

Hidden Actions and Preferences for Timing of Resolution of Uncertainty*

Haluk Ergin[†]

Todd Sarver[‡]

May 3, 2011

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 a general class of utility representations called hidden action representations, which interpret an intrinsic preference for timing of resolution of uncertainty as if an unobservable action is taken between the resolution of the two periods of information arrival. These representations permit a richer class of preferences for timing than was possible in Kreps and Porteus (1978) by incorporating a preference for flexibility. Our model contains several special cases where this hidden action can be given a novel economic interpretation, including a subjective-state-space model of ambiguity aversion and a model of costly contemplation.

*We thank the National Science Foundation (Grants SES-0850980 and SES-0851066) for financial support, and Eddie Dekel and Pablo Schenone for helpful comments. These results were previously distributed under the title “A Subjective Model of Temporal Preferences”.

[†]Correspondence address: Duke University, Department of Economics, 213 Social Sciences Building, 419 Chapel Drive, Box 90097, Durham, NC 27708-0097. Email: haluk.ergin@duke.edu.

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

1 Introduction

This paper considers several new classes of dynamic preferences, providing representations for preferences for both early and late resolution of uncertainty. The first purpose of this analysis is to unite two strands of the literature: We consider a model in which an individual may have an intrinsic preference for timing of resolution of uncertainty (as in Kreps and Porteus (1978)) while at the same time exhibiting a preference for flexibility (as in Kreps (1979) and Dekel, Lipman, and Rustichini (2001, henceforth DLR)). The second purpose of this analysis is to provide a simple and intuitive interpretation for such preferences: We provide a representation that suggests that intrinsic preferences for timing of resolution of uncertainty can be interpreted as being the result of some interim action that is not observable to the modeler (a *hidden* action). Thus, intrinsic preference for timing can be understood as an extrinsic (or instrumental) preference for timing arising due to some unobserved action.

1.1 Intrinsic Versus Extrinsic Preferences for Timing

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 optimally 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 model that enriches the additive dynamic expected-utility model by allowing 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 does not change with the timing of its resolution. 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 con-

sidered by Kreps and Porteus (1978) allow the individual to have a strict preference for that uncertainty to resolve earlier (or later), which we refer to as an intrinsic preference for the timing of the resolution of uncertainty. The additional flexibility of their model has proved useful in applications to macroeconomic models of asset pricing (Epstein and Zin (1989, 1991)), precautionary savings (Weil (1993)), and business cycles (Tallarini (2000)) (see Backus, Routledge, and Zin (2004) for a survey of these and related papers).

While an intrinsic preference for early resolution of uncertainty occurs by definition in the absence of any directly observable payoff-relevant action, it is possible that the individual does in fact take payoff-relevant action that is simply unobservable to the economic modeler. For example, suppose the individual described above is not permitted to save or borrow, yet still exhibits a preference for early resolution of uncertainty about income. It may be the case that this individual has some additional unobserved payoff-relevant action that she would like to condition on the resolution of this uncertainty. Thus, her apparent intrinsic preference for early resolution of uncertainty could in fact be an extrinsic preference arising due to an unobserved action. Kreps and Porteus (1979) provided an interpretation along these lines to the preferences considered in their 1978 paper. Our main representation theorem provides a similar *hidden action* interpretation for a broader class of preferences, which, in particular, permits us to consider some special cases where this hidden action can be given a novel economic interpretation.

1.2 A New Class of Preferences

We examine dynamic preferences in a simple menu-choice setting with two-stage objective uncertainty. This framework is a two-stage version of the environment considered by Kreps and Porteus (1978). However, we allow for more general axioms, which permits us to model a richer set of preferences for early or late resolution of uncertainty. In particular, we relax the strategic rationality axiom imposed implicitly by Kreps and Porteus (1978) (see Axiom 7 and the surrounding discussion) to allow for a preference for flexibility as in Kreps (1979) and DLR (2001).

One important feature of the preferences considered by Kreps and Porteus (1978) is that a preference for timing of resolution of uncertainty can occur when the individual has an interim choice only if the individual also exhibits the same preference for timing of resolution of uncertainty when she has no interim choice. In contrast, we allow for more general patterns of preference for timing, such as an individual who exhibits a preference for early resolution of uncertainty only for decision problems that offer the possibility of non-degenerate interim choices. Therefore, we not only allow for preference for timing together with preference for flexibility, but we allow the two to be connected

in a nontrivial way. As a result, we are able to incorporate some interesting and useful special cases that cannot be included in the original model of Kreps and Porteus (1978).

1.3 Overview of Results

We show that our general class of preferences for early and late resolution of uncertainty can be represented 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 (a 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 nature's ability to harm her.¹

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-state-space versions of a number of well-known representations in the literature. After describing the setting for our model in Section 2 and establishing our hidden action representation and uniqueness theorems in Section 3, we consider some of these special cases in Section 4. In Section 4.1, we show that subjective-state-space versions of the multiple priors model of Gilboa and Schmeidler (1989) and the variational preferences model of Maccheroni, Marinacci, and Rustichini (2006) overlap with the class of hidden action preferences exhibiting a preference for late resolution of uncertainty. In Section 4.2, we characterize the costly contemplation model of Ergin and Sarver (2010a) as a special case of the class of hidden action preferences exhibiting a preference for early resolution of uncertainty. The general framework in this paper provides a unification of these well-known representations and provides simple axiomatizations.

In Section 4.3, we formally connect our results to those of Kreps and Porteus (1978) and several other related papers. In Section 4.3.1, we discuss the axioms and representation result from Kreps and Porteus (1978), which allows us to formally illustrate the types of preferences for early and late resolution of uncertainty that are possible in our model but not in theirs. For example, we show that (non-trivial) costly contemplation

¹While 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. See, for example, Maccheroni, Marinacci, and Rustichini (2006) for a related discussion.

can only be modeled using preferences that lie outside of the class considered by Kreps and Porteus (1978). In Section 4.3.2, we present a generalization of the Kreps and Porteus (1978) representation that allows for subjective uncertainty as in DLR (2001). In Section 4.3.3, we provide an equivalence result that shows how these representations can be expressed as special cases of our hidden action representation, therefore generalizing some of the equivalence results of Kreps and Porteus (1979). We conclude Section 4.3.3 by also characterizing a related representation considered by Machina (1984) as a special case of our hidden action representation.

2 Choice Setting

Let Z be a finite set of alternatives, and let $\Delta(Z)$ denote the set of all probability distributions on Z , endowed with the Euclidean metric d and with generic elements denoted p, q, r . Let \mathcal{A} denote the set of all nonempty and closed subsets of $\Delta(Z)$, 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, with generic menus denoted A, B, C . Let $\Delta(\mathcal{A})$ denote the set of all Borel probability measures on \mathcal{A} , endowed with the weak* topology and with generic elements denoted P, Q, R .² The primitive of the model is a binary relation \succsim on $\Delta(\mathcal{A})$, 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 period 1, the individual first chooses a lottery P over menus. Then, the uncertainty associated with this chosen lottery P resolves, returning a menu A . In the (unmodeled) period 2, the individual chooses a lottery p out of A and this lottery resolves, returning an alternative z . We will refer to the uncertainty associated with the resolution of P as the *first-stage uncertainty* and 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.

For any $A \in \mathcal{A}$, let $\delta_A \in \Delta(\mathcal{A})$ 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 . Figure 1 illustrates such a

²Given 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 $\int_X f \mu_d(dx) \rightarrow \int_X f \mu(dx)$ for every bounded continuous function $f : X \rightarrow \mathbb{R}$.

lottery $P = \alpha\delta_A + (1 - \alpha)\delta_B$ for the case of $A = \{p^1, p^2\}$ and $B = \{q^1, q^2\}$, where $p^i = \beta_i\delta_{z_i} + (1 - \beta_i)\delta_{z'_i}$ and $q^i = \gamma_i\delta_{\tilde{z}_i} + (1 - \gamma_i)\delta_{\tilde{z}'_i}$. In this figure, nodes with rounded edges are those at which nature acts, and square nodes are those at which the individual makes a decision.

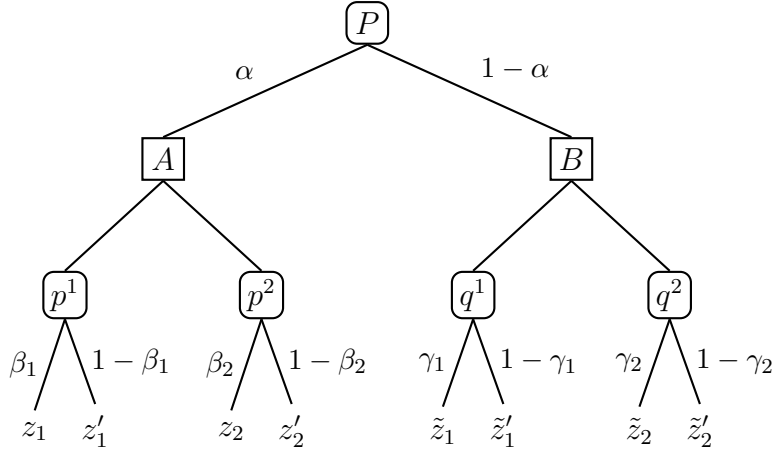


Figure 1: Decision Tree for the Lottery P

Our framework is a special case of that of Kreps and Porteus (1978), with only two periods and no consumption in period 1.³ As in Kreps and Porteus (1978), we refer to a lottery $P \in \Delta(\mathcal{A})$ over menus as a *temporal lottery* 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 $\Delta(\Delta(Z))$.

For any $A, B \in \mathcal{A}$ and $\alpha \in [0, 1]$, the convex combination of these two menus is defined by $\alpha A + (1 - \alpha)B \equiv \{\alpha p + (1 - \alpha)q : p \in A \text{ and } q \in B\}$. Let $\text{co}(A)$ denote the convex hull of the menu A . Finally, for any continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ and $P \in \Delta(\mathcal{A})$, we let $\mathbb{E}_P[V]$ denote the expected value of V under the lottery P , i.e., $\mathbb{E}_P[V] = \int_{\mathcal{A}} V(A) P(dA)$.

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

3 General Representations

3.1 Axioms

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.

Axiom 1

1. (Weak Order): \succsim is complete and transitive.
2. (Continuity): The upper and lower contour sets, $\{P \in \Delta(\mathcal{A}) : P \succsim Q\}$ and $\{P \in \Delta(\mathcal{A}) : P \precsim Q\}$, are closed in the weak* topology.
3. (First-Stage Independence): For any $P, Q, R \in \Delta(\mathcal{A})$ 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 $\alpha \geq Md_h(A, B)$,

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

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 $V : \mathcal{A} \rightarrow \mathbb{R}$ 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.

⁴In 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 (2010a).

Kreps and Porteus (1978) defined preference for early and late resolution of uncertainty using temporal lotteries. Formally, their preference for early resolution of uncertainty (PERU) axiom states that for any $p, q \in \Delta(Z)$ and $\alpha \in [0, 1]$,

$$\alpha\delta_{\{p\}} + (1 - \alpha)\delta_{\{q\}} \succeq \delta_{\{\alpha p + (1-\alpha)q\}}. \quad (1)$$

Their preference for late resolution of uncertainty (PLRU) axiom is defined similarly. In the temporal lottery $\alpha\delta_{\{p\}} + (1 - \alpha)\delta_{\{q\}}$, uncertainty regarding whether lottery p or q is selected resolves in period 1. In the temporal lottery $\delta_{\{\alpha p + (1-\alpha)q\}}$, 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.

