangle(Z)$ and $\alpha \in (0, 1)$. Letting $\beta = 1/(2 - \alpha)$ and $A = \{p\}$, RDD implies

$$\beta \delta_{\{p\}} + (1 - \beta) \delta_{\{q\}} \sim \beta \delta_{\{\alpha p + (1-\alpha)q\}} + (1 - \beta) \delta_{\{p\}}.$$

Since $\beta = 1/(2 - \alpha)$ implies that $\beta = 1 - \beta + \alpha\beta$ and $1 - \beta = (1 - \alpha)\beta$, the left side of this expression is equal to $(1 - \beta) \delta_{\{p\}} + \alpha\beta \delta_{\{p\}} + (1 - \alpha)\beta \delta_{\{q\}}$. Hence,

$$\beta \big[\alpha \delta_{\{p\}} + (1 - \alpha) \delta_{\{q\}} \big] + (1 - \beta) \delta_{\{p\}} \sim \beta \delta_{\{\alpha p + (1-\alpha)q\}} + (1 - \beta) \delta_{\{p\}},$$

which, by first-stage independence, implies $\alpha \delta_{\{p\}} + (1 - \alpha) \delta_{\{q\}} \sim \delta_{\{\alpha p + (1-\alpha)q\}}$.

Axiom 1, PERU, RDD, and monotonicity.

4.3 Kreps and Porteus (1978) and DLR (2001)

In this section, we will introduce a generalization of Kreps and Porteus (1978) and DLR (2001) additive representations, and characterize the intersection of this class with HA preferences. The first axiom we consider is the standard von Neumann-Morgenstern independence axiom imposed on the second-stage uncertainty. It is satisfied by RFCC preferences, but not by SSMP and SSV preferences, since in the latter contexts the individual may benefit from hedging even when she makes no choice in period 2.

Axiom 6 (Second-Stage Independence) *For any $p, q, r \in \triangle(Z)$, and $\alpha \in (0, 1)$,*

$$\delta_{\{p\}} \succ \delta_{\{q\}} \quad \Rightarrow \quad \delta_{\{\alpha p + (1-\alpha)r\}} \succ \delta_{\{\alpha q + (1-\alpha)r\}}.$$

Under weak order and continuity, the following axiom from Kreps (1979) guarantees that the individual is indifferent between any menu and its best singleton subset. Kreps and Porteus (1978) assume the same relationship between the individual's ranking of menus and alternatives.[15]

Axiom 7 (Strategic Rationality) *For any $A, B \in \mathcal{A}$, $\delta_A \succsim \delta_B$ implies $\delta_A \sim \delta_{A \cup B}$.*

Suppose there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Then, the preference $\succsim$ satisfies strategic rationality if and only if

$$V(A) = \max_{p \in A} V(\{p\}). \tag{7}$$

Thus, strategic rationality implies that the restriction of the individual's preference to temporal lotteries determines her entire preference. In particular, if such an individual is indifferent to timing of resolution of uncertainty when choosing among temporal

---

[15]To be precise, Kreps and Porteus (1978) consider both a period 1 preference $\succsim$ over first-stage lotteries in $\triangle(\mathcal{A})$ and a period 2 preference $\succsim_2$ over second-stage lotteries in $\triangle(Z)$. It is easy to show that imposing their temporal consistency axiom (Axiom 3.1 in their paper) on this pair of preferences $(\succsim, \succsim_2)$ implies that the period 1 preference $\succsim$ satisfies strategic rationality. Conversely, if the period 1 preference $\succsim$ satisfies strategic rationality along with continuity, then there exists *some* period 2 preference $\succsim_2$ such that the pair $(\succsim, \succsim_2)$ satisfies their temporal consistency axiom. Moreover, in this case, the period 1 preference $\succsim$ satisfies our second-stage independence axiom if and only if this period 2 preference $\succsim_2$ satisfies the substitution axiom of Kreps and Porteus (1978, Axiom 2.3).

lotteries, then Equation (7) implies that she must always be indifferent to timing of resolution of uncertainty. This is in contrast with RFCC preferences, where the individual is indifferent to timing of resolution of uncertainty when choosing among temporal lotteries, but may exhibit a strict PERU when she faces non-degenerate choices in period 2. Although the setup of Kreps and Porteus (1978) is rich enough to distinguish between attitudes towards timing of resolution of uncertainty for temporal lotteries (without period 2 choice) and more general lotteries over menus (with period 2 choice), the fact that they implicitly impose Axiom 7 throughout their analysis prevents them from doing so.

A second implication of strategic rationality is that it rules out a strict preference for flexibility (Kreps (1979)), i.e., situations where the union of two menus is strictly better than each menu separately: $\delta_{A \cup B} \succ \delta_A$ and $\delta_{A \cup B} \succ \delta_B$. The reason is that, by Equation (7), the individual behaves as if she has no uncertainty in period 1 about her ex post preference ranking over $\triangle(Z)$.

Kreps and Porteus (1978) considered the following representation:[16]

Definition 5 A *Kreps-Porteus representation* is a pair (ϕ, v), where v is an expected-utility function on $\triangle(Z)$ and $\phi : [a, b] \to \mathbb{R}$ is a Lipschitz continuous and strictly increasing function on the bounded interval $[a, b] = \{v(p) : p \in \triangle(Z)\}$, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by:

$$V(A) = \phi\Big(\max_{p \in A} v(p) \Big). \tag{8}$$

The following representation result is a special case of Theorem 1 in Kreps and Porteus (1978):[17]

Theorem 5 *The preference $\succsim$ has a Kreps-Porteus representation if and only if it satisfies Axiom 1, second-stage independence, and strategic rationality.*

The intuition behind Theorem 5 is relatively straightforward. Axiom 1 implies there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$.

---

[16]While this two-period model captures the essence of their representation, Kreps and Porteus (1978) allowed for uncertainty to resolve in a finite number of periods and allowed for consumption in each period. An infinite-horizon recursive formulation of this representation was considered by Epstein and Zin (1989), who also considered several non-expected-utility generalizations.

[17]The only difference is that Kreps and Porteus (1978) only require ϕ to be continuous. We additionally require Lipschitz continuity of ϕ since we impose the L–continuity axiom throughout the paper.

Together with continuity, second-stage independence implies there exists an expected-utility function v on $\triangle(Z)$ such that $v(p) \geq v(q) \iff \delta_{\{p\}} \succsim \delta_{\{q\}}$. Since we also have $V(\{p\}) \geq V(\{q\}) \iff \delta_{\{p\}} \succsim \delta_{\{q\}}$, this implies there exists a strictly increasing function ϕ such that $V(\{p\}) = \phi(v(p))$. By strategic rationality, V satisfies Equation (7), and hence

$$V(A) = \max_{p \in A} \phi\Big(v(p)\Big) = \phi\Big(\max_{p \in A} v(p)\Big).$$

The following axiom from DLR (2001) is the standard independence requirement applied to convex combinations of menus when there is no first-stage uncertainty:

Axiom 8 (Mixture Independence) *For any $A, B, C \in \mathcal{A}$ and $\alpha \in (0,1)$,*

$$\delta_A \succ \delta_B \quad \Rightarrow \quad \delta_{\alpha A + (1-\alpha)C} \succ \delta_{\alpha B + (1-\alpha)C}.$$

It is easy to see that mixture independence is stronger than second-period independence, but in the presence of Axiom 1, it is weaker than the combination of second-period independence and strategic rationality.

We next consider a class of representations that generalize the DLR (2001) additive representation where there is no objective first-stage uncertainty, and Kreps and Porteus (1978) representation where there is no subjective second-stage uncertainty. Unlike the Kreps and Porteus (1978) representation, the following class of representations are compatible with a strict preference for flexibility:

Definition 6 A *Kreps-Porteus-Dekel-Lipman-Rustichini (KPDLR) representation* is a pair (ϕ, μ), where μ is a finite signed Borel measure on $\mathcal{U}$ and $\phi : [a,b] \to \mathbb{R}$ is a Lipschitz continuous and strictly increasing function on the bounded interval $[a,b] = \{\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) : A \in \mathcal{A}\}$, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by:

$$V(A) = \phi\left(\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du)\right). \tag{9}$$

Note that the Kreps-Porteus representation corresponds to the special case of the KPDLR representation in which $\mu = \alpha \delta_u$ for some $u \in \mathcal{U}$ and $\alpha \geq 0$.[18] The following is our KPDLR representation result:

---

[18] It is immediate that any KPDLR representation (ϕ, μ) in which $\mu = \alpha \delta_u$ for $\alpha \geq 0$ can be written as a Kreps-Porteus representation (ϕ, v) where $v = \alpha u$. Conversely, for any Kreps-Porteus representation (ϕ, v), there exist $u \in \mathcal{U}$, $\alpha \geq 0$, and $\beta \in \mathbb{R}$ such that $v(p) = \alpha u(p) + \beta$ for all $p \in \triangle(Z)$. Let $\mu = \alpha \delta_u$. Define constants $\hat{a} = a - \beta$ and $\hat{b} = b - \beta$, and define a function $\hat{\phi} : [\hat{a}, \hat{b}] \to \mathbb{R}$ by $\hat{\phi}(t) = \phi(t + \beta)$. Then,

Theorem 6 *A. The preference $\succsim$ has a KPDLR representation if and only if it satisfies Axiom 1 and mixture independence.*[19]

B. If the preference $\succsim$ has the KPDLR representation (ϕ, μ), then $\succsim$ satisfies PERU [PLRU] if and only if ϕ is convex [concave].

It is easy to see that two KPDLR representations (ϕ, μ) and (ψ, ν) induce the same preference if and only if there exists $\lambda, \alpha > 0$ and $\beta \in \mathbb{R}$ such that $\mu = \lambda \nu$ and $\phi(t) = \alpha \psi(t/\lambda) + \beta$. Since it is possible to have $\int_{\mathcal{U}} u(p)\,\mu(du) = \lambda \int_{\mathcal{U}} u(p)\,\nu(du)$ for all $p \in \triangle(Z)$ even when $\mu \neq \lambda \nu$, this implies in particular that in KPDLR representations, the preference restricted to temporal lotteries does *not* determine the entire preference. This is in contrast to Kreps-Porteus preferences which, by strategic rationality, are determined entirely by their restriction to temporal lotteries.

However, like Kreps-Porteus preferences, KPDLR preferences impose a certain consistency between attitudes towards timing of resolution of uncertainty for temporal lotteries and more general lotteries. Specifically, Equation (9) implies that a preference $\succsim$ with a KPDLR representation exhibits a PERU [PLRU] for temporal lotteries if and only if it exhibits a PERU [PLRU] for lotteries over menus taken from the set $\{A \in \mathcal{A} : \delta_A \sim \delta_{\{p\}} \text{ for some } p \in \triangle(Z)\}$. In particular, if an individual's preference has a KPDLR representation and has the property that for every $A \in \mathcal{A}$ there exists a $p \in \triangle(Z)$ such that $\delta_A \sim \delta_{\{p\}}$, then, as in the case of Kreps-Porteus preferences, her attitude towards timing of resolution of uncertainty is determined entirely by her attitude towards timing of resolution of uncertainty for temporal lotteries.

Not every KPDLR preference will have an HA representation, but the subclass satisfying PERU or PLRU will. We next characterize the subclass of KPDLR preferences which satisfies PERU and PLRU within the class of HA preferences.

Theorem 7 *Let $V : \mathcal{A} \to \mathbb{R}$ and let μ be a nonzero finite signed Borel measure on $\mathcal{U}$. Then, the following are equivalent:*

1. *There exists a KPDLR representation (ϕ, μ) with convex [concave] ϕ such that V is given by Equation (9).*

---

the KPDLR representation $(\hat{\phi}, \mu)$ gives the same value function for menus V as the Kreps-Porteus representation (ϕ, v).

[19]It is not necessary to include indifference to randomization (IR) explicitly in this result since it is implied by mixture independence. Similarly, since mixture independence is implied by the combination of second-stage independence and strategic rationality, it is also not necessary to include IR explicitly in Theorem 5.

2. *There exists a max-HA [min-HA] representation $(\mathcal{M}, c)$ such that V is given by Equation* (1) *[(2)] where:*

(a) $\mathcal{M} \subset \{\lambda\mu : \lambda \in \mathbb{R}_+\}$.

(b) 0 *is not an isolated point of $\mathcal{M}$ and if $0 \in \mathcal{M}$ then*

$$\lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}} \frac{c(\lambda\mu) - c(0)}{\lambda} = \min_{A \in \mathcal{A}} \int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du)$$

$$\left[\lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}} \frac{c(\lambda\mu) - c(0)}{\lambda} = - \max_{A \in \mathcal{A}} \int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) \right].$$

The $(1) \Rightarrow (2.a)$ part of Theorem 7 suggests that if the KPDLR representation satisfies PERU or PLRU, then it is possible to rewrite the function V in Equation (9) as an HA representation where all measures are multiples of the fixed measure μ. In the $(2) \Rightarrow (1)$ part of Theorem 7, condition (2.a) ensures that the transformation ϕ in the KPDLR representation is nondecreasing. Condition (2.b) is merely a technical regularity condition on the derivative of the cost function c at 0 which ensures that ϕ is strictly increasing.

To interpret the HA representation in part (2) of Theorem 7, consider the case of a positive measure μ, and consider again the probability measure π on $\mathcal{V} = \mu(\mathcal{U})\mathcal{U}$ that (heuristically) puts weight $\pi(v) = \mu(u)/\mu(\mathcal{U})$ on each $v = \mu(\mathcal{U})u \in \mathcal{V}$. One interpretation of part (2.a) is that all actions lead to the same distribution π over ex post utilities in $\mathcal{V}$, but each action $\lambda\mu \in \mathcal{M}$ changes the magnitude of the ex post utilities by a common scalar multiple λ.[20] In the special case of Kreps-Porteus preferences, all measures are degenerate and put their weight on the same ex post von Neumann-Morgenstern utility function over $\triangle(Z)$. In this case, the individual has no uncertainty about her ex post preference ranking over lotteries in $\triangle(Z)$, and different actions only affect the strength of her ex post preference.

If a Kreps-Porteus preference $\succsim$ is non-trivial and satisfies PERU, then Theorem 7 implies that it has a max-HA representation $(\mathcal{M}, c)$ such that $\mathcal{M} \subset \{\lambda\delta_u : \lambda \in \mathbb{R}_+\}$ for some $u \in \mathcal{U}$. Following the interpretation given above, define a set of actions by $\Lambda = \{\lambda \in \mathbb{R}_+ : \lambda\delta_u \in \mathcal{M}\}$ and a cost function $\tilde{c} : \Lambda \to \mathbb{R}$ by $\tilde{c}(\lambda) = c(\lambda\delta_u)$. Then,

---

[20]Suppose without loss of generality that $\max\{\lambda : \lambda\mu \in \mathcal{M}\} = 1$. Then, one can also interpret each action $\lambda\mu \in \mathcal{M}$ to lead to the distribution π over $\mathcal{V}$ with probability λ, and to the ex post preference $0 \in \mathbb{R}^Z$ with probability $1 - \lambda$. Under this interpretation, the choice of action affects the probability of the 0 ex post preference, but not the conditional probability distribution π over $\mathcal{V}$.

