A PARAMETRIC MODEL OF AMBIGUITY HEDGING

Igor Kopylov*

Institute for Mathematical Behavioral Sciences, and

Department of Economics, University of California, Irvine, CA 92697

Abstract

The Ellsberg Paradox demonstrates that many people like to *hedge* ambiguity in a way that contradicts expected utility maximization. We postulate roughly that *all* violations of the expected utility theory are due to *affinity* for ambiguity hedging. The associated utility representation is a special case of maxmin expected utility (Gilboa and Schmeidler, 1989), where the set of priors Π has a well-known parametric structure called ε-*contamination*. Both the weight $\varepsilon \in [0, 1]$ and the contaminated prior $p \in \Pi$ are uniquely derived from preference. The parameter ε is interpretable as a degree of affinity for hedging. A model of *complete ignorance* is obtained as a special case where the prior p is contaminated by the simplex of all probability measures.

1 Introduction

Recall the famous paradox of Ellsberg [7], where a decision maker is told that (i) a ball will be drawn randomly from an urn that contains balls of three possible colors (red, green, and blue), and (ii) the proportion of red balls in the urn is $\frac{1}{3}$. Then the typical preference is

(1) to bet on the event $\{R\}$, whose chance is known to be $\frac{1}{3}$, rather than on $\{G\}$, whose chance is not known precisely and lies between 0 and $\frac{2}{3}$;

(2) to bet on the event $\{G, B\}$, whose chance is known to be $\frac{2}{3}$, rather than on $\{R, B\}$, whose chance is not known precisely and lies between $\frac{1}{3}$ and 1.

*I am especially grateful to Larry Epstein for providing comments at all stages of this research. I have also benefitted from discussions with Mark Machina, Jean-Claude Falmagne, and participants of the IMBS and UCSD research seminars. This draft is preliminary and incomplete. All mistakes are my own.

Informally, the above preference reversal occurs because a side bet on the event $\{B\}$ *hedges* the ambiguous bet on $\{G\}$, but does not hedge the unambiguous bet on $\{R\}$. The resulting choice behavior does not comply with expected utility and violates the standard separability conditions—the Sure Thing Principle and Independence—that are formulated by Savage [22] and by Anscombe–Aumann [1] respectively.

In this paper we postulate roughly that *all* violations of the expected utility theory are due to *affinity* for ambiguity hedging. This postulate strengthens Convexity in the multiple priors model (Gilboa–Schmeidler [14]) and delivers the following utility functional form:[1]

$$U(f) = (1-\varepsilon)\int_S u(f(s))\mathrm{d}p + \varepsilon \min_{q\in\Delta} \int_S u(f(s))\mathrm{d}q. \tag{1.1}$$

Here Δ is the set of all probability measures on S that agree with the information that the decision maker has a priori. (For example in the Ellsberg Paradox, Δ is the family of all measures that assign probability $\frac{1}{3}$ to the event $\{R\}$.) The set Δ is exogenous, but all the other components of representation (1.1)—the expected utility index u, the parameter $\varepsilon \in [0,1]$, the probability measure $p \in conv\ \Delta$—are derived from preference in an essentially unique way.

The representation (1.1) is a special case of maxmin expected utility, where the set of priors

$$\Pi = (1-\varepsilon)\{p\} + \varepsilon\ conv\ \Delta \tag{1.2}$$

has the parametric structure of *ε-contamination* besides the standard convexity and closedness. The added structure seems beneficial for applied modelers for whom the freedom in Gilboa-Schmeidler's specification of Π may be an embarrassment of riches.

One can adapt the representations (1.1) and (1.2) to model situations of *complete ignorance* where any probabilistic scenario on S agrees with a priori information given to the decision maker. Then Δ equals the simplex Δ_S of all probability measures on S, and

$$\Pi = (1-\varepsilon)\{p\} + \varepsilon\Delta_S. \tag{1.3}$$

This structure is convenient in applications, such as models of asset pricing (Epstein and Wang [8]) and search (Nishimura and Ozaki [19]).

The functional form (1.1) can be interpreted in a way different from the multiple priors model. The decision maker, as portrayed by (1.1), evaluates every act f as an ε-mixture of the lottery that f induces via the least favorable scenario in Δ and the lottery that f induces via the probability measure $p \in conv\ \Delta$. In this interpretation, the measure p reflects a subjective probabilistic *opinion*—a state

[1] This representation is defined for Anscombe–Aumann's *acts* f that map states of the world $s \in S$ into *lotteries*—objective distributions over deterministic outcomes. In principle, any model of choice on Anscombe–Aumann's domain can be translated into a Savage-style framework that does not rely on objective randomization (Ghirardato *et al.* [11], Casadeus-Masanell *et al.* [5]).

of mind less rigid than belief—and the parameter $\varepsilon \in [0,1]$ reflects a subjective willingness to compromise this opinion by hedging against unfavorable scenarios in Δ. For example, when $\varepsilon = 0$, the decision maker uses p to rank all acts via expected utility, but ignores hedging. Alternatively, when $\varepsilon = 1$, she evaluates every act f via the least favorable scenario in Δ regardless of p.

1.1 Related Literature

Ellsberg [7] adopts the functional form (1.1) to accommodate his paradox in an *ad hoc* fashion. He interprets $1 - \varepsilon$ as a degree of confidence that the decision maker has in the probabilistic scenario p.

A number of recent models (Gajdos, Tallon, and Vergnaud [9], Hayashi [15], Damiano [6], Olszewski [20]) adopt an exogenous set Δ to describe a priori information given to the decision maker. Some of these authors (Gajdos *et al.* and Hayashi) obtain ε-contamination as a special case in their representations. Moreover, Hayashi [15] and Damiano [6] derive an additional structure for p (the Steiner point and the nucleolus of Δ respectively). However, all of these researchers include the set Δ into objects of choice and allow comparisons between any "act-information" pairs (f, Δ) and (g, Δ'). For instance, Aversion towards Imprecision—the key axiom in Gajdos *et al.*—requires that (f, Δ) is preferred to (f, Δ') whenever Δ is suitably nested in Δ'. In contrast, we follow the more conventional approach and study preferences over acts; moreover, our axioms are relatively few and arguably more appealing.

Nishimura and Ozaki [18] axiomatize the special case (1.3) for preferences over Anscombe–Aumann's acts. However, they take the index ε as a primitive instead of deriving it from preference. Their model has many other differences from ours.

The ε-contamination structure (1.2) and its special case (1.3) have been extensively studied in statistics literature (Hodges and Lehmann [16], Blum and Rosenblatt [4], Bickel [3], Berger and Berliner [2], Wasserman and Kadane [23]). In this literature, the parameter ε is commonly interpreted as the amount of error that is deemed possible for the prior p. This interpretation differs from ours because it uses ε to describe the imprecision of a priori knowledge rather than the effect of this imprecision on decision making.

1.2 Outline

This paper proceeds as follows. Section 2 contains preliminaries. Section 3 presents the main representation results and a sketch of proofs. Discussion is relegated to Section 4, and formal proofs to the appendices.

2 Preliminaries

Adopt a simple version of Anscombe–Aumann's framework. Denote by

- X a set of deterministic *outcomes*;
- S a finite set of of *states* of the world;[2]
- $\mathcal{L}$ the set of all probability distributions that have a finite support in X;
- $\mathcal{H}$ the set of all functions $f : S \to \mathcal{L}$;
- $\succeq$ a binary relation on $\mathcal{H}$.

Call subsets of S *events*, elements of $\mathcal{L}$ *lotteries*, and elements of $\mathcal{H}$ *acts*. Identify lotteries with corresponding constant acts. Given acts f and g and a weight $\alpha \in [0, 1]$, define a *mixture* $\alpha f + (1 - \alpha)\, g$ as an act that satisfies

$$[\alpha f + (1 - \alpha)\, g](s) = \alpha f(s) + (1 - \alpha)\, g(s) \qquad \text{for all } s \in S.$$

Interpret any act $f \in \mathcal{H}$ as a physical action that delivers the lottery $f(s)$ contingent on the state of the world $s \in S$. The deterministic outcome of this action is then resolved in X via the distribution $f(s)$ generated by an objective randomizing device, such as a fair coin or a balanced roulette wheel.