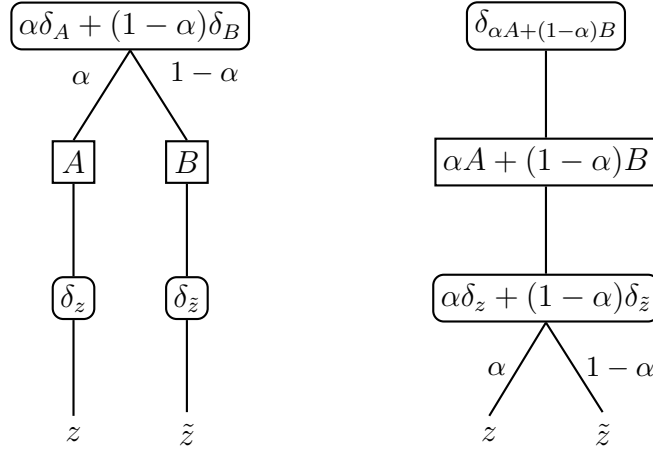


Figure 2: Illustration of Timing of Resolution of Uncertainty for Temporal Lotteries: $A = \{\delta_z\}$ and $B = \{\delta_{\tilde{z}}\}$

Figure 2 illustrates such temporal lotteries in the special case where $p = \delta_z$ and $q = \delta_{\tilde{z}}$ for some $z, \tilde{z} \in Z$. In this figure, nodes with rounded edges are those at which nature acts, and rectangular nodes are those at which the individual makes a decision. Since the trees in this figure correspond to temporal lotteries, the action nodes for the individual are always degenerate. The temporal lottery $\alpha\delta_{\{\delta_z\}} + (1 - \alpha)\delta_{\{\delta_{\tilde{z}}\}}$ corresponds to the first tree in Figure 2, in which the uncertainty about whether alternative z or \tilde{z} will be selected resolves in period 1. The temporal lottery $\delta_{\{\alpha\delta_z + (1-\alpha)\delta_{\tilde{z}}\}}$ corresponds to the second tree in Figure 2, in which the uncertainty about whether z or \tilde{z} will be selected resolves in period 2.

⁵In 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 also resolved in period 2.

Kreps and Porteus (1978) impose other axioms that tie the preference for timing for general 2-stage decision problems to the preference for timing on temporal lotteries. Since we make weaker overall assumptions on preferences, we adapt their preference for timing axioms to be explicit about the preferences being imposed on lotteries involving non-degenerate choices.⁶

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 $\alpha \in (0, 1)$,*

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

In the general case where A and B need not be singletons, $\delta_{\alpha A + (1 - \alpha)B}$ and $\alpha\delta_A + (1 - \alpha)\delta_B$ may 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 equivalent to a pair of choices $p \in A$ and $q \in B$, where after the individual chooses (p, q) , lottery p is selected with probability α and q is selected with probability $1 - \alpha$. Therefore, the period 2 choice from the menu $\alpha A + (1 - \alpha)B$ can be interpreted as a complete contingent plan out of the menus A and B .

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

⁶Note that while our preference for timing axioms are stronger than those explicitly stated by Kreps and Porteus (1978), Axioms 2 and 3 are implied by their temporal lottery counterparts when the other axioms of Kreps and Porteus (1978) are imposed.

It is also worth noting that other authors have used stronger versions of these preference for timing axioms in order to relax other assumptions on the preferences. For example, to study recursive non-expected utility models over temporal lotteries, Grant, Kajii, and Polak (1998, 2000) introduced a stronger version Equation (1) which, roughly speaking, requires agents to prefer when the resolution of the first-stage uncertainty is more informative in the sense of Blackwell.

preference to learn which menu is selected prior to making a choice out of the menus, whereas PLRU is the individual's preference to learn which menu is selected after making her contingent choice out of the menus. Since these axioms permit the individual to make contingent plans in the case of late resolution of uncertainty, her ability to condition her choice from menus on the realized menu is unaffected by the timing of resolution of uncertainty, ensuring that our axioms are only capturing intrinsic preferences for timing.

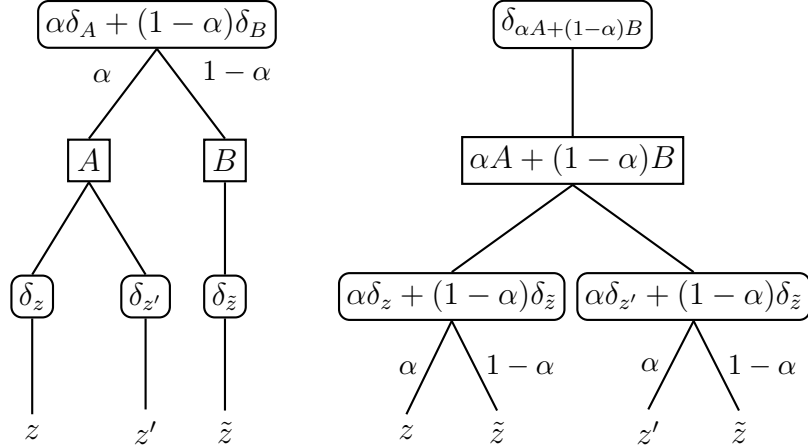


Figure 3: Illustration of Timing of Resolution of Uncertainty for non-Temporal Lotteries: $A = \{\delta_z, \delta_{z'}\}$ and $B = \{\delta_{\tilde{z}}\}$

Figure 3 illustrates timing of resolution of uncertainty in the case where $A = \{\delta_z, \delta_{z'}\}$ and $B = \{\delta_{\tilde{z}}\}$. The lottery $\alpha\delta_A + (1 - \alpha)\delta_B$ corresponds to the first tree in Figure 3, in which the uncertainty about whether the choice set will be A or B resolves in period 1, before the individual makes her choice from the realized menu. The lottery $\delta_{\alpha A + (1 - \alpha)B}$ corresponds to the second tree in Figure 3, in which the individual's period 2 choice is made prior to the resolution of uncertainty regarding whether her choice from A or B will be implemented. In this tree, the lottery $\alpha\delta_z + (1 - \alpha)\delta_{\tilde{z}}$ can be interpreted as a contingent plan where the individual commits to choosing δ_z if A is the realized choice set and $\delta_{\tilde{z}}$ if B is the realized choice set. Similarly, $\alpha\delta_{z'} + (1 - \alpha)\delta_{\tilde{z}}$ corresponds to making a contingent choice of $\delta_{z'}$ from the menu A .

The final axiom for our general model is a standard monotonicity axiom, which requires a weak preferences for larger menus.

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

Kreps (1979) and DLR (2001) use this axiom to capture a preference for flexibility. For example, if the individual is uncertain of whether she will prefer to choose lottery

p or q in period 2, then in period 1 she may strictly prefer the flexibility of $\delta_{\{p,q\}}$ to committing to either $\delta_{\{p\}}$ or $\delta_{\{q\}}$.⁷

3.2 Hidden Action Representations

Since expected-utility functions on $\Delta(Z)$ 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 $\Delta(Z)$ to be

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

We are ready to introduce our general representations:⁸

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 Borel measures \mathcal{M} on \mathcal{U} and a lower semi-continuous function $c : \mathcal{M} \rightarrow \mathbb{R}$ such that:

1. $P \succeq Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by Equation (2) [(3)]:

$$V(A) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) - c(\mu) \right) \quad (2)$$

$$V(A) = \min_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) + c(\mu) \right). \quad (3)$$

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 (2) [(3)] is different from V .

⁷It has been suggested that preferences for early or late resolution of uncertainty could also arise due to anticipatory feelings or anxiety. While our hidden action representation is in principle consistent with such an interpretation (for example, anticipating a particular level of consumption could be thought of as a hidden action on the part of the individual), our axioms are inconsistent with several of the well-known models of anticipatory feels in the literature (e.g., Caplin and Leahy (2001) and Epstein (2008)) because of our assumption of monotonicity. Loosely speaking, Caplin and Leahy (2001) and Epstein (2008) assume that anticipation has a greater impact on utility in early stages than in later, which causes the individual's ranking of lotteries to change over time. If the individual correctly foresees that she will be dynamically inconsistent in this way, then she will strictly prefer to commit herself to a particular lottery at the first stage and, hence, will violate monotonicity. We leave as an interesting question for future research whether an alternative (monotone) model of anticipatory feelings could be formulated as a special case of our hidden action representation.

⁸We endow the set of all finite Borel measures on \mathcal{U} with the weak* topology (see footnote 2).

The pair (\mathcal{M}, c) is an *HA representation* if it is a max-HA or a min-HA representation.

The following lemma shows that after appropriately renormalizing the set of ex post utility functions, one can reinterpret the integral term in Equations (2) and (3) as an expectation. Therefore, the HA representation can be interpreted as a normalized version of a representation in which the individual has subjective uncertainty about her ex post (period 2) utility function over $\Delta(Z)$.⁹

Lemma 1 *For any finite Borel measure μ on \mathcal{U} , there exists a probability measure π on the set $\mathcal{V} \equiv \mu(\mathcal{U})\mathcal{U}$ such that for all $A \in \mathcal{A}$,*

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

Conversely, for any compact set $\mathcal{V} \subset \mathbb{R}^Z$ and any probability measure π on \mathcal{V} , there exists a unique finite Borel measure μ on \mathcal{U} and scalar β such that for all $A \in \mathcal{A}$,

$$\int_{\mathcal{V}} \max_{p \in A} v(p) \pi(dv) = \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) + \beta.$$

Proof: Since this lemma follows from the same arguments used to prove Lemma 1 in Ergin and Sarver (2010a), we only provide the key steps. To prove the first claim, let $\lambda \equiv \mu(\mathcal{U}) \geq 0$ and let $\mathcal{V} \equiv \lambda\mathcal{U}$. If $\lambda = 0$, define π by $\pi(\{0\}) = 0$. Otherwise, define π for any measurable set $E \subset \mathcal{V}$ by $\pi(E) = \mu(\frac{1}{\lambda}E)/\lambda$. Heuristically, π puts weight $\mu(u)/\lambda$ on each $v = \lambda u \in \mathcal{V}$. Therefore,

$$\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) = \frac{1}{\lambda} \int_{\mathcal{U}} \max_{p \in A} \lambda u(p) \mu(du) = \int_{\mathcal{V}} \max_{p \in A} v(p) \pi(dv).$$

To prove the converse, note that for every $v \in \mathcal{V}$, there exist $a_v \geq 0$, $b_v \in \mathbb{R}$, and $u_v \in \mathcal{U}$ such that $v = a_v u_v + b_v$. Let $\beta = \int_{\mathcal{V}} b_v \pi(dv)$, and define a Borel measure μ by $\mu(E) = \int_{\{v \in \mathcal{V}: u_v \in E\}} a_v \pi(dv)$ for a measurable set $E \subset \mathcal{U}$. Using a standard change of variables, it follows that for every $A \in \mathcal{A}$,

$$\begin{aligned} \int_{\mathcal{V}} \max_{p \in A} v(p) \pi(dv) &= \int_{\mathcal{V}} a_v \max_{p \in A} u_v(p) \pi(dv) + \int_{\mathcal{V}} b_v \pi(dv) \\ &= \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) + \beta. \end{aligned}$$

⁹Note that the constant β in the second part of the lemma can be absorbed into the function c in the HA representation.

Intuitively, the magnitude of each utility function v is incorporated into the measure of the corresponding u_v . ■

Although the measures in our representation can be given a probabilistic interpretation, we prefer to formulate our general representation using non-probability measures μ that capture the combination of the probability and magnitude (cardinality) of ex post utility. This formulation has the important benefit of allowing for the unique identification of the parameters in our representation¹⁰, and it also simplifies the mathematical statement of some results.

We next interpret Equation (2). 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 $\Delta(Z)$, at cost $c(\mu)$. As argued above, the integral in Equation (2) can be interpreted as a reduced-form representation for the value of the action μ when the individual chooses from menu A . For each menu A , the individual maximizes the value minus cost of her action.

The interpretation of Equation (3) is dual. In this case, the individual is pessimistic about the measure μ that she will face in period 2. One way to interpret such preferences in terms of a hidden action is the following: In period 1, the individual anticipates that after the first-stage uncertainty is resolved but before she makes her choice in period 2, (a malevolent) nature will select an action μ from a set \mathcal{M} . The individual anticipates that nature will choose an action which minimizes the value to the individual plus a cost term. The function c can be interpreted as capturing the pessimism attitude of the individual. For constant c , she expects nature to choose 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.