Equation (1) for this max-HA representation simplifies to the following:[21]

$$V(A) = \max_{\lambda \in \Lambda} \left(\max_{p \in A} \lambda u(p) - \tilde{c}(\lambda) \right).$$ (10)

This formulation of the max-HA representation allows for a direct comparison to the results of Kreps and Porteus (1979). They considered a certain class of hidden action representations and determined the conditions on the representation that result in the corresponding preference $\succsim$ being a Kreps-Porteus preference (i.e., satisfying the axioms of Kreps and Porteus (1978)). Specifically, Propositions 5 and 6 in Kreps and Porteus (1979) show that a hidden action representation represents a Kreps-Porteus preference if and only if it takes a functional form that is essentially equivalent to the one in Equation (10).[22] Their results (and, more generally, our Theorem 7) are useful for determining the instances in which the Kreps-Porteus representation (or KPDLR representation) can be used as a reduced-form representation for a hidden action model.[23] The following example from Kreps and Porteus (1979) illustrates:

Example 1 Consider a consumption-savings problem in which the individual faces lotteries over future (period 2) income. Let Z be a finite subset of $\mathbb{R}$, denoting the possible levels of period 2 income. Let c_1 and c_2 denote the levels of consumption in periods 1 and 2, respectively. If the individual consumes c_1 in period 1, then her period 2 consumption when her realized period 2 income is z is $c_2 = z - c_1$. Suppose the individual has a continuous and additively-separable von Neumann-Morgenstern utility function for consumption $U(c_1, c_2) = U_1(c_1) + U_2(c_2)$. Suppose that c_1 is chosen from some compact interval C after the realization period 1 uncertainty but before the realization of period 2 uncertainty about income z. Defining $\mathcal{A}$ in the usual way for this set Z, the induced preferences for lotteries over income have the following representation: For any two lotteries $P, Q \in \triangle(\mathcal{A})$, $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is

---

[21]Theorem 7 also implies that if $0 \in \Lambda$, then 0 is not an isolated point and $\tilde{c}'(0) = \min_{p \in \triangle(Z)} u(p)$. As noted above, this condition implies that the mapping $r \mapsto \max_{\lambda \in \Lambda} (\lambda r - \tilde{c}(\lambda))$ is strictly increasing for all $r \in \{u(p) : p \in \triangle(Z)\}$.

[22]Kreps and Porteus (1979) restricted attention to temporal lotteries, and therefore their representation did not include a period 2 choice of lottery p from a menu A. Kreps and Porteus (1979) also allowed for consumption in period 1; this is possible in our model as well by making some minor changes to the functional form of the representation.

[23]It has also been suggested that a preference for early resolution of uncertainty could arise from anticipatory feelings. However, Epstein (2008) develops a model of anticipatory feelings and shows that while such preferences can lead to non-indifference to temporal resolution of uncertainty, they are not consistent with the model of Kreps and Porteus (1978) except in the trivial case of indifference to timing of resolution of uncertainty.

defined by

$$V(A) = \max_{c_1 \in C} \left(U_1(c_1) + \max_{p \in A} \sum_{z \in Z} U_2(z - c_1)p_z \right).$$

Kreps and Porteus (1979) observed that a sufficient condition for this representation to take the form described in Equation (10) is that $U_2(c_2) = -\exp(-\alpha c_2)$ for some $\alpha > 0$. This can easily be seen by letting $\Lambda = \{\exp(\alpha c_1) : c_1 \in C\}$, defining $\tilde{c} : \Lambda \to \mathbb{R}$ by $\tilde{c}(\lambda) = -U_1(\frac{1}{\alpha} \ln(\lambda))$, and defining $u \in \mathbb{R}^Z$ by $u(z) = -\exp(-\alpha z)$.[24] Intuitively, if U_2 has constant absolute risk aversion, then the choice of c_1 does not affect the individual's ranking of lotteries p over Z. This implies that the ex post expected-utility functions on $\triangle(Z)$ resulting from the various choices of c_1 must be affine transformations of a single fixed utility function, which is precisely the content of Equation (10).

One could ask a dual question to the one considered in Example 1: Instead of determining what conditions on U_2 are needed to ensure that the resulting preference over $\triangle(\mathcal{A})$ is a Kreps-Porteus preference, we could ask whether there is a class of preferences that can accommodate any choice of U_2. Clearly, the class of all max-HA preferences is sufficiently general for this purpose. However, notice that the induced preference in this example will satisfy strategic rationality regardless of the choice of U_2. The following theorem illustrates that maintaining strategic rationality while relaxing second-stage independence results in a natural generalization of Kreps-Porteus preferences that can accommodate the preferences in Example 1 for any choice of parameters:[25]

Theorem 8 *The preference $\succsim$ satisfies Axiom 1, strategic rationality, and PERU if and only if it has a max-HA representation $(\mathcal{M}, c)$ such that $\mathcal{M} \subset \{\lambda \delta_u : \lambda \in \mathbb{R}_+, u \in \mathcal{U}\}$.*[26]

We now sketch the proof Theorem 8. Axiom 1 implies there exists a Lipschitz continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Strategic rationality implies that V satisfies Equation (7), and PERU implies that V is convex. Define $f : \triangle(Z) \to \mathbb{R}$ by $f(p) = V(\{p\})$. Then, f is convex and $V(A) = \max_{p \in A} f(p)$. By the same duality results as those used to prove Theorem 1, f can be expressed as the maximum of the set of all affine function lying below it. Since affine functions on

---

[24]Although this definition implies that $u \notin \mathcal{U}$, u can be normalized to be in $\mathcal{U}$.

[25]It is well-known that second-stage independence will in general be violated if the individual takes a payoff-relevant action prior to the resolution of uncertainty; for instance, see Markowitz (1959, Chapters 10–11), Mossin (1969), and Spence and Zeckhauser (1972). The results of Kreps and Porteus (1979) discussed above characterize precisely those special cases in which independence is not violated.

[26]It is not necessary to include IR explicitly in this result because it is implied by the combination of weak order, continuity, first-stage independence, PERU, and strategic rationality.

$\triangle(Z)$ are precisely expected-utility functions, this implies there exists a set $\mathcal{V} \subset \mathbb{R}^Z$ such that $f(p) = \max_{v \in \mathcal{V}} v(p)$. The observation that f can be given dual representation of this kind is well-known; for example, see Machina (1984, Theorem 2).[27] Therefore, if $\succsim$ satisfies the axioms of Theorem 8, then V takes the following form:

$$V(A) = \max_{p \in A} f(p) = \max_{v \in \mathcal{V}} \left(\max_{p \in A} v(p) \right). \tag{11}$$

The final step of the proof is to transform the functional form in Equation (11) into a max-HA representation. By the definition of $\mathcal{U}$, for each $v \in \mathcal{V}$, there exist $\bar{u} \in \mathcal{U}$, $\lambda \geq 0$, and $\beta \in \mathbb{R}$ such that $v(p) = \lambda \bar{u}(p) + \beta$ for all $p \in \triangle(Z)$. Thus, defining a measure $\mu_v = \lambda \delta_{\bar{u}}$ and letting $\tilde{c}(v) = -\beta$, we have

$$\max_{p \in A} v(p) = \int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu_v(du) - \tilde{c}(v).$$

Let $\mathcal{M} = \{\mu_v : v \in \mathcal{V}\}$ and define $c : \mathcal{M} \to \mathbb{R}$ by $c(\mu) = \inf\{\tilde{c}(v) : \mu_v = \mu\}$. Then, Equation (11) can be written as

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

Therefore, $(\mathcal{M}, c)$ is a max-HA representation for $\succsim$.[28]

---

[27] Our motivation is also very similar to that of Machina (1984), who used this dual representation to perform local expected-utility analysis in a model of induced preferences over temporal lotteries over levels of future wealth.

[28] A complete proof, including the arguments for obtaining a compact and minimal set $\mathcal{M}$ and a lower semi-continuous function c, are contained in the appendix.

Appendix

A   Mathematical Preliminaries

In this section, we present 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. Throughout this section, let X be a real Banach space, and let X^* denote the space of all continuous linear functionals on X.

Definition 7 Suppose $C \subset X$. A function $f : C \to \mathbb{R}$ is said to be *Lipschitz continuous* if there is some real number K such that $|f(x) - f(y)| \leq K\|x - y\|$ for every $x, y \in C$. The number K is called a *Lipschitz constant* of f.

We now introduce the standard definition of the subdifferential of a function.

Definition 8 Suppose $C \subset X$ and $f : C \to \mathbb{R}$. 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 useful for the approximation of convex functions by affine functions. It is straightforward to show that $x^* \in \partial f(x)$ if and only if the affine function $h : X \to \mathbb{R}$ defined by $h(y) = f(x) + \langle y - x, x^* \rangle$ 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 result shows that a Lipschitz continuous and convex function always has a nonempty subdifferential:

Lemma 1 (Ergin and Sarver (2008)) *Suppose C is a convex subset of a Banach space X. If $f : C \to \mathbb{R}$ is Lipschitz continuous and convex, then $\partial f(x) \neq \emptyset$ for all $x \in C$.*

We now introduce the definition of the conjugate of a function.

Definition 9 Suppose $C \subset X$ and $f : C \to \mathbb{R}$. The *conjugate* (or *Fenchel conjugate*) of f is the function $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$ defined by

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

There is an important duality between f and f^*. Lemma 2 summarizes certain properties of f^* that are useful in establishing this duality.[29] We include a proof for completeness.

---

[29]For a complete discussion of the relationship between f and f^*, see Ekeland and Turnbull (1983) or Holmes (1975). A finite-dimensional treatment can be found in Rockafellar (1970).

Lemma 2 *Suppose $C \subset X$ and $f : C \to \mathbb{R}$. Then,*

1. *f^* is lower semi-continuous 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 semi-continuous.

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

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

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

$$\langle y, x^* \rangle - f(y) \leq \langle x, x^* \rangle - f(x). \tag{12}$$

for all $y \in C$. By the definition of the conjugate, Equation (12) holds if and only if $f^*(x^*) = \langle x, x^* \rangle - f(x)$, which is equivalent to $f(x) = \langle x, x^* \rangle - f^*(x^*)$. ∎

Suppose that $C \subset X$ is convex and $f : C \to \mathbb{R}$ is Lipschitz continuous and convex. As noted above, this implies that $\partial f(x) \neq \emptyset$ for all $x \in C$. Therefore, by parts 2 and 3 of Lemma 2, we have

$$f(x) = \max_{x^* \in X^*} \left[\langle x, x^* \rangle - f^*(x^*) \right] \tag{13}$$

for all $x \in C$.[30] In order to establish the existence of a minimal set of measures in the proof of Theorem 1, it is useful to establish that under certain assumptions, there is a minimal compact subset of X^* for which Equation (13) holds. Let C_f denote the set of all $x \in C$ for which the subdifferential of f at x is a singleton:

$$C_f = \{x \in C : \partial f(x) \text{ is a singleton}\}. \tag{14}$$

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

$$\mathcal{N}_f = \{x^* \in X^* : x^* \in \partial f(x) \text{ for some } x \in C_f\}. \tag{15}$$

---

[30]This is a slight variation of the classic Fenchel-Moreau theorem. The standard version of this theorem states that if $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous 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).

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

$$\mathcal{M}_f = \overline{\mathcal{N}_f}. \tag{16}$$

Before stating our first main result, recall that the *affine hull* of a set $C \subset X$, denoted $\mathrm{aff}(C)$, is defined to be the smallest affine subspace of X that contains C. Also, a set $C \subset X$ is said to be a *Baire space* if every countable intersection of dense open subsets of C is dense.

Theorem 9 (Ergin and Sarver (2008)) *Suppose (i) X is a separable Banach space, (ii) C is a convex subset of X that is a Baire space (when endowed with the relative topology) such that $\mathrm{aff}(C)$ is dense in X,[31] and (iii) $f : C \to \mathbb{R}$ is Lipschitz continuous and convex. Then, $\mathcal{M}_f$ is weak* compact, and for any weak* compact $\mathcal{M} \subset X^*$,*

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

The intuition for Theorem 9 is fairly simple. We already know from Lemma 2 that for any $x \in C_f$, $f(x) = \max_{x^* \in \mathcal{N}_f} [\langle x, x^* \rangle - f^*(x^*)]$. Ergin and Sarver (2008) show that under the assumptions of Theorem 9, C_f is dense in C. Therefore, it can be shown that for any $x \in C$,

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

In addition, if $\mathcal{M}$ is a weak* compact subset of X^* and $\mathcal{M}_f$ is not a subset of $\mathcal{M}$, then there exists $x^* \in \mathcal{N}_f$ such that $x^* \notin \mathcal{M}$. That is, there exists $x \in C_f$ such that $\partial f(x) = \{x^*\}$ and $x^* \notin \mathcal{M}$. Therefore, Lemma 2 implies $f(x) > \max_{x^* \in \mathcal{M}} [\langle x, x^* \rangle - f^*(x^*)]$.

In the proof of Theorem 1, we will construct an HA representation in which $\mathcal{M}_f$, for a certain function f, is the set of measures. In the proof of part B of Theorem 1, we will use the following result to establish that monotonicity leads to a positive set of measures. For this next result, assume that X is a Banach lattice.[32] Let $X_+ = \{x \in X : x \geq 0\}$ denote the *positive cone* of X. A function $f : C \to \mathbb{R}$ 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 $x \in X_+$.

Theorem 10 (Ergin and Sarver (2008)) *Suppose C is a convex subset of a Banach lattice X, such that at least one of the following conditions holds:*

1. $x \vee x' \in C$ for any $x, x' \in C$, or

2. $x \wedge x' \in C$ for any $x, x' \in C$.

Let $f : C \to \mathbb{R}$ be Lipschitz continuous, convex, and monotone. Then, the functionals in $\mathcal{M}_f$ are positive.

---

[31]In particular, if C is closed, then by the Baire Category theorem, then C is a Baire space. Also, note that if C contains the origin, then the affine hull of C is equal to the span of C.

[32]See Aliprantis and Border (1999, p302) for a definition of Banach lattices.

Finally, the following result will be used in the proof of Theorem 2 to establish the uniqueness of the HA representation.

Theorem 11 (Ergin and Sarver (2008)) *Suppose X is a Banach space and C is a convex subset of X. Let $\mathcal{M}$ be a weak* compact subset of X^*, and let $c : \mathcal{M} \to \mathbb{R}$ be weak* lower semi-continuous. Define $f : C \to \mathbb{R}$ by*

$$f(x) = \max_{x^* \in \mathcal{M}} \left[\langle x, x^* \rangle - c(x^*) \right]. \tag{17}$$

Then,

1. *The function f is Lipschitz continuous and convex.*

2. *For all $x \in C$, there exists $x^* \in \partial f(x)$ such that $x^* \in \mathcal{M}$ and $f^*(x^*) = c(x^*)$. In particular, this implies $\mathcal{N}_f \subset \mathcal{M}$, $\mathcal{M}_f \subset \mathcal{M}$, and $f^*(x^*) = c(x^*)$ for all $x^* \in \mathcal{N}_f$.*

3. *If C is also compact (in the norm topology), then $f^*(x^*) = c(x^*)$ for all $x^* \in \mathcal{M}_f$.*

B   Proof of Theorem 1

Note that $\mathcal{A}$ is a compact metric space since $\triangle(Z)$ is a compact metric space (see, e.g., Munkres (2000, p280–281) or Theorem 1.8.3 in Schneider (1993, p49)). We begin by showing that weak order, continuity, and first-stage independence imply that $\succsim$ has an expected-utility representation.

Lemma 3 *A preference $\succsim$ over $\triangle(\mathcal{A})$ satisfies weak order, continuity, and first-stage independence if and only if there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $\succsim$ is represented by $\mathbb{E}_P[V]$. Furthermore, if $V : \mathcal{A} \to \mathbb{R}$ and $V' : \mathcal{A} \to \mathbb{R}$ are continuous functions such that $\mathbb{E}_P[V]$ and $\mathbb{E}_P[V']$ represent the same preference over $\triangle(\mathcal{A})$, then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V + \beta$.*

Proof: This is a standard result. For example, it is asserted without proof in Corollary 5.22 of Kreps (1988). Alternatively, it can be verified that $\triangle(\mathcal{A})$ and $\succsim$ satisfy the conditions of Theorem 10.1 of Fishburn (1970), and hence there exists a bounded V such that $\succsim$ is represented by $\mathbb{E}_P[V]$. Since the mapping $P \mapsto \mathbb{E}_P[V]$ is affine, it is straightforward to show that the continuity axiom implies this mapping is weak* continuous. Therefore, for any sequence $\{A_n\} \subset \mathcal{A}$ and any $A \in \mathcal{A}$,

$$A_n \to A \implies \delta_{A_n} \xrightarrow{w^*} \delta_A \implies V(A_n) = \mathbb{E}_{\delta_{A_n}}[V] \to \mathbb{E}_{\delta_A}[V] = V(A),$$

which implies that V is continuous. Uniqueness of V follows from the uniqueness part of the mixture-space theorem (see Kreps (1988, Theorem 5.11) or Fishburn (1970, Theorem 8.4)). ∎

27

Let $\mathcal{A}^c \subset \mathcal{A}$ denote the collection of all convex menus. 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 (see Theorem 1.8.5 in Schneider (1993, p50)). Our strategy for proving the sufficiency of the axioms will be to show that the function V described in Lemma 3 satisfies the max-HA [min-HA] formula on $\mathcal{A}^c$. Using the IR axiom, it will then be straightforward to show that V satisfies the max-HA [min-HA] formula on all of $\mathcal{A}$.

