Kellogg School of Management

CMS-EMS
Center for Mathematical Studies in Economics
And Management Science

Discussion Paper #1567

Hidden Actions and Preferences for Timing of Resolution of Uncertainty

Haluk Ergin
University of California, Berkeley

Todd Sarver
Northwestern University

September 14, 2012

JEL Classification: D81

Keywords: temporal preferences, preference for flexibility, hidden action, subjective uncertainty

NORTHWESTERN UNIVERSITY

Hidden Actions and Preferences for Timing of Resolution of Uncertainty*

Haluk Ergin† Todd Sarver‡

September 14, 2012

Abstract

We study preferences for timing of resolution of objective uncertainty in a 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 the model of 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 David Ahn, Eddie Dekel, Pablo Schenone, and Marciano Siniscalchi for helpful comments. These results were previously distributed under the title "A Subjective Model of Temporal Preferences".

†Correspondence address: Department of Economics, University of California, Berkeley, 508-1 Evans Hall #3880, Berkeley, CA 94720-3880. Email: hie@econ.berkeley.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)). As we discuss later in the introduction, preferences exhibiting this combination are quite plausible in a variety of economic environments and can have important implications. 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 is unable to increase her current consumption. Even in this case, the preferences considered by Kreps and Porteus (1978) allow the individual to have a strict preference for that uncertainty to resolve earlier (or later), which we refer to as an intrinsic preference for the timing of the resolution of uncertainty. The additional flexibility of their model has proven 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 a 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 for the preferences considered in their 1978 paper, and Machina (1984) considered a related representation for slightly more general preferences. The actions that induce preferences for timing in these models are typically interpreted as physical choices, e.g., induced preferences for future income resulting from decisions about consumption-savings, consumption-investment, or consumption bundles (see Kreps and Porteus (1979, pages 86–87)). Our main representation theorem provides a similar *hidden action* interpretation for a broader class of preferences. This generalization not only allows us to model some useful preferences that have not been previously considered, but also permits novel psychological interpretations for the hidden action, such as ambiguity aversion or costly decision making.

1.2 Overview of Results

We examine dynamic preferences in a simple menu-choice setting with two stages of 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 incorporate a preference for flexibility as in Kreps (1979) and DLR (2001) into their temporal model. In addition, we allow for preference for timing to interact with preference for flexibility in a nontrivial way. For example, we allow for the possibility that preferences for timing are stronger when facing decision problems that offer flexibility in future choices. To illustrate the usefulness of these generalizations, we show in Section 1.3 that these new features have important implications for a broad class of mechanism design problems.

After describing the setting for our model in Section 2, we present our axioms and main results in Section 3. 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. 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 (2006a) 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.

Finally, in Section 4.3, we describe what is perhaps the simplest extension of the model of Kreps and Porteus (1978) that can accommodate a preference for flexibility. In this special case of our general model, the preference for timing depends only on the utility values of the possible menus that could result from a two-stage lottery, not on the actual content of those menus. In particular, the presence or absence of intermediate choice has no direct effect on the preference for timing. We will see that this restriction rules out many of the behaviors we would like to capture with our general model, such as the preferences to avoid contingent planning described in the following section.

[1] While we do not suggest that there literally exists a malevolent nature, it is a useful way to interpret a pessimistic or ambiguity-averse attitude on the part of the decision-maker. See, for example, Maccheroni, Marinacci, and Rustichini (2006a) for a related discussion.

1.3 A Motivating Example from Mechanism Design

The preferences modeled in this paper can be motivated by and applied to a number of issues in mechanism design. To illustrate, consider the problem of school choice. The student-optimal stable matching mechanism, which is based on the deferred acceptance algorithm of Gale and Shapley, is used to match students with high schools in New York City and Boston. This mechanism is dominant strategy incentive compatible for students, and therefore no student with "standard" preferences can benefit from learning the reports of the other students prior to submitting her own ranking. However, we will argue that it is quite reasonable to expect that students would prefer to learn the rankings of others prior to submitting their rankings. While this type of preference cannot be accommodated by standard models, even models of preferences for timing such as Kreps and Porteus (1978), such preferences can easily be accommodated within our framework.

To make ideas concrete, focus on a simple example with two schools, a and b, and two students, 1 and 2. Suppose student 1 has higher priority at school a and student 2 has higher priority at school b. The student-optimal matching mechanism gives the following matches based on the reported preferences of the students:

1's ranking	2's ranking	matching
$a \succ_1 b$	$a \succ_2 b$	$(1, a)$, $(2, b)$
$a \succ_1 b$	$b \succ_2 a$	$(1, a)$, $(2, b)$
$b \succ_1 a$	$a \succ_2 b$	$(1, b)$, $(2, a)$
$b \succ_1 a$	$b \succ_2 a$	$(1, a)$, $(2, b)$

Notice in particular that if student 2 reports a preference for school b, then given the priorities of the schools, student 1 is assigned to school a regardless of her preference. On the other hand, if student 2 reports a preference for school a, then student 1 is assigned to whichever school she ranks higher. Therefore, depending on the reported rankings of the other students and the priorities of the schools, there is a feasible set of schools for student 1 and she is assigned to her highest ranked school from this feasible set. The feasible sets for student 1 based on the reports of student 2 are summarized as follows:

2's ranking	1's feasible set
$a \succ_2 b$	$\{a, b\}$
$b \succ_2 a$	$\{a\}$

This table concisely illustrates a key property of this mechanism: The ranking that student 1 submits has an impact on her outcome in some instances (when $a \succ_2 b$), but not in others (when $b \succ_2 a$). Of course, if student 1 finds school a clearly superior to school b or vice versa, then this feature of the mechanism is not pertinent — she will simply submit her true ranking regardless of the likelihood that it will be implemented. However, a more realistic scenario is one in which each school has different strengths and weaknesses, and consequently student 1 finds it difficult to rank the two. In this case, knowing whether her feasible choice set is $\{a, b\}$ or $\{a\}$ prior to submitting her ranking is valuable to student 1, since she can put more effort into her decision when her submitted ranking is actually relevant and less when it is not.

The preferences described in this example are easily formalized within the framework of this paper, and are at the heart of our analysis. Suppose student 1 believes that with probability α student 2 will submit the ranking $a \succ_2 b$ and with probability $1-\alpha$ student 2 will submit the ranking $b \succ_2 a$. Then, in the student-optimal stable matching mechanism, student 1 submits her ranking of a and b with the foresight that her choice of a versus b will be implemented with probability α, and with probability $1-\alpha$ she will be assigned a regardless of her reported ranking. Thus, submitting the ranking $a \succ_1 b$ results in a (for certain) and the ranking $b \succ_1 a$ results in the lottery $\alpha b + (1-\alpha)a$. This implies that her ranking of the alternatives (or equivalently, her contingent plan from the sets $\{a, b\}$ and $\{a\}$) can be expressed within our framework as a choice from a set of distributions over outcomes $\{a, \alpha b + (1-\alpha)a\}$. In contrast, if student 1 learns the ranking of student 2 prior to submitting her own ranking, then with probability α she chooses from the set $\{a, b\}$ and with probability $1-\alpha$ she chooses from the set $\{a\}$. When the decision problems for student 1 are formulated in this manner, her preference to learn the ranking of student 2 prior to submitting her own ranking corresponds precisely to our axiom for preference for early resolution of uncertainty.

In this example, the preference for timing was motivated by the following consideration: Students may wish to avoid unnecessary contingent planning, i.e., investing the effort to rank schools that turn out not to be feasible. Our model provides a tractable way to analyze preferences for timing that are driven partly or entirely by a desire to avoid (or, alternatively, to engage in) contingent planning; these preferences can be accommodated by allowing for complementarities between the content of menus and the hidden action in our utility representation. That is, if there is greater variation in the optimal hidden action among nonsingleton choice sets than there is among singleton choice sets, then the preference for timing will manifest more strongly for decision problems involving nontrivial contingent planning (contingent choice from nonsingleton sets). For example, one specific case of our representation were hidden actions have this property is the costly contemplation model described in Section 4.2, which provides a simple and tractable functional form for modeling costly decision making.

The connection between contingent planning and preference for timing is absent from the previous literature on the subject, and drawing this connection is one of the main contributions of this paper. Kreps and Porteus (1978) made assumptions on the preferences that made the presence or absence of intermediate choice inconsequential for the preference for timing. For example, while their model could also allow for a preference to learn whether the feasible set is $\{a, b\}$ or $\{a\}$ prior to submitting a ranking, it would then impose the same preference for timing when the possible feasible sets are $\{b\}$ and $\{a\}$ (see Section 4.3 for a detailed description of their model).[2] However, in the second situation, student 1's ranking is completely irrelevant, so her contingent planning problem is trivial. Therefore, if she expresses the same preference for timing in both situations, we can infer that a factor other than contingent planning must be driving her preference.

There are other explanations for why student 1 may prefer to learn her feasible choice set sooner that match well with the Kreps and Porteus (1978) model. For example, she may prefer to have this uncertainty resolved sooner in order to reduce her anxiety about the outcome, or because there are other decisions in her life that she would like to condition on the outcome of the school match. These sources of preferences for information are certainly plausible, and our model allows for them as well; however, these causes of preference for timing have very different implications, both for the overall structure of preferences and for the design of optimal mechanisms. If students just want earlier arrival of information, then efficiency can be improved simply by running the mechanism at an earlier date. In other words, for the preferences considered by Kreps and Porteus (1978), it is efficient to run the static mechanism at some optimal date (determined by the precise preferences for timing).