Interpret $\succeq$ as the decision maker's weak preference over acts. Say that $\succeq$ is *non-degenerate* if its strict part $\succ$ is not empty.

Gilboa–Schmeidler [14] formulate the following axioms for $\succeq$.

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

Axiom 2 (Continuity). *For any acts $f \succ g \succ h$, there exist $\alpha, \gamma \in (0, 1)$ such that $\alpha f + (1 - \alpha)h \succ g \succ \gamma f + (1 - \gamma)h$.*

Axiom 3 (Monotonicity). *If acts f and g satisfy $f(s) \succeq g(s)$ for all $s \in S$, then $f \succeq g$.*

Axiom 4 (Certainty Independence). *For all $\alpha \in (0, 1)$, acts $f, g \in \mathcal{H}$, and lotteries $l \in \mathcal{L}$,*

$$f \succeq g \quad \Leftrightarrow \quad \alpha f + (1 - \alpha)l \succeq \alpha g + (1 - \alpha)l.$$

Axiom 5 (Convexity). *For all $\alpha \in (0, 1)$ and acts $f, g \in \mathcal{H}$,*

$$f \sim g \quad \Rightarrow \quad \alpha f + (1 - \alpha)g \succeq g.$$

[2] The assumption that S is finite is not essential for our representation results and will be dropped later.

Order, Continuity, and Monotonicity are standard conditions of rationality: in particular, Monotonicity asserts state independence for the ranking of lotteries. The other two axioms, Certainty Independence and Convexity, are weak separability conditions that seem consistent with the Ellsberg-type affinity for ambiguity hedging.

Let $\Delta_S = \{p, q \dots\}$ be the simplex of all probability measures on S. Given a probability measure $q \in \Delta_S$ and an act $f \in \mathcal{H}$, define a lottery

$$f(q) = \sum_{s \in S} q(s) f(s),$$

and say that f *induces* $f(q)$ under the scenario q.

Gilboa–Schmeidler's main result asserts that Axioms 1–5 are necessary and sufficient for the preference $\succeq$ to be represented by maxmin expected utility

$$U(f) = \min_{q \in \Pi} u\left(f(q)\right), \tag{2.1}$$

where $u : \mathcal{L} \to \mathbb{R}$ is an affine function, and $\Pi \subset \Delta_S$ is a non-empty, convex, and closed set of probability measures. Moreover, if $\succeq$ is non-degenerate, then u is unique up to a positive linear transformation, and Π is unique.

3 Representation Results

The following example illustrates a simple model of ambiguity hedging and motivates a refinement of Gilboa–Schmeidler's Convexity.

Example 1. Let the state of the world $s \in S = \{G, R, B\}$ be determined by the color of a ball drawn randomly from an urn. Confront the decision maker with acts

	f_*	g_*	h_*	h'_*	$\frac{f_*+h_*}{2}$	$\frac{g_*+h_*}{2}$	$\frac{f_*+h'_*}{2}$	$\frac{g_*+h'_*}{2}$
G	1	0	0.4	0	0.7	0.2	0.5	0
R	0	1	0	0.4	0	0.5	0.2	0.7
B	0	0	1	0.4	0.5	0.5	0.2	0.2

(3.1)

where outcomes are specified in von Neumann–Morgenstern's utils.[3] Suppose that the decision maker is not told anything about the composition of the urn (except that each ball is either red, or green, or blue). Then both mixtures $\frac{f_*+h'_*}{2}$ and $\frac{g_*+h_*}{2}$ guarantee her at least 0.2 utils in any state of the world, but neither $\frac{f_*+h_*}{2}$ nor $\frac{g_*+h'_*}{2}$ guarantees her more than 0. Intuitively this means that h'_* *hedges* f_* better than h_* does, but h_* hedges g_* better than h'_* does. Affinity for such hedging,

[3]Formally, one can identify any number π in this table with a lottery that pays a positive monetary prize (say \$100) with probability π, and nothing (\$0) otherwise.

together with the standard separability argument, should motivate the decision maker to comply with the invariance[4]

$$\tfrac{f_*+h_*}{2} \succeq \tfrac{g_*+h_*}{2} \quad \Rightarrow \quad \tfrac{f_*+h'_*}{2} \succeq \tfrac{g_*+h'_*}{2}. \tag{3.2}$$

Intuitively, in order to rank the mixtures $\frac{f_*+h_*}{2}$ and $\frac{g_*+h_*}{2}$, the decision maker should (i) compare the distinct components f_* and g_*, and (ii) evaluate hedging that the common component h_* provides for the ambiguous acts f_* and g_*. When h_* is replaced by h'_* in these mixtures, the components f_* and g_* do not change, while hedging improves in $\frac{f_*+h'_*}{2}$ and deteriorates in $\frac{g_*+h'_*}{2}$.

Next we generalize the above approach to the abstract framework with an arbitrary finite state space S. For any act $f \in \mathcal{H}$, let f_S be the worst lottery that f may yield on S. Formally, $f_S \in \mathcal{L}$ is such that $f(s) \succeq f_S$ for all $s \in S$, and $f(s') = f_S$ for some $s' \in S$. Write $f \succeq_S g$ or $f \succ_S g$ if $f_S \succeq g_S$ or $f_S \succ g_S$ respectively.

Impose a mild version of *complete ignorance*: assume that any probabilistic scenario in Δ_S agrees with a priori information given to the decision maker.[5] For any acts f, h, h' such that $\frac{f+h}{2} \succ_S \frac{f+h'}{2}$, the mixture $\frac{f+h}{2}$ guarantees at least the lottery $\left(\frac{f+h}{2}\right)_S$ in any state of the world. In contrast, the mixture $\frac{f+h'}{2}$ provides no such guarantee as it may yield a lottery $\left(\frac{f+h'}{2}\right)_S$ worse than $\left(\frac{f+h}{2}\right)_S$. Being completely ignorant, the decision maker should conclude that h hedges f better than h' does. This conclusion is no longer intuitive if $\frac{f+h'}{2} \succeq_S \frac{f+h}{2}$.

Axiom 6 (Affinity for Hedging). *For all acts $f, g, h, h' \in \mathcal{H}$,*

$$\tfrac{f+h'}{2} \succeq_S \tfrac{f+h}{2} \succeq \tfrac{g+h}{2} \succeq_S \tfrac{g+h'}{2} \quad \Rightarrow \quad \tfrac{f+h'}{2} \succeq \tfrac{g+h'}{2}. \tag{3.3}$$

This axiom asserts roughly that the decision maker likes to hedge the payoff that she receives in the most unfavorable state of the world. Accordingly, she can motivate a reversal of preference from $\frac{f+h}{2} \succeq \frac{g+h}{2}$ to $\frac{f+h'}{2} \prec \frac{g+h'}{2}$, but only if h hedges f better than h' does, or if h' hedges g better than h does. Neither condition holds when $\frac{f+h'}{2} \succeq_S \frac{f+h}{2}$ and $\frac{g+h}{2} \succeq_S \frac{g+h'}{2}$.

Given Axioms 1–4, Affinity for Hedging implies Convexity. We show this directly for symmetric mixtures. For any acts $f, g \in \mathcal{H}$, by Certainty Independence,

$$\begin{aligned}
&\tfrac{f+g_S}{2} \succeq \tfrac{g+g_S}{2}, \\
&\left(\tfrac{f+g}{2}\right)_S \succeq \tfrac{f_S+g_S}{2} \sim \left(\tfrac{f+g_S}{2}\right)_S, \quad \text{and} \\
&\left(\tfrac{g+g_S}{2}\right)_S \sim \tfrac{g_S+g_S}{2} \sim \left(\tfrac{g+g}{2}\right)_S.
\end{aligned}$$

[4] The decision maker might be still concerned about the invariance (3.2), for example, because the upside potential is 0.7 for payoffs in the mixtures $\frac{f_*+h_*}{2}$ and $\frac{g_*+h'_*}{2}$, but only 0.5 for payoffs in $\frac{f_*+h'_*}{2}$ and $\frac{g_*+h_*}{2}$. Our model assumes that such concerns should not affect choice behavior.