The following lemma shows the implications of our other axioms.

Lemma 4 *Suppose that $V : \mathcal{A} \to \mathbb{R}$ is a continuous function such that $\mathbb{E}_P[V]$ represents the preference $\succsim$ over $\triangle(\mathcal{A})$. Then:*

1. *If $\succsim$ satisfies L–continuity, then V is Lipschitz continuous on $\mathcal{A}^c$, i.e., there exists $K \geq 0$ such that $|V(A) - V(B)| \leq Kd_h(A, B)$ for any $A, B \in \mathcal{A}^c$.[33]*

2. *If V is Lipschitz continuous (on $\mathcal{A}$), then $\succsim$ satisfies L–continuity.*

3. *The preference $\succsim$ satisfies PERU [PLRU] if and only if V is convex [concave].*

4. *The preference $\succsim$ satisfies monotonicity if and only if V is monotone (i.e., $A \subset B$ implies $V(B) \geq V(A)$ for any $A, B \in \mathcal{A}$).*

Proof: Claims 3 and 4 follow immediately from the definitions. To prove claim 1, we use the arguments in the proof of Lemma 13 in Ergin and Sarver (2009). Suppose that $\succsim$ satisfies L–continuity for $M \geq 0$ and $A^*, A_* \in \mathcal{A}$. First, note that if $M = 0$, then L–continuity implies that $V(A) = V(B)$ for all $A, B \in \mathcal{A}$, i.e., V is Lipschitz continuous with a Lipschitz constant $K = 0$. If $M > 0$, then let $K \equiv 2M[V(A^*) - V(A_*)] \geq 0$. We first show that for any $A, B \in \mathcal{A}^c$:

$$d_h(A, B) \leq \tfrac{1}{2M} \implies |V(A) - V(B)| \leq Kd_h(A, B). \tag{18}$$

Suppose that $d_h(A, B) \leq \frac{1}{2M}$ and let $\alpha \equiv Md_h(A, B)$. Then, $\alpha \leq 1/2$ and

$$V(B) - V(A) \leq \tfrac{\alpha}{1-\alpha}[V(A^*) - V(A_*)] \leq 2\alpha[V(A^*) - V(A_*)] = Kd_h(A, B),$$

where the first inequality follows from L–continuity, the second inequality follows from $\alpha \leq 1/2$, and the equality follows form the definitions of α and K. Interchanging the roles of A and B above, we also have that $V(A) - V(B) \leq Kd_h(A, B)$, proving Equation (18).

Next, we use the argument in the proof of Lemma 8 in the supplementary appendix of DLRS (2007) to show that for any $A, B \in \mathcal{A}^c$:

$$|V(A) - V(B)| \leq Kd_h(A, B), \tag{19}$$

i.e., the requirement $d_h(A, B) \leq \frac{1}{2M}$ in Equation (18) is not necessary. To see this, take any sequence $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_n < \lambda_{n+1} = 1$ such that $(\lambda_{i+1} - \lambda_i)d_h(A, B) \leq \frac{1}{2M}$. Let

---

[33]If $\succsim$ also satisfies IR, then it can be shown that V is Lipschitz continuous on $\mathcal{A}$.

$A_i = \lambda_i A + (1 - \lambda_i)B$. It is straightforward to verify that:[34]

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

Combining this with the triangle inequality and Equation (18), we obtain

$$|V(A) - V(B)| \leq \sum_{i=0}^{n} |V(A_{i+1}) - V(A_i)|$$

$$\leq K \sum_{i=0}^{n} d_h(A_{i+1}, A_i) = K \sum_{i=0}^{n} (\lambda_{i+1} - \lambda_i)d_h(A, B) = K d_h(A, B).$$

This establishes Equation (19), which implies V is Lipschitz continuous on $\mathcal{A}^c$ with a Lipschitz constant K.

To prove claim 2, suppose that V is Lipschitz continuous, and let $K > 0$ be a Lipschitz constant of V. Let A^* be a maximizer of V on $\mathcal{A}$ and let A_* be a minimizer of V on $\mathcal{A}$. If $V(A^*) = V(A_*)$, then $P \sim Q$ for any $P, Q \in \triangle(\mathcal{A})$, implying that L–continuity holds trivially for A^*, A_*, and $M = 0$. If $V(A^*) > V(A_*)$, then let $M \equiv K/[V(A^*) - V(A_*)] > 0$. For any $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$ with $\alpha \geq M d_h(A, B)$, we have

$$(1 - \alpha)[V(B) - V(A)] \leq V(B) - V(A) \leq K d_h(A, B) \leq K\alpha/M = \alpha[V(A^*) - V(A_*)],$$

which implies the conclusion of L–continuity. ∎

We now 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. As in the text, let

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

For any $A \in \mathcal{A}^c$, the support function $\sigma_A : \mathcal{U} \to \mathbb{R}$ 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

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

Lemma 5   *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.*

---

[34]Note that the convexity of the menus A and B is needed for the first equality.

2. *For all $A, B \in \mathcal{A}^c$ and any $\lambda \in [0,1]$, $\sigma_{\lambda A + (1-\lambda)B} = \lambda \sigma_A + (1-\lambda)\sigma_B$.*

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

4. *Σ is convex and compact, and $0 \in \Sigma$.*

Proof: Parts 1–3 are standard results that can be found in Rockafellar (1970) or Schneider (1993).[35] For instance, in Schneider (1993), part 1 follows from Theorem 1.7.1, part 2 follows from Theorem 1.7.5, and part 3 follows from Theorem 1.8.11.

For part 4, note that the set Σ is convex by the convexity of $\mathcal{A}^c$ and part 2 of this lemma. As discussed above, the set $\mathcal{A}^c$ is compact, and hence by parts 1 and 3 of this lemma, Σ is a compact subset of the Banach space $C(\mathcal{U})$. Also, if we take $q = (1/|Z|, \ldots, 1/|Z|) \in \triangle(Z)$, then $u \cdot q = 0$ for all $u \in \mathcal{U}$. This implies $\sigma_{\{q\}} = 0$, and hence $0 \in \Sigma$. ∎

The following lemma shows that a function defined on $\mathcal{A}^c$ can be transformed into a function on Σ.

Lemma 6 *Suppose $V : \mathcal{A}^c \to \mathbb{R}$, and define a function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$. Then:*

1. *$V(A) = W(\sigma_A)$ for all $A \in \mathcal{A}^c$.*

2. *V is Lipschitz continuous if and only if W is Lipschitz continuous.*

3. *If V is convex [concave] if and only if W is convex [concave].*

4. *V is monotone if and only if W is monotone (i.e., $\sigma \leq \sigma'$ implies $W(\sigma) \leq W(\sigma')$ for any $\sigma, \sigma' \in \Sigma$).*

Proof: (1): This follows immediately from part 1 of Lemma 5.

(2): If V is Lipschitz continuous with a Lipschitz constant $K \geq 0$, then by parts 1 and 3 of Lemma 5, for any $A, B \in \mathcal{A}^c$,

$$|W(\sigma_A) - W(\sigma_B)| = |V(A) - V(B)| \leq K d_h(A, B) = K \|\sigma_A - \sigma_B\|_\infty.$$

A similar argument shows if W Lipschitz continuous, then V is Lipschitz continuous.

(3): If V is convex, then by parts 1 and 2 of Lemma 5, for any $A, B \in \mathcal{A}^c$ and $\lambda \in [0,1]$,

$$W(\lambda \sigma_A + (1-\lambda)\sigma_B) = W(\sigma_{\lambda A + (1-\lambda)B}) = V(\lambda A + (1-\lambda)B)$$
$$\leq \lambda V(A) + (1-\lambda)V(B) = \lambda W(\sigma_A) + (1-\lambda)W(\sigma_B).$$

Similar arguments can be used to show that convexity of W implies convexity of V and that V is concave if and only if W is concave.

---

[35]The standard setting for support functions is the set of nonempty closed and convex subsets of $\mathbb{R}^n$. However, by imposing our normalizations on the domain of the support functions $\mathcal{U}$, the standard results are easily adapted to our setting of nonempty closed and convex subsets of $\triangle(Z)$.

(4): This is an immediate consequence of the following fact, which is easy to see from part 1 of Lemma 5 and the definitions of σ_A and A_σ: For all $A, B \in \mathcal{A}^c$, $A \subset B$ if and only if $\sigma_A \leq \sigma_B$. ∎

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 $\mu \in C(\mathcal{U})^*$.[36]

For any function $W : \Sigma \to \mathbb{R}$, define the subdifferential ∂W and the conjugate W^* as in Appendix A. Also, define $\Sigma_W, \mathcal{N}_W$, and $\mathcal{M}_W$ as in Equations (14), (15), and (16), respectively:

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

where the closure is taken with respect to the weak* topology. We now apply Theorem 9 to the current setting.

Lemma 7 *Suppose $W : \Sigma \to \mathbb{R}$ is Lipschitz continuous and convex. Then, $\mathcal{M}_W$ is weak* compact, and for any weak* compact $\mathcal{M} \subset C(\mathcal{U})^*$,*

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

Proof: We simply need to verify that $C(\mathcal{U})$, Σ, and W satisfy the assumptions of Theorem 9. Since $\mathcal{U}$ is a compact metric space, $C(\mathcal{U})$ is separable (see Theorem 8.48 of Aliprantis and Border (1999)). By part 4 of Lemma 5, Σ is a closed and convex subset of $C(\mathcal{U})$ containing the origin. Since Σ is a closed subset of a Banach space, it is a Baire space by the Baire Category theorem. Although the result is stated slightly differently, it is shown in Hörmander (1954) that span(Σ) is dense in $C(\mathcal{U})$. This result is also proved in DLR (2001). Since $0 \in \Sigma$ implies that aff(Σ) = span(Σ), the affine hull of Σ is therefore dense in $C(\mathcal{U})$. Finally, W is Lipschitz continuous and convex by assumption. ∎

---

[36]Since $\mathcal{U}$ 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.

B.1 Sufficiency of the axioms for the max-HA representations

To prove the sufficiency of the axioms for the max-HA representation in part A, suppose that $\succsim$ satisfies Axiom 1 and PERU. By Lemma 3, there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $\mathbb{E}_P[V]$ represents $\succsim$. Moreover, by Lemma 4, the restriction of V to the set $\mathcal{A}^c$ of convex menus is Lipschitz continuous and convex. With slight abuse of notation, we also denote this restriction by V. By Lemma 6, the function $W : \Sigma \to \mathbb{R}$ defined by $W(\sigma) = V(A_\sigma)$ is Lipschitz continuous and convex. Therefore, by Lemma 7, for all $\sigma \in \Sigma$,

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

This implies that for all $A \in \mathcal{A}$,

$$
\begin{aligned}
V(A) = V(\mathrm{co}(A)) &= W(\sigma_{\mathrm{co}(A)}) \\
&= \max_{\mu \in \mathcal{M}_W} \left(\int_{\mathcal{U}} \max_{p \in \mathrm{co}(A)} u(p) \, \mu(du) - W^*(\mu) \right) \\
&= \max_{\mu \in \mathcal{M}_W} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) - W^*(\mu) \right),
\end{aligned}
$$

where the first equality follows from IR and the second equality follows from part 1 of Lemma 6. The function W^* is lower semi-continuous by part 1 of Lemma 2, and $\mathcal{M}_W$ is compact by Lemma 7. It is also immediate from Lemma 7 that $\mathcal{M}_W$ satisfies the minimality condition in Definition 1. Therefore, $(\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ is a max-HA representation for $\succsim$.

To prove the sufficiency of the axioms for the monotone max-HA representation in part B, suppose that, in addition, $\succsim$ satisfies monotonicity. Then, by Lemmas 4 and 6, the function W is monotone. Also, note that for any $A, B \in \mathcal{A}^c$, $\sigma_A \vee \sigma_B = \sigma_{A \cup B}$. Hence, $\sigma \vee \sigma' \in \Sigma$ for any $\sigma, \sigma' \in \Sigma$. Therefore, by Theorem 10, the measures in $\mathcal{M}_W$ are positive.

B.2 Sufficiency of the axioms for the min-HA representations

To prove the sufficiency of the axioms for the min-HA representation in part A, suppose that $\succsim$ satisfies Axiom 1 and PLRU. By Lemma 3, there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $\mathbb{E}_P[V]$ represents $\succsim$. By Lemmas 4 and 6, the function $W : \Sigma \to \mathbb{R}$ defined by $W(\sigma) = V(A_\sigma)$ is Lipschitz continuous and concave. Define a function $\bar{W} : \Sigma \to \mathbb{R}$ by $\bar{W}(\sigma) = -W(\sigma)$. Then, $\bar{W}$ is Lipschitz continuous and convex, so by Lemma 7, for all $\sigma \in \Sigma$,

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

Let $\mathcal{M} \equiv -\mathcal{M}_{\bar{W}} = \{-\mu : \mu \in \mathcal{M}_{\bar{W}}\}$, and define $c : \mathcal{M} \to \mathbb{R}$ by $c(\mu) = \bar{W}^*(-\mu)$. Then, for any $\sigma \in \Sigma$,

$$W(\sigma) = -\bar{W}(\sigma) = \min_{\mu \in \mathcal{M}_{\bar{W}}} \left[-\langle \sigma, \mu \rangle + \bar{W}^*(\mu) \right]$$
$$= \min_{\mu \in \mathcal{M}} \left[-\langle \sigma, -\mu \rangle + \bar{W}^*(-\mu) \right]$$
$$= \min_{\mu \in \mathcal{M}} \left[\langle \sigma, \mu \rangle + c(\mu) \right].$$

This implies that for all $A \in \mathcal{A}$,

$$V(A) = V(\mathrm{co}(A)) = W(\sigma_{\mathrm{co}(A)})$$
$$= \min_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in \mathrm{co}(A)} u(p)\, \mu(du) + c(\mu) \right)$$
$$= \min_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) + c(\mu) \right),$$

where the first equality follows from IR and the second equality follows from part 1 of Lemma 6. The function $\bar{W}^*$ is lower semi-continuous by part 1 of Lemma 2, which implies that c is lower semi-continuous. The compactness of $\mathcal{M}$ follows from the compactness of $\mathcal{M}_{\bar{W}}$, which follows from Lemma 7. Also, by Lemma 7 and the above construction, it is immediate that $\mathcal{M}$ satisfies the minimality condition in Definition 1. Therefore, $(\mathcal{M}, c)$ is a min-HA representation for $\succsim$.

To prove the sufficiency of the axioms for the monotone min-HA representation in part B, suppose that, in addition, $\succsim$ satisfies monotonicity. Then, by Lemmas 4 and 6, the function W is monotone. Let $\hat{\Sigma} \equiv -\Sigma = \{-\sigma : \sigma \in \Sigma\}$, and define a function $\hat{W} : \hat{\Sigma} \to \mathbb{R}$ by $\hat{W}(\sigma) \equiv \bar{W}(-\sigma) = -W(-\sigma)$. Notice that $\hat{W}$ is monotone and convex: By the monotonicity of W, for any $\sigma, \sigma' \in \hat{\Sigma}$,

$$\sigma \le \sigma' \implies -\sigma \ge -\sigma' \implies \hat{W}(\sigma) = -W(-\sigma) \le -W(-\sigma') = \hat{W}(\sigma).$$

By the concavity of W, for any $\sigma, \sigma' \in \hat{\Sigma}$ and $\lambda \in [0, 1]$,

$$\hat{W}(\lambda \sigma + (1 - \lambda)\sigma') = -W(\lambda(-\sigma) + (1 - \lambda)(-\sigma'))$$
$$\le -\lambda W(-\sigma) - (1 - \lambda)W(-\sigma') = \lambda \hat{W}(\sigma) + (1 - \lambda)\hat{W}(\sigma').$$

Also, for any $A, B \in \mathcal{A}^c$, $(-\sigma_A) \wedge (-\sigma_B) = -(\sigma_A \vee \sigma_B) = -\sigma_{A \cup B}$. Hence, $\sigma \wedge \sigma' \in \hat{\Sigma}$ for any $\sigma, \sigma' \in \hat{\Sigma}$. Therefore, by Theorem 10, the measures in $\mathcal{M}_{\hat{W}}$ are positive. For any $\mu \in C(\mathcal{U})^*$ and $\sigma, \sigma' \in \hat{\Sigma}$, note that

$$\hat{W}(\sigma') - \hat{W}(\sigma) \ge \langle \sigma' - \sigma, \mu \rangle \iff \bar{W}(-\sigma') - \bar{W}(-\sigma) \ge \langle \sigma' - \sigma, \mu \rangle = \langle -\sigma' + \sigma, -\mu \rangle,$$

and hence $\mu \in \partial \hat{W}(\sigma) \iff -\mu \in \partial \bar{W}(-\sigma)$. In particular, $\hat{\Sigma}_{\hat{W}} = -\Sigma_{\bar{W}}$ and $\mathcal{N}_{\hat{W}} = -\mathcal{N}_{\bar{W}}$.