On the other hand, if the students want earlier arrival of information in order to avoid contingent planning, then running the same mechanism sooner is of no benefit to them. Instead, there could be efficiency gains associated with the use of dynamic mechanisms. Having agents act sequentially allows them to utilize information about the past actions of other agents and may help them to avoid planning for unrealized contingencies. Thus, the incorporation of preferences regarding contingent planning into our model has important implications for the design of optimal mechanisms.

The applications of our model are not limited to the problem of school choice. The interpretation of the agents' reports as complete contingent plans naturally carries over to all economic environments in which a dominant strategy incentive compatible direct revelation mechanism (like a VCG mechanism) is employed. Just as in the school choice example, for these mechanisms every report of the other agents translates into a feasible set of outcomes that a particular agent can obtain from her different reports. If the mechanism is dominant strategy incentive compatible, then it must always select the best feasible outcome for the agent according to her reported type. Therefore, submitting a report at the same time as other agents is equivalent to forming a contingent plan from the possible feasible sets, and agents with the preferences considered in this paper may again benefit from learning the reports of the other agents in advance of submitting their

[2] These feasible sets could arise if student 2 has higher priority than student 1 at both schools instead of just school b.

own.[3]

These observations suggest that there are benefits associated with using dynamic mechanisms in these settings as well. In fact, in environments with monetary transfers, there are already results in the applied literature to support this claim for several special cases of our general model. For example, when agents can engage in costly information acquisition about their private values (which corresponds to the special case of our model described in Section 4.2), Compte and Jehiel (2007) show that multistage auctions lead to higher revenues than sealed-bid auctions. Athey and Segal (2007) provide an elegant construction of an efficient, budget-balanced, and Bayesian incentive-compatible dynamic mechanism in a setting where agents could have a very general set of hidden actions (including information acquisition).[4] There are also many other settings where the preferences considered in this paper could be applied. In particular, one important open research question is finding optimal dynamic mechanisms in environments without transferable utility such as the school match problem described in this section. By clarifying the basic structure of these individual preferences, the axiomatic analysis in this paper can make it easier to approach applications involving multiple agents in complex environments.

2 Choice Setting

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

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

Elements of $\mathcal{A}$ are called menus, with generic menus denoted A, B, C. Let $\triangle(\mathcal{A})$ denote the set of all Borel probability measures on $\mathcal{A}$, endowed with the weak* topology and

[3] A simple example is the sealed-bid second-price auction with independent private values. Let b_i and b_{-i} denote the bid of agent i and the highest bid of the other agents, respectively. From the perspective of agent i, each bid/report b_i corresponds to a complete contingent plan where she commits to buy the object at price b_{-i} if $b_i > b_{-i}$, and commits not to buy it if $b_i < b_{-i}$. A preference for timing driven by a desire to avoid contingent planning implies a preference by agent i to learn the price b_{-i} she faces before deciding whether or not to buy the object.

[4] These models involve agents who can take hidden actions in multiple stages, and the optimal dynamic mechanisms therefore also involve several stages. While our axiomatic analysis is restricted to two stages for tractability and expositional simplicity, the basic insights uncovered here are also useful for understanding multistage settings.

with generic elements denoted P, Q, R.[5] The primitive of the model is a binary relation $\succsim$ on $\triangle(\mathcal{A})$, representing the individual's preferences over lotteries over menus.

We interpret $\succsim$ as corresponding to the individual's choices in the first period of a two-period decision problem. In 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.[6]

For any $A \in \mathcal{A}$, let $\delta_A \in \triangle(\mathcal{A})$ denote the degenerate lottery that puts probability 1 on the menu A. Then, $\alpha\delta_A + (1-\alpha)\delta_B$ denotes the lottery that puts probability α on the menu A and probability $1-\alpha$ on the menu B. Figure 1 illustrates such a 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.

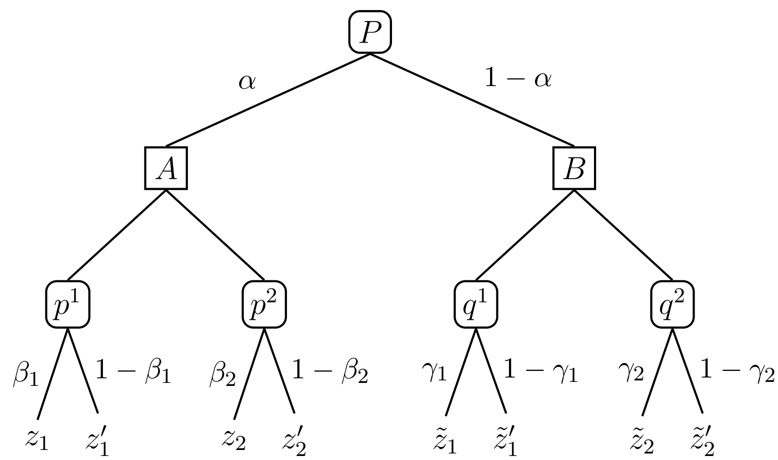

Figure 1: Decision Tree for the Lottery P

Our framework is a special case of that of Kreps and Porteus (1978), with only two periods and no consumption in period 1.[7] As in Kreps and Porteus (1978), we refer to a

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

[6] Since period 2 choice in our model is stochastic, incorporating it explicitly into the framework would involve a number technical complications. Ahn and Sarver (2011) analyze a model that combines choice of menus with stochastic choice from menus. Similar techniques may allow period 2 choice to be incorporated formally in our temporal framework in future research.

[7] This framework was also used in Epstein and Seo (2009) and in Section 4 of Epstein, Marinacci, and

lottery $P \in \triangle(\mathcal{A})$ 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 $\triangle(\triangle(Z))$.

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

3 General Representations

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

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

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

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

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

Axioms 1.1 and 1.2 are standard. Axiom 1.3 is the von Neumann-Morgenstern independence axiom imposed with respect to the first-stage uncertainty. Axioms 1.1–1.3 ensure that there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. Given Axioms 1.1–1.3, Axiom 1.4 is a technical condition implying the

Seo (2007).

Lipschitz continuity of V.[8] 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 $\mathrm{co}(A)$ offer the same set of options, and hence they are identical from the perspective of the individual.

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 $p, q \in \triangle(Z)$ and $\alpha \in [0,1]$,

$$\alpha\delta_{\{p\}} + (1-\alpha)\delta_{\{q\}} \succsim \delta_{\{\alpha p+(1-\alpha)q\}}. \tag{1}$$

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.[9] PERU requires a weak preference the first temporal lottery, whereas PLRU requires a weak preference for the second temporal lottery.

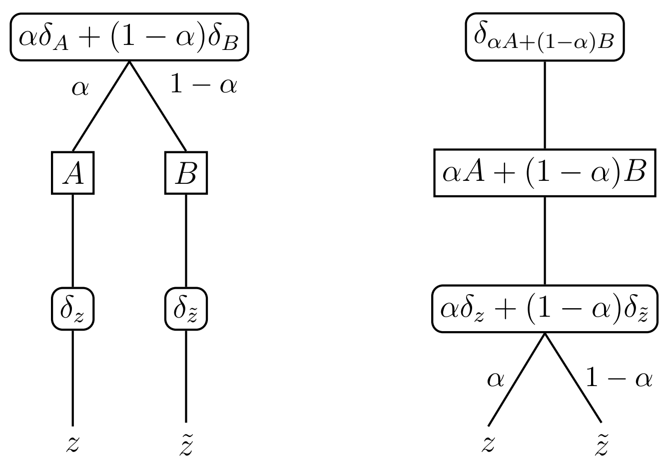

Figure 2: Illustration of Timing of Resolution of Uncertainty for Temporal Lotteries: $A = \{\delta_z\}$ and $B = \{\delta_{\tilde{z}}\}$

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

[8] In models with preferences over menus of lotteries, related L–continuity axioms were used by Dekel, Lipman, Rustichini, and Sarver (2007, henceforth DLRS), Sarver (2008), and Ergin and Sarver (2010a).

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

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.

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.[10]

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 early resolution lottery $\alpha\delta_A + (1-\alpha)\delta_B$, any uncertainty regarding the feasible set resolves in period 1, in particular, before the individual makes a choice from the realized menu. In the late resolution lottery $\delta_{\alpha A+(1-\alpha)B}$, the individual learns nothing in period 1 and then makes a choice form the menu $\alpha A + (1-\alpha)B$. We interpret this menu as the set of all contingent plans from the menus A and B (or, more precisely, the distributions over outcomes resulting from those contingent plans). To understand this interpretation, suppose the individual is asked to make a contingent plan $(p, q) \in A \times B$, where p will be implemented if the realized menu is A and q will be implemented in the case of B. Since A will be the relevant menu with probability α, this contingent plan induces the distribution over outcomes $\alpha p + (1-\alpha)q \in \alpha A + (1-\alpha)B$.

With this interpretation in mind, late resolution of uncertainty corresponds to learning nothing in the period 1 and then making a contingent plan (from the yet unrealized choice

[10] Note 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 individuals to prefer when the resolution of the first-stage uncertainty is more informative in the sense of Blackwell.

sets) that will be carried out after this uncertainty resolves in period 2. Therefore, timing of resolution of uncertainty can be broken into two components in our model:

1. *Absolute timing:* Whether the individual gets information sooner or later.
2. *Relative timing:* Whether the individual gets information prior to committing to a plan of action or not.