In the above representations, both the set of available actions 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 (2) and (3), 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} .

¹⁰There are many pairs (\mathcal{V}, π) that give the same integral expression as the measure μ on \mathcal{U} . The lack of identification of probabilities is a common issue in models with state-dependent utility. See Kreps (1988) for a general discussion of the state-dependence issue, and Section 3 of Ergin and Sarver (2010a) for discussion specific to this setting.

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

Theorem 1 *The preference \succsim has a max-HA [min-HA] representation if and only if it satisfies Axiom 1, PERU [PLRU], and monotonicity.*^{11,12}

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 from Equations (2) and (3), leading to an analogue of DLR (2001)'s additive representation in which the individual reduces compound lotteries.

We next give a brief intuition about Theorem 1. Axiom 1 guarantees the existence of a Lipschitz continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ such that $V(\text{co}(A)) = V(A)$ 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 a set of continuous functions Σ known as the support functions, preserving the metric and the linear operations. Therefore, by using the property $V(\text{co}(A)) = V(A)$ 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} . This allows us to 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.¹³ Then, we apply the Riesz representation theorem to write each such continuous affine function as an integral against a measure μ minus [plus] a scalar $c(\mu)$. Finally, 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 from the theory of conjugate convex functions (see Theorem 12 in Appendix A). A similar application of the duality and uniqueness results can be found in Ergin and Sarver (2010a).

¹¹IR 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.

¹²It is possible to relax the assumption of monotonicity if signed measures are permitted in the HA representations. However, since our main focus in this paper is on preferences that satisfy the monotonicity axiom, we relegate this representation result for non-monotone preferences to Appendix B.

¹³See Rockafellar (1970), Phelps (1993), and Appendix A of the current paper for variations of this duality result.

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 Special Cases

4.1 Ambiguity Aversion and Robustness

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 min-HA model is equivalent to two representations that have natural interpretations in terms of ambiguity-aversion and robustness. 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} , $U : \Omega \rightarrow \mathbb{R}^Z$ is a Z -dimensional, \mathcal{F} -measurable, and bounded random vector, and Π is a set of probability measures on (Ω, \mathcal{F}) , such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by

$$V(A) = \min_{\pi \in \Pi} \int_{\Omega} \max_{p \in A} U(\omega) \cdot p \pi(d\omega), \quad (4)$$

and the minimization in Equation (4) 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} , $U : \Omega \rightarrow \mathbb{R}^Z$ is a Z -dimensional, \mathcal{F} -measurable, and bounded random vector, Π is a set of probability measures on (Ω, \mathcal{F}) , and $c : \Pi \rightarrow \mathbb{R}$ is a function, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by

$$V(A) = \min_{\pi \in \Pi} \left(\int_{\Omega} \max_{p \in A} U(\omega) \cdot p \pi(d\omega) + c(\pi) \right), \quad (5)$$

and the minimization in Equation (5) has a solution for every $A \in \mathcal{A}$.¹⁴

The SSV representation generalizes the SSMP representation by allowing a “cost” $c(\pi)$ to be assigned to each measure π in the representation. Like the SSMP representation, the SSV representation has an ambiguity-aversion interpretation; however, special cases of the function c can also be interpreted in terms of robustness to model misspecification. Specifically, a subjective-state-space version of the multiplier preferences considered by Hansen and Sargent (2001) can be obtained by taking $c(\pi) = \theta R(\pi \parallel \eta)$ for some scalar $\theta > 0$ and reference probability measure η , where R is the *relative entropy* of π with respect to η ,

$$R(\pi \parallel \eta) = \begin{cases} \int_{\Omega} \left(\log \frac{d\pi}{d\eta}(\omega) \right) \pi(d\omega), & \text{if } \pi \ll \eta, \\ +\infty, & \text{otherwise.} \end{cases}$$

See Maccheroni, Marinacci, and Rustichini (2006) and Strzalecki (2011) for additional discussion and axiomatic foundations in an objective-state-space setting.

In the Anscombe-Aumann framework, 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-state-space versions of the representations is the state-dependence of the utility functions. The following theorem formalizes this claim and, moreover, shows that a preference \succsim can be represented by one of these representations if and only if it has a min-HA representation.

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

1. *There exists a min-HA representation such that V is given by Equation (3).*
2. *There exists an SSMP representation such that V is given by Equation (4).*

¹⁴Note that for simplicity, we directly assume in the SSMP and SSV representations that the minimizations in Equations (4) and (5) have solutions. One alternative approach that does not require this indirect assumption on the parameters would be to replace the minimums in Equations (4) and (5) 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 bounded and continuous, Π is weak*-compact, and c is lower semi-continuous.

3. *There exists an SSV representation such that V is given by Equation (5).*

Proof Sketch: The complete proof of Theorem 3 in the appendix includes arguments related to measurability, the minimality and compactness of \mathcal{M} in the min-HA representation, existence of a minimizing measure in the SSMP and SSV representations, and other technical details. We sketch the main ideas here.

Given Lemma 1, the equivalence of (1) and (3) is not surprising; the SSV representation simply makes our probabilistic interpretation of the min-HA representation literal.¹⁵ Also, (2) \Rightarrow (3) is immediate since the SSMP representation is a special case of the SSV representation. Therefore, the substantive part of this result is (3) \Rightarrow (2).

To see the intuition for (3) \Rightarrow (2), consider an SSV representation $((\Omega, \mathcal{F}), U, \Pi, c)$, and define V by Equation (5). For simplicity, suppose Ω and Π are finite. Let $\tilde{\Omega} = \Omega \times \Pi$ be the state space for the constructed SSMP representation. Letting $\mathbf{1} \in \mathbb{R}^Z$ denote the vector whose coordinates are equal to 1, define $\tilde{U} : \tilde{\Omega} \rightarrow \mathbb{R}^Z$ by $\tilde{U}(\omega, \pi) = U(\omega) + c(\pi)\mathbf{1}$ for $\tilde{\omega} = (\omega, \pi) \in \tilde{\Omega}$. For any probability measure $\pi \in \Pi$, 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 $\tilde{\Omega}$. Also, for any $A \in \mathcal{A}$,

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

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 a consequence of state-dependence of utility, which allows any constant to be absorbed in the utility function. Above, integration against the probability measure π plus the constant $c(\pi)$ is re-expressed as integration against the probability measure ρ_π whose support is a subset of states in $\Omega \times \{\pi\}$ where the utility function is “shifted” by $c(\pi)\mathbf{1}$. ■

The parameters of the SSMP and SSV representations cannot be uniquely identified from the preference. Theorem 3 illustrates the extent of this non-uniqueness: Since

¹⁵The only other difference is that the SSV representation is formulated using a set of priors over a state space Ω and a random variable U . Taking the distribution of U for each of these priors gives a set of probability measures over \mathbb{R}^Z . Then, Lemma 1 can be applied to write these as measures over \mathcal{U} .

the SSMP representation is the special case of the SSV representation where the cost function is identically equal to 0, the equivalence of the SSMP and SSV representations implies, in particular, that it cannot be determined from the preference whether or not c in the SSV representation takes non-zero values. However, given our uniqueness result for the min-HA representation and Theorem 3, it follows that the min-HA representation identifies the equivalence classes of SSV representations that lead to the same choice behavior. Therefore, when considering these models, working with the equivalent min-HA representation is desirable since its parameters are uniquely identified and therefore have behavioral meaning.

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 provide an intuition for why the SSMP and SSV representations satisfy PLRU in our model. In the SSMP and SSV representations, the agent’s subjective uncertainty about the state $\omega \in \Omega$ is resolved in between the two stages of objective uncertainty. Therefore, the timing of resolution of uncertainty determines the order in which the agent faces objective and subjective uncertainty: In the case of early resolution of uncertainty, objective uncertainty precedes subjective uncertainty; in the case of late resolution of uncertainty, subjective uncertainty precedes objective uncertainty. Since ambiguity averse agents prefer objective risk to follow subjective risk for hedging purposes, they show a preference for late resolution of uncertainty.¹⁶

¹⁶Strzalecki (2009) also studied attitudes toward timing of resolution of uncertainty implied by different ambiguity aversion models. The main difference between his analysis and ours is that Strzalecki (2009) studied the timing of resolution of *subjective* uncertainty, whereas we study the timing of resolution of *objective* uncertainty. While the general principle that ambiguity averse agents prefer subjective uncertainty to resolve prior to objective uncertainty for hedging purposes remains valid in his model, the setting of his model leads to some conclusions that may appear inconsistent with our results. For example, in Strzalecki (2009)’s model, the dynamic multiple-priors preferences exhibit indifference to timing of resolution of subjective uncertainty.

To understand these apparently contradictory results, note that Strzalecki (2009) adopted a multi-period Anscombe–Aumann model that allows for objective uncertainty about the consumption in each period separately, but does not allow for objective uncertainty about the consumption at future dates. Therefore, in his model, changing the timing of resolution of (subjective) uncertainty does not effect the order of subjective and objective uncertainty. If Strzalecki (2009)’s model were extended to allow for temporal timing of objective uncertainty as well as temporal timing of subjective uncertainty, then the multiple-priors preferences would show a preference for early resolution of (subjective) uncertainty for the hedging purposes described above. This is precisely the same reason the SSMP and SSV agents show a preference for late resolution of (objective) uncertainty in our model.

Our SSMP representation bears some similarity to a representation considered by Epstein, Marinacci, and Seo (2007, Theorem 1). Moreover, our motivation for PLRU in terms of ambiguity aversion and hedging (or a malevolent nature) parallels their discussion (see page 361). However, since they work with the simpler setting of menus of lotteries, it is necessary for them to make some auxiliary assumptions on the preference in order to obtain their representation.¹⁷ As a consequence, their representation is also more restrictive than our SSMP representation; specifically, their representation imposes a normalization on the state-dependent utility functions.¹⁸

As the arguments in the proof sketch 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 min-HA representation, it would not be possible to write every 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 by 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

Recall that a choice out of the convex combination menu $\alpha A + (1 - \alpha)B$ can be interpreted as a complete contingent plan out of the two menus A and B : Each lottery $\alpha p + (1 - \alpha)q \in \alpha A + (1 - \alpha)B$ 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, PERU can be naturally attributed to a desire to avoid making complete contingent plans. Note however that a pure desire to avoid contingent planning is a special kind of PERU. For instance, when the menus A and B are singletons so that the contingent planning problem faced in $\alpha A + (1 - \alpha)B$ is trivial, there is no reason

¹⁷We are referring to *Worst* and *Certainty Independence* axioms used in their Theorem 1. Epstein, Marinacci, and Seo (2007) acknowledge that these two axioms are “excess baggage” (page 363).

¹⁸Kraus and Sagi (2006, Theorem 5.1) also studied a representation that bears some similarity to a multi-utility version of SSMP for incomplete preferences. They too imposed a normalization (different than that of Epstein, Marinacci, and Seo (2007)) on the state-dependent utility functions.

for an individual who is averse to contingent planning to prefer $\alpha\delta_A + (1 - \alpha)\delta_B$ over $\delta_{\alpha A + (1-\alpha)B}$. In particular, if the driving force underlying an individual's PERU is solely an aversion to contingent planning, then it is natural to observe indifference to timing of resolution of uncertainty over temporal lotteries.

In Ergin and Sarver (2010a), we study preferences exhibiting aversion to contingent planning in the simpler framework of preferences over menus of lotteries. We obtain a representation for such preferences that can be interpreted in terms of costly contemplation. The following is the natural extension of that representation to the current framework of lotteries over menus.