Taking closures, we have $\mathcal{M}_{\hat{W}} = -\mathcal{M}_{\bar{W}} = \mathcal{M}$. Thus, the measures in $\mathcal{M}$ are positive.

B.3 Necessity of the axioms

We begin by demonstrating some of the properties of the function V defined by an HA representation.

Lemma 8 *Suppose $(\mathcal{M}, c)$ is an HA representation.*

1. *If $(\mathcal{M}, c)$ is a max-HA representation and $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (1), then V is Lipschitz continuous and convex. In addition, defining the function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$, we have $\mathcal{M} = \mathcal{M}_W$ and $c(\mu) = W^*(\mu)$ for all $\mu \in \mathcal{M}$.*

2. *If $(\mathcal{M}, c)$ is a min-HA representation and $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (2), then V is Lipschitz continuous and concave. In addition, defining the function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$, we have $\mathcal{M} = -\mathcal{M}_{-W}$ and $c(\mu) = [-W]^*(-\mu)$ for all $\mu \in \mathcal{M}$.*

Proof: (1): By the definitions of V and W, we have

$$W(\sigma) = \max_{\mu \in \mathcal{M}} \big[\langle \sigma, \mu \rangle - c(\mu) \big], \quad \forall \sigma \in \Sigma.$$

By part 1 of Theorem 11, W is Lipschitz continuous and convex. Therefore, the restriction of V to $\mathcal{A}^c$ is Lipschitz continuous and convex by Lemma 6. Let $K \geq 0$ be any Lipschitz constant of $V|_{\mathcal{A}^c}$, and take any $A, B \in \mathcal{A}$. It is easily verified that $V(A) = V(\mathrm{co}(A))$, $V(B) = V(\mathrm{co}(B))$, and $d_h(\mathrm{co}(A), \mathrm{co}(B)) \leq d_h(A, B)$. Hence,

$$|V(A) - V(B)| = |V(\mathrm{co}(A)) - V(\mathrm{co}(B))| \leq K d_h(\mathrm{co}(A), \mathrm{co}(B)) \leq K d_h(A, B),$$

which implies that V is Lipschitz continuous on all of $\mathcal{A}$ with the same Lipschitz constant K. Also, for any $A, B \in \mathcal{A}$ and $\lambda \in [0, 1]$,

$$V(\lambda A + (1 - \lambda) B) = V(\mathrm{co}(\lambda A + (1 - \lambda) B)) = V(\lambda \mathrm{co}(A) + (1 - \lambda) \mathrm{co}(B))$$
$$\leq \lambda V(\mathrm{co}(A)) + (1 - \lambda) V(\mathrm{co}(B)) = \lambda V(A) + (1 - \lambda) V(B),$$

which implies that V is convex on $\mathcal{A}$. Also, by parts 2 and 3 of Theorem 11 and the compactness of Σ, $\mathcal{M}_W \subset \mathcal{M}$ and $W^*(\mu) = c(\mu)$ for all $\mu \in \mathcal{M}_W$. By Lemma 7 and the minimality of $\mathcal{M}$, this implies $\mathcal{M} = \mathcal{M}_W$, and hence $c(\mu) = W^*(\mu)$ for all $\mu \in \mathcal{M}$.

(2): Define a function $\bar{W} : \Sigma \to \mathbb{R}$ by $\bar{W}(\sigma) = -W(\sigma)$. Then, for any $\sigma \in \Sigma$,

$$\bar{W}(\sigma) = -W(\sigma) = -\min_{\mu \in \mathcal{M}} \big[\langle \sigma, \mu \rangle + c(\mu) \big]$$
$$= \max_{\mu \in \mathcal{M}} \big[\langle \sigma, -\mu \rangle - c(\mu) \big]$$
$$= \max_{\mu \in -\mathcal{M}} \big[\langle \sigma, \mu \rangle - c(-\mu) \big].$$

34

By the same arguments used above, this implies that $\bar{W}$ is Lipschitz continuous and convex, which in turn implies that V is Lipschitz continuous and concave. Moreover, the above arguments imply that $-\mathcal{M} = \mathcal{M}_{\bar{W}}$ and $c(-\mu) = \bar{W}^*(\mu)$ for all $\mu \in -\mathcal{M}$. Thus, $\mathcal{M} = -\mathcal{M}_{\bar{W}} = -\mathcal{M}_{-W}$ and $c(\mu) = \bar{W}^*(-\mu) = [-W]^*(-\mu)$ for all $\mu \in \mathcal{M}$. ∎

Suppose that $\succsim$ has a max-HA [min-HA] representation $(\mathcal{M}, c)$, and suppose $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (1) [(2)]. Since $\mathbb{E}_P[V]$ represents $\succsim$ and V is continuous (by Lemma 8), $\succsim$ satisfies weak order, continuity, and first-stage independence by Lemma 3. Since V is Lipschitz continuous and convex [concave] by Lemma 8, $\succsim$ satisfies L–continuity and PERU [PLRU] by Lemma 4. Since $V(A) = V(\mathrm{co}(A))$ for all $A \in \mathcal{A}$, it is immediate that $\succsim$ satisfies IR. Finally, if the measures in $\mathcal{M}$ are positive, then it is obvious that V is monotone, which implies that $\succsim$ satisfies monotonicity.

C   Proof of Theorem 2

Throughout this section, we will continue to use notation and results for support functions that were established in Appendix B. Suppose $(\mathcal{M}, c)$ and $(\mathcal{M}', c')$ are two max-HA representations $\succsim$. Define $V : \mathcal{A} \to \mathbb{R}$ and $V' : \mathcal{A} \to \mathbb{R}$ for these respective representations, and define $W : \Sigma \to \mathbb{R}$ and $W' : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$ and $W'(\sigma) = V'(A_\sigma)$. By part 1 of Lemma 8, $\mathcal{M} = \mathcal{M}_W$ and $c(\mu) = W^*(\mu)$ for all $\mu \in \mathcal{M}$. Similarly, $\mathcal{M}' = \mathcal{M}_{W'}$ and $c'(\mu) = W'^*(\mu)$ for all $\mu \in \mathcal{M}'$.

Since V is continuous (by Lemma 8), the uniqueness part of Lemma 3 implies that there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V - \beta$. This implies that $W' = \alpha W - \beta$. Therefore, for any $\sigma, \sigma' \in \Sigma$,

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

and 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,

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

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

The proof of the uniqueness of the min-HA representation is similar and involves an application of part 2 of Lemma 8.

D   Proof of Theorem 3

$(1 \Rightarrow 3)$: Fix a monotone min-HA representation $(\mathcal{M}, c)$, and define V by Equation (2). Since $\mathcal{M}$ is compact, there is $\kappa > 0$ such that $\mu(\mathcal{U}) \leq \kappa$ for all $\mu \in \mathcal{M}$. Let $\Omega = \cup_{\lambda \in [0,\kappa]} \lambda \mathcal{U}$ and let $\mathcal{F}$ be the Borel σ-algebra generated by the relative topology of Ω in $\mathbb{R}^Z$. Define $U : \Omega \to \mathbb{R}^Z$ by $U(\omega) = \omega$.

For each $\mu \in \mathcal{M}$, define the probability measure π_μ on $(\Omega, \mathcal{F})$ as follows. If $\mu(\mathcal{U}) = 0$, let π_μ be the degenerate probability measure that puts probability one on $0 \in \Omega$, i.e., for any $E \in \mathcal{F}$, $\pi_\mu(E) = 1$ if $0 \in E$, and $\pi_\mu(E) = 0$ otherwise. If $\mu(\mathcal{U}) > 0$, then define the probability measure $\tilde{\mu}$ on $\mathcal{U}$ and its Borel σ-algebra by $\tilde{\mu}(E) = \frac{1}{\mu(\mathcal{U})}\mu(E)$ for any measurable $E \subset \mathcal{U}$. Define the function $f_\mu : \mathcal{U} \to \Omega$ by $f_\mu(u) = \mu(\mathcal{U})u$. Note that f is measurable because it is continuous. Finally, let π_μ be defined by $\pi_\mu = \tilde{\mu} \circ f_\mu^{-1}$. Then,

$$\int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi_\mu(d\omega) = \int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du)$$

for any $A \in \mathcal{A}$.[37] Let $\Pi = \{\pi_\mu : \mu \in \mathcal{M}\}$ and $\tilde{c}(\pi_\mu) = c(\mu)$. Then, V can be expressed in the following SSV form:

$$V(A) = \min_{\pi \in \Pi} \left(\int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega) + \tilde{c}(\pi) \right).$$

$(3 \Rightarrow 2)$: Let $((\Omega, \mathcal{F}), U, \Pi, c)$ be an SSV representation, and define V by Equation (4). Let the subset $\Pi' \subset \Pi$ stand for the set of $\pi \in \Pi$ such that there exists $A \in \mathcal{A}$ for which π solves the minimization problem in Equation (4). Note that Equation (4) continues to hold when Π is replaced by Π', i.e.,

$$V(A) = \min_{\pi \in \Pi'} \left(\int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega) + c(\pi) \right) \tag{20}$$

for all $A \in \mathcal{A}$.

We first show that c is bounded on Π'. Note that since U is bounded, there exists $\kappa > 0$ such that the absolute value of the integral term in Equation (20) is bounded by κ for every

---

[37] This is easy to see if $\mu(\mathcal{U}) = 0$. If $\mu(\mathcal{U}) > 0$, then define the function $g : \Omega \to \mathbb{R}$ by $g(\omega) = \max_{p \in A} U(\omega) \cdot p$. To see that g is $\mathcal{F}$-measurable, let B be a countable dense subset of A. At each $\omega \in \Omega$, $\max_{p \in A} \tilde{U}(\omega) \cdot p$ exists and is equal to $\sup_{p \in B} \tilde{U}(\omega) \cdot p$. For each $p \in B$, $\tilde{U} \cdot p$ is $\mathcal{F}$-measurable as a convex combination of $\mathcal{F}$-measurable random variables. Hence, g an $\mathcal{F}$-measurable as the pointwise supremum of countably many $\mathcal{F}$-measurable random variables (see Billingsley (1995, p184), Theorem 13.4(i)). Then,

$$\int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi_\mu(d\omega) = \int_{\mathcal{U}} \max_{p \in A} \mu(\mathcal{U})u(p) \, \tilde{\mu}(du) = \int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du),$$

where the first equality follows from the change of variables formula $\int_\Omega g(\omega) \, (\tilde{\mu} \circ f_\mu^{-1})(d\omega) = \int_{\mathcal{U}} g(f_\mu(u)) \, \tilde{\mu}(du)$.

menu $A \in \mathcal{A}$ and probability measure in $\pi \in \Pi'$. Take any $\pi, \pi' \in \Pi'$, and suppose that they solve the minimization in Equation (20) for menus A and A', respectively. Then, optimality of π at A implies:

$$c(\pi) - c(\pi') \leq \int_\Omega \max_{p \in A} U(\omega) \cdot p \ \pi'(d\omega) - \int_\Omega \max_{p \in A} U(\omega) \cdot p \ \pi(d\omega) \leq 2\kappa.$$

Similarly, optimality of π' at A' implies:

$$-2\kappa \leq \int_\Omega \max_{p \in A'} U(\omega) \cdot p \ \pi'(d\omega) - \int_\Omega \max_{p \in A'} U(\omega) \cdot p \ \pi(d\omega) \leq c(\pi) - c(\pi').$$

Therefore, $|c(\pi) - c(\pi')| \leq 2\kappa$ for any $\pi, \pi' \in \Pi'$, implying that c is bounded on Π'.

Let $\tilde{\Omega} = \Omega \times \Pi'$. Let $\mathcal{G}$ be any σ-algebra on Π' that contains all singletons and such that $c|_{\Pi'} : \Pi' \to \mathbb{R}$ is $\mathcal{G}$-measurable (e.g. $\mathcal{G} = 2^{\Pi'}$). Let $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{G}$ be the product σ-algebra on $\tilde{\Omega}$. Let $\mathbf{1} \in \mathbb{R}^Z$ denote the vector whose coordinates are equal to 1, and define $\tilde{U} : \tilde{\Omega} \to \mathbb{R}^Z$ by $\tilde{U}(\omega, \pi) = U(\omega) + c(\pi)\mathbf{1}$ for any $\tilde{\omega} = (\omega, \pi) \in \tilde{\Omega}$. Note that $\tilde{U}$ is $\tilde{\mathcal{F}}$-measurable and bounded.[38]

For any $\pi \in \Pi'$, define the function $f_\pi : \Omega \to \tilde{\Omega}$ by $f_\pi(\omega) = (\omega, \pi)$. Note that f_π is measurable.[39] Define the probability measure ρ_π on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by $\rho_\pi = \pi \circ f_\pi^{-1}$. For any $A \in \mathcal{A}$,

$$\int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \ \rho_\pi(d\tilde{\omega}) = \int_\Omega \max_{p \in A} \tilde{U}(f_\pi(\omega)) \cdot p \ \pi(d\omega)$$

$$= \int_\Omega \left[\max_{p \in A} U(\omega) \cdot p + c(\pi) \right] \pi(d\omega)$$

$$= \int_\Omega \max_{p \in A} U(\omega) \cdot p \ \pi(d\omega) + c(\pi),$$

where the first equality above follows from the change of variables formula.[40]

Letting $\tilde{\Pi}' = \{\rho_\pi : \pi \in \Pi'\}$, by Equation (20), we see that V can be expressed in the following SSMP form:

$$V(A) = \min_{\rho \in \tilde{\Pi}'} \int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \ \rho(d\tilde{\omega}).$$

---

[38]$\tilde{U}$ is bounded because U is bounded on Ω and c is bounded on Π'. To see that $\tilde{U}$ is $\tilde{\mathcal{F}}$-measurable, note that since U is $\mathcal{F}$-measurable and $\tilde{\mathcal{F}}$ is the product of the σ-algebras $\mathcal{F}$ and $\mathcal{G}$, the function $f : \tilde{\Omega} \to \mathbb{R}^Z$ defined by $f(\omega, \pi) = U(\omega)$ is $\tilde{\mathcal{F}}$-measurable. Also note that since $c|_{\Pi'}$ is $\mathcal{G}$-measurable, and $\tilde{\mathcal{F}}$ is the product of the σ-algebras $\mathcal{F}$ and $\mathcal{G}$, the function $g : \tilde{\Omega} \to \mathbb{R}^Z$ defined by $g(\omega, \pi) = c(\pi)\mathbf{1}$ is also $\tilde{\mathcal{F}}$-measurable. Therefore, $\tilde{U}$ is $\tilde{\mathcal{F}}$-measurable as the sum of the two $\tilde{\mathcal{F}}$-measurable functions f and g.