If the presence or absence of intermediate choice is inconsequential for the preference for timing (as in the model of Kreps and Porteus (1978)), we can infer that only absolute timing is important to the individual. On the other hand, if the preference for timing changes in the presence of intermediate choice, then relative timing is also relevant. By taking into account both absolute and relative timing of uncertainty, we can model novel issues such as difficulty in making complex contingent decisions.

When considering the timing of uncertainty relative to choice, it is important to keep in mind the potential instrumental value of information. As illustrated by the simple consumption/savings example in Section 1.1, changing the timing of information relative to choice has the potential to alter the individual's ability to condition her actions on the realization of uncertainty. However, this well-understood interaction between information and choice is not at work in our preference for timing axioms. The use of contingent plans in our comparison of late versus early resolution of uncertainty ensures that the individual's ability to condition her choices on the realized set is unaffected by the timing of resolution of uncertainty. Since the distributions over final outcomes available to the individual are the same in the case of early or late resolution, the only difference is whether she must commit to a plan of action prior to the resolution of uncertainty.

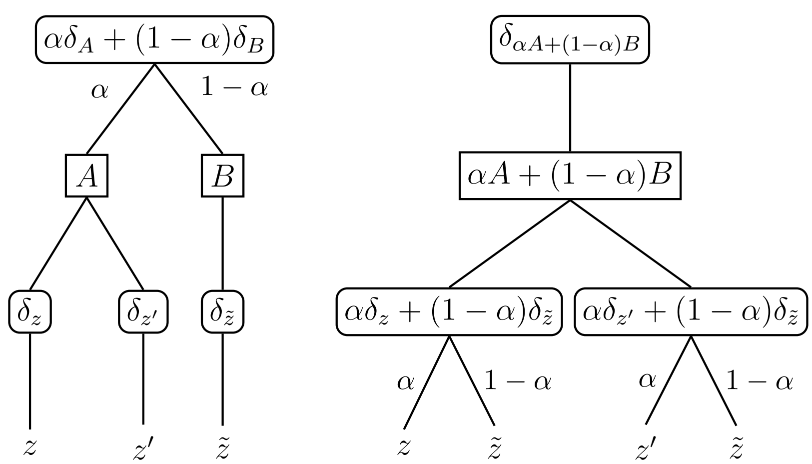

Figure 3: Illustration of Timing of Resolution of Uncertainty for non-Temporal Lotteries: $A = \{\delta_z, \delta_{z'}\}$ and $B = \{\delta_{\tilde{z}}\}$

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 preference 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) used 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 to retain the flexibility of $\delta_{\{p,q\}}$ rather than committing to either $\delta_{\{p\}}$ or $\delta_{\{q\}}$.

Axiom 4 allows for uncertainty about future tastes, but still imposes dynamic consistency on the part of the individual. In contrast, if the individual anticipates that her future choices will be inconsistent with her current preferences, she may strictly prefer to commit to a smaller menu. For example, Gul and Pesendorfer (2001) and DLR (2009) relaxed monotonicity in a menu-choice setting in order to model temptation and costly self-control. Our focus is instead on the interaction between preferences for flexibility and timing, so we impose monotonicity throughout the main text. However, in Appendix B we describe a generalization of our main representation to non-monotone preferences, which can be used as a starting point for future research on incorporating temptation into our temporal model.

It has also been suggested that preferences for early or late resolution of uncertainty could also arise due to anticipatory feelings or anxiety. 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. However, our axioms are inconsistent with several of the well-known models of anticipatory feelings 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/anxiety 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.

3.2 Hidden Action Representations

Note that expected-utility functions on $\triangle(Z)$ are equivalent to vectors in $\mathbb{R}^Z$, by associating each expected-utility function with its values for sure outcomes. We therefore use the notation $u(p)$ and $u \cdot p$ interchangeably for any $u \in \mathbb{R}^Z$. We define the set of *normalized (non-constant) expected-utility functions* on $\triangle(Z)$ to be

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

We are ready to introduce our general representations:[11]

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 $c : \mathcal{M} \to \mathbb{R}$ such that:

1. $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (2) [(3)]:
$$V(A) = \max_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) - c(\mu) \right) \tag{2}$$
$$V(A) = \min_{\mu \in \mathcal{M}} \left(\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du) + c(\mu) \right). \tag{3}$$

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 *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 $\triangle(Z)$.

[11] We endow the set of all finite Borel measures on $\mathcal{U}$ with the weak* topology (see footnote 5).

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 $A \in \mathcal{A}$,[12]

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

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 $\pi(\{0\}) = 1$. 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 $A \in \mathcal{A}$,

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

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 representation[13], and it also simplifies the mathematical

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

[13]There are many pairs $(\mathcal{V}, \pi)$ that give the same integral expression as the measure μ on $\mathcal{U}$. The lack

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 $\triangle(Z)$, 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 chose an action that outright minimizes her utility from a menu. Different cost functions put different restrictions on the individual's perception of the malevolent nature's objective.

In the above representations, both the set of available actions and and their costs are subjective in that they are part of the representation. Therefore, $\mathcal{M}$ and c are not directly observable to the modeler and need to be identified from the individual's preferences. Note that in both Equations (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 $\mathcal{M}$.

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.*[14,15]

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.

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

[15]It 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.

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 $V : \mathcal{A} \to \mathbb{R}$ such that $V(co(A)) = V(A)$ and $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$. In terms of this expected utility representation, it is easy to see that PERU corresponds to convexity of V and PLRU corresponds to concavity of V. The set $\mathcal{A}^c$ of convex menus can be mapped one-to-one to a set of continuous functions Σ known as the support functions, preserving the metric and the linear operations. Therefore, by using the property $V(co(A)) = V(A)$ and mimicking the construction in DLR (2001), V can be thought of as a function defined on the subset Σ of the Banach space $C(\mathcal{U})$ of continuous real-valued functions on $\mathcal{U}$. 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.[16] 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 10 in Appendix A). A similar application of the duality and uniqueness results can be found in Ergin and Sarver (2010a).

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

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

multiple-priors representation allows for ambiguity regarding the distribution over ex post subjective states and is intuitively similar to the multiple-priors representation proposed by Gilboa and Schmeidler (1989) in the Anscombe-Aumann setting.

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

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

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

In this representation, the individual has a subjective state space Ω, and her tastes over lotteries in $\triangle(Z)$ are summarized by the random vector U representing her state-dependent expected-utility function. Her utility from a lottery $p \in \triangle(Z)$ conditional on the subjective state $\omega \in \Omega$ is therefore given by $U(\omega) \cdot p = \sum_{z \in Z} p_z U_z(\omega)$. Each prior π represents a different distribution of the subjective states (tastes), and multiple priors in the set Π captures ambiguity about which is the correct prior. This representation is similar to one considered by Epstein, Marinacci, and Seo (2007, Theorem 1) in the setting of menus of lotteries; we will describe the connection in more detail after presenting the main results of this section.

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

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

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

and the minimization in Equation (5) has a solution for every $A \in \mathcal{A}$.[17]

[17] Note 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

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 $c(\pi) = \theta R(\pi \| \eta)$ for some scalar $\theta > 0$ and reference probability measure η, where R is the *relative entropy* of π with respect to η,

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

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

The following theorem show that in our setting a preference $\succsim$ is represented by an SSMP or SSV representation if and only if it has a min-HA representation.

Theorem 3 *Let $V : \mathcal{A} \to \mathbb{R}$. 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)*.*
3. *There exists an SSV representation such that V is given by Equation* (5)*.*

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.[18] 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). This equivalence of the SSMP and SSV representations is somewhat surprising since in the Anscombe-Aumann framework, the class of variational preferences considered by Maccheroni, Marinacci, and Rustichini (2006a) is strictly larger than the class of multiple-prior expected-utility preferences considered by Gilboa and Schmeidler (1989). The distinguishing feature of our SSMP and SSV representations that results in their equivalence is the subjectivity of the state spaces and the state-dependence of the utility functions.

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.

[18] The 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 $\mathcal{U}$.

To illustrate the role of the subjective state space in obtaining the equivalence of the SSMP and SSV representations, consider the simple case of an SSV representation $((\Omega, \mathcal{F}), U, \Pi, c)$ with a finite state space and just two priors, $\Pi = \{\pi_1, \pi_2\}$. Notice first that if c were constant, then we could simply add the value $c(\pi_1) = c(\pi_2)$ to the utility function U to transform this SSV representation into an SSMP representation. However, since c is not constant in general, we will need a more sophisticated approach. To construct a corresponding SSMP representation, we will use a richer state space that contains one copy of Ω for each measure in Π. Formally, since we are considering the case where Π contains only two priors, let $\tilde{\Omega} = \Omega \times \{1, 2\}$. Then, letting $\mathbf{1} \in \mathbb{R}^Z$ denote the vector whose coordinates are all equal to 1 (equivalently, the expected-utility function that takes a constant value 1), define $\tilde{U} : \tilde{\Omega} \to \mathbb{R}^Z$ as follows for $\tilde{\omega} = (\omega, i) \in \tilde{\Omega}$:

$$\tilde{U}(\omega, 1) = U(\omega) + c(\pi_1)\mathbf{1}$$
$$\tilde{U}(\omega, 2) = U(\omega) + c(\pi_2)\mathbf{1}$$

Thus, we add $c(\pi_1)$ to the state-dependent expected-utility function U for every state in the first copy of Ω and add $c(\pi_2)$ for every state in the second copy.