[5] Here it is sufficient to assume that any *extreme* scenario in Δ_S agrees with a priori information. So in Example 1, the decision maker might be told that all balls in the urn have the same (but unspecified) color.

Thus by Affinity for Hedging,

$$f \sim g \quad \Rightarrow \quad \tfrac{f+g}{2} \succeq_S \tfrac{f+g_S}{2} \succeq \tfrac{g+g_S}{2} \succeq_S \tfrac{g+g}{2} = g \quad \Rightarrow \quad \tfrac{f+g}{2} \succeq g.$$

Note that Gilboa–Schmeidler's representation result remains valid if Convexity is imposed only for symmetric mixtures $\frac{f+g}{2}$.

On the other hand, if S has at least three elements, then Affinity for Hedging need not hold in the multiple priors model. For example, let $S = \{G, R, B\}$. Then the preference represented by maxmin expected utility with a set of priors $\Pi = \{p \in \Delta_S \,:\, p(\{R\}) = \frac{1}{3}\}$ violates the invariance (3.2).

The following result employs Affinity for Hedging to characterize a parametric utility representation for the preference $\succeq$.

Theorem 3.1. *$\succeq$ satisfies Order, Continuity, Monotonicity, Certainty Independence, and Affinity for Hedging iff $\succeq$ is represented by*

$$U(f) = (1-\varepsilon)u(f(p)) + \varepsilon \min_{s \in S} u(f(s)), \tag{3.4}$$

where $u : \mathcal{L} \to \mathbb{R}$ is an affine index, $\varepsilon \in [0,1]$, and $p \in \Delta_S$.

Moreover, if $\succeq$ is non-degenerate and S is non-singleton, then u is unique up to a positive linear transformation, ε is unique, and p is unique whenever $\varepsilon < 1$.

Obviously, the function U is a special case of maxmin expected utility with a set of priors[6]

$$\Pi = (1-\varepsilon)p + \varepsilon \Delta_S.$$

Thus, Theorem 3.1 refines the multiple priors model and delivers a parametric structure for the convex and closed set Π—this set is an *ε-contamination* of the probability measure $p \in \Delta_S$ by the entire simplex Δ_S. Note that both components of this structure, ε and p, are uniquely derived from preference. Figure 1 provides an example of the set Π in a three-color setting.

Alternatively, one can interpret the functional form (3.4) in a way that does not use the set Π altogether. The decision maker, as portrayed by (3.4), evaluates every act f via an ε-mixture of the worst lottery f_S that f may possibly yield, and the lottery $f(p)$ that f induces via p. Here the measure $p \in \Delta_S$ reflects her subjective probabilistic *opinion*, and the parameter $\varepsilon \in [0,1]$ her subjective willingness to compromise this opinion by ambiguity hedging. For example, when $\varepsilon = 0$, the decision maker relies on p to rank all acts via expected utility, and ignores ambiguity hedging. Alternatively, when $\varepsilon = 1$, she ignores p and focuses only on hedging; in this case, she is indifferent between an act f and the worst lottery f_S that f might yield.

[6]The function U is also a special case of Choquet expected utility with a capacity ν such that $\nu(E) = (1-\varepsilon)p(E)$ for all proper subsets $E \subset S$.

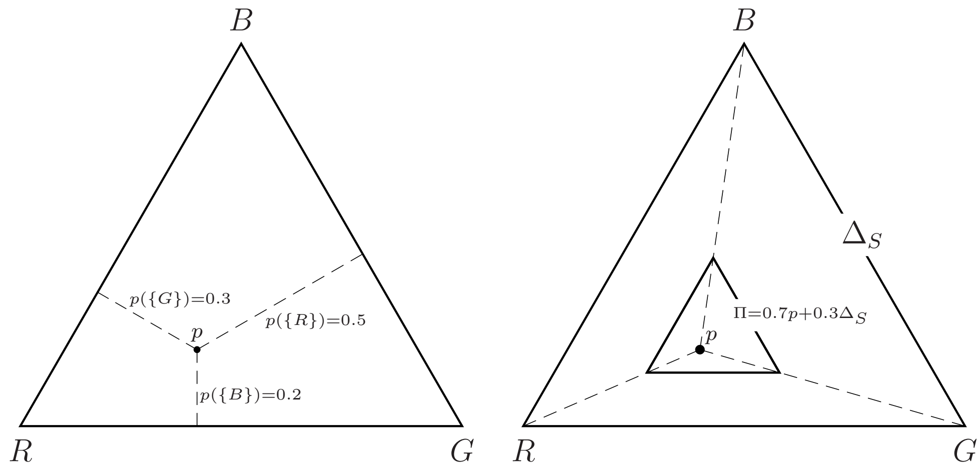

Figure 1: ε-contamination under complete ignorance

Affinity for Hedging and the associated utility representation (3.4) become problematic if some probabilistic scenarios in Δ_S are inconsitent with a priori information given to the decision maker.

Example 2. Recall the Ellsberg Paradox where $S = \{G, R, B\}$ and the decision maker is told only that the proportion of red balls in the urn is $\frac{1}{3}$. Then it is plausible that she ranks all bets on the event $\{R\}$ via expected utility, prefers to bet on $\{R\}$ rather than on $\{G\}$, but also on $\{G, B\}$ rather than on $\{R, B\}$. In this case, *any* maxmin expected utility representation for her preference must violate the ε-contamination structure $\Pi = (1-\varepsilon)p + \varepsilon\Delta_S$ because the set Π must be non-singleton, yet all priors in Π must agree on $\{R\}$.

Here in contrast with Example 1, the mixture $\frac{f_*+h_*}{2}$ guarantees the decision maker at least $\frac{1}{3}$ utils because it yields at least 0.5 utils on the event $\{G, B\}$ whose chance is known to be $\frac{2}{3}$. However, the mixture $\frac{f_*+h'_*}{2}$ does not guarantee more than 0.2 utils because it seems possible a priori that the urn contains no green balls. Thus the decision maker should conclude that the act h_* hedges f_* better than h'_* does. Affinity for such hedging could motivate her to violate the invariance (3.2).

To relax the assumption of complete ignorance, add another primitive to Anscombe–Aumann's tuple $(S, X, \succeq)$. Let $\Delta \subset \Delta_S$ be the set of all probability measures on S that agree with a priori information given to the decision maker. For example, let $\Delta = \{q \in \Delta_S : q(\{R\}) = \frac{1}{3}\}$ in the Ellsberg setting. Note that the decision maker does not choose Δ, and her preference $\succeq$ is still defined over acts in $\mathcal{H}$.

In general, we assume only that Δ is non-empty and closed, but not necessarily convex. For example, let $S = \prod_{i=1}^{n}\{G, R, B\}$ be a sequence space generated by n independent drawings from a three-color urn. Suppose that the decision maker is not told anything about the composition of this urn, except that this composition does not vary across drawings. Then Δ is the family of all Bernoulli distributions on the sequence state space S—this family is not convex. (Note that in order to derive the set Δ endogenously from the preference $\succeq$, we will assume in Section 4 that Δ has a mathematical structure that is even stronger than convexity.)

Next we adapt the axioms of Monotonicity and Affinity for Hedging for the given set Δ.

Axiom 7 (Monotonicity on Δ). *If acts f and g satisfy $f(q) \succeq g(q)$ for all $q \in \Delta$, then $f \succeq g$.*

This axiom states roughly that the decision maker should (i) compare any acts f and g only under probabilistic scenarios $q \in \Delta$ that are consistent with a priori information, and (ii) rank the induced lotteries $f(q)$ and $g(q)$ independently of the underlying scenario q. Note that if the preference $\succeq$ complies with expected utility on $\mathcal{L}$, then Monotonicity on Δ implies Monotonicity.