[39]To see this, note that the collection $\tilde{\mathcal{F}}'$ of sets $E \subset \tilde{\Omega}$ satisfying $\{\omega \in \Omega : (\omega, \pi') \in E\} \in \mathcal{F}$ for every $\pi' \in \Pi'$, is a σ-algebra. Since $\tilde{\mathcal{F}}'$ contains both $F \times \Pi'$ and $\Omega \times G$ for every $F \in \mathcal{F}$ and $G \in \mathcal{G}$, we have that $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{G} \subset \tilde{\mathcal{F}}'$. It is easy to see that f_π would be measurable if $\tilde{\Omega}$ were endowed with the σ-algebra $\tilde{\mathcal{F}}'$. Therefore, f_π is measurable since $\tilde{\Omega}$ is endowed with the coarser σ-algebra $\tilde{\mathcal{F}}$.

[40]To see this, define the function $g : \tilde{\Omega} \to \mathbb{R}$ by $g(\tilde{\omega}) = \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p$. By a similar argument as in Footnote 37, g is $\tilde{\mathcal{F}}$-measurable. Then, the change of variables formula is $\int_{\tilde{\Omega}} g(\tilde{\omega}) (\pi \circ f_\pi^{-1})(d\tilde{\omega}) = \int_\Omega g(f_\pi(\omega)) \pi(d\omega)$.

$(2 \Rightarrow 1)$: Let $((\Omega, \mathcal{F}), U, \Pi)$ be an SSMP representation, and define V by Equation (3). It is easy to see that V is monotone and concave. We next show that V is Lipschitz continuous. For every $\pi \in \Pi$, define $f_\pi : \mathcal{A} \to \mathbb{R}$ by

$$f_\pi(A) = \int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega).$$

Since U is bounded, there exists $\kappa > 0$ such that $\|U(\omega)\| \le \kappa$ for all $\omega \in \Omega$. Let $A, B \in \mathcal{A}$. Given a state $\omega \in \Omega$, let p^* be a solution of $\max_{p \in A} U(\omega) \cdot p$. By definition of Hausdorff distance, there exists $q^* \in B$ such that $\|p^* - q^*\| \le d_h(A, B)$. Then,

$$\max_{p \in A} U(\omega) \cdot p - \max_{q \in B} U(\omega) \cdot q = U(\omega) \cdot p^* - \max_{q \in B} U(\omega) \cdot q$$
$$\le U(\omega) \cdot p^* - U(\omega) \cdot q^* \le \|U(\omega)\| \, \|p^* - q^*\| \le \kappa \, d_h(A, B).$$

Taking the expectation of the above inequality with respect to π, we obtain:

$$f_\pi(A) - f_\pi(B) \le \kappa \, d_h(A, B).$$

Hence f_π is Lipschitz continuous with a Lipschitz constant κ that does not depend on $\pi \in \Pi$. Since V is the pointwise minimum of f_π over $\pi \in \Pi$, it is also Lipschitz continuous with the same Lipschitz constant κ.

Since $V : \mathcal{A} \to \mathbb{R}$ is monotone, concave, Lipschitz continuous, and it satisfies the IR condition $V(A) = V(\mathrm{co}(A))$ for all $A \in \mathcal{A}$, the construction in Appendix B.2 implies that there exists a monotone min-HA representation such that V is given by Equation (2).

E   Proof of Theorem 4

We define the set of *translations* to be

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

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 $p, q \in \triangle(Z)$, we have $p - q \in \Theta$.

Definition 10 A function $V : \mathcal{A} \to \mathbb{R}$ is *translation linear* if there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}$, we have $V(A + \theta) = V(A) + v \cdot \theta$.

Lemma 9 *Suppose that $V : \mathcal{A} \to \mathbb{R}$ is a function such that $\mathbb{E}_P[V]$ represents the preference $\succsim$ over $\triangle(\mathcal{A})$. Then, V is translation linear if and only if $\succsim$ satisfies RDD.*

Proof: Assume that $\mathbb{E}_P[V]$ represents the preference $\succsim$. Then, it is easy to see that $\succsim$ satisfies RDD if and only if

$$V(\alpha A + (1-\alpha)\{p\}) - V(\alpha A + (1-\alpha)\{q\}) = (1-\alpha)[V(\{p\}) - V(\{q\})] \tag{21}$$

for any $\alpha \in [0,1]$, $A \in \mathcal{A}$, and $p, q \in \triangle(Z)$.

If there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}$, we have $V(A + \theta) = V(A) + v \cdot \theta$, then both sides of Equation (21) are equal to $(1 - \alpha)v \cdot (p - q)$, showing that $\succsim$ satisfies RDD.

If $\succsim$ satisfies RDD, then define the function $f : \triangle(Z) \to \mathbb{R}$ by $f(p) = V(\{p\})$ for all $p \in \triangle(Z)$. Let $\alpha \in [0,1]$ and $p, q \in \triangle(Z)$, then

$$
\begin{aligned}
2f(\alpha p + (1-\alpha)q) &= [f(\alpha p + (1-\alpha)q) - f(\alpha p + (1-\alpha)p)] \\
&\quad + [f(\alpha p + (1-\alpha)q) - f(\alpha q + (1-\alpha)q)] + f(p) + f(q) \\
&= (1-\alpha)[f(q) - f(p)] + \alpha[f(p) - f(q)] + f(p) + f(q) \\
&= 2[\alpha f(p) + (1-\alpha)f(q)],
\end{aligned}
$$

where the second equality follows from Equation (21) and the definition of f. Therefore, $f(\alpha p + (1-\alpha)q) = \alpha f(p) + (1-\alpha)f(q)$ for any $\alpha \in [0,1]$ and $p, q \in \triangle(Z)$. It is standard to show that this implies that there exists $v \in \mathbb{R}^Z$ such that $f(p) = v \cdot p$ for all $p \in \triangle(Z)$.

To see that V is translation linear, let $A \in \mathcal{A}$ and $\theta \in \Theta$ be such that $A + \theta \in \mathcal{A}$. If $\theta = 0$, then the conclusion of translation linearity follows trivially, so without loss of generality assume that $\theta \neq 0$. Ergin and Sarver (2009) show in the proof of their Lemma 4 that if $A \in \mathcal{A}$ and $A + \theta \in \mathcal{A}$ for some $\theta \in \Theta \setminus \{0\}$, then there exist $A' \in \mathcal{A}$, $p, q \in \triangle(Z)$, and $\alpha \in (0,1]$ such that $A = (1-\alpha)A' + \alpha\{p\}$, $A + \theta = (1-\alpha)A' + \alpha\{q\}$, and $\theta = \alpha(p - q)$. Then

$$
\begin{aligned}
V(A + \theta) - V(A) &= V((1-\alpha)A' + \alpha\{p\}) - V((1-\alpha)A' + \alpha\{q\}) \\
&= \alpha[V(\{p\}) - V(\{q\})] \\
&= \alpha[v \cdot p - v \cdot q] \\
&= v \cdot \theta,
\end{aligned}
$$

where the second equality follows from Equation (21) and the third equality follows from the expected utility form of f. Therefore, V is translation linear. ∎

We are now ready to prove Theorem 4. The necessity of RDD in parts A and B are straightforward and left to the reader. In the rest of this section, we will continue to use the notation and results from Appendix B. For the sufficiency direction of part A, suppose that $\succsim$ satisfies Axiom 1, PERU, and RDD. Then, $(\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ constructed in Appendix B is a max-HA representation for $\succsim$. Since $\succsim$ satisfies RDD, by Lemma 9, V is translation linear. Let $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}$, we have $V(A + \theta) = V(A) + v \cdot \theta$.

Let $q = (1/|Z|, \ldots, 1/|Z|) \in \triangle(Z)$. By Lemma 22 of Ergin and Sarver (2009), for all $\mu \in \mathcal{M}_W$ and $p \in \triangle(Z)$, $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$. The consistency of $\mathcal{M}_W$ follows immediately from this fact because for any $\mu, \mu' \in \mathcal{M}_W$ and $p \in \triangle(Z)$, we have

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

If $\succsim$ additionally satisfies monotonicity, then $(\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ above is a monotone HA representation for $\succsim$. Therefore, the sufficiency direction of part B also follows from the consistency of $\mathcal{M}_W$ established above.

F   Proof of Theorem 5

The necessity of the axioms is straightforward. For sufficiency, suppose that $\succsim$ satisfies Axiom 1, second-stage independence, and strategic rationality. By Lemma 3, there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$.

   We first show that since $\succsim$ satisfies strategic rationality, $V(A) = \max_{p \in A} V(\{p\})$ for all $A \in \mathcal{A}$. To see this, fix any $A \in \mathcal{A}$, and let p^* be a solution to $\max_{p \in A} V(\{p\})$. Since $\succsim$ satisfies weak order and strategic rationality, it also satisfies monotonicity, implying that $A \succsim \{p^*\}$. We next show that for any finite subset B of A, $\{p^*\} \succsim B$. To see this, let B be a finite subset of A where $|B| = n$. Let $B = \{p_1, \ldots, p_n\}$, where $\{p_1\} \succsim \{p_2\} \succsim \ldots \succsim \{p_n\}$. For any $k = 2, \ldots, n$,

$$\{p_k\} \sim \{p_k, \ldots, p_n\} \implies \{p_{k-1}\} \sim \{p_{k-1}, \ldots, p_n\}, \tag{22}$$

since by strategic rationality, $\{p_{k-1}\} \succsim \{p_k\} \sim \{p_k, \ldots, p_n\}$ implies that $\{p_{k-1}\} \sim \{p_{k-1}\} \cup \{p_k, \ldots, p_n\}$. By applying backwards induction on $k = 2, \ldots, n$ using Equation (22), we have that $\{p_1\} \succsim B$. Since $\{p^*\} \succsim \{p\}$ for all $p \in A$, we also have that $\{p^*\} \succsim \{p_1\} \succsim B$. Therefore, $\{p^*\} \succsim B$ for any finite $B \subset A$. Since $\triangle(Z)$ is a compact metric space, there exists a sequence of finite subsets $\{B_m\}$ of A such that B_m converge to A in Hausdorff topology. Since $\{p^*\} \succsim B_m$ for all m, by continuity of $\succsim$ we have that $\{p^*\} \succsim A$. This proves that $\{p^*\} \sim A$. Therefore, $V(A) = V(\{p^*\}) = \max_{p \in A} V(\{p\})$, as desired.

   Define a preference $\succsim'$ on $\triangle(Z)$ by $p \succsim' q \iff \delta_{\{p\}} \succsim \delta_{\{q\}}$ (or, equivalently, $p \succsim' q \iff V(\{p\}) \geq V(\{q\})$). Continuity of $\succsim$ implies continuity of $\succsim'$, and second-stage independence implies that $\succsim'$ satisfies independence. Therefore, by the standard von Neumann-Morgenstern expected-utility theorem, there exists $v \in \mathbb{R}^Z$ such that $p \succsim' q \iff v(p) \geq v(q)$. Note that $\{v(p) : p \in \triangle(Z)\} = [a, b]$ for some $-\infty < a \leq b < +\infty$. Since $V(\{p\}) \geq V(\{q\}) \iff v(p) \geq v(q)$, there exists a strictly increasing function $\phi : [a, b] \to \mathbb{R}$ such that for all $p \in \triangle(Z)$,

$$V(\{p\}) = \phi\Big(v(p)\Big).$$

Therefore, for any menu $A \in \mathcal{A}$,

$$V(A) = \max_{p \in A} V(\{p\}) = \max_{p \in A} \phi\Big(v(p)\Big) = \phi\Big(\max_{p \in A} v(p)\Big).$$

To establish the Lipschitz continuity of ϕ, first recall that by Lemma 4, L–continuity implies there exists $K \geq 0$ such that $|V(A) - V(B)| \leq K d_h(A, B)$ for any $A, B \in \mathcal{A}^c$. In particular, for any $p, q \in \triangle(Z)$, $|V(\{p\}) - V(\{q\})| \leq K d_h(\{p\}, \{q\}) = K\|p - q\|$. If $a = b$, then ϕ is trivially Lipschitz continuous. Next, suppose that $a < b$. Take $p_*, p^* \in \triangle(Z)$ such that $v(p_*) = a$ and $v(p^*) = b$. For any $t \in [a, b]$, let $\alpha(t) \equiv (t - a)/(b - a) \in [0, 1]$, which implies $v(\alpha(t)p^* + (1 - \alpha(t))p_*) = t$. Then, for any $s, t \in [a, b]$,

$$
\begin{aligned}
|\phi(t) - \phi(s)| &= \Big| \phi\Big(v(\alpha(t)p^* + (1 - \alpha(t))p_*)\Big) - \phi\Big(v(\alpha(s)p^* + (1 - \alpha(s))p_*)\Big)\Big| \\
&= \Big|V(\{\alpha(t)p^* + (1 - \alpha(t))p_*\}) - V(\{\alpha(s)p^* + (1 - \alpha(s))p_*\})\Big| \\
&\leq K|\alpha(t) - \alpha(s)| \, \|p^* - p_*\| \\
&= K|t - s| \, \|p^* - p_*\|/(b - a),
\end{aligned}
$$

which implies ϕ is Lipschitz continuous with a Lipschitz constant of $K\|p^* - p_*\|/(b - a)$.

G Proof of Theorem 6

G.1 Proof of Theorem 6.A

The necessity of the axioms is straightforward. For sufficiency, suppose that $\succsim$ satisfies Axiom 1 and mixture independence. By Lemma 3, there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Since $\succsim$ satisfies IR, $V(A) = V(\mathrm{co}(A))$ for all $A \in \mathcal{A}$. It therefore suffices to show the existence of a finite signed Borel measure μ of $\mathcal{U}$ and a Lipschitz continuous and strictly increasing function $\phi : [a, b] \to \mathbb{R}$ such that for all $A \in \mathcal{A}^c$ (the set of all convex menus),

$$V(A) = \phi \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) \right),$$

where $[a, b] = \{\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) : A \in \mathcal{A}^c\}$.

Define a preference $\succsim'$ on $\mathcal{A}^c$ by $A \succsim' B \iff \delta_A \succsim \delta_B$ (or, equivalently, $A \succsim' B \iff V(A) \geq V(B)$). Continuity of $\succsim$ implies continuity of $\succsim'$, and mixture independence implies that $\succsim'$ satisfies independence. Therefore, by the Hernstein-Milnor Theorem, there exists an affine function $U : \mathcal{A}^c \to \mathbb{R}$ such that $A \succsim' B \iff U(A) \geq U(B)$. Moreover, U is continuous by the continuity of $\succsim'$, which by the compactness of $\mathcal{A}^c$ implies the existence of $-\infty < a \leq b < +\infty$ such that $[a, b] = \{U(A) : A \in \mathcal{A}^c\}$. Since $V(A) \geq V(B) \iff U(A) \geq U(B)$, there

exists a strictly increasing function $\phi : [a, b] \to \mathbb{R}$ such that

$$V(A) = \phi\Big(U(A)\Big).$$

To establish the Lipschitz continuity of ϕ, first recall that by Lemma 4, L–continuity implies there exists $K \geq 0$ such that $|V(A) - V(B)| \leq Kd_h(A, B)$ for any $A, B \in \mathcal{A}^c$. If $a = b$, then ϕ is trivially Lipschitz continuous. Next, suppose that $a < b$. Take $A_*, A^* \in \mathcal{A}^c$ such that $U(A_*) = a$ and $U(A^*) = b$. Note that for any $\alpha, \beta \in [0, 1]$,

$$d_h(\alpha A^* + (1 - \alpha)A_*, \beta A^* + (1 - \beta)A_*) = |\alpha - \beta|d_h(A^*, A_*).$$

Therefore, by analogous arguments to those in Appendix F, the linearity of U implies that ϕ is Lipschitz continuous with a Lipschitz constant of $Kd_h(A^*, A_*)/(b - a)$.