We then take the probability measures for our SSMP representation to be precisely those from the SSV representation when restricted to the appropriate copy of the original state space. Formally, define $\tilde{\pi}_1$ and $\tilde{\pi}_2$ as follows for $\tilde{\omega} = (\omega, i) \in \tilde{\Omega}$:

$$\tilde{\pi}_1(\omega, 1) = \pi_1(\omega) \qquad \tilde{\pi}_2(\omega, 1) = 0$$
$$\tilde{\pi}_1(\omega, 2) = 0 \qquad \tilde{\pi}_2(\omega, 2) = \pi_2(\omega)$$

By this construction, for $i = 1, 2$ and for any $A \in \mathcal{A}$,

$$\begin{aligned} \int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \; \tilde{\pi}_i(d\tilde{\omega}) &= \int_{\Omega} \max_{p \in A} \tilde{U}(\omega, i) \cdot p \; \pi_i(d\omega) \\ &= \int_{\Omega} \max_{p \in A} U(\omega) \cdot p \; \pi_i(d\omega) + c(\pi_i). \end{aligned}$$

Therefore, we have constructed an SSMP representation that gives the same value function for menus as the original SSV representation:

$$\min_{\tilde{\pi} \in \{\tilde{\pi}_1, \tilde{\pi}_2\}} \int_{\tilde{\Omega}} \max_{p \in A} \tilde{U}(\tilde{\omega}) \cdot p \; \tilde{\pi}(d\tilde{\omega}) = \min_{\pi \in \{\pi_1, \pi_2\}} \left(\int_{\Omega} \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega) + c(\pi) \right).$$

As this example illustrates, the key to our equivalence result is the flexibility afforded by the subjectivity of the state space in the SSMP representation. In the case of a fixed objective state space, a construction that enriches the state space as described above is no longer possible. This explains why a result like Theorem 3 is possible for our subjective-state-space representations, even though the multiple-priors representation considered by Gilboa and Schmeidler (1989) is strictly more restrictive than the variational representation considered by Maccheroni, Marinacci, and Rustichini (2006a).

Another consequence of the subjectivity of the state space is that the parameters of the SSMP and SSV representations cannot be uniquely identified from the preference. In fact, Theorem 3 illustrates the extent of this non-uniqueness: Since 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.

Combining Theorems 1 and 3 yields the following axiomatic characterization of 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.*

There is a simple intuition underlying the connection between the SSMP and SSV representations and preferences for late resolution of uncertainty. A standard motivation used in static models of ambiguity aversion is that objective uncertainty resolving after the state is realized can hedge against ambiguity. In our SSMP and SSV models, second-stage objective uncertainty can hedge in precisely the same way. However, our model also permits first-stage objective uncertainty, which occurs prior to the realization of the subjective state $\omega \in \Omega$. Since the outcome of first-stage uncertainty is known at the time the individual faces subjective uncertainty, it does not provide the same hedging benefit as second-stage randomization. This implies an ambiguity averse individual would prefer to delay the resolution of objective uncertainty until the second period.

The connection between ambiguity aversion and preference for late resolution of objective uncertainty is quite general, and extends well beyond our specific model. For example, this preference for timing also arises in models of ambiguity aversion that use an objective state space. In the dynamic models of multiple priors and variational preferences considered by Epstein and Schneider (2003) and Maccheroni, Marinacci, and Rustichini (2006b), respectively, this type of preference does not appear simply because the framework they used is not rich enough to describe preferences for timing — their framework only includes objective uncertainty about consumption in the current period. However, when these models are nested in a richer domain that allows for objective uncertainty about both current consumption and continuation acts, it is again possible to describe preferences for timing, and ambiguity aversion results in a preference for late resolution of objective uncertainty just as in our model. Several recent papers, including Hayashi (2005), Seo (2009), and Saito (2011), have used a richer framework with objective uncertainty about both current and future consumption and made this connection between ambiguity aversion and preferences for timing explicit.

Strzalecki (2011b) took an alternative perspective on the issue of timing in dynamic models of ambiguity aversion (with an objective state space). Instead of studying preferences for timing of *objective* uncertainty, he considered preferences for timing of resolution of *subjective* uncertainty. In principle, the same forces that lead to preference for late resolution of objective uncertainty would lead to a preference for early resolution of subjective uncertainty; both cause subjective uncertainty to resolve sooner relative to objective, which allows for hedging. However, like Epstein and Schneider (2003) and Maccheroni, Marinacci, and Rustichini (2006b), the framework adopted by Strzalecki (2011b) allows for objective uncertainty about the consumption in each period separately, but does not permit objective uncertainty about consumption at future dates. Therefore, changing the timing of subjective uncertainty does not allow for additional hedging of ambiguity; instead, his results demonstrate the role of discounting and consumption in intermediate periods in determining preferences for timing in models of ambiguity aversion.

As noted above, our SSMP representation is related to a representation considered by Epstein, Marinacci, and Seo (2007) in the setting of menus of lotteries. Moreover, our motivation for PLRU in terms of ambiguity aversion and hedging (or a malevolent nature) parallels their discussion (see page 361). However, their are two main distinctions between our models: First, our framework allows objective uncertainty to resolve in multiple stages, and we are therefore able to explicitly model the preferences for timing associated with the SSMP model. The second and more significant difference is that they require a normalization on the state-dependent utility functions — they permit only one possible utility function for each ex post preference.[19] Although such a normalization is without loss of generality for our HA representations due to the use of non-probability measures (see Lemma 1), for representations that require probability measures such as the SSMP representation, this requirement places nontrivial additional restrictions on the preference. This manifests in the Epstein, Marinacci, and Seo (2007) model as two auxiliary axioms needed to obtain their representation.[20] Moreover, under their normalization, since ex post utilities cannot be transformed by adding a constant, the equivalence of multiple-priors and variation representations fails.

[19] Kraus 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.

[20] We 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).

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 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 studied preferences exhibiting aversion to contingent planning in the simpler framework of preferences over menus of lotteries. We obtained 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 $c : \mathbf{G} \to \mathbb{R}$ is a function, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by

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

and the maximization in Equation (6) has a solution for every $A \in \mathcal{A}$.[21]

The interpretation of the CC representation is as follows. The individual is uncertain about her tastes over $\triangle(Z)$. This uncertainty is modeled by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a state-dependent expected-utility function U over $\triangle(Z)$. 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

[21] We showed in Ergin and Sarver (2010a) that the integrability of U implies that the term $\mathbb{E}_{\mathbb{P}}\left[\max_{p\in A} \mathbb{E}_{\mathbb{P}}[U|\mathcal{G}] \cdot p\right]$ 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.

or, more compactly, 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 $\mathbb{E}_{\mathbb{P}}\left[U\middle|\mathcal{G}\right] \cdot p$. 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 are subjective in the sense that they are not directly observable, but instead must be elicited from the individual's preferences.[22]

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:[23]

Theorem 4 (Ergin and Sarver (2010a)) *Let $V : \mathcal{A} \to \mathbb{R}$. 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*:*

$$\forall \mu, \nu \in \mathcal{M} \text{ and } \forall p \in \triangle(Z): \quad \int_{\mathcal{U}} u(p)\, \mu(du) = \int_{\mathcal{U}} u(p)\, \nu(du)$$

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 $\mathcal{G}$. Thus, each contemplation strategy $\mathcal{G}$ can be associated with the distribution over ex post utility functions over $\triangle(Z)$ 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

[22] The 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".

[23] Although 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.

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 $(\mathcal{M}, c)$.

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

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

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

We will call a choice out of a singleton menu 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 $q = \delta_{\hat{z}}$.

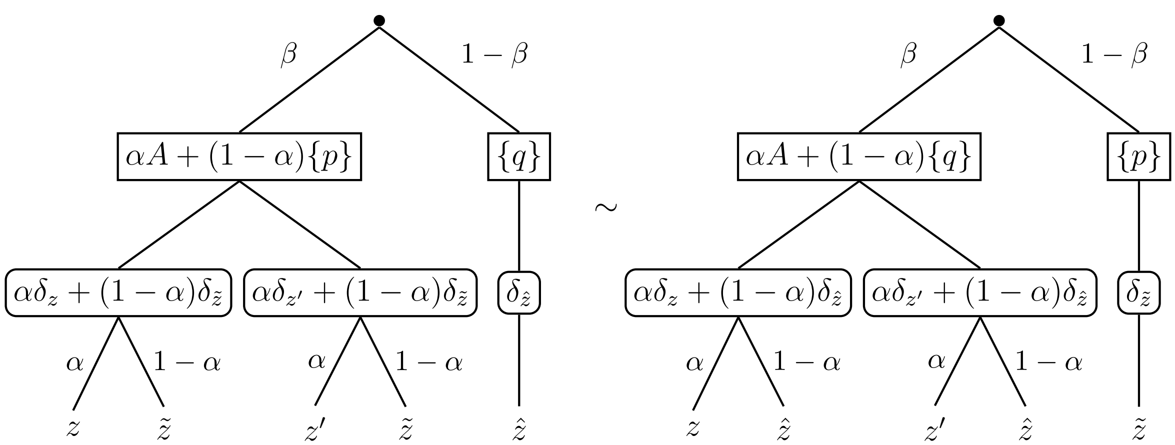

Figure 4: Reversibility of Degenerate Decisions when $A = \{\delta_z, \delta_{z'}\}$, $p = \delta_{\tilde{z}}$, and $q = \delta_{\hat{z}}$

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.*

The following CC representation theorem is obtained from 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 $p, q \in \triangle(Z)$ and $\alpha \in (0,1)$:

$$\alpha\delta_{\{p\}} + (1-\alpha)\delta_{\{q\}} \sim \delta_{\{\alpha p+(1-\alpha)q\}}.$$ [24]

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.