For any acts $f \in \mathcal{H}$, let f_Δ be the lottery that f induces under the most unfavorable scenario in Δ. Formally, $f_\Delta \in \mathcal{L}$ is such that $f(q) \succeq f_\Delta$ for all $q \in \Delta$, and $f(q') = f_\Delta$ for some $q' \in \Delta$. This lottery exists because Δ is closed, and $\succeq$ is continuous. Write $f \succeq_\Delta g$ or $f \succ_\Delta g$ if $f_\Delta \succeq g_\Delta$ or $f_\Delta \succ g_\Delta$ respectively.

Then for any acts f, h, h' such that $\frac{f+h}{2} \succ_\Delta \frac{f+h'}{2}$, the mixture $\frac{f+h}{2}$ guarantees a lottery $\left(\frac{f+h}{2}\right)_\Delta$ for any scenario $q \in \Delta$. In contrast, $\frac{f+h'}{2}$ induces a lottery $\left(\frac{f+h'}{2}\right)_\Delta$ worse than $\left(\frac{f+h}{2}\right)_\Delta$ under some scenario $q' \in \Delta$. Thus the decision maker should conclude that h *hedges* f on Δ better than h' does. This conclusion is no longer intuitive if $\frac{f+h'}{2} \succeq_\Delta \frac{f+h}{2}$.

Axiom 8 (Affinity for Hedging on Δ). *For all acts $f, g, h, h' \in \mathcal{H}$,*

$$\tfrac{f+h'}{2} \succeq_\Delta \tfrac{f+h}{2} \succeq \tfrac{g+h}{2} \succeq_\Delta \tfrac{g+h'}{2} \quad \Rightarrow \quad \tfrac{f+h'}{2} \succeq \tfrac{g+h'}{2}. \tag{3.5}$$

This axiom assert roughly that the decision maker likes to hedge the payoff that she receives under the most unfavorable scenario from the set Δ. Accordingly, she can motivate a reversal of preference from $\frac{f+h}{2} \succeq \frac{g+h}{2}$ to $\frac{f+h'}{2} \prec \frac{g+h'}{2}$, but only if h hedges f on Δ better than h' does, or if h' hedges g on Δ better than h does. Neither condition holds if $\frac{f+h'}{2} \succeq_\Delta \frac{f+h}{2}$ and $\frac{g+h}{2} \succeq_\Delta \frac{g+h'}{2}$.

Note that if Δ equals Δ_S or at least contains all of the extreme points of Δ_S, then the relations $\succeq_\Delta$ and $\succeq_S$ coincide, and the above axiom reduces to Affinity for Hedging. In general, the two axioms rely on different concepts of hedging and have no direct logical relation. Either of these conditions, together with Axioms 1–4, implies Convexity.

Given Axioms 1–4, Affinity for Hedging on Δ implies a more general invariance

$$\alpha f + (1-\alpha)h' \succeq_\Delta \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h \succeq_\Delta \alpha g + (1-\alpha)h' \quad \Rightarrow \\ \alpha f + (1-\alpha)h' \succeq \alpha g + (1-\alpha)h'$$

for all $\alpha \in [0,1]$. This invariance (which follows from Theorem 3.2 below) asserts that *all* reversals of preference from $\alpha f+(1-\alpha)h \succeq \alpha g+(1-\alpha)h$ to $\alpha f+(1-\alpha)h' \prec \alpha g + (1-\alpha)h'$ are due to affinity for ambiguity hedging.

Next we formulate our main representation result.

Theorem 3.2. $\succeq$ *satisfies Order, Continuity, Monotonicity on* Δ*, Certainty Independence, and Affinity for Hedging on* Δ *iff* $\succeq$ *is represented by*

$$U(f) = (1-\varepsilon)u\left(f(p)\right) + \varepsilon \min_{q\in\Delta} u\left(f(q)\right), \tag{3.6}$$

where $u : \mathcal{L} \to \mathbb{R}$ *is an affine index,* $\varepsilon \in [0,1]$*, and* $p \in conv\ \Delta$*.*

Moreover, if $\succeq$ *is non-degenerate and* Δ *is non-singleton, then* u *is unique up to a positive linear transformation,* ε *is unique, and* p *is unique whenever* $\varepsilon < 1$*.*

The above representation U is a special case of maxmin expected utility with the set of priors

$$\Pi = (1-\varepsilon)p + \varepsilon conv\ \Delta.$$

(Note that U need not comply with Choquet expected utility.) Thus, Theorem 3.2 refines the multiple priors model and delivers a parametric structure for the convex and closed set Π—this set is an ε-contamination of the probability measure $p \in conv\ \Delta$ by the convex hull $conv\ \Delta$ of probabilistic scenarios in Δ. In this structure, the set Δ is exogenous, but the components ε and p are uniquely derived from preference.

Again, one can interpret the functional form (3.6) in a way that does not use the set Π. The decision maker, as portrayed by (3.6), evaluates every act f via an ε-mixture of the lottery f_Δ that f induces under the least favorable scenario in Δ, and the lottery $f(p)$ that f induces under p. Here the measure $p \in conv\ \Delta$ reflects her probabilistic opinion, which is less rigid than a belief, and the parameter $\varepsilon \in [0,1]$ reflects her willingness to compromise this opinion by ambiguity hedging. Following Ellsberg [7], one could also interpret $1-\varepsilon$ as a degree of *confidence* that the decision maker has in the probabilistic opinion p.

Sketch of Proof. Theorem 3.1 is a corollary of Theorem 3.2 for $\Delta = \Delta_S$. The non-trivial part in the proof of Theorem 3.2 is to show that the axioms are sufficient for the existence of the utility representation (3.6). Note that all the axioms and the utility representation in Theorem 3.2 are unaffected if Δ is replaced by $conv\ \Delta$. Thus wlog assume that Δ is convex.

Let $a \cdot b$ denote the inner product of vectors a and b in the Euclidean space $\mathbb{R}^S$. Let $\vec{1} = (1,1,\ldots,1) \in \mathbb{R}^S$. Given an act $f \in \mathcal{H}$ and an affine function $u : \mathcal{L} \to \mathbb{R}$, let $u(f) \in \mathbb{R}^S$ denote the superposition of f and u. Note that for every $q \in \Delta$, $u(f(q)) = q \cdot u(f)$. We show sufficiency in three steps.

Step 1. Adapt the analysis in Gilboa–Schmeidler [14] to conclude that the preference $\succeq$ is represented by

$$U(f) = \min_{q\in\Pi} q \cdot u(f),$$

where $u : \mathcal{L} \to \mathbb{R}$ is affine and $\Pi \subset \Delta_S$ is convex and closed. Monotonicity on Δ and the separation theorem on $\mathbb{R}^S$ implies that $\Pi \subset \Delta$.

Step 2. For every $a \in \mathbb{R}^S$, let

$$V(a) = \min_{q \in \Delta} q \cdot a \quad \text{and} \quad W(a) = \min_{q \in \Pi} q \cdot a.$$

Derive from Affinity for Hedging on Δ that for all $a, b, c \in \mathbb{R}^S$ such that $V(a + c) \geq V(a)$ and $V(b + c) \leq V(b)$,

$$W(a) \geq W(b) \quad \Rightarrow \quad W(a + c) \geq W(b + c). \tag{3.7}$$

Let $\mathbb{D}$ be the set of all points $a \in \mathbb{R}^S$ at which the concave functions V and W are both differentiable. For every $a \in \mathbb{D}$, let

$$v(a) = \nabla V(a) \quad \text{and} \quad w(a) = \nabla W(a).$$

By [21, Theorem 25.5], the functions w and v are continuous, and the complement of the set $\mathbb{D}$ has measure zero; hence, the set $\mathbb{D}$ is dense in all of $\mathbb{R}^S$.