It remains only to show that there exists a finite signed Borel measure μ on $\mathcal{U}$ such that for every $A \in \mathcal{A}^c$,

$$U(A) = \int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du).$$

To establish the existence of such a measure μ, it suffices to show that U is Lipschitz continuous (see the arguments used in the construction of the additive EU representation in the supplementary appendix of DLRS (2007)). We prove that U is Lipschitz continuous by contradiction; we will argue that if U is not Lipschitz continuous, then ϕ cannot be strictly increasing.[41] In particular, we show that if U is not Lipschitz continuous, then $\phi(b) - \phi(a) < \varepsilon \cdot (b - a)$ for any $\varepsilon > 0$, which implies $\phi(b) = \phi(a)$, a contradiction. These arguments are completed in three steps:

Step 1 — For any $t \in (a, b)$ and $n \in \mathbb{N}$, there exist $A_n, B_n \in \mathcal{A}^c$ such that $U(A_n) < t < U(B_n)$ and $|U(B_n) - U(A_n)| > n \cdot d_h(A_n, B_n)$: Since U is not Lipschitz continuous, for any $n \in \mathbb{N}$, there must exist $A'_n, B'_n \in \mathcal{A}^c$ such that $|U(B'_n) - U(A'_n)| > n \cdot d_h(A'_n, B'_n)$. Without loss of generality, suppose $U(A'_n) < U(B'_n)$. If $U(A'_n) < t < U(B'_n)$, then let $A_n = A'_n$ and $B_n = B'_n$. If $t \leq U(A'_n)$, then take $\alpha \in (0, 1)$ such that $\alpha a + (1 - \alpha)U(A'_n) < t < \alpha a + (1 - \alpha)U(B'_n)$. Let $A_n = \alpha A_* + (1 - \alpha)A'_n$ and $B_n = \alpha A_* + (1 - \alpha)B'_n$. Then, $U(A_n) < t < U(B_n)$ and

$$\begin{aligned}
|U(B_n) - U(A_n)| &= (1 - \alpha)[U(B'_n) - U(A'_n)] \\
&> (1 - \alpha) \cdot n \cdot d_h(A'_n, B'_n) \\
&= n \cdot d_h(A_n, B_n).
\end{aligned}$$

---

[41]The Lipschitz continuity of U does not follow immediately from the arguments in DLRS (2007) because we impose a different L–continuity axiom than the one used in their paper. The statement of their L–continuity axiom in our framework would be the following axiom, which we do not explicitly assume:

Axiom 9 (Mixture L–Continuity) *There exist $A^*, A_* \in \mathcal{A}$ and $M \geq 0$ such that for every $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$ with $\alpha \geq Md_h(A, B)$, $\delta_{(1-\alpha)A+\alpha A^*} \succsim \delta_{(1-\alpha)B+\alpha A_*}$.*

The case of $U(B'_n) \leq t$ is similar.

Step 2 — For any $t \in (a, b)$ and $\varepsilon > 0$, there exists an open interval $(a_t, b_t) \subset (a, b)$ containing t such that $|\phi(s) - \phi(s')| \leq \varepsilon \cdot |s - s'|$ for any $s, s' \in (a_t, b_t)$: Recall that there exists $K \geq 0$ such that $|V(A) - V(B)| \leq K d_h(A, B)$ for any $A, B \in \mathcal{A}^c$. Choose $n \in \mathbb{N}$ such that $K/n < \varepsilon$. From Step 1, we know there exist $A_n, B_n \in \mathcal{A}^c$ such that $U(A_n) < t < U(B_n)$ and $|U(B_n) - U(A_n)| > n \cdot d_h(A_n, B_n)$. Let $a_t = U(A_n)$ and $b_t = U(B_n)$. Fix any $s, s' \in (a_t, b_t)$. Then, there exist $\alpha, \alpha' \in (0, 1)$ such that

$$s = U(\alpha A_n + (1 - \alpha) B_n) \quad \text{and} \quad s' = U(\alpha' A_n + (1 - \alpha') B_n).$$

Therefore,

$$\begin{aligned}
|s - s'| &= \left|(\alpha - \alpha')[U(A_n) - U(B_n)]\right| \\
&> |\alpha - \alpha'| \cdot n \cdot d_h(A_n, B_n) \\
&= n \cdot d_h(\alpha A_n + (1 - \alpha) B_n, \alpha' A_n + (1 - \alpha') B_n),
\end{aligned}$$

and hence

$$\begin{aligned}
|\phi(s) - \phi(s')| &= \left|V(\alpha A_n + (1 - \alpha) B_n) - V(\alpha' A_n + (1 - \alpha') B_n)\right| \\
&\leq K d_h(\alpha A_n + (1 - \alpha) B_n, \alpha' A_n + (1 - \alpha') B_n) \\
&< \frac{K}{n} \cdot |s - s'| < \varepsilon \cdot |s - s'|.
\end{aligned}$$

Step 3 — For any $\varepsilon > 0$, $\phi(b) - \phi(a) \leq \varepsilon \cdot (b - a)$: Fix any $\varepsilon > 0$. Since we established above that ϕ is continuous, it suffices to show that for any $\underline{t}, \overline{t} \in (a, b)$, $\underline{t} < \overline{t}$, we have $\phi(\overline{t}) - \phi(\underline{t}) \leq \varepsilon \cdot (\overline{t} - \underline{t})$. To see that this is true, note that the collection of intervals $\{(a_t, b_t) : t \in (a, b)\}$ defined in Step 2 (for this ε) forms an open cover of the closed interval $[\underline{t}, \overline{t}]$. Therefore, there exists a finite subcover taken from this collection of intervals that also covers $[\underline{t}, \overline{t}]$. The finiteness of this subcover implies the existence of a finite set of numbers $\underline{t} = s_1 \leq \cdots \leq s_k = \overline{t}$ such that for any $i \in \{1, \ldots, k - 1\}$ there exists $t \in (a, b)$ such that $s_i, s_{i+1} \in (a_t, b_t)$. By the definition of (a_t, b_t), we have

$$\phi(\overline{t}) - \phi(\underline{t}) = \sum_{i=1}^{k-1} [\phi(s_{i+1}) - \phi(s_i)] \leq \varepsilon \sum_{i=1}^{k-1} (s_{i+1} - s_i) = \varepsilon \cdot (\overline{t} - \underline{t}).$$

This completes the proof.

G.2 Proof of Theorem 6.B

Suppose $\succsim$ has a KPDLR representation (ϕ, μ). First, note that for any $A, B \in \mathcal{A}$ and any $\alpha \in (0, 1)$,

$$\int_{\mathcal{U}} \max_{p \in \alpha A + (1-\alpha)B} u(p)\, \mu(du) = \alpha \int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) + (1-\alpha) \int_{\mathcal{U}} \max_{p \in B} u(p)\, \mu(du).$$

For any $s, t \in [a, b]$, let $A, B \in \mathcal{A}$ be such that $s = \int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du)$ and $t = \int_{\mathcal{U}} \max_{p \in B} u(p)\, \mu(du)$. Then, for any $\alpha \in (0, 1)$,

$$\begin{aligned}
&\alpha \delta_A + (1-\alpha)\delta_B \succsim \delta_{\alpha A + (1-\alpha)B} \\
\iff\ &\alpha V(A) + (1-\alpha)V(B) \geq V(\alpha A + (1-\alpha)B) \\
\iff\ &\alpha \phi\left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du)\right) + (1-\alpha)\phi\left(\int_{\mathcal{U}} \max_{p \in B} u(p)\, \mu(du)\right) \\
&\geq \phi\left(\alpha \int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) + (1-\alpha)\int_{\mathcal{U}} \max_{p \in B} u(p)\, \mu(du)\right) \\
\iff\ &\alpha\phi(s) + (1-\alpha)\phi(t) \geq \phi(\alpha s + (1-\alpha)t).
\end{aligned}$$

Thus, $\succsim$ satisfies PERU if and only if ϕ is convex. A similar argument shows that $\succsim$ satisfies PLRU if and only if ϕ is concave.

H Proof of Theorem 7

Throughout this section, we use the notation ∂f, f^*, $\mathcal{N}_f$, and $\mathcal{M}_f$ introduced in Appendix A.

Lemma 10 *Let $a, b \in \mathbb{R}$ with $a < b$ and let $\phi : [a, b] \to \mathbb{R}$ be Lipschitz continuous and convex. Then, $1 \Leftrightarrow 2 \Rightarrow 3$:*

1. *ϕ is strictly increasing.*

2. *(a) $\mathcal{M}_\phi \subset \mathbb{R}_+$.*

   (b) The right-derivative of ϕ^* at 0, $\frac{d\phi^*}{d\lambda^+}(0)$, exists and is equal to a.

3. *0 is not an isolated point of $\mathcal{M}_\phi$.*

Proof: $(1 \Rightarrow 2)$ Part a follows from Theorem 10.

To see part b, it is enough to show that for all $t \in (a, b]$, there exists $\lambda > 0$ such that

$$\lambda' a \leq \phi^*(\lambda') - \phi^*(0) \leq \lambda' t \quad \forall \lambda' \in (0, \lambda). \tag{23}$$

Since ϕ is nondecreasing, $0 \in \partial\phi(a)$. Along with Lemma 2, this implies that $-\phi^*(0) = \phi(a) \geq \lambda' a - \phi^*(\lambda')$ for any $\lambda' \geq 0$, establishing the first inequality in Equation (23). Take any

$t \in (a, b]$. By Lemma 1 there exists $\lambda \in \partial\phi(t)$. Note that $\lambda > 0$. Otherwise, if $\lambda \leq 0$, then by Lemma 2,

$$\phi(a) \geq \lambda a - \phi^*(\lambda) \geq \lambda t - \phi^*(\lambda) = \phi(t),$$

a contradiction to ϕ being strictly increasing. Let $\lambda' \in (0, \lambda)$. Since ϕ is continuous and its domain is compact, there exists $t' \in [a, b]$ such that $\phi^*(\lambda') = t'\lambda' - \phi(t')$. By Lemma 2, this implies that $\lambda' \in \partial\phi(t')$. Monotonicity of the subdifferential $\partial\phi$ implies that $t' \leq t$.[42] Then, by Lemma 2 and ϕ being nondecreasing,

$$-\phi^*(0) = \phi(a) \leq \phi(t') = \lambda't' - \phi^*(\lambda') \leq \lambda't - \phi^*(\lambda'),$$

which implies the second inequality in Equation (23).

$(2 \Rightarrow 1)$ Theorem 9 and part a imply that ϕ is nondecreasing. Therefore, $0 \in \partial\phi(a)$, implying $\phi(a) = -\phi^*(0)$ by Lemma 2.

We will first show that $\phi(a) < \phi(t)$ for any $t \in (a, b]$. Suppose for a contradiction that $\phi(a) = \phi(t)$ for some $t \in (a, b]$. Then, for any $\lambda > 0$,

$$\phi^*(\lambda) \geq \lambda t - \phi(t) = \lambda t - \phi(a) = \lambda t + \phi^*(0)$$

implying that $\frac{\phi^*(\lambda) - \phi^*(0)}{\lambda} \geq t > a$ for any $\lambda > 0$, a contradiction to $\frac{d\phi^*}{d\lambda^+}(0) = a$.

To conclude that ϕ is strictly increasing, it remains to show that $\phi(t) < \phi(t')$ for any $t, t' \in (a, b]$ such that $t < t'$. By Lemma 1, there exists $\lambda \in \partial\phi(t)$. If $\lambda \leq 0$, then

$$\phi(a) \geq \lambda a - \phi^*(\lambda) \geq \lambda t - \phi^*(\lambda) = \phi(t)$$

by Lemma 2, contradicting $\phi(a) < \phi(t)$. Therefore, $\lambda > 0$, implying

$$\phi(t) = \lambda t - \phi^*(\lambda) < \lambda t' - \phi^*(\lambda) \leq \phi(t'),$$

by Lemma 2, as desired.

$(1 \Rightarrow 3)$ Suppose for a contradiction that 0 is an isolated point of $\mathcal{M}_\phi$. Then, $0 \in \mathcal{N}_\phi$, i.e., there exists $t \in [a, b]$ such that $\partial\phi(t) = \{0\}$. Then, Lemma 2 implies

$$-\phi^*(0) = \phi(t) > \lambda t - \phi^*(\lambda) \quad \forall\lambda \in \mathcal{M}_\phi \setminus \{0\}.$$

Since 0 is an isolated point of $\mathcal{M}_\phi$ and $\mathcal{M}_\phi$ is compact by Theorem 9, $\mathcal{M}_\phi \setminus \{0\}$ is also compact. Therefore, the above inequality implies that

$$-\phi^*(0) > \max_{\lambda \in \mathcal{M}_\phi \setminus \{0\}} \left[\lambda t - \phi^*(\lambda) \right]. \tag{24}$$

---

[42]To see that $\partial\phi$ is monotone, note that by the definition of the subdifferential, $\lambda \in \partial\phi(t)$ implies $\lambda(t' - t) \leq \phi(t') - \phi(t)$ and $\lambda' \in \partial\phi(t')$ implies $\lambda'(t - t') \leq \phi(t) - \phi(t')$. Summing these inequalities, we have $(\lambda - \lambda')(t - t') \geq 0$.

Let $\Delta > 0$ be the difference of the left hand side and the right hand side in Equation (24) and let $M > 0$ be such that $\mathcal{M}_\phi \subset [0, M]$. Take any $s \in [a, b]$ such that $|t - s| < \frac{\Delta}{M}$. Then, $|\lambda t - \lambda s| < \Delta$ for any $\lambda \in \mathcal{M}_\phi \setminus \{0\}$, implying that Equation (24) continues to hold if t is replaced by s. Therefore,

$$-\phi^*(0) = \max_{\lambda \in \mathcal{M}_\phi} \left[\lambda s - \phi^*(\lambda) \right] = \phi(s),$$

where the second equality follows from Theorem 9. This implies that ϕ is constant at a $\frac{\Delta}{M}$ neighborhood of t, contradicting the assumption that ϕ is strictly increasing. ∎

Corollary 2 *Let $a, b \in \mathbb{R}$ with $a < b$ and let $\phi : [a, b] \to \mathbb{R}$ be Lipschitz continuous and convex. Then, $1 \Leftrightarrow 2 \Rightarrow 3$:*

1. *ϕ is strictly decreasing.*

2. *(a) $\mathcal{M}_\phi \subset \mathbb{R}_-$.*

   (b) The left-derivative of ϕ^* at 0, $\frac{d\phi^*}{d\lambda^-}(0)$, exists and is equal to b.

3. *0 is not an isolated point of $\mathcal{M}_\phi$.*

Proof: Define $\hat{\phi} : [-b, -a] \to \mathbb{R}$ by $\hat{\phi}(t) = \phi(-t)$. Note that $\hat{\phi}$ is Lipschitz continuous and convex. For any $\lambda \in \mathbb{R}$ and $t \in [-b, -a]$, we have

$$
\begin{aligned}
\lambda \in \partial\hat{\phi}(t) &\iff \hat{\phi}(s) - \hat{\phi}(t) \geq \lambda(s - t) \ \forall s \in [-b, -a] \\
&\iff \phi(-s) - \phi(-t) \geq -\lambda(-s - (-t)) \ \forall s \in [-b, -a] \\
&\iff -\lambda \in \partial\phi(-t).
\end{aligned}
$$

Therefore, $\mathcal{N}_{\hat{\phi}} = -\mathcal{N}_\phi$ implying that $\mathcal{M}_{\hat{\phi}} = -\mathcal{M}_\phi$. For any $\lambda \in \mathbb{R}$,

$$\hat{\phi}^*(\lambda) = \max_{t \in [-b, -a]} [\lambda t - \hat{\phi}(t)] = \max_{t' \in [a,b]} [-\lambda t' - \hat{\phi}(-t')] = \max_{t' \in [a,b]} [-\lambda t' - \phi(t')] = \phi^*(-\lambda).$$

implying,

$$\frac{d\hat{\phi}^*}{d\lambda^+}(0) = \lim_{\lambda \searrow 0} \frac{\hat{\phi}^*(\lambda) - \hat{\phi}^*(0)}{\lambda} = -\lim_{\lambda \searrow 0} \frac{\phi^*(0) - \phi^*(-\lambda)}{\lambda} = -\frac{d\phi^*}{d\lambda^-}(0).$$

Therefore, conditions 1, 2.a, 2,b, and 3 are equivalent to:

1' $\hat{\phi}$ is strictly increasing,

2' (a') $\mathcal{M}_{\hat{\phi}} \subset \mathbb{R}_+$,

   (b') The right-derivative of $\hat{\phi}^*$ at 0, $\frac{d\hat{\phi}^*}{d\lambda^+}(0)$, exists and is equal to $-b$, and

3' 0 is not an isolated point of $\mathcal{M}_{\hat{\phi}}$,

respectively. Applying Lemma 10 to $\hat{\phi}$, we have $1' \Leftrightarrow 2' \Rightarrow 3'$, as desired. ∎

In the next lemma, Σ denotes the set of support functions defined in Appendix B.