Definition 4 A *Costly Contemplation (CC) representation* is a tuple $((\Omega, \mathcal{F}, \mathbb{P}), \mathbf{G}, U, c)$ where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, \mathbf{G} is a collection of sub- σ -algebras of \mathcal{F} , U is a Z -dimensional, \mathcal{F} -measurable, and integrable random vector, and $c : \mathbf{G} \rightarrow \mathbb{R}$ is a function, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by

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

and the minimization in Equation (6) has a solution for every $A \in \mathcal{A}$.¹⁹

The interpretation of the CC representation is as follows. The individual is uncertain about her tastes over $\Delta(Z)$. This uncertainty is modeled by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a state-dependent expected-utility function U over $\Delta(Z)$. Before making a choice out of a menu A , the individual is able to engage in contemplation in order to resolve some of this uncertainty. Contemplation strategies are modeled as signals about the state or, more compactly, where as a collection \mathbf{G} of σ -algebras generated by these signals. If the individual carries out the contemplation strategy \mathcal{G} , she is able to update her expected-utility function using her information \mathcal{G} and choose a lottery p in A maximizing her conditional expected-utility $\mathbb{E}_{\mathbb{P}}[U | \mathcal{G}] \cdot p$. Faced with the menu A , the individual chooses her contemplation strategy optimally by maximizing the ex ante value minus the cost $c(\mathcal{G})$ of contemplation, giving Equation (6). Note that the CC formula is mathematically identical to a standard costly information acquisition problem. The difference is that the parameters $((\Omega, \mathcal{F}, \mathbb{P}), \mathbf{G}, U, c)$ of the CC representation

¹⁹We showed in Ergin and Sarver (2010a) that the integrability of U implies that the term $\mathbb{E}_{\mathbb{P}} [\max_{p \in A} \mathbb{E}_{\mathbb{P}} [U | \mathcal{G}] \cdot p]$ is well-defined and finite for every $A \in \mathcal{A}$ and $\mathcal{G} \in \mathbf{G}$. For simplicity, we directly assume that the outer maximization in Equation (6) has a solution instead of making topological assumptions on \mathbf{G} to guarantee the existence of a maximum. An alternative approach that does not require this indirect assumption on the parameters of the representation would be to replace the outer maximization in Equation (6) with a supremum, in which case all of our results would carry over.

are subjective in the sense that they are not directly observable, but instead must be elicited from the individual’s preferences.²⁰

Theorem 2 from Ergin and Sarver (2010a) can be applied to the current setting to show that a CC representation can be written in reduced form as a max-HA representation satisfying a consistency condition.²¹

Theorem 4 (Ergin and Sarver (2010a)) *Let $V : \mathcal{A} \rightarrow \mathbb{R}$. Then, the following are equivalent:*

1. *There exists a CC representation such that V is given by Equation (6).*
2. *There exists a max-HA representation (\mathcal{M}, c) such that V is given by Equation (2), and \mathcal{M} satisfies consistency:*

$$\forall \mu, \nu \in \mathcal{M} \text{ and } \forall p \in \Delta(Z) : \int_U u(p) \mu(du) = \int_U u(p) \nu(du)$$

Therefore, consistency is key for the interpretation of the max-HA representation as a subjective information acquisition problem. The intuition for how a CC representation can be transformed into a consistent max-HA representation is as follows. In the CC representation, each contemplation strategy \mathcal{G} leads to the random variable $\mathbb{E}_{\mathbb{P}}[U|\mathcal{G}]$ denoting the individual’s ex post expected-utility function after acquiring signal \mathcal{G} . Therefore, each contemplation strategy \mathcal{G} can be associated with the probability measure over ex post utility functions over $\Delta(Z)$ that it induces. Moreover, the law of iterated expectations implies that for any contemplation strategy \mathcal{G} , the ex ante expected value of the ex post utility function $\mathbb{E}_{\mathbb{P}}[U|\mathcal{G}]$ must agree with the utility function prior to acquiring any information, $\mathbb{E}_{\mathbb{P}}[U]$, which implies the consistency condition on the corresponding set of measures.

Given a max-HA representation (\mathcal{M}, c) , we will show that the following axiom captures consistency of (\mathcal{M}, c) .

²⁰The costly contemplation representation in Equation (6) is similar to the functional form considered in Ergin (2003), where the primitive is a preference over menus taken from a finite set of alternatives. Ortoleva (2009) also considered a related model of costly thinking using slightly different primitives. The main conceptual distinction from our model is that Ortoleva considered an individual who may choose her contemplation strategy suboptimally. The individual’s anticipation of possible over-thinking when choosing from a menu in the future leads to a violation of the monotonicity axiom that Ortoleva referred to as “thinking aversion”.

²¹Although there are some minor differences in the assumptions imposed on the representations in this paper and Ergin and Sarver (2010a), adapting the result to the current context is straightforward.

Axiom 5 (Reversibility of Degenerate Decisions (RDD)) For any $A \in \mathcal{A}$, $p, q \in \Delta(Z)$, 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 a degenerate decision. To interpret Axiom 5, consider first the lottery $\beta\delta_{\alpha A + (1-\alpha)\{p\}} + (1 - \beta)\delta_{\{q\}}$. Under this lottery, 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 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. The lottery $\beta\delta_{\alpha A + (1-\alpha)\{q\}} + (1 - \beta)\delta_{\{p\}}$ has a similar interpretation with the roles of p and q reversed. Figure 4 illustrates these two lotteries for the case where $A = \{\delta_z, \delta_{z'}\}$, $p = \delta_{\tilde{z}}$, and $q = \delta_{\hat{z}}$.

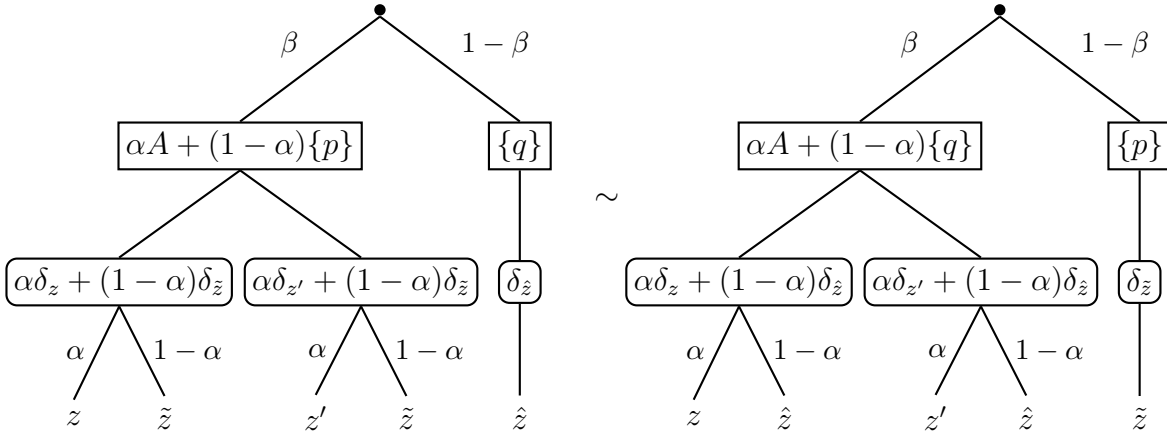


Figure 4: Reversibility of Degenerate Decisions when $A = \{\delta_z, \delta_{z'}\}$, $p = \delta_{\tilde{z}}$, and $q = \delta_{\hat{z}}$

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. Still, 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 $\delta_{\{q\}}$. Similarly, depending on her preference between $\delta_{\{p\}}$ and $\delta_{\{q\}}$, she need not be indifferent between the lotteries $\beta\delta_{\alpha A+(1-\alpha)\{p\}}+(1-\beta)\delta_{\{q\}}$ and $\beta\delta_{\alpha A+(1-\alpha)\{q\}}+(1-\beta)\delta_{\{p\}}$ if the probabilities of the paths leading to p and q , i.e., $\beta(1-\alpha)$ and $1-\beta$ are different. The RDD axiom requires the individual to be indifferent between these two lotteries when the probabilities of these paths are the same, i.e., when $\beta(1-\alpha) = 1-\beta$ or, equivalently, $\beta = 1/(2-\alpha)$. In the example illustrated in Figure 4, in both trees, the probabilities of the paths leading to \tilde{z} and \hat{z} are the same when $\beta = 1/(2-\alpha)$.

We next present the main result of this section. Given a max-HA representation (\mathcal{M}, c) , we show that RDD is equivalent to consistency of (\mathcal{M}, c) .

Theorem 5 *Suppose that the preference \succsim has a max-HA representation (\mathcal{M}, c) . Then, (\mathcal{M}, c) satisfies consistency if and only if \succsim satisfies RDD.*

We obtain the following CC representation theorem as an application of Theorems 1, 4, and 5.

Corollary 2 *The preference \succsim has a CC representation if and only if it satisfies Axiom 1, PERU, RDD, and monotonicity.*

By Corollary 2, a preference with a CC representation satisfies PERU. However, it is immediate from the representation that such a preference always satisfies indifference to timing of resolution of uncertainty when restricted to temporal lotteries, i.e., for all $p, q \in \Delta(Z)$ and $\alpha \in (0, 1)$:

$$\alpha\delta_{\{p\}} + (1-\alpha)\delta_{\{q\}} \sim \delta_{\{\alpha p+(1-\alpha)q\}}.^{22}$$

Therefore, as suggested at the beginning of this section, an individual with CC preferences never has a strict PERU unless she has non-degenerate choices in period 2.

²²This property can also be established directly as a consequence of RDD and first-stage independence. Fix any $p, q \in \Delta(Z)$ 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,

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

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

4.3 Kreps and Porteus Temporal Preferences

In this section, we formally connect our results to those of Kreps and Porteus (1978) and several other related papers. We first discuss the axioms and representation result from Kreps and Porteus (1978). Then, we present a generalization of their representation that allows for subjective uncertainty as in DLR (2001). Finally, we show that when either PERU or PLRU is satisfied, these representations can be expressed as special cases of our HA representation.

4.3.1 The Kreps and Porteus (1978) Representation Result

We now describe the axioms and representation result from Kreps and Porteus (1978), and we explain why their axioms will be violated in some of the special cases of our HA representation that have been considered in previous sections. The first axiom we consider is the standard von Neumann-Morgenstern independence axiom imposed on the second-stage uncertainty. It is satisfied by CC 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 $p, q, r \in \Delta(Z)$, 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.²³

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

²³To be precise, Kreps and Porteus (1978) considered both a period 1 preference \succsim over first-stage lotteries in $\Delta(\mathcal{A})$ and a period 2 preference \succsim_2 over second-stage lotteries in $\Delta(Z)$. 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).

Suppose there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ 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

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

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 lotteries, then Equation (7) implies that she must always be indifferent to timing of resolution of uncertainty. This is in contrast with CC 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 toward 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 $\Delta(Z)$.

Kreps and Porteus (1978) considered the following representation:²⁴

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

$$V(A) = \phi\left(\max_{p \in A} v(p)\right). \quad (8)$$

The following representation result is a special case of Theorem 1 in Kreps and Porteus (1978):²⁵

²⁴While 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.

²⁵The only difference is that Kreps and Porteus (1978) only required ϕ to be continuous. We additionally require Lipschitz continuity of ϕ since we impose the L-continuity axiom throughout the

Theorem 6 *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 6 is relatively straightforward. Axiom 1 implies there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Together with continuity, second-stage independence implies there exists an expected-utility function v on $\Delta(Z)$ 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

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

4.3.2 A Generalization

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 now consider a class of representations that generalize the DLR (2001) additive representation (where there is no objective first-stage uncertainty) and the Kreps-Porteus representation (where there is no subjective second-stage uncertainty).²⁶ Unlike the Kreps-Porteus 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 Borel measure on \mathcal{U} and $\phi : [a, b] \rightarrow \mathbb{R}$ is a Lipschitz continuous and strictly increasing function on the bounded interval $[a, b] =$*

paper.

²⁶Kraus and Sagi (2006, Theorem 5.2) also studied a similar generalization of the DLR (2001) additive representation and the Kreps-Porteus representation for incomplete preferences.