Show that for any $a, b \in \mathbb{D}$ such that $v(a) \neq v(b)$, there exists a unique $\varepsilon \geq 0$ such that

$$w(a) - w(b) = \varepsilon(v(a) - v(b)). \tag{3.8}$$

Assume that no such ε exists. Wlog $W(a) = W(b)$; otherwise, replace a by $a - W(a)\vec{1}$ and b by $b - W(b)\vec{1}$. Separate the open rays generated by the non-zero vectors $v(a) - v(b)$ and $w(a) - w(b)$: take $c \in \mathbb{R}^S$ such that

$$c \cdot (v(a) - v(b)) > 0 > c \cdot (w(a) - w(b)).$$

Wlog $c \cdot v(a) > 0 > c \cdot v(b)$; otherwise, replace c by $c - \frac{c \cdot v(a) + c \cdot v(b)}{2}\vec{1}$. Then for sufficiently small $\delta > 0$,

$$V(a + \delta c) > V(a), \quad V(b + \delta c) < V(b), \quad W(a + \delta c) - W(b + \delta c) < 0.$$

This contradicts (3.7).

Step 3. Show that ε in (3.8) is invariant of a and b. Take any $a_1, a_2, a_3 \in \mathbb{R}^S$ and let $w_i = w(a_i)$ and $v_i = v(a_i)$ for $i = 1 \dots 3$. Then the triangles $v_1v_2v_3$ and $w_1w_2w_3$ in Figure 2 are similar because each segment v_iv_j is parallel to the corresponding segment w_iw_j. It follows that the triangle $w_1w_2w_3$ is homothetic to $v_1v_2v_3$ with a unique coefficient $\varepsilon \geq 0$. Thus $\Pi = \varepsilon\Delta + (1 - \varepsilon)p$ for some $p \in \mathbb{R}^S$. Finally, as $\Pi \subset \Delta$, then $\varepsilon \leq 1$ and $p \in \mathit{conv}\ \Delta$ by the separation theorem in the Euclidean space $\mathbb{R}^S$.

4 Discussion

Suppose that the preference $\succeq$ satisfies the axioms in Theorem 3.2 and hence, has a utility representation (3.6) with components (u, ε, p). In this section we seek a

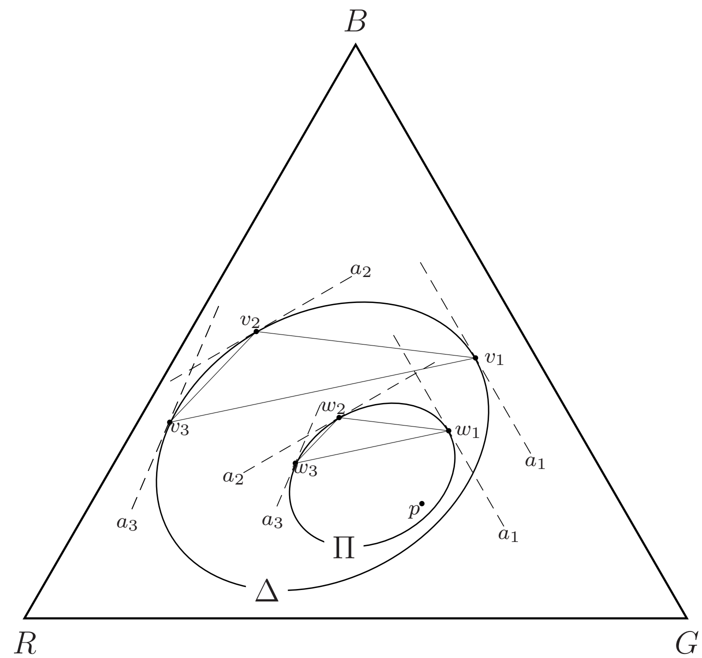

Figure 2: The existence of ε and p in Theorem 3.2

behavioral interpretation for these components. We have already noted that for u: this index represents the decision maker's risk attitude, that is, her ranking of lotteries. Turn to ε and p.

For any lotteries $l, l' \in \mathcal{L}$ and any event A, let lAl' denote a binary act that yields l for $s \in A$ and l' for $s \notin A$. Rewrite the utility representation over binary acts in a *biseparable* form (see Ghirardato and Marinacci [12, 13, 10]): for all $l \succeq l'$,

$$U(lAl') = u(l)\rho(A) + u(l')(1 - \rho(A)). \tag{4.1}$$

Here $\rho : 2^S \to [0, 1]$ is a capacity that reflects the decision maker's willingness to bet on events in S.

Next for each event A, let

$$\rho_\Delta(A) = \min_{q \in \Delta} q(A). \tag{4.2}$$

This function provides a priori lower bound for probabilities of events in S. It follows from (3.6) that for all events A

$$p(A) \geq \rho(A) = \varepsilon\rho_\Delta(A) + (1 - \varepsilon)p(A) \geq \rho_\Delta(A). \tag{4.3}$$

Then for all events A such that $\rho(A)+\rho(A^c) > \rho_\Delta(A)+\rho_\Delta(A^c)$

$$\varepsilon = \frac{1-\rho(A)-\rho(A^c)}{1-\rho_\Delta(A)-\rho_\Delta(A^c)} \tag{4.4}$$

$$p(A) = \frac{\rho(A)-\rho_\Delta(A)}{\rho(A)-\rho_\Delta(A)+\rho(A^c)-\rho_\Delta(A^c)}. \tag{4.5}$$

Thus the parameter ε is proportional to the gain in utility that the decision maker obtains by hedging a binary act lAl' by $l'Al$:

$$\begin{aligned} \tfrac{l+l'}{2} &= \tfrac{lAl'+l'Al}{2} \\ U\left(\tfrac{l+l'}{2}\right) - \tfrac{U(lAl')+U(l'Al)}{2} &= \tfrac{|u(l)-u(l')|}{2}(1-\rho(A)-\rho(A^c)). \end{aligned}$$

Moreover, one can use (4.1)–(4.5) to extend the results in Theorem 3.2 to the general setting with an infinite state space S. In order to do so, fix an algebra of events Σ in S, and let $\mathcal{H}$ be the set of all Σ-measurable acts that have a finite range in $\mathcal{L}$. Let Δ be a non-empty closed set of probability measures on (S,Σ). Suppose that the preference $\succeq$ over $\mathcal{H}$ satisfies the axioms in Theorem 3.2. Let $u:\mathcal{L}\to\mathbb{R}$ be an affine utility representation on $\mathcal{L}$, and construct $\varepsilon\in[0,1]$ and $p\in conv\ \Delta$ as follows.

Case 1. For all events $A\in\Sigma$, $\rho(A)=\rho_\Delta(A)$. Then for any finite subalgebra $\Sigma'\subset\Sigma$, Theorem 3.2 implies that the utility function $U(f)=\min_{q\in\Delta} f(q)$ represents the preference $\succeq$ restricted to Σ' measurable acts. Thus U represents $\succeq$ on all of $\mathcal{H}$.

Case 2. There exists an event $A\in\Sigma$ such that $\rho(A)>\rho_\Delta(A)$. Define $\varepsilon<1$ by (4.4). For every event $B\in\Sigma$, let

$$p(B) = \frac{\rho(B)-\rho_\Delta(B)}{\rho(B)-\rho_\Delta(B)+\rho(B^c)-\rho_\Delta(B^c)}$$

if $\rho(B)+\rho(B^c)>\rho_\Delta(B)+\rho_\Delta(B^c)$; let $p(B)=\rho(B)=\rho_\Delta(B)$ otherwise. Then for any finite subalgebra $\Sigma'\subset\Sigma$ such that $A\in\Sigma'$, Theorem 3.2 implies that p has a finitely additive restriction to Σ', and the triple (u,ε,p) represents the preference $\succeq$ restricted to Σ' measurable acts. It follows that p is finitely additive on all of Σ, and the tuple (u,ε,p) provides the required representation on all of $\mathcal{H}$.

Our main representation result Theorem 3.2 is formulated for a given set Δ of probabilistic scenarios. The exogenous specification of Δ is natural in some settings, but it can be problematic in others. Indeed, it may be unclear what information is available to the decision maker a priori, and which probabilistic scenarios "agree with" this information. To alleviate these concerns, one could attempt to define the class Δ in terms of preference. However, one cannot do so uniquely. In particular, if $\varepsilon<\varepsilon'\le 1$ and $\Delta'=\varepsilon'\Delta+(1-\varepsilon')p$, then

$$\varepsilon\Delta+(1-\varepsilon)p = \tfrac{\varepsilon}{\varepsilon'}\Delta' + \left(1-\tfrac{\varepsilon}{\varepsilon'}\right)p.$$

Therefore, the set Δ in Theorem 3.2 can be replaced by Δ'.