Lemma 11 *Let μ be a nonzero finite signed Borel measure on $\mathcal{U}$ and $[a,b] = \{\langle \sigma, \mu \rangle : \sigma \in \Sigma\}$. Let $\phi : [a,b] \to \mathbb{R}$ be Lipschitz continuous and convex and define $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = \phi(\langle \sigma, \mu \rangle)$ for any $\sigma \in \Sigma$. Then,*

1. *W is Lipschitz continuous and convex.*

2. *$W^*(\lambda\mu) = \phi^*(\lambda)$ for any $\lambda \in \mathbb{R}$.*

3. *$\mathcal{M}_W = \{\lambda\mu : \lambda \in \mathcal{M}_\phi\}$.*

Proof: 1. Let $K \geq 0$ be a Lipschitz constant for ϕ. Then, for any $\sigma, \sigma' \in \Sigma$,

$$|W(\sigma) - W(\sigma')| = |\phi(\langle \sigma, \mu \rangle) - \phi(\langle \sigma', \mu \rangle)| \leq K|\langle \sigma, \mu \rangle - \langle \sigma', \mu \rangle| \leq K\|\mu\|\|\sigma - \sigma'\|,$$

implying that W is Lipschitz continuous with a Lipschitz constant $K\|\mu\|$. W is convex as the composition of a linear and a convex function.

2. Let $\lambda \in \mathbb{R}$. Then,

$$
\begin{aligned}
W^*(\lambda\mu) &= \max_{\sigma \in \Sigma} \left[\langle \sigma, \lambda\mu \rangle - W(\sigma) \right] \\
&= \max_{\sigma \in \Sigma} \left[\lambda\langle \sigma, \mu \rangle - \phi(\langle \sigma, \mu \rangle) \right] \\
&= \max_{t \in [a,b]} \left[\lambda t - \phi(t) \right] \\
&= \phi^*(\lambda).
\end{aligned}
$$

3. We will first show that $\mathcal{N}_W \subset \{\lambda\mu : \lambda \in \mathcal{M}_\phi\}$. This will imply that $\mathcal{M}_W = \overline{\mathcal{N}_W} \subset \{\lambda\mu : \lambda \in \mathcal{M}_\phi\}$ since $\mathcal{M}_\phi$ is closed. Let $\nu \in \mathcal{N}_W$, then there exists $\sigma \in \Sigma$ such that $\partial W(\sigma) = \{\nu\}$. For any $\lambda \in \partial\phi(\langle \sigma, \mu \rangle)$,

$$W(\sigma') - W(\sigma) = \phi(\langle \sigma', \mu \rangle) - \phi(\langle \sigma, \mu \rangle) \geq \lambda[\langle \sigma', \mu \rangle - \langle \sigma, \mu \rangle] = \langle \sigma' - \sigma, \lambda\mu \rangle) \quad \forall \sigma' \in \Sigma,$$

implying $\lambda\mu \in \partial W(\sigma) = \{\nu\}$. Therefore, $\{\lambda\mu : \lambda \in \partial\phi(\langle \sigma, \mu \rangle)\} \subset \{\nu\}$. Since μ is nonzero and $\partial\phi(\langle \sigma, \mu \rangle) \neq \emptyset$ by Lemma 1, there exists a unique $\lambda \in \mathbb{R}$ such that $\partial\phi(\langle \sigma, \mu \rangle) = \{\lambda\}$. Note that $\lambda \in \mathcal{N}_\phi \subset \mathcal{M}_\phi$ and $\nu = \lambda\mu$, as desired.

Let $\mathcal{M} = \{\lambda \in \mathbb{R} : \lambda\mu \in \mathcal{M}_W\}$. We will next show that $\mathcal{M}_\phi \subset \mathcal{M}$, which will imply $\{\lambda\mu : \lambda \in \mathcal{M}_\phi\} \subset \mathcal{M}_W$. Since μ is nonzero and $\mathcal{M}_W$ is compact by part 1 and Theorem 9, $\mathcal{M}$ is also compact. Let $t \in [a,b]$, and $\sigma \in \Sigma$ be such that $t = \langle \sigma, \mu \rangle$. Then,

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

47

where the second equality follows from part 1 and Theorem 9, the third equality follows from $\mathcal{M}_W \subset \{\lambda\mu : \lambda \in \mathbb{R}\}$, and the last equality follows from part 2. Therefore, by Theorem 9, $\mathcal{M}_\phi \subset \mathcal{M}$. ∎

Proof of Theorem 7, the convex case: In the following, let $W : \Sigma \to \mathbb{R}$ be defined by $W(\sigma) = V(A_\sigma)$. Also, let $[a, b] = \{\int_{\mathcal{U}} \max_{p \in A} u(p)\,\mu(du) : A \in \mathcal{A}\}$.

$(1 \Rightarrow 2)$ For any $\sigma \in \Sigma$,

$$W(\sigma) = V(A_\sigma) = \phi(\langle \sigma_{(A_\sigma)}, \mu \rangle) = \phi(\langle \sigma, \mu \rangle),$$

where the last equality follows from part 1 of Lemma 5. Since W is Lipschitz continuous and convex by Lemma 11, $V(A) = V(\mathrm{co}(A))$ for all $A \in \mathcal{A}$, and $W(\sigma) = V(A_\sigma)$ for all $\sigma \in \Sigma$, the construction in Section B.1 implies that $(\mathcal{M}, c) := (\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ is a max-HA representation such that V is given by Equation (1). By part 2.a of Lemma 10 and part 3 of Lemma 11, $\mathcal{M}_W \subset \{\lambda\mu : \lambda \in \mathbb{R}_+\}$. By part 2.b of Lemma 10 and part 2 of Lemma 11,

$$\lim_{\lambda \searrow 0} \frac{W^*(\lambda\mu) - W^*(0)}{\lambda} = \frac{d\phi^*}{d\lambda^+}(0) = a \equiv \min_{A \in \mathcal{A}} \int_{\mathcal{U}} \max_{p \in A} u(p)\,\mu(du). \tag{25}$$

By part 3 of Lemma 10, part 3 of Lemma 11, and μ being nonzero, 0 is not an isolated point of $\mathcal{M}_W$. Therefore, if $0 \in \mathcal{M}_W$, then the limit term in Equation (25) agrees with $\lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}_W} \frac{c(\lambda\mu) - c(0)}{\lambda}$.

$(2 \Rightarrow 1)$ The mapping $\lambda \mapsto c(\lambda\mu)$ is lower semi-continuous since c is lower semi-continuous, and $\{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}\}$ is nonempty by part a, and it is compact since $\mathcal{M}$ is compact and μ is nonzero. Therefore, we can define $\phi : [a, b] \to \mathbb{R}$ by

$$\phi(t) = \max_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [\lambda t - c(\lambda\mu)] \quad \forall t \in [a, b].$$

By Theorem 11, ϕ is Lipschitz continuous and convex. Furthermore, for any $A \in \mathcal{A}$,

$$V(A) = \max_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [\langle \sigma_A, \lambda\mu \rangle - c(\lambda\mu)] = \max_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [\lambda \langle \sigma_A, \mu \rangle - c(\lambda\mu)] = \phi(\langle \sigma_A, \mu \rangle),$$

where the first equality follows from Equation (1) and part a. Therefore, it only remains to show that ϕ is strictly increasing.

By Lemma 8, $\mathcal{M} = \mathcal{M}_W$ and $c(\nu) = W^*(\nu)$ for all $\nu \in \mathcal{M}$. Note that

$$W(\sigma) = V(A_\sigma) = \phi(\langle \sigma_{(A_\sigma)}, \mu \rangle) = \phi(\langle \sigma, \mu \rangle) \quad \forall \sigma \in \Sigma,$$

where the last equality follows from part 1 of Lemma 5. By part 3 of Lemma 11, $\mathcal{M} = \{\lambda\mu : \lambda \in \mathcal{M}_\phi\}$. Therefore, since μ is nonzero: $0 \in \mathcal{M}$ if and only if $0 \in \mathcal{M}_\phi$; part a implies $\mathcal{M}_\phi \subset \mathbb{R}_+$; and the first part of b implies that 0 is not an isolated point of $\mathcal{M}_\phi$.

First suppose that $0 \notin \mathcal{M}$, implying $0 \notin \mathcal{M}_\phi$. Let $t, t' \in [a, b]$ be such that $t < t'$. By Theorem 9,

$$\phi(s) = \max_{\lambda \in \mathcal{M}_\phi} [\lambda s - \phi^*(\lambda)] \quad \forall s \in [a, b].$$

Let $\hat{\lambda} > 0$ be a solution of the above maximization at $s = t$. Then,

$$\phi(t) = \hat{\lambda} t - \phi^*(\hat{\lambda}) < \hat{\lambda} t' - \phi^*(\hat{\lambda}) \le \max_{\lambda \in \mathcal{M}_\phi} [\lambda t' - \phi^*(\lambda)] = \phi(t').$$

Next suppose that $0 \in \mathcal{M}$, implying that $0 \in \mathcal{M}_\phi$. Then,

$$a = \lim_{\lambda \searrow 0 : \lambda \mu \in \mathcal{M}} \frac{c(\lambda \mu) - c(0)}{\lambda} = \lim_{\lambda \searrow 0 : \lambda \in \mathcal{M}_\phi} \frac{\phi^*(\lambda) - \phi^*(0)}{\lambda} \tag{26}$$

where the first equality follows from part b and the second equality follows from $\mathcal{M} = \{\lambda \mu : \lambda \in \mathcal{M}_\phi\}$, μ being nonzero, $c = W^*|_{\mathcal{M}}$, and part 2 of Lemma 11. For any $\lambda \in (0, \infty)$, define $a_\lambda \in \mathbb{R}$ by $a_\lambda = \frac{\phi^*(\lambda) - \phi^*(0)}{\lambda}$. Since $0 \in \mathcal{M}_\phi$ is not an isolated point of $\mathcal{M}_\phi$, Equation (26) implies that there exists a sequence λ_n in $\mathcal{M}_\phi \setminus \{0\}$ such that $\lambda_n \searrow 0$ and $\lim_n a_{\lambda_n} = a$. Since ϕ^* is convex, a_λ is nondecreasing in $\lambda \in (0, \infty)$. Therefore, for any sequence λ'_n in $(0, \infty)$ such that $\lambda'_n \searrow 0$, $\lim_n a_{\lambda'_n} = \lim_n a_{\lambda_n}$. This implies that the limit on the right hand side of Equation (26) is equal to $\frac{d\phi^*}{d\lambda^+}(0)$. By Lemma 10, ϕ is strictly increasing. $\blacksquare$

Proof of Theorem 7, the concave case: In the following, let $W : \Sigma \to \mathbb{R}$ and $\bar{W} : \Sigma \to \mathbb{R}$ be defined by $W(\sigma) = V(A_\sigma)$ and $\bar{W}(\sigma) = -W(\sigma)$. Also, let $[a, b] = \{\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) : A \in \mathcal{A}\}$.

$(1 \Rightarrow 2)$ For any $\sigma \in \Sigma$,

$$W(\sigma) = V(A_\sigma) = \phi(\langle \sigma_{(A_\sigma)}, \mu \rangle) = \phi(\langle \sigma, \mu \rangle),$$

where the last equality follows from part 1 of Lemma 5. Define $\bar{\phi} : [a, b] \to \mathbb{R}$ by $\bar{\phi}(t) = -\phi(t)$. Then, $\bar{\phi}$ is Lipschitz continuous, convex, and strictly decreasing, and $\bar{W}(\sigma) = \bar{\phi}(\langle \sigma, \mu \rangle)$ for all $\sigma \in \Sigma$. Let $\mathcal{M} = -\mathcal{M}_{\bar{W}}$ and define $c : \mathcal{M} \to \mathbb{R}$ by $c(\nu) = \bar{W}^*(-\nu)$. Since $\bar{W}$ is Lipschitz continuous and convex by Lemma 11, $V(A) = V(co(A))$ for all $A \in \mathcal{A}$, and $W(\sigma) = V(A_\sigma)$ for all $\sigma \in \Sigma$, the construction in Section B.2 implies that $(\mathcal{M}, c)$ is a min-HA representation such that V is given by Equation (2). By part 2.a of Corollary 2 and part 3 of Lemma 11, $\mathcal{M}_{\bar{W}} \subset \{\lambda \mu : \lambda \in \mathbb{R}_-\}$, implying that $\mathcal{M} \subset \{\lambda \mu : \lambda \in \mathbb{R}_+\}$. By part 2.b of Corollary 2 and part 2 of Lemma 11,

$$\lim_{\lambda \searrow 0} \frac{\bar{W}^*(0) - \bar{W}^*(-\lambda \mu)}{\lambda} = \frac{d\bar{\phi}^*}{d\lambda^-}(0) = b \equiv \max_{A \in \mathcal{A}} \int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du). \tag{27}$$

By part 3 of Corollary 2, part 3 of Lemma 11, and μ being nonzero, 0 is not an isolated point of $\mathcal{M}_{\bar{W}}$. Therefore, 0 is also not an isolated point of $\mathcal{M} = -\mathcal{M}_{\bar{W}}$. Therefore, if $0 \in \mathcal{M}$, then

the limit term in Equation (27) agrees with $\lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}} \frac{c(0) - c(\lambda\mu)}{\lambda}$.

$(2 \Rightarrow 1)$ The mapping $\lambda \mapsto c(\lambda\mu)$ is lower semi-continuous since c is lower semi-continuous, and $\{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}\}$ is nonempty by part a, and it is compact since $\mathcal{M}$ is compact and μ is nonzero. Therefore, we can define $\phi : [a, b] \to \mathbb{R}$ by

$$\phi(t) = \min_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [\lambda t + c(\lambda\mu)] \quad \forall t \in [a, b].$$

Define $\bar{\phi} : [a, b] \to \mathbb{R}$ by $\bar{\phi}(t) = -\phi(t)$. Then,

$$\bar{\phi}(t) = -\phi(t) = \max_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [-\lambda t - c(\lambda\mu)] = \max_{\lambda' \in \mathbb{R}_- : \lambda'\mu \in -\mathcal{M}} [\lambda' t - c(-\lambda'\mu)] \quad \forall t \in [a, b].$$

Since $\{\lambda' \in \mathbb{R}_- : \lambda'\mu \in -\mathcal{M}\} = -\{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}\}$ is nonempty and compact, and the mapping $\lambda' \mapsto c(-\lambda'\mu)$ is lower semi-continuous, by Theorem 11, $\bar{\phi}$ is Lipschitz continuous and convex. Therefore, ϕ is Lipschitz continuous and concave. Furthermore, for any $A \in \mathcal{A}$,

$$V(A) = \min_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [\langle \sigma_A, \lambda\mu \rangle + c(\lambda\mu)] = \min_{\lambda \in \mathbb{R}_+ : \lambda\mu \in \mathcal{M}} [\lambda \langle \sigma_A, \mu \rangle + c(\lambda\mu)] = \phi(\langle \sigma_A, \mu \rangle),$$

where the first equality follows from Equation (2) and part a. Therefore, it only remains to show that $\bar{\phi}$ is strictly decreasing which will imply that ϕ is strictly increasing.

By Lemma 8, $\mathcal{M} = -\mathcal{M}_{\bar{W}}$ and $c(\nu) = \bar{W}^*(-\nu)$ for all $\nu \in \mathcal{M}$. Note that

$$W(\sigma) = V(A_\sigma) = \phi(\langle \sigma_{(A_\sigma)}, \mu \rangle) = \phi(\langle \sigma, \mu \rangle) \quad \forall \sigma \in \Sigma,$$

where the last equality follows from part 1 of Lemma 5. Therefore, $\bar{W}(\sigma) = \bar{\phi}(\langle \sigma, \mu \rangle)$ for all $\sigma \in \Sigma$. By part 3 of Lemma 11, $-\mathcal{M} = \{\lambda\mu : \lambda \in \mathcal{M}_{\bar{\phi}}\}$. Therefore, since μ is nonzero: $0 \in \mathcal{M}$ if and only if $0 \in \mathcal{M}_{\bar{\phi}}$; part a implies $\mathcal{M}_{\bar{\phi}} \subset \mathbb{R}_-$; and the first part of b implies that 0 is not an isolated point of $\mathcal{M}_{\bar{\phi}}$.