$\{\int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du) : A \in \mathcal{A}\}$, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by:

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

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 $\alpha \geq 0$.²⁷ The following is our KPDLR representation result:

Theorem 7 *A. The preference \succsim has a KPDLR representation if and only if it satisfies Axiom 1, monotonicity, and mixture independence.²⁸*

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 $p \in \Delta(Z)$ 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 toward 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 $\{A \in \mathcal{A} : \delta_A \sim \delta_{\{p\}} \text{ for some } p \in \Delta(Z)\}$. In particular, if an individual's preference has a KPDLR representation and has the property that for every $A \in \mathcal{A}$ there exists

²⁷It 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 $p \in \Delta(Z)$. Let $\mu = \alpha \delta_u$. Define constants $\hat{a} = a - \beta$ and $\hat{b} = b - \beta$, and define a function $\hat{\phi} : [\hat{a}, \hat{b}] \rightarrow \mathbb{R}$ by $\hat{\phi}(t) = \phi(t + \beta)$. Then, the KPDLR representation $(\hat{\phi}, \mu)$ gives the same value function for menus V as the Kreps-Porteus representation (ϕ, v) .

²⁸It 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 6.

a $p \in \Delta(Z)$ such that $\delta_A \sim \delta_{\{p\}}$, then, as in the case of Kreps-Porteus preferences, her attitude toward timing of resolution of uncertainty is determined entirely by her attitude toward timing of resolution of uncertainty for temporal lotteries.

4.3.3 Hidden Action Interpretation

Note that KPDLR preferences need not be a subset of HA preferences, since they may violate PERU or PLRU. Thus, not every KPDLR preference will have an HA representation, but the subclass satisfying PERU or PLRU will. We next characterize this subclass of KPDLR preferences within the class of HA preferences.

Theorem 8 *Let $V : \mathcal{A} \rightarrow \mathbb{R}$ and let μ be a nonzero finite 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).*
2. *There exists a max-HA [min-HA] representation (\mathcal{M}, c) such that V is given by Equation (2) [(3)] 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 8 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 8, 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 8, consider any 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 .²⁹ 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 $\Delta(Z)$. In this case, the individual has no uncertainty about her ex post preference ranking over lotteries in $\Delta(Z)$, 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 8 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 $\tilde{c} : \Lambda \rightarrow \mathbb{R}$ by $\tilde{c}(\lambda) = c(\lambda\delta_u)$. Then, Equation (2) for this max-HA representation simplifies to the following:³⁰

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

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).³¹ Their results (and, more generally, our Theorem 8) 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. 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

²⁹Suppose without loss of generality that $\max\{\lambda : \lambda\mu \in \mathcal{M}\} = 1$. Then, one can also interpret each action $\lambda\mu \in \mathcal{M}$ as leading 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} .

³⁰Theorem 8 also implies that if $0 \in \Lambda$, then 0 is not an isolated point and $\tilde{c}'(0) = \min_{p \in \Delta(Z)} u(p)$. 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 $r \in \{u(p) : p \in \Delta(Z)\}$.

³¹Kreps 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.

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 of 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 $P, Q \in \Delta(\mathcal{A})$, $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is 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 $\tilde{c} : \Lambda \rightarrow \mathbb{R}$ by $\tilde{c}(\lambda) = -U_1(\frac{1}{\alpha} \ln(\lambda))$, and defining $u \in \mathbb{R}^Z$ by $u(z) = -\exp(-\alpha z)$.³² 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 $\Delta(Z)$ 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 $\Delta(\mathcal{A})$ 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:³³

Theorem 9 *The preference \succsim satisfies Axiom 1, strategic rationality, and PERU if and*

³²Although this definition implies that $u \notin \mathcal{U}$, u can be normalized to be in \mathcal{U} .

³³It 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.

only if it has a max-HA representation (\mathcal{M}, c) such that $\mathcal{M} \subset \{\lambda\delta_u : \lambda \in \mathbb{R}_+, u \in \mathcal{U}\}$.³⁴

We now sketch the proof Theorem 9. Axiom 1 implies there exists a Lipschitz continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ 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 $f : \Delta(Z) \rightarrow \mathbb{R}$ 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 $\Delta(Z)$ 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).³⁵ Therefore, if \succsim satisfies the axioms of Theorem 9, then V takes the following form:

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

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 $p \in \Delta(Z)$. 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 $c : \mathcal{M} \rightarrow \mathbb{R}$ 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 .³⁶

³⁴It 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.

³⁵Our 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.

³⁶A 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 $f : C \rightarrow \mathbb{R}$ 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 $f : C \rightarrow \mathbb{R}$. For $x \in C$, the *subdifferential* of f at x is defined to be

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

The subdifferential is 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 $h : X \rightarrow \mathbb{R}$ 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 2 (Ergin and Sarver (2010b)) *Suppose C is a convex subset of a Banach space X . If $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous and convex, then $\partial f(x) \neq \emptyset$ for all $x \in C$.*

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

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

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

There is an important duality between f and f^* . Lemma 3 summarizes certain properties of f^* that are useful in establishing this duality.³⁷ The proof is standard and can be found, for example, in the supplementary appendix of Ergin and Sarver (2010a).

³⁷For 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 3 *Suppose $C \subset X$ and $f : C \rightarrow \mathbb{R}$. 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 $x^* \in \partial f(x)$.

Suppose that $C \subset X$ is convex and $f : C \rightarrow \mathbb{R}$ 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 3, we have

$$f(x) = \max_{x^* \in X^*} [\langle x, x^* \rangle - f^*(x^*)] \quad (12)$$

for all $x \in C$.³⁸ 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 (12) holds. Let C_f denote the set of all $x \in C$ for which the subdifferential of f at x is a singleton:

$$C_f = \{x \in C : \partial f(x) \text{ is a singleton}\}. \quad (13)$$

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

$$\mathcal{N}_f = \{x^* \in X^* : x^* \in \partial f(x) \text{ for some } x \in C_f\}. \quad (14)$$

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

$$\mathcal{M}_f = \overline{\mathcal{N}_f}. \quad (15)$$

Before stating our first main result, recall that the *affine hull* of a set $C \subset X$, denoted $\text{aff}(C)$, is defined to be the smallest affine subspace of X that contains C . 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 10 (Ergin and Sarver (2010b)) *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 $\text{aff}(C)$ is dense in X ,³⁹ and (iii) $f : C \rightarrow \mathbb{R}$ is Lipschitz continuous and convex. Then, \mathcal{M}_f is weak* compact, and for any weak* compact $\mathcal{M} \subset X^*$,*

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

³⁸This is a slight variation of the classic Fenchel-Moreau theorem. The standard version of this theorem states that if $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ 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).

³⁹In 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 .

The intuition for Theorem 10 is fairly simple. We already know from Lemma 3 that for any $x \in C_f$, $f(x) = \max_{x^* \in \mathcal{N}_f} [\langle x, x^* \rangle - f^*(x^*)]$. Ergin and Sarver (2010b) show that under the assumptions of Theorem 10, C_f is dense in C . Therefore, it can be shown that for any $x \in C$,

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

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 3 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. We will then 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.⁴⁰ Let $X_+ = \{x \in X : x \geq 0\}$ denote the *positive cone* of X . A function $f : C \rightarrow \mathbb{R}$ on a subset C of X is *monotone* if $f(x) \geq f(y)$ whenever $x, y \in C$ are such that $x \geq y$. A continuous linear functional $x^* \in X^*$ is *positive* if $\langle x, x^* \rangle \geq 0$ for all $x \in X_+$.

Theorem 11 (Ergin and Sarver (2010a, Supplementary Appendix)) *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 $f : C \rightarrow \mathbb{R}$ be Lipschitz continuous, convex, and monotone. Then, the functionals in \mathcal{M}_f are positive.

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

Theorem 12 (Ergin and Sarver (2010a, Supplementary Appendix)) *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 $c : \mathcal{M} \rightarrow \mathbb{R}$ be weak* lower semi-continuous. Define $f : C \rightarrow \mathbb{R}$ by*

$$f(x) = \max_{x^* \in \mathcal{M}} [\langle x, x^* \rangle - c(x^*)]. \tag{16}$$

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 $x^* \in \mathcal{M}_f$.*

⁴⁰See Aliprantis and Border (1999, p302) for a definition of Banach lattices.

B Proof of Theorem 1

In this section, we prove two results. We first prove a general representation theorem for preferences that may violate monotonicity and subsequently establish Theorem 1 as a special case. The following is a generalization of the HA representation to allow for signed measures:

Definition 10 A *signed 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 $c : \mathcal{M} \rightarrow \mathbb{R}$ such that:

1. $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by Equation (2) [(3)].
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 (2) [(3)] is different from V .

The pair (\mathcal{M}, c) is an *signed HA representation* if it is a signed max-HA or a signed min-HA representation.

In this section, we prove the following theorem:

Theorem 13 *A. The preference \succsim has a signed max-HA [min-HA] representation if and only if it satisfies Axiom 1 and PERU [PLRU].*

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

Theorem 13.B is simply a restatement of Theorem 1, and Theorem 13.A characterizes the signed HA representation. It has been shown that an individual's preferences may violate monotonicity, referred to as a *preference for commitment*, due to psychological features such as regret and temptation (see, e.g., Sarver (2008), Gul and Pesendorfer (2001), and DLR (2009)). Therefore, Theorem 13.A may be a useful starting point for incorporating regret and temptation into our model of temporal preferences. However, we leave the study of specific violations of monotonicity that correspond to these phenomena within our model as a subject for future research.

The remainder of this section is devoted to the proof of Theorem 13. Note that \mathcal{A} is a compact metric space since $\Delta(Z)$ 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 4 *A preference \succsim over $\Delta(\mathcal{A})$ satisfies weak order, continuity, and first-stage independence if and only if there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ such that \succsim is represented by $\mathbb{E}_P[V]$. Furthermore, if $V : \mathcal{A} \rightarrow \mathbb{R}$ and $V' : \mathcal{A} \rightarrow \mathbb{R}$ are continuous functions such that $\mathbb{E}_P[V]$ and $\mathbb{E}_P[V']$ represent the same preference over $\Delta(\mathcal{A})$, 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 $\Delta(\mathcal{A})$ 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 $A \in \mathcal{A}$,

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

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)). ■

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 4 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 5 *Suppose that $V : \mathcal{A} \rightarrow \mathbb{R}$ is a continuous function such that $\mathbb{E}_P[V]$ represents the preference \succsim over $\Delta(\mathcal{A})$. 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 $A, B \in \mathcal{A}^c$.⁴¹*
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 (2010a). 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 $A, B \in \mathcal{A}^c$:

$$d_h(A, B) \leq \frac{1}{2M} \implies |V(A) - V(B)| \leq Kd_h(A, B). \quad (17)$$

⁴¹If \succsim also satisfies IR, then it can be shown that V is Lipschitz continuous on \mathcal{A} .

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

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

where the first inequality follows from L-continuity, the second inequality follows from $\alpha \leq 1/2$, and the equality follows from 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 (17).

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

$$|V(A) - V(B)| \leq Kd_h(A, B), \quad (18)$$

i.e., the requirement $d_h(A, B) \leq \frac{1}{2M}$ in Equation (17) is not necessary. To see this, take any sequence $0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \lambda_{n+1} = 1$ such that $(\lambda_{i+1} - \lambda_i)d_h(A, B) \leq \frac{1}{2M}$. Let $A_i = \lambda_i A + (1 - \lambda_i)B$. It is straightforward to verify that:⁴²

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

Combining this with the triangle inequality and Equation (17), we obtain

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

This establishes Equation (18), 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 $P, Q \in \Delta(\mathcal{A})$, 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 $\alpha \geq Md_h(A, B)$, we have

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

which implies the conclusion of L-continuity. ■

⁴²Note that the convexity of the menus A and B is needed for the first equality.