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

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

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

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

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

4.3 Simple Models of Preference for Flexibility and Timing

In this section, we describe a simple model that allows for both preference for flexibility and preference for timing, but does not allow the preference for timing to depend on the content of the menu per se. The approach we follow here parallels that of Kreps and Porteus (1978). They were able to incorporate preferences for timing into a standard expected-utility model by taking a nonlinear transformation of second-stage expected utility before taking expectations with respect to first-stage uncertainty. Formally, period 2 choice in their model maximizes some expected-utility function v, and thus menus in the second period are evaluated by $\max_{p\in A} v(p)$. This utility value is then transformed by some function ϕ to obtain the Bernoulli utility index for first-stage uncertainty:

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

Notice that $\alpha V(A) + (1-\alpha)V(B) \geq V(\alpha A + (1-\alpha)B)$ for all $A, B \in \mathcal{A}$ if and only if ϕ is convex. Therefore, for the first-stage expected-utility representation $\mathbb{E}_P[V]$, PERU corresponds to convexity of ϕ, and PLRU to concavity.

This approach of using a nonlinear transformation to alter preferences for timing can be applied to many models beyond standard expected utility. For example, to allow for preference for flexibility, suppose menus in the second period are evaluated by the DLR (2001) additive representation $\int_{\mathcal{U}} \max_{p\in A} u(p)\, \mu(du)$ for some measure μ on the set of expected-utility functions $\mathcal{U}$. As in Kreps and Porteus (1978), we can transform this utility value by a function ϕ to incorporate preferences for early or late resolution of uncertainty. This suggests the following representation, which includes the (two-stage) Kreps-Porteus representation as a special case.[25]

Definition 5 A *Kreps-Porteus-Dekel-Lipman-Rustichini (KPDLR) representation* is a pair (ϕ, μ), where μ is a finite Borel measure on $\mathcal{U}$ and $\phi : [a, b] \to \mathbb{R}$ is a Lipschitz continuous and strictly increasing function on the bounded interval $[a, b] = \{\int_{\mathcal{U}} \max_{p\in A} u(p)\, \mu(du) : A \in \mathcal{A}\}$, such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by:

$$V(A) = \phi\bigg(\int_{\mathcal{U}} \max_{p\in A} u(p)\, \mu(du)\bigg). \tag{8}$$

A *Kreps-Porteus representation* is a KPDLR representation where $\mu = \alpha\delta_u$ for some $u \in \mathcal{U}$ and $\alpha \geq 0$.[26]

[25] Kraus 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.

[26] A KPDLR representation (ϕ, μ) in which $\mu = \alpha\delta_u$ for $\alpha \geq 0$ corresponds to the Kreps-Porteus formulation in Equation (7) for the expected-utility function $v = \alpha u$.

While the KPDLR representation (and the Kreps-Porteus representation in particular) has the virtue of being relatively parsimonious, its drawback is that it places nontrivial restrictions the possible preferences for timing. To illustrate, consider any two menus A and B such that $V(A) = V(B)$. It then follows from Equation (8) that $V(\alpha A+(1-\alpha)C) = V(\alpha B + (1-\alpha)C)$ for any other menu C. This implies that the preference for early or late resolution of uncertainty exhibited for the menus A and C must be the same as that exhibited for the menus B and C. In fact, the utility difference between early and late resolution of uncertainty must be the same in both cases:

$$\alpha V(A) + (1-\alpha)V(C) - V(\alpha A + (1-\alpha)C)$$
$$= \alpha V(B) + (1-\alpha)V(C) - V(\alpha B + (1-\alpha)C).$$

This shows that the preference for timing does not depend directly on the content of the menus, only on the resulting utility values.

In particular, if a nonsingleton menu A satisfies $V(A) = V(\{p\})$ for some lottery p, the preference for timing for two-stage lotteries involving A is the same as for lotteries where $\{p\}$ takes the place of A. This illustrates why in the Kreps-Porteus representation, the preference for timing for general two-stage lotteries is completely determined by the preference for timing for temporal lotteries (without period 2 choice). This feature of their model is in contrast with the motivating example in Section 1.3 (and the costly contemplation representation from the previous section), 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.

To better illustrate the connection with our general model, we now provide an axiomatic treatment of the KPDLR model and describe how it can be formulated as a special case of our HA representation. Since the value function in Equation (8) is a monotone transformation of the DLR (2001) additive representation, it follows that for any menus A and B,

$$\delta_A \succsim \delta_B \iff \int_{\mathcal{U}} \max_{p\in A} u(p)\,\mu(du) \geq \int_{\mathcal{U}} \max_{p\in B} u(p)\,\mu(du)$$

Therefore, the KPDLR representation must satisfy the DLR (2001) axioms on the restricted domain of degenerate lotteries over menus. In fact, since V is determined up to a monotonic transformation by the ranking of lotteries δ_A for $A \in \mathcal{A}$, their axioms are also sufficient for the KPDLR representation (when combined with Axiom 1). Aside from weak order, continuity, and monotonicity, which were stated above, the key axiom for their representation is an independence axiom for menus. The following is a translation of their axiom to our two-stage setting.[27]

[27] The restrictions on preferences for timing in the KPDLR representation are also easily established

Axiom 6 (Mixture Independence) *For any $A, B, C \in \mathcal{A}$ and $\alpha \in (0,1)$,*

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

To obtain the more specialized Kreps-Porteus representation, we need to strengthen the monotonicity axiom to ensure there is no strict preference for flexibility. This is accomplished by the following axiom from Kreps (1979), which guarantees that the individual is indifferent between any menu and its best singleton subset. Kreps and Porteus (1978) implicitly assume the same relationship between the individual's ranking of menus and alternatives.[28]

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

The following result formalizes the connection between these axioms and the KPDLR representation.

Theorem 6 *A. The preference $\succsim$ has a KPDLR representation if and only if it satisfies Axiom 1, mixture independence, and monotonicity.*[29]

B. (Kreps and Porteus (1978)) The preference $\succsim$ has a Kreps-Porteus representation if and only if it satisfies Axiom 1, mixture independence, and strategic rationality.[30]

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

Note that KPDLR preferences need not be a subset of HA preferences, since they may violate PERU or PLRU. However, Theorem 6.C shows that a consistent preference for timing corresponds to convexity or concavity of ϕ. The following theorem describes how this subclass of KPDLR representations can be expressed in as special cases of our HA representation.

as direct implications of the axioms. By continuity and mixture independence, if $\delta_A \sim \delta_B$, then $\delta_{\alpha A+(1-\alpha)C} \sim \delta_{\alpha B+(1-\alpha)C}$ for any menu C. By continuity and first-stage independence, $\alpha\delta_A+(1-\alpha)\delta_C \sim \alpha\delta_B+(1-\alpha)\delta_C$. Thus, the preference for early or late resolution of uncertainty must be the same for A and C as for B and C.

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

[29] It is not necessary to include indifference to randomization (IR) explicitly in this result since it is implied by mixture independence.

[30] Kreps and Porteus (1978) only required that the transformation ϕ be continuous. We additionally require Lipschitz continuity of ϕ since we impose the L–continuity axiom throughout the paper.

Theorem 7 *Let $V : \mathcal{A} \to \mathbb{R}$ 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 (8).*
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*

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

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

Theorem 7 shows that for any KPDLR representation satisfying PERU or PLRU, the hidden actions in the corresponding HA representation can be indexed by a real number λ. In particular, condition (2.a) shows that every hidden action is a scalar multiple of a fixed measure μ. 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.

The form of the HA representation in condition (2) suggests the following interpretation: The distribution of possible tastes (determined by μ) is the same for every hidden action, and changing the action simply changes he magnitude of the ex post utilities by a common scalar multiple λ. The optimal action for a given menu A is therefore determined entirely by the value $\int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du)$ and the shape of the cost function c. This implies that if this integral expression takes the same value for any two menus A and B, the optimal hidden action must be the same for both menus. Consequently, the preference for timing in any two-stage lottery involving A must be the same if B takes the place of A in the lottery. Thus, condition (2) corresponds to the same restrictions on preferences for timing that were described already for the KPDLR representation, but expresses them in a different way: Condition (2) restricts the possible preferences for timing by placing strong restrictions on the complementarities between menus and hidden actions.

As noted above, one benefit of the KPDLR representation is that it is relatively parsimonious. Taking this perspective, one can think of Theorem 7 as describing the instances in which the KPDLR representation can be used as a reduced-form representation for a hidden action model. However, since condition (2) is quite restrictive, one implication of this result is that the KPDLR representation will only be appropriate in a rather limited set of circumstances.[31] Since the Kreps-Porteus representation is a special case — where $\mu = \alpha\delta_u$ and the corresponding hidden action representation exhibits no uncertainty about the ex post preference ranking — this argument applies to that model a fortiori.