To identify the set Δ endogenously, impose a suitable maximality constraint on Δ. Require that for all $q, q' \in \Delta$ and $\gamma \in \mathbb{R}$, if $\gamma q + (1-\gamma)q' \in \Delta_S$, then $\gamma q + (1-\gamma)q' \in \Delta$. Call such Δ *superconvex*. This condition is stronger than convexity because the weight γ can vary beyond the interval $[0,1]$ as long as the linear combination $\gamma q + (1-\gamma)q'$ remains in the probability simplex.

For example, given a probability measure $q' \in \Delta_S$ and a subclass $\Gamma \subset 2^S$ of events, the set

$$\Delta(q', \Gamma) = \{q \in \Delta_S : q(A) = q'(A) \quad \text{for all } A \in \Gamma\}$$

is superconvex. This set is a natural candidate for Δ when objective probabilities are given but only for events in the subclass Γ. For example in the Ellsberg Paradox, $\Gamma = \{\{R\}, \{G, B\}\}$ and q' is such that $q'(\{R\}) = \frac{1}{3}$.

More generally, take a subclass $\Theta \subset 2^S \times 2^S$ of pairs of events. Then

$$\Delta(q', \Theta) = \{q \in \Delta_S : q(A \cap B)q'(B) = q'(A \cap B)q(B) \quad \text{for all } (A, B) \in \Theta\}$$

is superconvex. This set is a natural candidate for Δ when objective *conditional* probabilities are given but only for all pairs $(A, B) \in \Theta$. For example, suppose that $S = \{G, R, B\}$, and the decision maker is told only that the numbers G and B are equal. Then $\Theta = \{(\{G\}, \{G, B\}), (\{G\}, \{G, B\})\}$ and q' is such that $q'(\{G, B\}) = 2q'(\{B\})$.

Call an act $r \in \mathcal{H}$ *crisp* if for all $\alpha \in [0,1]$ and for all acts $f, g \in \mathcal{H}$,

$$f \succeq g \quad \Rightarrow \quad \alpha f + (1-\alpha)r \succeq \alpha g + (1-\alpha)r. \tag{4.6}$$

This definition is proposed by Kopylov [17] and by Ghiradato *et al.* [10], whose terminology we adopt.

Let $\mathcal{R}$ denote the class of all crisp acts r.

Corollary 4.1. *If Δ is non-singleton, closed, and superconvex set, and $\succeq$ has a utility representation* (3.6) *with $\varepsilon > 0$, then*

$$\Delta = \{q \in \Delta_S : \textit{for all } r \in \mathcal{R},\ r \sim r(q).\}$$

In this case, Δ can be derived from preference as the set of all probability measures that agree with an expected utility representation over crisp acts.

A APPENDIX: PROOF OF THEOREM 3.2

Let $\Delta \subset \Delta_S$ be a non-empty closed set. Note that all the axioms and the utility representation in Theorem 3.2 are unaffected if Δ is replaced by *conv* Δ. Thus wlog assume that Δ is convex.

For every $a \in \mathbb{R}^S$, let

$$V\left(a\right) = \min_{q \in \Delta} q \cdot a. \tag{A.1}$$

Then the function $V : \mathbb{R}^S \to \mathbb{R}$ is continuous, concave, homogeneous, and C-independent. The last two properties mean that for all vectors $a \in \mathbb{R}^S$, for all scalars $\alpha \geq 0$ and $\gamma \in \mathbb{R}$,

$$V\left(\alpha a + \gamma \vec{1}\right) = \alpha V\left(a\right) + \gamma. \tag{A.2}$$

Moreover, if an affine function $u : \mathcal{L} \to \mathbb{R}$ represents $\succeq$ on $\mathcal{L}$, then

$$u(f_\Delta) = \min_{q \in \Delta} u(f(q)) = \min_{q \in \Delta} q \cdot u(f) = V(u(f)) \tag{A.3}$$

for all acts $f \in \mathcal{H}$.

Show the necessity of the axioms in Theorem 3.2. Suppose that the preference $\succeq$ has a utility representation (3.6):

$$U(f) = (1-\varepsilon)p \cdot u(f) + \varepsilon V(u(f)),$$

where $u : \mathcal{L} \to \mathbb{R}$ is affine, $\varepsilon \in [0,1]$, and $p \in conv\ \Delta$. Then Order, Mixure Continuity and Certainty Independence are straighforward to show.

Show Monotonicity on Δ. Fix any acts f and g such that $f(q) \succeq g(q)$ for every $q \in \Delta$. Then $q \cdot u(f) \geq q \cdot u(g)$ for every $q \in \Delta$ and hence for every $q \in conv\ \Delta$. Thus $U(f) \geq U(g)$ because $p \cdot u(f) \geq p \cdot u(g))$ and $V(u(f)) \geq V(u(g))$.

Show Affinity for Hedging on Δ. Fix any acts $f, g, h, h' \in \mathcal{H}$ such that

$$\tfrac{f+h'}{2} \succeq_\Delta \tfrac{f+h}{2} \succeq \tfrac{g+h}{2} \succeq_\Delta \tfrac{g+h'}{2}.$$

By (A.3), $V\left(u\left(\frac{f+h'}{2}\right)\right) \geq V\left(u\left(\frac{f+h}{2}\right)\right)$ and $V\left(u\left(\frac{g+h}{2}\right)\right) \geq V\left(u\left(\frac{g+h'}{2}\right)\right)$. Then

$$U\left(\tfrac{f+h'}{2}\right) - U\left(\tfrac{g+h'}{2}\right) = (1-\varepsilon)p \cdot \left(u\left(\tfrac{f+h'}{2}\right) - u\left(\tfrac{g+h'}{2}\right)\right) +$$
$$\varepsilon\left(V\left(u\left(\tfrac{f+h'}{2}\right)\right) - V\left(u\left(\tfrac{g+h'}{2}\right)\right)\right) \geq (1-\varepsilon)p \cdot \left(u\left(\tfrac{f+h}{2}\right) - u\left(\tfrac{g+h}{2}\right)\right) +$$
$$\varepsilon\left(V\left(u\left(\tfrac{f+h}{2}\right)\right) - V\left(u\left(\tfrac{g+h}{2}\right)\right)\right) = U\left(\tfrac{f+h}{2}\right) - U\left(\tfrac{g+h}{2}\right) \geq 0.$$

Thus $\frac{f+h'}{2} \succeq \frac{g+h'}{2}$.

Turn to sufficiency. Suppose that Δ is convex (we drop this assumption later), and that $\succeq$ satisfies Order, Mixture Continuity, Monotonicity on Δ, Certainty Independence, and Affinity for Hedging on Δ. By the von Neumann–Morgenstern Theorem, there exists an affine utility index $u : \mathcal{L} \to \mathbb{R}$ that represents the preference $\succeq$ restricted to lotteries $l \in \mathcal{L}$. If u is constant, then by Monotonicity on Δ, $f \succeq g$ for all $f, g \in \mathcal{H}$. Wlog assume that u is non-constant, and its range covers $[-1, 1]$.

Gilboa–Schmeidler [14] show that for all $f, g \in \mathcal{H}$,

$$f \succeq g \quad \Leftrightarrow \quad W(u(f)) \geq W(u(g)), \tag{A.4}$$

where the function $W : \mathbb{R}^S \to \mathbb{R}$ is continuous, homogeneous and C-independent. The construction of such W requires only Order, Mixture Continuity, Monotonicity and Certainty Independence.

Lemma A.1. *The functions W and V satisfy the following properties:*

(i) For all $a, b \in \mathbb{R}^S$, if $q \cdot a \geq q \cdot b$ for every $q \in \Delta$, then $W(a) \geq W(b)$.