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

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

Lemma 6 1. For all $A \in \mathcal{A}^c$ and $\sigma \in \Sigma$, $A_{(\sigma_A)} = A$ and $\sigma_{(A_\sigma)} = \sigma$. Hence, σ is a bijection from \mathcal{A}^c to Σ .

2. For all $A, B \in \mathcal{A}^c$ 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).⁴³ 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 $q = (1/|Z|, \dots, 1/|Z|) \in \Delta(Z)$, 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 7 Suppose $V : \mathcal{A}^c \rightarrow \mathbb{R}$, and define a function $W : \Sigma \rightarrow \mathbb{R}$ 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.

⁴³The 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 $\Delta(Z)$.

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 6.

(2): If V is Lipschitz continuous with a Lipschitz constant $K \geq 0$, then by parts 1 and 3 of Lemma 6, 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 6, for any $A, B \in \mathcal{A}^c$ and $\lambda \in [0, 1]$,

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

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.

(4): This is an immediate consequence of the following fact, which is easy to see from part 1 of Lemma 6 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 $\mu \in C(\mathcal{U})^*$.⁴⁴

For any function $W : \Sigma \rightarrow \mathbb{R}$, define the subdifferential ∂W and the conjugate W^* as in Appendix A. Also, define Σ_W , \mathcal{N}_W , and \mathcal{M}_W as in Equations (13), (14), and (15), respectively:

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

where the closure is taken with respect to the weak* topology. We now apply Theorem 10 to the current setting.

⁴⁴Since \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.

Lemma 8 *Suppose $W : \Sigma \rightarrow \mathbb{R}$ is Lipschitz continuous and convex. Then, \mathcal{M}_W is weak* compact, and for any weak* compact $\mathcal{M} \subset C(\mathcal{U})^*$,*

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

Proof: We simply need to verify that $C(\mathcal{U})$, Σ , and W satisfy the assumptions of Theorem 10. 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 6, 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 $\text{span}(\Sigma)$ is dense in $C(\mathcal{U})$. This result is also proved in DLR (2001). Since $0 \in \Sigma$ implies that $\text{aff}(\Sigma) = \text{span}(\Sigma)$, the affine hull of Σ is therefore dense in $C(\mathcal{U})$. Finally, W is Lipschitz continuous and convex by assumption. \blacksquare

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

To prove the sufficiency of the axioms for the signed max-HA representation in part A, suppose that \succsim satisfies Axiom 1 and PERU. By Lemma 4, there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ such that $\mathbb{E}_P[V]$ represents \succsim . Moreover, by Lemma 5, 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 7, the function $W : \Sigma \rightarrow \mathbb{R}$ defined by $W(\sigma) = V(A_\sigma)$ is Lipschitz continuous and convex. Therefore, by Lemma 8, for all $\sigma \in \Sigma$,

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

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

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

where the first equality follows from IR and the second equality follows from part 1 of Lemma 7. The function W^* is lower semi-continuous by part 1 of Lemma 3, and \mathcal{M}_W is compact by Lemma 8. It is also immediate from Lemma 8 that \mathcal{M}_W satisfies the minimality condition in Definition 10. Therefore, $(\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ is a signed 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 5 and 7, 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 11, 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 signed min-HA representation in part A, suppose that \succsim satisfies Axiom 1 and PLRU. By Lemma 4, there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ such that $\mathbb{E}_P[V]$ represents \succsim . By Lemmas 5 and 7, the function $W : \Sigma \rightarrow \mathbb{R}$ defined by $W(\sigma) = V(A_\sigma)$ is Lipschitz continuous and concave. Define a function $\bar{W} : \Sigma \rightarrow \mathbb{R}$ by $\bar{W}(\sigma) = -W(\sigma)$. Then, \bar{W} is Lipschitz continuous and convex, so by Lemma 8, for all $\sigma \in \Sigma$,

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

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

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

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

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

where the first equality follows from IR and the second equality follows from part 1 of Lemma 7. The function \bar{W}^* is lower semi-continuous by part 1 of Lemma 3, 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 8. Also, by Lemma 8 and the above construction, it is immediate that \mathcal{M} satisfies the minimality condition in Definition 10. Therefore, (\mathcal{M}, c) is a signed 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 5 and 7, the function W is monotone. Let $\hat{\Sigma} \equiv -\Sigma = \{-\sigma : \sigma \in \Sigma\}$, and define a function $\hat{W} : \hat{\Sigma} \rightarrow \mathbb{R}$ 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 $\sigma, \sigma' \in \hat{\Sigma}$,

$$\sigma \leq \sigma' \implies -\sigma \geq -\sigma' \implies \hat{W}(\sigma) = -W(-\sigma) \leq -W(-\sigma') = \hat{W}(\sigma').$$

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

$$\begin{aligned} \hat{W}(\lambda\sigma + (1-\lambda)\sigma') &= -W(\lambda(-\sigma) + (1-\lambda)(-\sigma')) \\ &\leq -\lambda W(-\sigma) - (1-\lambda)W(-\sigma') = \lambda\hat{W}(\sigma) + (1-\lambda)\hat{W}(\sigma'). \end{aligned}$$

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 11, the measures in $\mathcal{M}_{\hat{W}}$ are positive. For any $\mu \in C(\mathcal{U})^*$ and $\sigma, \sigma' \in \hat{\Sigma}$, note that

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

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 a signed HA representation.

Lemma 9 *Suppose (\mathcal{M}, c) is a signed HA representation.*

1. *If (\mathcal{M}, c) is a signed max-HA representation and $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by Equation (2), then V is Lipschitz continuous and convex. In addition, defining the function $W : \Sigma \rightarrow \mathbb{R}$ 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 signed min-HA representation and $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by Equation (3), then V is Lipschitz continuous and concave. In addition, defining the function $W : \Sigma \rightarrow \mathbb{R}$ 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

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

By part 1 of Theorem 12, W is Lipschitz continuous and convex. Therefore, the restriction of V to \mathcal{A}^c is Lipschitz continuous and convex by Lemma 7. 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 $V(A) = V(\text{co}(A))$, $V(B) = V(\text{co}(B))$, and $d_h(\text{co}(A), \text{co}(B)) \leq d_h(A, B)$. Hence,

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

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

$$\begin{aligned} V(\lambda A + (1 - \lambda)B) &= V(\text{co}(\lambda A + (1 - \lambda)B)) = V(\lambda \text{co}(A) + (1 - \lambda)\text{co}(B)) \\ &\leq \lambda V(\text{co}(A)) + (1 - \lambda)V(\text{co}(B)) = \lambda V(A) + (1 - \lambda)V(B), \end{aligned}$$

which implies that V is convex on \mathcal{A} . Also, by parts 2 and 3 of Theorem 12 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 8 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 $\bar{W} : \Sigma \rightarrow \mathbb{R}$ by $\bar{W}(\sigma) = -W(\sigma)$. Then, for any $\sigma \in \Sigma$,

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

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 $\mu \in \mathcal{M}$. \blacksquare

Suppose that \succsim has a signed max-HA [signed min-HA] representation (\mathcal{M}, c) , and suppose $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by Equation (2) [(3)]. Since $\mathbb{E}_P[V]$ represents \succsim and V is continuous (by Lemma 9), \succsim satisfies weak order, continuity, and first-stage independence by Lemma 4. Since V is Lipschitz continuous and convex [concave] by Lemma 9, \succsim satisfies L-continuity and PERU [PLRU] by Lemma 5. Since $V(A) = V(\text{co}(A))$ 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

We next state and prove a generalization of Theorem 2 to signed-HA representations (see Definition 10). Theorem 2 is a special case of Theorem 14, and therefore follows directly.

Theorem 14 *If (\mathcal{M}, c) and (\mathcal{M}', c') are two signed max-HA [signed 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 $\mu \in \mathcal{M}$.*

Proof: Throughout the proof, 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 signed

max-HA representations for \succsim . Define $V : \mathcal{A} \rightarrow \mathbb{R}$ and $V' : \mathcal{A} \rightarrow \mathbb{R}$ for these respective representations, and define $W : \Sigma \rightarrow \mathbb{R}$ and $W' : \Sigma \rightarrow \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$ and $W'(\sigma) = V'(A_\sigma)$. By part 1 of Lemma 9, $\mathcal{M} = \mathcal{M}_W$ 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 9), the uniqueness part of Lemma 4 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,

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

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

The proof of the uniqueness of the signed min-HA representation is similar and involves an application of part 2 of Lemma 9. ■

D Proof of Theorem 3

(1 \Rightarrow 3): Fix a min-HA representation (\mathcal{M}, c) , and define V by Equation (3). 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 $U : \Omega \rightarrow \mathbb{R}^Z$ by $U(\omega) = \omega$.

For each $\mu \in \mathcal{M}$, define the probability measure π_μ on (Ω, \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 $f_\mu : \mathcal{U} \rightarrow \Omega$ 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 $A \in \mathcal{A}$.⁴⁵ Let $\Pi = \{\pi_\mu : \mu \in \mathcal{M}\}$ and $\tilde{c}(\pi_\mu) = c(\mu)$. Then, V can be expressed in the

⁴⁵ This is easy to see if $\mu(\mathcal{U}) = 0$. If $\mu(\mathcal{U}) > 0$, then define the function $g : \Omega \rightarrow \mathbb{R}$ 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

following SSV form:

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

(3 \Rightarrow 2): Let $((\Omega, \mathcal{F}), U, \Pi, c)$ be an SSV representation, and define V by Equation (5). 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 (5). Note that Equation (5) continues to hold when Π is replaced by Π' , i.e.,

$$V(A) = \min_{\pi \in \Pi'} \left(\int_{\Omega} \max_{p \in A} U(\omega) \cdot p \pi(d\omega) + c(\pi) \right) \quad (19)$$

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 (19) is bounded by κ for every 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 (19) 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 $c|_{\Pi'} : \Pi' \rightarrow \mathbb{R}$ 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 $\tilde{U} : \tilde{\Omega} \rightarrow \mathbb{R}^Z$ 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.⁴⁶

$\omega \in \Omega$, $\max_{p \in A} \tilde{U}(\omega) \cdot p$ 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 is \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 $\int_{\Omega} g(\omega) (\tilde{\mu} \circ f_{\mu}^{-1})(d\omega) = \int_{\mathcal{U}} g(f_{\mu}(u)) \tilde{\mu}(du)$.

⁴⁶ \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 $f : \tilde{\Omega} \rightarrow \mathbb{R}^Z$ defined by $f(\omega, \pi) = U(\omega)$ is $\tilde{\mathcal{F}}$ -measurable. Also note that since $c|_{\Pi'}$ is \mathcal{G} -measurable, and

For any $\pi \in \Pi'$, define the function $f_\pi : \Omega \rightarrow \tilde{\Omega}$ by $f_\pi(\omega) = (\omega, \pi)$. Note that f_π is measurable.⁴⁷ Define the probability measure ρ_π on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by $\rho_\pi = \pi \circ f_\pi^{-1}$. For any $A \in \mathcal{A}$,

$$\begin{aligned} \int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \rho_\pi(d\tilde{\omega}) &= \int_{\Omega} \max_{p \in A} \tilde{U}(f_\pi(\omega)) \cdot p \pi(d\omega) \\ &= \int_{\Omega} \left[\max_{p \in A} U(\omega) \cdot p + c(\pi) \right] \pi(d\omega) \\ &= \int_{\Omega} \max_{p \in A} U(\omega) \cdot p \pi(d\omega) + c(\pi), \end{aligned}$$

where the first equality above follows from the change of variables formula.⁴⁸

Letting $\tilde{\Pi}' = \{\rho_\pi : \pi \in \Pi'\}$, by Equation (19), we see that V can be expressed in the following SSMP form:

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

(2 \Rightarrow 1): Let $((\Omega, \mathcal{F}), U, \Pi)$ be an SSMP representation, and define V by Equation (4). 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 $f_\pi : \mathcal{A} \rightarrow \mathbb{R}$ 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 $\|U(\omega)\| \leq \kappa$ 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 $\|p^* - q^*\| \leq d_h(A, B)$. Then,

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

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

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

Hence f_π is Lipschitz continuous with a Lipschitz constant κ that does not depend on $\pi \in \Pi$.