On a final note, Theorem 7 generalizes several results from Kreps and Porteus (1979), who considered a class of hidden action representations and determined the conditions on the representation under which the resulting preference satisfies the axioms of Kreps and Porteus (1978). Specifically, Propositions 5 and 6 in Kreps and Porteus (1979) show that a hidden action representation corresponds to a Kreps-Porteus preference if and only if it takes a functional form that is essentially equivalent to the one described by condition (2) for a measure of the form $\mu = \alpha\delta_u$. Thus, their results follow when Theorem 7 is applied to measures taking the Kreps-Porteus form described in Definition 5.

[31] It is well-known that 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). Theorem 7 characterizes precisely those special cases in which independence is not violated.

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

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

Definition 7 Suppose $C \subset X$ and $f : C \to \mathbb{R}$. For $x \in C$, the *subdifferential* of f at x is defined to be

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

The subdifferential is useful for the approximation of convex functions by affine functions. It is straightforward to show that $x^* \in \partial f(x)$ if and only if the affine function $h : X \to \mathbb{R}$ defined by $h(y) = f(x) + \langle y - x, x^* \rangle$ satisfies $h \leq f$ and $h(x) = f(x)$. It should also be noted that when X is infinite-dimensional it is possible to have $\partial f(x) = \emptyset$ for some $x \in C$, even if f is convex. However, the following result shows that a Lipschitz continuous and convex function always has a nonempty subdifferential:

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

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

Definition 8 Suppose $C \subset X$ and $f : C \to \mathbb{R}$. The *conjugate* (or *Fenchel conjugate*) of f is the function $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$ defined by

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

There is an important duality between f and f^*. Lemma 3 summarizes certain properties of f^* that are useful in establishing this duality.[32] The proof is standard and can be found, for example, in the supplementary appendix of Ergin and Sarver (2010a).

Lemma 3 *Suppose $C \subset X$ and $f : C \to \mathbb{R}$. Then,*

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

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

$$f(x) = \max_{x^* \in X^*} \left[\langle x, x^* \rangle - f^*(x^*) \right] \tag{9}$$

for all $x \in C$.[33] 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 (9) holds. Let C_f denote the set of all $x \in C$ for which the subdifferential of f at x is a singleton:

$$C_f = \{x \in C : \partial f(x) \text{ is a singleton}\}. \tag{10}$$

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

$$\mathcal{N}_f = \{x^* \in X^* : x^* \in \partial f(x) \text{ for some } x \in C_f\}. \tag{11}$$

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

$$\mathcal{M}_f = \overline{\mathcal{N}_f}. \tag{12}$$

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 8 (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,[34] and (iii) $f : C \to \mathbb{R}$ is Lipschitz continuous and convex. Then, $\mathcal{M}_f$ is weak* compact, and for any weak* compact $\mathcal{M} \subset X^*$,*

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

The intuition for Theorem 8 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 8, C_f is dense in C. Therefore, it can be shown that for any $x \in C$,

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

[33] This is a slight variation of the classic Fenchel-Moreau theorem. The standard version of this theorem states that if $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and convex, then $f(x) = f^{**}(x) \equiv \sup_{x^* \in X^*} [\langle x, x^* \rangle - f^*(x^*)]$. See, e.g., Proposition 1 in Ekeland and Turnbull (1983, p97).

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

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

Theorem 9 (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 $f : C \to \mathbb{R}$ 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 10 (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 $c : \mathcal{M} \to \mathbb{R}$ be weak* lower semi-continuous. Define $f : C \to \mathbb{R}$ by*

$$f(x) = \max_{x^* \in \mathcal{M}} \big[\langle x, x^* \rangle - c(x^*)\big]. \tag{13}$$

Then,

1. *The function f is Lipschitz continuous and convex.*
2. *For all $x \in C$, there exists $x^* \in \partial f(x)$ such that $x^* \in \mathcal{M}$ and $f^*(x^*) = c(x^*)$. In particular, this implies $\mathcal{N}_f \subset \mathcal{M}$, $\mathcal{M}_f \subset \mathcal{M}$, and $f^*(x^*) = c(x^*)$ for all $x^* \in \mathcal{N}_f$.*
3. *If C is also compact (in the norm topology), then $f^*(x^*) = c(x^*)$ for all $x^* \in \mathcal{M}_f$.*

B Proof of Theorem 1

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:

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

Definition 9 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 $c : \mathcal{M} \to \mathbb{R}$ such that:

1. $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$, where $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (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 11 *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 11.B is simply a restatement of Theorem 1, and Theorem 11.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 11.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 11. Note that $\mathcal{A}$ is a compact metric space since $\triangle(Z)$ is a compact metric space (see, e.g., Munkres (2000, p280–281) or Theorem 1.8.3 in Schneider (1993, p49)). We begin by showing that weak order, continuity, and first-stage independence imply that $\succsim$ has an expected-utility representation.

Lemma 4 *A preference $\succsim$ over $\triangle(\mathcal{A})$ satisfies weak order, continuity, and first-stage independence if and only if there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $\succsim$ is represented by $\mathbb{E}_P[V]$. Furthermore, if $V : \mathcal{A} \to \mathbb{R}$ and $V' : \mathcal{A} \to \mathbb{R}$ are continuous functions such that $\mathbb{E}_P[V]$ and $\mathbb{E}_P[V']$ represent the same preference over $\triangle(\mathcal{A})$, then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $V' = \alpha V + \beta$.*

Proof: This is a standard result. For example, it is asserted without proof in Corollary 5.22 of Kreps (1988). It is also a consequence of Theorem 10.1 of Fishburn (1970) together with some simple arguments to establish continuity of V from the continuity axiom. ■

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

1. *If $\succsim$ satisfies L–continuity, then V is Lipschitz continuous on $\mathcal{A}^c$, i.e., there exists $K \geq 0$ such that $|V(A) - V(B)| \leq K d_h(A, B)$ for any $A, B \in \mathcal{A}^c$.*[36]

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, suppose $\succsim$ satisfies L–continuity for $M \geq 0$ and $A^*, A_* \in \mathcal{A}$. First, note that if $M = 0$, then L–continuity implies that $V(A) = V(B)$ for all $A, B \in \mathcal{A}$, i.e., V is Lipschitz continuous with a Lipschitz constant $K = 0$. If $M > 0$, then let $K \equiv 2M[V(A^*) - V(A_*)] \geq 0$. We first show that for any $A, B \in \mathcal{A}^c$:

$$d_h(A, B) \leq \tfrac{1}{2M} \Longrightarrow |V(A) - V(B)| \leq K d_h(A, B). \tag{14}$$

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

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

where the first inequality follows from L–continuity, the second inequality follows from $\alpha \leq 1/2$, and the equality follows form the definitions of α and K. Interchanging the roles of A and B above, we also have that $V(A) - V(B) \leq K d_h(A, B)$, proving Equation (14).

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

$$|V(A) - V(B)| \leq K d_h(A, B), \tag{15}$$

i.e., the requirement $d_h(A, B) \leq \frac{1}{2M}$ in Equation (14) is not necessary. To see this, take any sequence $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_n < \lambda_{n+1} = 1$ such that $(\lambda_{i+1} - \lambda_i) d_h(A, B) \leq \frac{1}{2M}$. Let $A_i = \lambda_i A + (1 - \lambda_i)B$. It is straightforward to verify that:[37]

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

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

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

Combining this with the triangle inequality and Equation (14), we obtain

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

This establishes Equation (15), which implies V is Lipschitz continuous on $\mathcal{A}^c$ with a Lipschitz constant K.

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

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

which implies the conclusion of L–continuity. ■

We now follow a construction similar to the one in DLR (2001) to obtain from V a function W whose domain is the set of support functions. As in the text, let

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

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

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

The next lemma summarizes some important properties of support functions. See DLR (2001) or Ergin and Sarver (2010a, Lemmas 5 and 6) for precise references and additional details.

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 $0 \in \Sigma$.*

Using the properties of support functions established in Lemma 6, the following result shows that a function defined on $\mathcal{A}^c$ can be transformed into a function on Σ, while maintaining the properties described in Lemma 5. For a proof, see Ergin and Sarver (2010a, Lemma 7).

Lemma 7 *Suppose $V : \mathcal{A}^c \to \mathbb{R}$, and define a function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$. Then:*

1. *$V(A) = W(\sigma_A)$ for all $A \in \mathcal{A}^c$.*
2. *V is Lipschitz continuous if and only if W is Lipschitz continuous.*
3. *If V is convex [concave] if and only if W is convex [concave].*
4. *V is monotone if and only if W is monotone (i.e., $\sigma \le \sigma'$ implies $W(\sigma) \le W(\sigma')$ for any $\sigma, \sigma' \in \Sigma$).*

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 $\mu \in C(\mathcal{U})^*$.[38]

For any function $W : \Sigma \to \mathbb{R}$, 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 (10), (11), and (12), 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 8 to the current setting.