(ii) For all $a \in \mathbb{R}^S$, $W(a) \geq V(a)$.

(iii) For all $a, b, c \in \mathbb{R}^S$ such that $V(a+c) \geq V(a)$ and $V(b+c) \leq V(b)$,

$$W(a) \geq W(b) \quad \Rightarrow \quad W(a+c) \geq W(b+c). \tag{A.5}$$

(iv) There exists a unique convex and closed set $\Pi \subset \Delta$ such that

$$W(a) = \min_{q \in \Pi} q \cdot a \qquad \text{for all } a \in \mathbb{R}^S. \tag{A.6}$$

Proof. For any $a, b \in \mathbb{R}^S$, take $\alpha > 0$ such that $|\alpha a(s)|, |\alpha b(s)| \leq 1$ for all $s \in S$. As the range of u covers $[-1, 1]$, there exist acts $f, g \in \mathcal{H}$ such that $u(f) = \alpha a$ and $u(g) = \alpha b$. Suppose that for every $q \in \Delta$, $q \cdot a \geq q \cdot b$. Then for every $q \in \Delta$, $f(q) \succeq g(q)$ because $q \cdot u(f) = \alpha(q \cdot a) \geq \alpha(q \cdot b) = q \cdot u(g)$. By Monotonicity on Δ, $f \succeq g$. By (A.4) and (A.2), $W(a) \geq W(b)$.

For (ii), let $a \in \mathbb{R}^S$ and $b = V(a)\vec{1}$. By (A.1), $q \cdot a \geq V(a) = q \cdot b$ for all $q \in \Delta$. By (i), $W(a) \geq W(b)$, that is, $W(a) \geq V(a)$.

Turn to (iii). For any $a, b, c \in \mathbb{R}^S$, take $\alpha > 0$ and acts $f, g, h, h' \in \mathcal{H}$ such that $u(f) = \alpha a$, $u(g) = \alpha b$, $u(h) = 0$, and $u(h') = \alpha c$. Then

$$\left.\begin{array}{lll}
V(a+c) \geq V(a) & \Rightarrow V\left(\frac{u(f)+u(h')}{2}\right) \geq V\left(\frac{u(f)+u(h)}{2}\right) & \Rightarrow_1 \frac{f+h'}{2} \succeq_\Delta \frac{f+h}{2} \\
W(a) \geq W(b) & \Rightarrow W\left(\frac{u(f)+u(h)}{2}\right) \geq W\left(\frac{u(g)+u(h)}{2}\right) & \Rightarrow_2 \frac{f+h}{2} \succeq \frac{g+h}{2} \\
V(b) \geq V(b+c) & \Rightarrow V\left(\frac{u(g)+u(h)}{2}\right) \geq V\left(\frac{u(g)+u(h')}{2}\right) & \Rightarrow_1 \frac{g+h}{2} \succeq_\Delta \frac{g+h'}{2}
\end{array}\right\} \Rightarrow_3$$

$$\frac{f+h'}{2} \succeq \frac{g+h'}{2} \Rightarrow_2 W\left(\frac{u(f)+u(h')}{2}\right) \geq W\left(\frac{u(g)+u(h')}{2}\right) \Rightarrow W(a+c) \geq W(b+c),$$

where $\Rightarrow_1$ follow from (A.3), $\Rightarrow_2$ from (A.4), $\Rightarrow_3$ from Affinity for Hedging on Δ, and all other implications from the homogeneity of V and W.

Turn to (iv). Show that W is concave. For any $\alpha \in [0,1]$ and $a, b \in \mathbb{R}^S$, let $a' = \alpha(a - W(a)\vec{1})$, $b' = \alpha(b - W(b)\vec{1})$, and $c = (1-\alpha)\left(b - V(b)\vec{1}\right)$. Then

$$V(a'+c) \geq V(a') + V(c) = V(a')$$
$$V(b'+c) = V\left(b - \alpha W(b)\vec{1} - (1-\alpha)V(b)\vec{1}\right) = \alpha(V(b) - W(b)) = V(b'),$$

and $W(a') = W(b') = 0$. By (iv), $W(a'+c) \geq W(b'+c)$, that is,

$$W(\alpha a + (1-\alpha)b) - \alpha W(a) - (1-\alpha)V(b) \geq W(b) - \alpha W(b) - (1-\alpha)V(b).$$

Thus $W(\alpha a + (1-\alpha)b) \geq \alpha W(a) + (1-\alpha)W(b)$.

The function W satisfies all the conditions of [14, Lemma 3.3], and hence, there exists a unique convex and closed set $\Pi \subset \Delta_S$ such that $W(a) = \min_{q \in \Pi} q \cdot a$ for all $a \in \mathbb{R}^S$. Show that $\Pi \subset \Delta$. Suppose that there exists $q' \in \Pi \setminus \Delta$. Then by the separation theorem, there exists $a \in \mathbb{R}^S$ such that for all $q \in \Delta$, $q \cdot a > 0 > q' \cdot a$. Then $V(a) > 0 > W(a)$, which contradicts (ii). $\square$

Let $\mathbb{D}$ be the set of all points $a \in \mathbb{R}^S$ at which the concave functions V and W are both differentiable. For every $a \in \mathbb{D}$, let

$$v(a) = \nabla V(a) \quad \text{and} \quad w(a) = \nabla W(a).$$

By [21, Theorem 25.5], the functions $v, w : \mathbb{D} \to \mathbb{R}^S$ are continuous, and the complement of the set $\mathbb{D}$ has measure zero; hence, the set $\mathbb{D}$ is dense.

Lemma A.2. *The functions v and w have the following properties:*

(i) For all $a \in \mathbb{D}$ and $q \in \Delta$, $q = v(a)$ iff $V(a) = q \cdot a$.

(ii) For all $a \in \mathbb{D}$ and $q \in \Pi$, $q = w(a)$ iff $W(a) = q \cdot a$.

(iii) For all $a \in \mathbb{D}$, $\alpha > 0$ and $\gamma \in \mathbb{R}$,

$$\alpha a + \gamma\vec{1} \in \mathbb{D}, \quad v(a) = v(\alpha a + \gamma\vec{1}), \quad w(a) = w(\alpha a + \gamma\vec{1}).$$

(iv) For any $a, b \in \mathbb{D}$ such that $v(a) \neq v(b)$, there exists $\varepsilon \geq 0$ such that

$$w(a) - w(b) = \varepsilon(v(a) - v(b)). \tag{A.7}$$

Proof. Fix $a \in \mathbb{D}$ and $q \in \Delta$ such that $V(a) = q \cdot a$. For all $b \in \mathbb{R}^S$ and $\delta \in \mathbb{R}$,

$$V(a) + \delta(q \cdot b) = q \cdot (a + \delta b) \geq V(a + \delta b) = V(a) + \delta(v(a) \cdot b) + o(\delta).$$

Then $q \cdot b = v(a) \cdot b$ for all $b \in \mathbb{R}^S$, that is, $q = v(a)$. Similarly for (ii).

Next, for any $a \in \mathbb{D}$, $\alpha > 0$ and $\gamma \in \mathbb{R}$, write the function $V(b)$ as a superposition $\alpha V\left(\frac{b-\gamma\vec{1}}{\alpha}\right) + \gamma$. By the chain rule, $V(b)$ is differentiable at $b = \alpha a + \gamma\vec{1}$, and $v(\alpha a + \gamma\vec{1}) = v(a)$. Similarly, for W and w.