$\tilde{\mathcal{F}}$ is the product of the σ -algebras \mathcal{F} and \mathcal{G} , the function $g : \tilde{\Omega} \rightarrow \mathbb{R}^Z$ 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 .

⁴⁷To 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 $\tilde{\mathcal{F}}$.

⁴⁸To see this, define the function $g : \tilde{\Omega} \rightarrow \mathbb{R}$ by $g(\tilde{\omega}) = \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p$. By a similar argument as in Footnote 45, g is $\tilde{\mathcal{F}}$ -measurable. Then, the change of variables formula is $\int_{\tilde{\Omega}} g(\tilde{\omega}) (\pi \circ f_\pi^{-1})(d\tilde{\omega}) = \int_{\Omega} g(f_\pi(\omega)) \pi(d\omega)$.

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 $V : \mathcal{A} \rightarrow \mathbb{R}$ is monotone, concave, Lipschitz continuous, and it satisfies the IR condition $V(A) = V(\text{co}(A))$ for all $A \in \mathcal{A}$, the construction in Appendix B.2 implies that there exists a min-HA representation such that V is given by Equation (3).

E Proof of Theorem 5

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

Definition 11 A function $V : \mathcal{A} \rightarrow \mathbb{R}$ 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 10 Suppose that $V : \mathcal{A} \rightarrow \mathbb{R}$ is a function such that $\mathbb{E}_P[V]$ represents the preference \succsim over $\Delta(\mathcal{A})$. 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

$$V(\alpha A + (1 - \alpha)\{p\}) - V(\alpha A + (1 - \alpha)\{q\}) = (1 - \alpha)[V(\{p\}) - V(\{q\})] \quad (20)$$

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

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 (20) are equal to $(1 - \alpha)v \cdot (p - q)$, showing that \succsim satisfies RDD.

If \succsim satisfies RDD, then define the function $f : \Delta(Z) \rightarrow \mathbb{R}$ by $f(p) = V(\{p\})$ for all $p \in \Delta(Z)$. Let $\alpha \in [0, 1]$ and $p, q \in \Delta(Z)$, 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 (20) 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 $p, q \in \Delta(Z)$. 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 $p \in \Delta(Z)$.

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 (2010a) 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 $A' \in \mathcal{A}$, $p, q \in \Delta(Z)$, 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 (20) and the third equality follows from the expected utility form of f . Therefore, V is translation linear. \blacksquare

We are now ready to prove Theorem 5. The necessity of RDD is straightforward and left to the reader. For the other direction, suppose that \succsim has a max-HA representation (\mathcal{M}, c) and that it satisfies RDD. In the rest of this section, we will continue to use the notation and results from Appendix B. By Theorem 1, \succsim satisfies Axiom 1 and PERU. Therefore, $(\mathcal{M}_W, W^*|_{\mathcal{M}_W})$ constructed in Appendix B is also a max-HA representation for \succsim . Since \succsim satisfies RDD, by Lemma 10, the value function V for this representation is translation linear. Let $v \in \mathbb{R}^Z$ be 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 $q = (1/|Z|, \dots, 1/|Z|) \in \Delta(Z)$. By Lemma 22 of Ergin and Sarver (2010a), for all $\mu \in \mathcal{M}_W$ and $p \in \Delta(Z)$, $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$. The consistency of \mathcal{M}_W follows immediately from this fact because for any $\mu, \mu' \in \mathcal{M}_W$ and $p \in \Delta(Z)$, we have

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

By Theorem 2, there exists $\alpha > 0$ such that $\mathcal{M} = \alpha\mathcal{M}_W$. Therefore, (\mathcal{M}, c) is also consistent.

F Proof of Theorem 6

The necessity of the axioms is straightforward. For sufficiency, suppose that \succsim satisfies Axiom 1, second-stage independence, and strategic rationality. By Lemma 4, there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ 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 $B = \{p_1, \dots, p_n\}$, where $\{p_1\} \succsim \{p_2\} \succsim \dots \succsim \{p_n\}$. For any

$k = 2, \dots, n,$

$$\{p_k\} \sim \{p_k, \dots, p_n\} \implies \{p_{k-1}\} \sim \{p_{k-1}, \dots, p_n\}, \quad (21)$$

since by strategic rationality, $\{p_{k-1}\} \succsim \{p_k\} \sim \{p_k, \dots, p_n\}$ implies that $\{p_{k-1}\} \sim \{p_{k-1}\} \cup \{p_k, \dots, p_n\}$. By applying backward induction on $k = 2, \dots, n$ using Equation (21), 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 $\Delta(Z)$ 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 $\Delta(Z)$ 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 $\{v(p) : p \in \Delta(Z)\} = [a, b]$ 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 $\phi : [a, b] \rightarrow \mathbb{R}$ such that for all $p \in \Delta(Z)$,

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

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

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

To establish the Lipschitz continuity of ϕ , first recall that by Lemma 5, 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 $p, q \in \Delta(Z)$, $|V(\{p\}) - V(\{q\})| \leq K d_h(\{p\}, \{q\}) = K \|p - q\|$. If $a = b$, then ϕ is trivially Lipschitz continuous. Next, suppose that $a < b$. Take $p_*, p^* \in \Delta(Z)$ 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 $s, t \in [a, b]$,

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

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

G Proof of Theorem 7

Consider the following generalization of the KPDLR representation allowing for signed measures:

Definition 12 A *signed Kreps-Porteus-Dekel-Lipman-Rustichini (KPDLR) representation* is a pair (ϕ, μ) , where μ is a finite signed Borel measure on \mathcal{U} and $\phi : [a, b] \rightarrow \mathbb{R}$ 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 $V : \mathcal{A} \rightarrow \mathbb{R}$ is defined by (9).

The following is a representation theorem for signed KPDLR representations. It is proved in subsections G.1 and G.2.

Theorem 15 *A. The preference \succsim has a signed KPDLR representation if and only if it satisfies Axiom 1 and mixture independence.*

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

We next argue how Theorem 7 follows from Theorem 15. Theorem 7.B is a special case of Theorem 15.B, therefore, it follows directly. The necessity of the axioms in Theorem 7.A are straightforward. To see the sufficiency direction of Theorem 7.A, suppose that \succsim satisfies Axiom 1, monotonicity, and mixture independence. Define the preference $\succsim_{\mathcal{A}}$ over \mathcal{A} by: $A \succsim_{\mathcal{A}} B$ if and only if $\delta_A \succsim \delta_B$. Since \succsim is monotone, $\succsim_{\mathcal{A}}$ is also monotone. By Theorem 15.A, \succsim has a signed KPDLR representation (ϕ, μ) . This implies that $\succsim_{\mathcal{A}}$ is represented by $V_{\mathcal{A}} : \mathcal{A} \rightarrow \mathbb{R}$ defined by $V_{\mathcal{A}}(A) = \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du)$. Therefore, $\succsim_{\mathcal{A}}$ satisfies the axioms of a monotone additive representation of DLR (2001) and DLRS (2007) (weak order, continuity, independence, and monotonicity). Therefore, the construction in DLR (2001) and DLRS (2007) implies that there is a (positive) measure $\hat{\mu}$ such that $\succsim_{\mathcal{A}}$ is also represented by $\hat{V}_{\mathcal{A}} : \mathcal{A} \rightarrow \mathbb{R}$ defined by $\hat{V}_{\mathcal{A}}(A) = \int_{\mathcal{U}} \max_{p \in A} u(p) \hat{\mu}(du)$, and that $\mu = c\hat{\mu}$ for some $c > 0$. This implies that μ is also a (positive) measure, hence, (ϕ, μ) is a KPDLR representation.

G.1 Proof of Theorem 15.A

The necessity of the axioms is straightforward. For sufficiency, suppose that \succsim satisfies Axiom 1 and mixture independence. By Lemma 4, there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Since \succsim satisfies IR, $V(A) = V(\text{co}(A))$ 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 $\phi : [a, b] \rightarrow \mathbb{R}$ 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 $A \succsim' B \iff \delta_A \succ \delta_B$ (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 Herstein-Milnor Theorem, there exists an affine function $U : \mathcal{A}^c \rightarrow \mathbb{R}$ 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 $\phi : [a, b] \rightarrow \mathbb{R}$ such that

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

To establish the Lipschitz continuity of ϕ , first recall that by Lemma 5, L-continuity implies there exists $K \geq 0$ such that $|V(A) - V(B)| \leq K d_h(A, B)$ for any $A, B \in \mathcal{A}^c$. 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 $K d_h(A^*, A_*) / (b - a)$.

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.⁴⁹ 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

⁴⁹The 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 $\alpha \geq M d_h(A, B)$, $\delta_{(1-\alpha)A + \alpha A^*} \succ \delta_{(1-\alpha)B + \alpha A_*}$.*

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

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

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,

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

and hence

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

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 $\underline{t}, \bar{t} \in (a, b)$, $\underline{t} < \bar{t}$, we have $\phi(\bar{t}) - \phi(\underline{t}) \leq \varepsilon \cdot (\bar{t} - \underline{t})$. 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 $[\underline{t}, \bar{t}]$. Therefore, there exists a finite subcover taken from this collection of intervals that also covers $[\underline{t}, \bar{t}]$. The finiteness of this subcover implies the existence of a finite set of numbers $\underline{t} = s_1 \leq \dots \leq s_k = \bar{t}$ such that for any $i \in \{1, \dots, k - 1\}$ 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

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

This completes the proof.

G.2 Proof of Theorem 15.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 8

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

Lemma 11 *Let $a, b \in \mathbb{R}$ with $a < b$ and let $\phi : [a, b] \rightarrow \mathbb{R}$ 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}_ϕ .

Proof: (1 \Rightarrow 2) Part a follows from Theorem 11.

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

$$\lambda'a \leq \phi^*(\lambda') - \phi^*(0) \leq \lambda't \quad \forall \lambda' \in (0, \lambda). \quad (22)$$

Since ϕ is nondecreasing, $0 \in \partial\phi(a)$. Along with Lemma 3, this implies that $-\phi^*(0) = \phi(a) \geq \lambda'a - \phi^*(\lambda')$ for any $\lambda' \geq 0$, establishing the first inequality in Equation (22). Take any $t \in (a, b]$. By Lemma 2 there exists $\lambda \in \partial\phi(t)$. Note that $\lambda > 0$. Otherwise, if $\lambda \leq 0$, then by Lemma 3,

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

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 3, this implies that $\lambda' \in \partial\phi(t')$. Monotonicity of the subdifferential $\partial\phi$ implies that $t' \leq t$.⁵⁰ Then, by Lemma 3 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 (22).

(2 \Rightarrow 1) Theorem 10 and part a imply that ϕ is nondecreasing. Therefore, $0 \in \partial\phi(a)$, implying $\phi(a) = -\phi^*(0)$ by Lemma 3.

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 2, 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 3, contradicting $\phi(a) < \phi(t)$. Therefore, $\lambda > 0$, implying

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

by Lemma 3, as desired.

⁵⁰To 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 $(\lambda - \lambda')(t - t') \geq 0$.

(1 \Rightarrow 3) Suppose for a contradiction that 0 is an isolated point of \mathcal{M}_ϕ . Then, $0 \in \mathcal{N}_\phi$, i.e., there exists $t \in [a, b]$ such that $\partial\phi(t) = \{0\}$. Then, Lemma 3 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}_ϕ and \mathcal{M}_ϕ is compact by Theorem 10, $\mathcal{M}_\phi \setminus \{0\}$ is also compact. Therefore, the above inequality implies that

$$-\phi^*(0) > \max_{\lambda \in \mathcal{M}_\phi \setminus \{0\}} [\lambda t - \phi^*(\lambda)]. \quad (23)$$

Let $\Delta > 0$ be the difference of the left hand side and the right hand side in Equation (23) 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 (23) continues to hold if t is replaced by s . Therefore,

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

where the second equality follows from Theorem 10. This implies that ϕ is constant at a $\frac{\Delta}{M}$ neighborhood of t , contradicting the assumption that ϕ is strictly increasing. \blacksquare

Corollary 3 *Let $a, b \in \mathbb{R}$ with $a < b$ and let $\phi : [a, b] \rightarrow \mathbb{R}$ 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}_ϕ .