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

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

Proof: We simply need to verify that $C(\mathcal{U})$, Σ, and W satisfy the assumptions of Theorem 8. 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)

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

that span(Σ) 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. ■

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 $V : \mathcal{A} \to \mathbb{R}$ 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 $W : \Sigma \to \mathbb{R}$ defined by $W(\sigma) = V(A_\sigma)$ is Lipschitz continuous and convex. Therefore, by Lemma 8, for all $\sigma \in \Sigma$,

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

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 9. 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 9, 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 $V : \mathcal{A} \to \mathbb{R}$ such that $\mathbb{E}_P[V]$ represents $\succsim$. By Lemmas 5 and 7, the function $W : \Sigma \to \mathbb{R}$ defined by $W(\sigma) = V(A_\sigma)$ is Lipschitz continuous and concave. Define a function $\bar{W} : \Sigma \to \mathbb{R}$ by $\bar{W}(\sigma) = -W(\sigma)$. Then, $\bar{W}$ is Lipschitz continuous and convex, so by Lemma 8, for all $\sigma \in \Sigma$,

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

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

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

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 9. 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 $\hat{W} : \hat{\Sigma} \to \mathbb{R}$ by $\hat{W}(\sigma) \equiv \bar{W}(-\sigma) = -W(-\sigma)$. Notice that $\hat{W}$ is monotone and convex: By the monotonicity of W, for any $\sigma, \sigma' \in \hat{\Sigma}$,

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

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 9, 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 $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (2), then V is Lipschitz continuous and convex. In addition, defining the function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$, we have $\mathcal{M} = \mathcal{M}_W$ and $c(\mu) = W^*(\mu)$ for all $\mu \in \mathcal{M}$.*
2. *If $(\mathcal{M}, c)$ is a signed min-HA representation and $V : \mathcal{A} \to \mathbb{R}$ is defined by Equation (3), then V is Lipschitz continuous and concave. In addition, defining the function $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$, we have $\mathcal{M} = -\mathcal{M}_{-W}$ and $c(\mu) = [-W]^*(-\mu)$ for all $\mu \in \mathcal{M}$.*

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

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

By part 1 of Theorem 10, 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 10 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 $\bar{W} : \Sigma \to \mathbb{R}$ by $\bar{W}(\sigma) = -W(\sigma)$. Then, for any $\sigma \in \Sigma$,

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

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 $\mu \in \mathcal{M}$. ■

Suppose that $\succsim$ has a signed max-HA [signed min-HA] representation $(\mathcal{M}, c)$, and suppose $V : \mathcal{A} \to \mathbb{R}$ 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 9). Theorem 2 is a special case of Theorem 12, and therefore follows directly.

Theorem 12 *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 $V : \mathcal{A} \to \mathbb{R}$ and $V' : \mathcal{A} \to \mathbb{R}$ for these respective representations, and define $W : \Sigma \to \mathbb{R}$ and $W' : \Sigma \to \mathbb{R}$ by $W(\sigma) = V(A_\sigma)$ and $W'(\sigma) = V'(A_\sigma)$. By part 1 of Lemma 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,

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

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 $U : \Omega \to \mathbb{R}^Z$ by $U(\omega) = \omega$.

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

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

for any $A \in \mathcal{A}$.[39] Let $\Pi = \{\pi_\mu : \mu \in \mathcal{M}\}$ and $\tilde{c}(\pi_\mu) = c(\mu)$. Then, V can be expressed in the following SSV form:

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

(3 $\Rightarrow$ 2): Let $((\Omega, \mathcal{F}), U, \Pi, c)$ be an SSV representation, and define V by Equation (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.,

$$V(A) = \min_{\pi \in \Pi'} \left(\int_\Omega \max_{p \in A} U(\omega) \cdot p \; \pi(d\omega) + c(\pi) \right) \tag{16}$$

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 (16) 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 (16) for menus A and A', respectively. Then, optimality of

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

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

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

π at A implies:

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

Similarly, optimality of π' at A' implies:

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

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

Let $\tilde{\Omega} = \Omega \times \Pi'$. Let $\mathcal{G}$ be any σ-algebra on Π' that contains all singletons and such that $c|_{\Pi'} : \Pi' \to \mathbb{R}$ is $\mathcal{G}$-measurable (e.g., $\mathcal{G} = 2^{\Pi'}$). Let $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{G}$ be the product σ-algebra on $\tilde{\Omega}$. Let $\mathbf{1} \in \mathbb{R}^Z$ denote the vector whose coordinates are equal to 1, and define $\tilde{U} : \tilde{\Omega} \to \mathbb{R}^Z$ by $\tilde{U}(\omega, \pi) = U(\omega) + c(\pi)\mathbf{1}$ for any $\tilde{\omega} = (\omega, \pi) \in \tilde{\Omega}$. Note that $\tilde{U}$ is $\tilde{\mathcal{F}}$-measurable and bounded.[40]

For any $\pi \in \Pi'$, define the function $f_\pi : \Omega \to \tilde{\Omega}$ by $f_\pi(\omega) = (\omega, \pi)$. Note that f_π is measurable.[41] 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.[42]

Letting $\tilde{\Pi}' = \{\rho_\pi : \pi \in \Pi'\}$, by Equation (16), 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

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

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

[42] To see this, define the function $g : \tilde{\Omega} \to \mathbb{R}$ by $g(\tilde{\omega}) = \max_{p\in A} \tilde{U}(\tilde{\omega})\cdot p$. By a similar argument as in Footnote 39, g is $\tilde{\mathcal{F}}$-measurable. Then, the change of variables formula is $\int_{\tilde{\Omega}} g(\tilde{\omega})\,(\pi \circ f_\pi^{-1})(d\tilde{\omega}) = \int_\Omega g(f_\pi(\omega))\,\pi(d\omega)$.

every $\pi \in \Pi$, define $f_\pi : \mathcal{A} \to \mathbb{R}$ by

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

Since U is bounded, there exists $\kappa > 0$ such that $\|U(\omega)\| \leq \kappa$ 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 $\pi \in \Pi$. Since V is the pointwise minimum of f_π over $\pi \in \Pi$, it is also Lipschitz continuous with the same Lipschitz constant κ.

Since $V : \mathcal{A} \to \mathbb{R}$ is monotone, concave, Lipschitz continuous, and it satisfies the IR condition $V(A) = V(\text{co}(A))$ 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 $p, q \in \triangle(Z)$, we have $p - q \in \Theta$.

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

Lemma 10 *Suppose that $V : \mathcal{A} \to \mathbb{R}$ is a function such that $\mathbb{E}_P[V]$ represents the preference $\succsim$ over $\triangle(\mathcal{A})$. Then, V is translation linear if and only if $\succsim$ satisfies RDD.*

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

$$V(\alpha A + (1-\alpha)\{p\}) - V(\alpha A + (1-\alpha)\{q\}) = (1-\alpha)[V(\{p\}) - V(\{q\})] \tag{17}$$

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

If there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}$ and $\theta \in \Theta$ with $A+\theta \in \mathcal{A}$, we have $V(A+\theta) = V(A) + v \cdot \theta$, then both sides of Equation (17) are equal to $(1-\alpha)v \cdot (p-q)$, showing that $\succsim$ satisfies RDD.

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

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

where the second equality follows from Equation (17) and the definition of f. Therefore, $f(\alpha p + (1-\alpha)q) = \alpha f(p) + (1-\alpha)f(q)$ for any $\alpha \in [0,1]$ and $p, q \in \triangle(Z)$. It is standard to show that this implies that there exists $v \in \mathbb{R}^Z$ such that $f(p) = v \cdot p$ for all $p \in \triangle(Z)$.

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

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

where the second equality follows from Equation (17) and the third equality follows from the expected utility form of f. Therefore, V is translation linear. ■

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 $q = (1/|Z|, \ldots, 1/|Z|) \in \triangle(Z)$. By Lemma 22 of Ergin and Sarver (2010a), for all $\mu \in \mathcal{M}_W$ and $p \in \triangle(Z)$, $\langle \sigma_{\{p\}}, \mu \rangle = v \cdot (p - q)$. The consistency of $\mathcal{M}_W$ follows immediately from this fact because for any $\mu, \mu' \in \mathcal{M}_W$ and $p \in \triangle(Z)$, we have

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

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

F.1 Proof of Theorem 6.A

The necessity of the axioms is straightforward. For sufficiency, suppose that $\succsim$ satisfies Axiom 1, mixture independence, and monotonicity. By Lemma 4, there exists a continuous function $V : \mathcal{A} \to \mathbb{R}$ such that $P \succsim Q$ if and only if $\mathbb{E}_P[V] \geq \mathbb{E}_Q[V]$.