Prove (iv) by contradiction. Suppose that there exist $a, b \in \mathbb{D}$ such that $v(a) \neq v(b)$, and for all $\varepsilon \geq 0$,

$$w(a) - w(b) \neq \varepsilon(v(a) - v(b)).$$

Wlog $W(a) = W(b)$; otherwise, replace a by $a - W(a)\vec{1}$ and b by $b - W(b)\vec{1}$. As $v(a) - v(b) \neq 0$, then $w(a) - w(b) \neq 0$, and the open rays generated by the non-zero vectors $v(a) - v(b)$ and $w(a) - w(b)$ are disjoint. By the separation theorem, there exists $c \in \mathbb{R}^S$ such that

$$c \cdot (v(a) - v(b)) > 0 > c \cdot (w(a) - w(b)).$$

Wlog $c \cdot v(a) > 0 > c \cdot v(b)$; otherwise, replace c by $c - \frac{c \cdot v(a) + c \cdot v(b)}{2}\vec{1}$. Take a sufficiently small $\delta > 0$ so that

$$\begin{aligned}
&V(a + \delta c) = V(a) + \delta c \cdot v(a) + o(\delta) > V(a) \\
&V(b + \delta c) = V(b) + \delta c \cdot v(b) + o(\delta) < V(b) \\
&W(a + \delta c) - W(b + \delta c) = \delta c \cdot (w(a) - w(b)) + o(\delta) < 0.
\end{aligned}$$

This is a contradiction with (A.5) because $W(a) = W(b)$. □

Lemma A.3. *There exists $\varepsilon \in [0, 1]$ and $p \in \Delta$ such that*

$$\Pi = \varepsilon\Delta + (1 - \varepsilon)p.$$

Moreover, if Δ is non-singleton, then ε is unique, and if $\varepsilon < 1$, then p is also unique.

Proof. Consider three cases.

Case 1. The function v is constant on $\mathbb{D}$. Take $p \in \Delta$ such that $v(a) = p$ and $V(a) = p \cdot a$ for all $a \in \mathbb{D}$. By continuity, $V(a) = p \cdot a$ for all $a \in \mathbb{R}^S$. Thus, both $\Delta = \{p\}$ and $\Pi \subset \Delta$ are singletons. In this case, $\varepsilon \in [0, 1]$ is arbitrary.

Case 2. The range of the function v consists of two distinct points $v_1, v_2 \in \Delta$. Then $V(a) = \min\{v_1 \cdot a, v_2 \cdot a\}$ for all $a \in \mathbb{D}$ and by continuity, for all $a \in \mathbb{R}^S$. Therefore, Δ is a segment with end points v_1 and v_2. Π is a closed and convex subset of Δ, and hence, a segment with end points $\alpha v_1 + (1-\alpha)v_2$ and $\gamma v_1 + (1-\gamma)v_2$ for some $\alpha \leq \gamma$. Let $\varepsilon = \gamma - \alpha$. If $\varepsilon = 1$, then $\Pi = \Delta$. If $\varepsilon < 1$, then $\Pi = \varepsilon\Delta + (1 - \varepsilon)p$, where $p = \frac{\alpha}{1-\varepsilon}v_1 + \frac{1-\gamma}{1-\varepsilon}v_2$ is determined uniquely in the representation (3.6).

Case 3. There exist $a_1, a_2, a_3 \in \mathbb{D}$ such that the points $v(a_1), v(a_2), v(a_3) \in \Delta$ are distinct. Let $v_i = v(a_i)$ and $w_i = w(a_i)$ for $i = 1, 2, 3$. By (A.7), there exist unique $\varepsilon_{12}, \varepsilon_{23}, \varepsilon_{31} \geq 0$ such that

$$\begin{aligned} w_1 - w_2 &= \varepsilon_{12}(v_1 - v_2) \\ w_2 - w_3 &= \varepsilon_{23}(v_2 - v_3) \\ w_3 - w_1 &= \varepsilon_{31}(v_3 - v_1). \end{aligned}$$

By Lemma A.2, $v_1 \cdot a_1 < v_2 \cdot a_2$ and $v_1 \cdot a_1 < v_3 \cdot a_3$, which implies that $v_1 \notin [v_2, v_3]$. Similarly, $v_2 \notin [v_1, v_3]$ and $v_3 \notin [v_1, v_2]$. Thus, v_1, v_2, v_3 do not lie on the same line and form a triangle. Thus the equality

$$\varepsilon_{12}(v_1 - v_2) + \varepsilon_{23}(v_2 - v_3) + \varepsilon_{31}(v_3 - v_1) = 0$$

implies $\varepsilon_{12} = \varepsilon_{23} = \varepsilon_{31}$. Let $\varepsilon = \varepsilon_{12}$ and

$$\hat{p} = w_1 - \varepsilon v_1 = w_2 - \varepsilon v_2 = w_3 - \varepsilon v_3.$$

Fix any $b \in \mathbb{D}$. Then wlog $v(b) \neq v_1$ and $v(b) \neq v_2$, and hence, $\hat{p} = w(b) - \varepsilon v(b)$. It follows that

$$W(b) = w(b) \cdot b = \varepsilon v(b) \cdot b + \hat{p} \cdot b = \varepsilon V(b) + \hat{p} \cdot b.$$

By continuity, $W(b) = V(b) + \hat{p} \cdot b$ for all $b \in \mathbb{R}^S$, that is,

$$\min_{q \in \Pi} q \cdot b = \min_{q \in \varepsilon\Delta + \hat{p}} q \cdot b.$$

By the separation theorem, $\Pi = \varepsilon\Delta + \hat{p}$. From $\Pi \subset \Delta$ it follows that $\varepsilon \in [0, 1]$. If $\varepsilon = 1$, then $\Pi = \Delta$, and p in the representation (3.6) is arbitrary. If $\varepsilon < 1$, then $p = \frac{1}{1-\varepsilon}\hat{p}$ is determined uniquely.

□

References

[1] F. Anscombe and R. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

[2] J. Berger and M. Berliner. Robust Bayes and empirical analysis with epsilon contaminated priors. *The Annals of Statistics*, 14:461–486, 1986.

[3] P. Bickel. Parametric robustness or small biases can be worthwhile. *The Annals of Statistics*, 12:864–879, 1984.

[4] J. Blume and J. Rosenblatt. On partial a priori information in statistical inference. *The Annals of Mathematical Statistics*, 38:1671–1678, 1967.

[5] R. Casadeus-Masanell, P. Klibanoff, and E. Ozdenoren. Maxmin expected utility over Savage acts with a set of priors. *Journal of Economic Theory*, 92:33–65, 2000.

[6] E. Damiano. Choice under limited uncertainty. Working Paper, Yale University, 1999.

[7] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quaterly Journal of Economics*, 75:643–669, 1961.

[8] L. Epstein and T. Wang. Intertemporal asset pricing and Knightian uncertainty. *Econometrica*, 62:283–382, 1994.

[9] T. Gajdos, J.-M. Tallon, and J.-C. Vergnaud. Decision making with imprecise probabilistic information. *Journal of Mathematical Economics*, 40(6):647–681, 2004. available at http://ideas.repec.org/a/eee/mateco/v40y2004i6p647-681.html.

[10] P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118:133–173, 2004.

[11] P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi. A subjective spin on roulette wheels. *Econometrica*, 71:1897–1908, 2003.

[12] P. Ghirardato and M. Marinacci. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research*, 26:864–890, 2001.

[13] P. Ghirardato and M. Marinacci. Ambiguity mase precise: A comparative foundation. *Journal of Economic Theory*, 102:251–289, 2002.

[14] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[15] T. Hayashi. Information, subjective belief and preference. Working Paper, University of Texas, 2005.

[16] J. L. Hodges and E. L. Lehmann. The use of previous experience in reaching statistical decisions. *Annals of Mathematical Statistics*, 23:396–407, 1952.

[17] I. Kopylov. Procedural rationality in the multiple priors model. Working Paper, University of Rochester, 2001.

[18] K. Nishimura and H. Ozaki. An axiomatic approach to epsilon contamination. Working Paper, University of Tokyo, 2003.

[19] K. Nishimura and H. Ozaki. Search and Knightian uncertainty. *Journal of Economic Theory*, 119:299–333, 2004.

[20] W. Olszewski. Preferences over sets of lotteries. Working Paper, Northwestern University, 2005.

[21] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

[22] L. J. Savage. *The Foundations of Statistics*. Dover Publications Inc., New York, second revised edition, 1972.

[23] L. Wasserman and J. Kadane. Bayes' Theorem for Choquet capacities. *The Annals of Statistics*, 18:1328–1339, 1990.