First suppose that $0 \notin \mathcal{M}$, implying $0 \notin \mathcal{M}_{\bar{\phi}}$. Let $t, t' \in [a, b]$ be such that $t < t'$. By Theorem 9,

$$\bar{\phi}(s) = \max_{\lambda \in \mathcal{M}_{\bar{\phi}}} [\lambda s - \bar{\phi}^*(\lambda)] \quad \forall s \in [a, b].$$

Let $\hat{\lambda} < 0$ be a solution of the above maximization at $s = t'$. Then,

$$\bar{\phi}(t') = \hat{\lambda} t' - \bar{\phi}^*(\hat{\lambda}) < \hat{\lambda} t - \bar{\phi}^*(\hat{\lambda}) \leq \max_{\lambda \in \mathcal{M}_{\bar{\phi}}} [\lambda t - \bar{\phi}^*(\lambda)] = \bar{\phi}(t).$$

Next suppose that $0 \in \mathcal{M}$, implying that $0 \in \mathcal{M}_{\bar{\phi}}$. Then,

$$b = \lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}} \frac{c(0) - c(\lambda\mu)}{\lambda} = \lim_{\lambda \searrow 0 : -\lambda \in \mathcal{M}_{\bar{\phi}}} \frac{\bar{\phi}^*(0) - \bar{\phi}^*(-\lambda)}{\lambda} \tag{28}$$

50

where the first equality follows from part b and the second equality follows from $\mathcal{M} = \{\lambda\mu : -\lambda \in \mathcal{M}_{\bar{\phi}}\}$, μ being nonzero, $c(\nu) = \bar{W}^*(-\nu)$ for all $\nu \in \mathcal{M}$, and part 2 of Lemma 11. For any $\lambda \in (0,\infty)$, define $b_\lambda \in \mathbb{R}$ by $b_\lambda = \frac{\bar{\phi}^*(0) - \bar{\phi}^*(-\lambda)}{\lambda}$. Since $0 \in \mathcal{M}_{\bar{\phi}}$ is not an isolated point of $\mathcal{M}_{\bar{\phi}}$, Equation (28) implies that there exists a sequence λ_n in $-\mathcal{M}_{\bar{\phi}} \setminus \{0\}$ such that $\lambda_n \searrow 0$ and $\lim_n b_{\lambda_n} = b$. Since $\bar{\phi}^*$ is convex, b_λ is nonincreasing in $\lambda \in (0,\infty)$. Therefore, for any sequence λ'_n in $(0,\infty)$ such that $\lambda'_n \searrow 0$, $\lim_n b_{\lambda'_n} = \lim_n b_{\lambda_n}$. This implies that the limit on the right hand side of Equation (28) is equal to $\frac{d\bar{\phi}^*}{d\lambda-}(0)$. By Corollary 2, $\bar{\phi}$ is strictly decreasing. ∎

I Proof of Theorem 8

The necessity of the axioms is straightforward. For sufficiency, suppose that $\succsim$ satisfies Axiom 1, strategic rationality, and PERU. By Lemma 3, there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. As we showed in Appendix F, since $\succsim$ satisfies strategic rationality, $V(A) = \max_{p \in A} V(\{p\})$ for all $A \in \mathcal{A}$. Define $f : \triangle(Z) \to \mathbb{R}$ by $f(p) = V(\{p\})$. By Lemma 4, f is Lipschitz continuous and convex. Let

$$H \equiv \left\{ u \in \mathbb{R}^Z : \sum_{z \in Z} u_z = 0 \right\}.$$

The following result identifies H with the set of all linear functions on $\triangle(Z)$ in order to prove a variation of Fenchel duality for the function f:

Lemma 12 *If $f : \triangle(Z) \to \mathbb{R}$ is Lipschitz continuous and convex, then there exists a nonempty compact set $\mathcal{V} \subset H$ and a lower semi-continuous function $c : \mathcal{V} \to \mathbb{R}$ such that*

$$f(p) = \max_{v \in \mathcal{V}} [v(p) - c(v)]$$

for all $p \in \triangle(Z)$. Moreover, this set $\mathcal{V}$ can chosen to be minimal in the following sense: If $\mathcal{V}'$ is a compact proper subset of $\mathcal{V}$, then there exists $p \in \triangle(Z)$ such that $f(p) > \max_{v \in \mathcal{V}'} [v(p) - c(v)]$.

Proof: Suppose that $f : \triangle(Z) \to \mathbb{R}$ is a Lipschitz continuous and convex function. Let $n = |Z|$. Without loss of generality, let $n \geq 2$ and $Z = \{1, \ldots, n\}$. Let

$$C = \left\{ s \in \mathbb{R}^{n-1} : -\frac{1}{n} \leq s_i \text{ and } \sum_{i=1}^{n-1} s_i \leq \frac{1}{n} \right\}.^{43}$$

---

[43]It is standard to represent the set of all probability distributions over a set of n prizes as the set $D = \{s \in \mathbb{R}^{n-1} : 0 \leq s_i \text{ and } \sum_{i=1}^{n-1} s_i \leq 1\}$. The set C is simply the translation of the set D so that the uniform distribution is represented by the 0 vector in $\mathbb{R}^{n-1}$ instead of the vector $(1/n, \ldots, 1/n) \in \mathbb{R}^{n-1}$.

Note that C is closed and convex. Define $\zeta : C \to \triangle(Z)$ by:

$$\zeta(s) = \left(s_1 + \frac{1}{n}, \ldots, s_{n-1} + \frac{1}{n}, -\sum_{j=1}^{n-1} s_j + \frac{1}{n} \right)$$

for any $s \in C$. Note that ζ is well defined (i.e., takes values in $\triangle(Z)$), one-to-one, onto $\triangle(Z)$, and satisfies $\zeta(\alpha s + (1-\alpha)s') = \alpha\zeta(s) + (1-\alpha)\zeta(s')$ for all $s, s' \in C$, $\alpha \in [0,1]$. Therefore, as an affine function defined on a finite-dimensional vector space, ζ is Lipschitz continuous. Also, since ζ is a bijection, the inverse function $\zeta^{-1} : \triangle(Z) \to C$ exists.

Define $\xi : \mathbb{R}^{n-1} \to H$ by:

$$\xi(t) = \left(t_1 - \frac{1}{n}\sum_{j=1}^{n-1} t_j, \ldots, t_{n-1} - \frac{1}{n}\sum_{j=1}^{n-1} t_j, -\frac{1}{n}\sum_{j=1}^{n-1} t_j \right)$$

for any $t = (t_1, \ldots, t_{n-1}) \in \mathbb{R}^{n-1}$. Note that ξ is well defined (i.e., takes values in H), linear, one-to-one, and onto H. Therefore, the inverse function $\xi^{-1} : H \to R^{n-1}$ exists and is linear. Since ξ and ξ^{-1} are linear, they are Lipschitz continuous. Note also that $t \cdot s = \xi(t) \cdot \zeta(s)$ for all $t \in \mathbb{R}^{n-1}$ and $s \in C$.

Define the function $g : C \to \mathbb{R}$ by $g = f \circ \zeta$. Then, g is Lipschitz continuous as the composition of two Lipschitz continuous functions, and g is convex since ζ is affine and f is convex. By Theorem 9, there exists a nonempty compact $T \subset \mathbb{R}^{n-1}$ and a lower semi-continuous function $d : T \to \mathbb{R}$ such that

$$g(s) = \max_{t \in T} [t \cdot s - d(t)]$$

for all $s \in C$ and, in addition, for any compact proper subset T' of T, there exists $s \in C$ such that $g(s) > \max_{t \in T'} [t \cdot s - d(t)]$. Let $\mathcal{V} = \xi(T)$ and define the function $c : \mathcal{V} \to \mathbb{R}$ by $c = d \circ \xi^{-1}$. Then, $\mathcal{V}$ is a compact subset of H and c is lower semi-continuous. For all $p \in \triangle(Z)$,

$$
\begin{aligned}
f(p) = g(\zeta^{-1}(p)) &= \max_{t \in T} [t \cdot \zeta^{-1}(p) - d(t)] \\
&= \max_{t \in T} [\xi(t) \cdot p - c(\xi(t))] \\
&= \max_{v \in \mathcal{V}} [v(p) - c(v)].
\end{aligned}
$$

Moreover, if $\mathcal{V}'$ is a compact proper subset of $\mathcal{V}$, then $T' \equiv \xi^{-1}(\mathcal{V}')$ is a compact proper subset of T. Thus, there exists $s \in C$ such that $g(s) > \max_{t \in T'} [t \cdot s - d(t)]$. Letting $p = \zeta(s)$, it follows that $f(p) > \max_{t \in \mathcal{V}'} [v(p) - c(v)]$. ∎

By Lemma 12, there exists a compact set $\mathcal{V} \subset H$ and a lower semi-continuous function $c : \mathcal{V} \to \mathbb{R}$ such that for any $A \in \mathcal{A}$,

$$V(A) = \max_{p \in A} f(p) = \max_{v \in \mathcal{V}} \left(\max_{p \in A} v(p) - c(v) \right).$$

Moreover, for any proper compact subset $\mathcal{V}'$ of $\mathcal{V}$, there exists $p \in \triangle(Z)$ such that $V(\{p\}) > \max_{v \in \mathcal{V}'} [v(p) - c(v)]$.

Define a function $\phi : H \to C(\mathcal{U})^*$ as follows:

$$\phi(v) = \begin{cases} \|v\| \delta_{\frac{v}{\|v\|}} & \text{if } \|v\| \neq 0 \\ 0 & \text{if } \|v\| = 0. \end{cases}$$

Note that ϕ is norm-to-weak* continuous. To see this, fix any $v \in H$ and any net $\{v_d\}_{d \in D} \subset H$ such that $v_d \to v$. Let $\mu_d = \phi(v_d)$ for $d \in D$ and let $\mu = \phi(v)$. We will show that $\mu_d \xrightarrow{w^*} \mu$. There are two cases to consider:

- *Case 1 — $v = 0$:* In this case, $\|v_d\| \to 0$ and $\mu = 0$. Fix any continuous function $g : \mathcal{U} \to \mathbb{R}$. Then, there exists $M > 0$ such that $|g(u)| \leq M$ for all $u \in \mathcal{U}$. Therefore,

$$\left| \int_{\mathcal{U}} g(u) \mu_d(du) \right| \leq \|v_d\| M \to 0 = \int_{\mathcal{U}} g(u) \mu(du).$$

  Since this is true for any continuous function g, we have $\mu_d \xrightarrow{w^*} \mu$.

- *Case 2 — $v \neq 0$:* In this case, since $v_d \to v$, there exists $d' \in D$ such that $\|v_d\| > 0$ for all $d \geq d'$. Without loss of generality, suppose $\|v_d\| > 0$ for all $d \in D$. Hence, $\frac{v_d}{\|v_d\|} \to \frac{v}{\|v\|}$. Fix any continuous function $g : \mathcal{U} \to \mathbb{R}$. Then,

$$\int_{\mathcal{U}} g(u) \mu_d(du) = \|v_d\| g\left(\frac{v_d}{\|v_d\|} \right) \to \|v\| g\left(\frac{v}{\|v\|} \right) = \int_{\mathcal{U}} g(u) \mu(du).$$

  Since this is true for any continuous function g, we have $\mu_d \xrightarrow{w^*} \mu$.

Let $\mathcal{M} = \phi(\mathcal{V})$. Then, $\mathcal{M}$ is weak* compact by the continuity of ϕ. Define $\tilde{c} : \mathcal{M} \to \mathbb{R}$ by $\tilde{c}(\mu) = c(\int_{\mathcal{U}} u \, \mu(du))$. Since c is lower semi-continuous and the mapping $\mu \mapsto \int_{\mathcal{U}} u \, \mu(du)$ is weak* continuous, $\tilde{c}$ is weak* lower semi-continuous. We claim that $(\mathcal{M}, \tilde{c})$ is a max-HA representation for $\succsim$. To see this, fix any $v \in \mathcal{V}$ and let $\mu = \phi(v)$. Then, $\int_{\mathcal{U}} u \, \mu(du) = v$, and hence $\tilde{c}(\mu) = c(v)$. Therefore,

$$\int_{\mathcal{U}} \max_{p \in A} u(p) \, \mu(du) - \tilde{c}(\mu) = \max_{p \in A} v(p) - c(v),$$

which implies that V satisfies Equation (1) for $(\mathcal{M}, \tilde{c})$:

$$V(A) = \max_{v \in \mathcal{V}} \left(\max_{p \in A} v(p) - c(v) \right) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) - \tilde{c}(\mu) \right).$$

To see that $\mathcal{M}$ is minimal, consider any proper subset $\mathcal{M}'$ of $\mathcal{M}$. Let $\mathcal{V}' = \{v \in \mathcal{V} : \phi(v) \in \mathcal{M}'\}$. Then, $\mathcal{V}'$ is a compact proper subset of $\mathcal{V}$ such that $\phi(\mathcal{V}') = \mathcal{M}'$. By the minimality of $\mathcal{V}$, there exists $p \in \triangle(Z)$ such that

$$V(\{p\}) > \max_{v \in \mathcal{V}'} (v(p) - c(v)) = \max_{\mu \in \mathcal{M}'} \left(\int_{\mathcal{U}} u(p)\, \mu(du) - \tilde{c}(\mu) \right).$$

Thus, $\mathcal{M}$ is minimal.

References

Aliprantis, C., and K. Border (1999): *Infinite Dimensional Analysis*. Berlin, Germany: Springer-Verlag.

Billingsley, P. (1995): *Probability and Measure*. New York: John Wiley and Sons, Inc.

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 (2009): "Temptation Driven Preferences," *Review of Economic Studies*, 76, 937–971.

Dekel, E., B. Lipman, A. Rustichini, and T. Sarver (2007): "Representing Preferences with a Unique Subjective State Space: Corrigendum," *Econometrica*, 75, 591–600.

Ekeland, I., and T. Turnbull (1983): *Infinite-Dimensional Optimization and Convexity*. Chicago: The University of Chicago Press.

Epstein, L. G. (2008): "Living with Risk," *Review of Economic Studies*, 75, 1121–1141.

Epstein, L. G., M. Marinacci, and K. Seo (2007): "Coarse Contingencies and Ambiguity," *Theoretical Economics*, 2, 355–394.

Epstein, L. G., and K. Seo (2007): "Subjective States: A More Robust Model," *Games and Economic Behavior*, forthcoming.

Epstein, L. G., and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Ergin, H. (2003): "Costly Contemplation," Mimeo.

Ergin, H., and T. Sarver (2009): "A Unique Costly Contemplation Representation," Mimeo.

Ergin, H., and T. Sarver (2008): "An Extension and Application of Mazur's Theorem," Mimeo.

Fishburn, P. C. (1970): *Utility Theory for Decision Making*. New York: John Wiley and Sons.

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.

Kreps, D. (1988): *Notes on the Theory of Choice.* Colorado: Westview Press.

Kreps, D., and E. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

Kreps, D., and E. Porteus (1979): "Temporal von Neumann-Morgenstern and Induced Preferences," *Journal of Economic Theory*, 20, 81–109.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447-1498.

Machina, M. (1984): "Temporal Risk and the Nature of Induced Preferences," *Journal of Economic Theory*, 33, 199–231.

Markowitz, H. M. (1959): *Portfolio Selection: Efficient Diversification of Investments.* New York: John Wiley and Sons.

Mossin, J. (1969): "A Note on Uncertainty and Preferences in a Temporal Context," *American Economic Review*, 59, 172–174.

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. (2008): "Anticipating Regret: Why Fewer Options May Be Better," *Econometrica*, 76, 263–305.

Schneider, R. (1993): *Convex Bodies: The Brunn-Minkowski Theory.* Cambridge: Cambridge University Press.

Spence, M., and R. Zeckhauser (1972): "The Effect of the Timing of Consumption Decisions and the Resolution of Lotteries on the Choice of Lotteries," *Econometrica*, 40, 401–403.