Define a preference $\succsim'$ on $\mathcal{A}^c$ by $A \succsim' B \iff \delta_A \succsim \delta_B$ (or, equivalently, $A \succsim' B \iff V(A) \geq V(B)$). The axioms assumed on $\succsim$ then imply that $\succsim'$ satisfies the DLR (2001) axioms: Continuity of $\succsim$ implies continuity of $\succsim'$; mixture independence implies that $\succsim'$ satisfies independence; and monotonicity of $\succsim$ implies that $\succsim'$ satisfies monotonicity. Therefore, by Theorem S2 in the supplementary appendix of DLRS (2007), there exists a finite Borel measure μ on $\mathcal{U}$ such that $U : \mathcal{A} \to \mathbb{R}$ defined by

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

represents $\succsim'$. Moreover, since U is continuous and $\mathcal{A}$ is compact, there exist $-\infty < a \leq b < +\infty$ such that $[a, b] = \{U(A) : A \in \mathcal{A}\}$. Since $V(A) \geq V(B) \iff U(A) \geq U(B)$, there exists a strictly increasing function $\phi : [a, b] \to \mathbb{R}$ such that

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

To establish the Lipschitz continuity of ϕ, first recall that by Lemma 5, L–continuity implies there exists $K \geq 0$ such that $|V(A) - V(B)| \leq K d_h(A, B)$ for any $A, B \in \mathcal{A}^c$ (the set of all convex menus). 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 $U(A^*) = b$. For any $t \in [a, b]$, let $\alpha(t) \equiv (t - a)/(b - a) \in [0, 1]$, which implies $U(\alpha(t)A^* + (1 - \alpha(t))A_*) = t$. Note that for any $\alpha, \beta \in [0, 1]$,

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

Thus, for any $s, t \in [a, b]$,

$$\begin{aligned}
|\phi(t) - \phi(s)| &= \Big|\phi\Big(U(\alpha(t)A^* + (1 - \alpha(t))A_*)\Big) - \phi\Big(U(\alpha(s)A^* + (1 - \alpha(s))A_*)\Big)\Big| \\
&= \big|V(\alpha(t)A^* + (1 - \alpha(t))A_*) - V(\alpha(s)A^* + (1 - \alpha(s))A_*)\big| \\
&\leq K|\alpha(t) - \alpha(s)| d_h(A^*, A_*) \\
&= K|t - s| d_h(A^*, A_*)/(b - a),
\end{aligned}$$

which implies ϕ is Lipschitz continuous with a Lipschitz constant of $K d_h(A^*, A_*)/(b - a)$.

F.2 Proof of Theorem 6.B

The necessity of the axioms is straightforward. For sufficiency, suppose that $\succsim$ satisfies Axiom 1, second-stage independence, and strategic rationality. By part A, $\succsim$ has a KPDLR representation (ϕ, μ). It therefore suffices to show that μ has singleton support.

Suppose for a contradiction that μ has more than one utility function in its support. Fix any $u', u'' \in \text{supp}(\mu)$. Choose lotteries $p, q \in \triangle(Z)$ such that $u'(p) > u'(q)$ and $u''(q) > u''(p)$. Then, since these inequalities also hold on any small neighborhoods of u' and u'', respectively, this implies

$$\int_{\mathcal{U}} \max_{r \in \{p,q\}} u(r)\, \mu(du) > \int_{\mathcal{U}} u(p)\, \mu(du) \quad \text{and} \quad \int_{\mathcal{U}} \max_{r \in \{p,q\}} u(r)\, \mu(du) > \int_{\mathcal{U}} u(q)\, \mu(du).$$

Therefore, $V(\{p,q\}) > V(\{p\})$ and $V(\{p,q\}) > V(\{q\})$, which constitutes a violation of the strategic rationality axiom. Thus, if $\succsim$ satisfies strategic rationality, $\text{supp}(\mu) = \{u\}$ for some $u \in \mathcal{U}$. Taking $\alpha = \mu(\{u\})$, we then have $\mu = \alpha \delta_u$, as desired.

F.3 Proof of Theorem 6.C

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 menus that satisfy $s = \int_{\mathcal{U}} \max_{p \in A} u(p)\, \mu(du)$ and $t = \int_{\mathcal{U}} \max_{p \in B} u(p)\, \mu(du)$. Then, for any $\alpha \in (0,1)$,

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

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.

G Proof of Theorem 7

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

Lemma 11 *Let $a, b \in \mathbb{R}$ with $a < b$ and let $\phi : [a,b] \to \mathbb{R}$ be Lipschitz continuous and convex. Then, $1 \Leftrightarrow 2 \Rightarrow 3$:*

1. *ϕ is strictly increasing.*
2. *(a) $\mathcal{M}_\phi \subset \mathbb{R}_+$.*

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

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

Proof: (1 ⇒ 2) Part a follows from Theorem 9.

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

$$\lambda' a \leq \phi^*(\lambda') - \phi^*(0) \leq \lambda' t \quad \forall \lambda' \in (0, \lambda). \tag{18}$$

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 (18). 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 $t' \leq t$.[43] 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 (18).

(2 ⇒ 1) Theorem 8 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.

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

$(1 \Rightarrow 3)$ Suppose for a contradiction that 0 is an isolated point of $\mathcal{M}_\phi$. Then, $0 \in \mathcal{N}_\phi$, i.e., there exists $t \in [a, b]$ such that $\partial\phi(t) = \{0\}$. Then, Lemma 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 $\mathcal{M}_\phi$ and $\mathcal{M}_\phi$ is compact by Theorem 8, $\mathcal{M}_\phi \setminus \{0\}$ is also compact. Therefore, the above inequality implies that

$$-\phi^*(0) > \max_{\lambda \in \mathcal{M}_\phi \setminus \{0\}} \big[\lambda t - \phi^*(\lambda)\big]. \tag{19}$$

Let $\Delta > 0$ be the difference of the left hand side and the right hand side in Equation (19) 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 (19) continues to hold if t is replaced by s. Therefore,

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

where the second equality follows from Theorem 8. This implies that ϕ is constant at a $\frac{\Delta}{M}$ neighborhood of t, contradicting the assumption that ϕ is strictly increasing. ■

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 $\phi : [a, b] \to \mathbb{R}$ be Lipschitz continuous and convex and define $W : \Sigma \to \mathbb{R}$ by $W(\sigma) = \phi(\langle\sigma, \mu\rangle)$ for any $\sigma \in \Sigma$. Then,*

1. *W is Lipschitz continuous and convex.*
2. *$W^*(\lambda\mu) = \phi^*(\lambda)$ for any $\lambda \in \mathbb{R}$.*
3. *$\mathcal{M}_W = \{\lambda\mu : \lambda \in \mathcal{M}_\phi\}$.*

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

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

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

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

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

(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 $\mathcal{M}_\phi$ is closed. Let $\nu \in \mathcal{N}_W$, then there exists $\sigma \in \Sigma$ such that $\partial W(\sigma) = \{\nu\}$. For any $\lambda \in \partial\phi(\langle\sigma, \mu\rangle)$,

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

implying $\lambda\mu \in \partial W(\sigma) = \{\nu\}$. Therefore, $\{\lambda\mu : \lambda \in \partial\phi(\langle\sigma, \mu\rangle)\} \subset \{\nu\}$. Since μ is nonzero and $\partial\phi(\langle\sigma, \mu\rangle) \neq \emptyset$ by Lemma 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 8, $\mathcal{M}$ 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 8, 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 8, $\mathcal{M}_\phi \subset \mathcal{M}$. ■

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

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

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

where the last equality follows from part 1 of Lemma 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 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,

$$\lim_{\lambda\searrow 0} \frac{W^*(\lambda\mu) - W^*(0)}{\lambda} = \frac{d\phi^*}{d\lambda^+}(0) = a \equiv \min_{A\in\mathcal{A}} \int_\mathcal{U} \max_{p\in A} u(p)\, \mu(du). \tag{20}$$

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 (20) agrees with $\lim_{\lambda\searrow 0 : \lambda\mu\in\mathcal{M}_W} \frac{c(\lambda\mu) - c(0)}{\lambda}$.

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

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

By Theorem 10, 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 $\mathcal{M}_\phi$.

First suppose that $0 \notin \mathcal{M}$, implying $0 \notin \mathcal{M}_\phi$. Let $t, t' \in [a, b]$ be such that $t < t'$. By Theorem 8,

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

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,

$$a = \lim_{\lambda \searrow 0 : \lambda\mu \in \mathcal{M}} \frac{c(\lambda\mu) - c(0)}{\lambda} = \lim_{\lambda \searrow 0 : \lambda \in \mathcal{M}_\phi} \frac{\phi^*(\lambda) - \phi^*(0)}{\lambda} \tag{21}$$

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 $\mathcal{M}_\phi$, Equation (21) 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 (21) is equal to $\frac{d\phi^*}{d\lambda^+}(0)$. By Lemma 11, ϕ is strictly increasing. ■

References

Ahn, D., and T. Sarver (2011): "Preference for Flexibility and Random Choice," Mimeo.

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

Athey, S., and I. Segal (2007): "An Efficient Dynamic Mechanism," Mimeo.

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.

Compte, O., and P. Jehiel (2007): "Auctions and Information Acquisition: Sealed Bid or Dynamic Formats?," *RAND Journal of Economics*, 38, 355–372.

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 M. Schneider (2003): "Recursive Multiple-Priors," *Journal of Economic Theory*, 113, 1–31.

Epstein, L. G., and K. Seo (2009): "Subjective States: A More Robust Model," *Games and Economic Behavior*, 67, 408–427.

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.

Hayashi, T. (2005): "Intertemporal Substitution, Risk Aversion, and Ambiguity Aversion," *Economic Theory*, 25, 933–956.

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 (2006a): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006b): "Dynamic Variational Preferences," *Journal of Economic Theory*, 128, 4–44.

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.

Saito, K. (2011): "Preference for Randomization and Ambiguity Aversion," Mimeo.

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.

Seo, K. (2009): "Ambiguity and Second-Order Belief," *Econometrica*, 77, 1575–1605.

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. (2011a): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79, 47–73.

Strzalecki, T. (2011b): "Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion," Mimeo.

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.