Proof: Define $\hat{\phi} : [-b, -a] \rightarrow \mathbb{R}$ 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

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

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 11 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 12 *Let μ be a nonzero finite signed Borel measure on \mathcal{U} and $[a, b] = \{\langle \sigma, \mu \rangle : \sigma \in \Sigma\}$. Let $\phi : [a, b] \rightarrow \mathbb{R}$ be Lipschitz continuous and convex and define $W : \Sigma \rightarrow \mathbb{R}$ 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,

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

(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}_ϕ 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 2, 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 10, \mathcal{M} is also compact. Let $t \in [a, b]$, and $\sigma \in \Sigma$ be such that $t = \langle\sigma, \mu\rangle$. Then,

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

where the second equality follows from part 1 and Theorem 10, 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 10, $\mathcal{M}_\phi \subset \mathcal{M}$. \blacksquare

We next state and prove a generalization of Theorem 8 to signed-HA representations. Theorem 8 is a special case of Theorem 16, therefore, it follows directly.

Theorem 16 *Let $V : \mathcal{A} \rightarrow \mathbb{R}$ and let μ be a nonzero finite signed Borel measure on \mathcal{U} . Then, the following are equivalent:*

1. *There exists a signed KPDLR representation (ϕ, μ) with convex [concave] ϕ such that V is given by Equation (9).*
2. *There exists a signed max-HA [signed min-HA] representation (\mathcal{M}, c) such that V is given by Equation (2) [(3)] 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

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

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

Proof of Theorem 16, the convex case: In the following, let $W : \Sigma \rightarrow \mathbb{R}$ 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 6. Since W is Lipschitz continuous and convex by Lemma 12, $V(A) = V(\text{co}(A))$ 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 signed max-HA representation such that V is given by Equation (2). By part 2.a of Lemma 11 and part 3 of Lemma 12, $\mathcal{M}_W \subset \{\lambda\mu : \lambda \in \mathbb{R}_+\}$. By part 2.b of Lemma 11 and part 2 of Lemma 12,

$$\lim_{\lambda \searrow 0} \frac{W^*(\lambda\mu) - W^*(0)}{\lambda} = \frac{d\phi^*}{d\lambda^+}(0) = a \equiv \min_{A \in \mathcal{A}} \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du). \quad (24)$$

By part 3 of Lemma 11, part 3 of Lemma 12, 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 (24) 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 $\phi : [a, b] \rightarrow \mathbb{R}$ 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 12, 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 (2) and part a. Therefore, it only remains to show that ϕ is strictly increasing.

By Lemma 9, $\mathcal{M} = \mathcal{M}_W$ 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 6. By part 3 of Lemma 12, $\mathcal{M} = \{\lambda\mu : \lambda \in \mathcal{M}_\phi\}$. 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}_ϕ .

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 10,

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

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

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

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

$$a = \lim_{\lambda \searrow 0: \lambda \mu \in \mathcal{M}} \frac{c(\lambda \mu) - c(0)}{\lambda} = \lim_{\lambda \searrow 0: \lambda \in \mathcal{M}_\phi} \frac{\phi^*(\lambda) - \phi^*(0)}{\lambda} \quad (25)$$

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 12. 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}_ϕ , Equation (25) 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 (25) is equal to $\frac{d\phi^*}{d\lambda^+}(0)$. By Lemma 11, ϕ is strictly increasing. ■

Proof of Theorem 16, the concave case: In the following, let $W : \Sigma \rightarrow \mathbb{R}$ and $\bar{W} : \Sigma \rightarrow \mathbb{R}$ 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 6. Define $\bar{\phi} : [a, b] \rightarrow \mathbb{R}$ 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 $c : \mathcal{M} \rightarrow \mathbb{R}$ by $c(\nu) = \bar{W}^*(-\nu)$. Since \bar{W} is Lipschitz continuous and convex by Lemma 12, $V(A) = V(\text{co}(A))$ 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 signed min-HA representation such that V is given by Equation (3). By part 2.a of Corollary 3 and part 3 of Lemma 12, $\mathcal{M}_{\bar{W}} \subset \{\lambda \mu : \lambda \in \mathbb{R}_-\}$, implying that $\mathcal{M} \subset \{\lambda \mu : \lambda \in \mathbb{R}_+\}$. By part 2.b of Corollary 3 and part 2 of Lemma 12,

$$\lim_{\lambda \searrow 0} \frac{\bar{W}^*(0) - \bar{W}^*(-\lambda \mu)}{\lambda} = \frac{d\bar{\phi}^*}{d\lambda^-}(0) = b \equiv \max_{A \in \mathcal{A}} \int_{\mathcal{U}} \max_{p \in A} u(p) \mu(du). \quad (26)$$

By part 3 of Corollary 3, part 3 of Lemma 12, 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 (26) 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 $\phi : [a, b] \rightarrow \mathbb{R}$ 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 $\bar{\phi} : [a, b] \rightarrow \mathbb{R}$ 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 12, $\bar{\phi}$ 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 (3) 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 9, $\mathcal{M} = -\mathcal{M}_{\bar{W}}$ 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 6. Therefore, $\bar{W}(\sigma) = \bar{\phi}(\langle \sigma, \mu \rangle)$ for all $\sigma \in \Sigma$. By part 3 of Lemma 12, $-\mathcal{M} = \{\lambda\mu : \lambda \in \mathcal{M}_{\bar{\phi}}\}$. 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 10,

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

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,

$$b = \lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}} \frac{c(0) - c(\lambda\mu)}{\lambda} = \lim_{\lambda \searrow 0 : -\lambda \in \mathcal{M}_{\bar{\phi}}} \frac{\bar{\phi}^*(0) - \bar{\phi}^*(-\lambda)}{\lambda} \quad (27)$$

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 12. 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 (27) 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 (27) is equal to $\frac{d\bar{\phi}^*}{d\lambda^-}(0)$. By Corollary 3, $\bar{\phi}$ is strictly decreasing. ■

I Proof of Theorem 9

The necessity of the axioms is straightforward. For sufficiency, suppose that \succsim satisfies Axiom 1, strategic rationality, and PERU. By Lemma 4, there exists a continuous function $V : \mathcal{A} \rightarrow \mathbb{R}$ 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 $f : \Delta(Z) \rightarrow \mathbb{R}$ by $f(p) = V(\{p\})$. By Lemma 5, 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 $\Delta(Z)$ in order to prove a variation of Fenchel duality for the function f :

Lemma 13 *If $f : \Delta(Z) \rightarrow \mathbb{R}$ is Lipschitz continuous and convex, then there exists a nonempty compact set $\mathcal{V} \subset H$ and a lower semi-continuous function $c : \mathcal{V} \rightarrow \mathbb{R}$ such that*

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

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

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

$$C = \left\{ s \in \mathbb{R}^{n-1} : -\frac{1}{n} \leq s_i \text{ and } \sum_{i=1}^{n-1} s_i \leq \frac{1}{n} \right\}.$$
⁵¹

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

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

for any $s \in C$. Note that ζ is well defined (i.e., takes values in $\Delta(Z)$), one-to-one, onto $\Delta(Z)$, 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 $\zeta^{-1} : \Delta(Z) \rightarrow C$ exists.

⁵¹It 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 $(1/n, \dots, 1/n) \in \mathbb{R}^{n-1}$.

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

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

for any $t = (t_1, \dots, t_{n-1}) \in \mathbb{R}^{n-1}$. Note that ξ is well defined (i.e., takes values in H), linear, one-to-one, and onto H . Therefore, the inverse function $\xi^{-1} : H \rightarrow \mathbb{R}^{n-1}$ 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 $g : C \rightarrow \mathbb{R}$ 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 10, there exists a nonempty compact $T \subset \mathbb{R}^{n-1}$ and a lower semi-continuous function $d : T \rightarrow \mathbb{R}$ 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 $c : \mathcal{V} \rightarrow \mathbb{R}$ by $c = d \circ \xi^{-1}$. Then, \mathcal{V} is a compact subset of H and c is lower semi-continuous. For all $p \in \Delta(Z)$,

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

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 $f(p) > \max_{v \in \mathcal{V}'} [v(p) - c(v)]$. \blacksquare

By Lemma 13, there exists a compact set $\mathcal{V} \subset H$ and a lower semi-continuous function $c : \mathcal{V} \rightarrow \mathbb{R}$ 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 $p \in \Delta(Z)$ such that $V(\{p\}) > \max_{v \in \mathcal{V}'} [v(p) - c(v)]$.

Define a function $\phi : H \rightarrow C(\mathcal{U})^*$ 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 $v_d \rightarrow v$. 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, $\|v_d\| \rightarrow 0$ and $\mu = 0$. Fix any continuous function $g : \mathcal{U} \rightarrow \mathbb{R}$. Then, there exists $M > 0$ such that $|g(u)| \leq M$ for all $u \in \mathcal{U}$. Therefore,

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

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 $v_d \rightarrow v$, 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, $\frac{v_d}{\|v_d\|} \rightarrow \frac{v}{\|v\|}$. Fix any continuous function $g : \mathcal{U} \rightarrow \mathbb{R}$. Then,

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

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 $\tilde{c} : \mathcal{M} \rightarrow \mathbb{R}$ 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 (2) 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 $p \in \Delta(Z)$ 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.
- Backus, D., B. Routledge, and S. Zin (2004): “Exotic Preferences for Macroeconomics,” *NBER Macroeconomics Annual*, 19, 319–390.
- Billingsley, P. (1995): *Probability and Measure*. New York: John Wiley and Sons, Inc.
- Caplin, A., and J. Leahy (2001): “Psychological Expected Utility Theory and Anticipatory Feelings,” *Quarterly Journal of Economics*, 116, 55–79.
- 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.
- Epstein, L. G., and S. Zin (1991): “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis,” *Journal of Political Economy*, 99, 263–286.
- Ergin, H. (2003): “Costly Contemplation,” Mimeo.
- Ergin, H., and T. Sarver (2010a): “A Unique Costly Contemplation Representation,” *Econometrica*, 78, 1285–1339.

- Ergin, H., and T. Sarver (2010b): “The Unique Minimal Dual Representation of a Convex Function,” *Journal of Mathematical Analysis and Applications*, 370, 600–606.
- 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.
- Grant, S., A. Kajii, and B. Polak (1998): “Intrinsic Preference for Information,” *Journal of Economic Theory*, 83, 233–259.
- Grant, S., A. Kajii, and B. Polak (2000): “Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models,” *Econometrica*, 68, 425–434.
- Hansen, L. P., and T. Sargent (2001): “Robust Control and Model Uncertainty,” *American Economic Review*, 91, 60–66.
- 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.
- Kraus, A., and J. S. Sagi (2006): “Inter-temporal Preference for Flexibility and Risky Choice,” *Journal of Mathematical Economics*, 42, 698–709.
- 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.
- Ortoleva, P. (2009): “The Price of Flexibility: Towards a Theory of Thinking Aversion,” Mimeo.
- 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.
- Strzalecki, T. (2009): “Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion,” Mimeo.
- Strzalecki, T. (2011): “Axiomatic Foundations of Multiplier Preferences,” *Econometrica*, 79, 47–73.
- Tallarini, T. (2000): “Risk-Sensitive Real Business Cycles,” *Journal of Monetary Economics*, 45, 507–532.
- Weil, P. (1993): “Precautionary Savings and the Permanent Income Hypothesis,” *Review of Economic Studies*, 60, 367–383.