Ordering Ambiguous Acts[1]

Ian Jewitt
Nuffield College, Oxford University.

Sujoy Mukerji
Department of Economics and University College, Oxford University

June 14, 2011

Abstract

We investigate what it means for one act to be more ambiguous than another. The question is evidently analogous to asking what makes one prospect riskier than another, but beliefs are neither objective nor representable by a unique probability. Our starting point is an abstract class of preferences constructed to be (strictly) partially ordered by a more ambiguity averse relation. We define two notions of more ambiguous with respect to such a class. A more ambiguous (I) act makes an ambiguity averse decision maker (DM) worse off but does not affect the welfare of an ambiguity neutral DM. A more ambiguous (II) act adversely affects a more ambiguity averse DM more, as measured by the compensation they require to switch acts. Unlike more ambiguous (I), more ambiguous (II) does not require indifference of ambiguity neutral elements to the acts being compared. Second, we implement the abstract definitions to characterize more ambiguous (I) and (II) for two explicit preference families: α-maxmin expected utility and smooth ambiguity. Our characterizations show that (the outcome of) a more ambiguous act is less robust to a perturbation in probability distribution governing the states. Third, the characterizations also establish important connections between more ambiguous and more informative as defined on statistical experiments by Blackwell (1953) and others. Fourthly, we give applications to defining ambiguity "in the small" and to the comparative statics of more ambiguous in a standard portfolio problem and a consumption-saving problem.

JEL Classification Numbers: C44, D800, D810, G11

Keywords: Ambiguity, Uncertainty, Knightian Uncertainty, Ambiguity Aversion, Uncertainty aversion, Ellsberg paradox, Comparative statics, Single-crossing, More ambiguous, Portfolio choice, More informative, Information, Garbling.

1   Introduction

Consider a decision maker (DM) choosing among *acts*, choices with contingent conse-
quences. Following intuitive arguments of Knight (1921) and Ellsberg (1961), pioneering
formalizations by Schmeidler (1989) and Gilboa and Schmeidler (1989), and a body of
subsequent work, modern decision theory distinguishes two categories of subjectively un-
certain belief: *unambiguous* and *ambiguous*. An ambiguous belief cannot be expressed
using a single probability distribution. Intuitively, an event is deemed (subjectively)
ambiguous if the DM's belief about the event, as revealed by his preferences, cannot be
expressed as a unique probability.[1] The usual interpretation is that the DM is uncertain
about the 'true' probability of the ambiguous event (and takes this uncertainty into ac-
count when making his choice). A DM considers an act to be unambiguous if, for each
set of consequences, its inverse image is unambiguous. Otherwise, the act is ambiguous.
In this paper we investigate what makes one act *more ambiguous* than another.

One focus of the recent literature applying ideas of ambiguity to economic contexts,
finance and macroeconomics in particular, is on how equilibrium trade in financial as-
sets is affected when agents seek assets that are, in a sense, 'robust' to the perceived
ambiguity. A comparative static question of interest in such models is, naturally, that
of more ambiguous.[2] We need concepts of more ambiguous just as concepts of orders of
riskiness were needed to facilitate comparative statics of 'more risky'. One challenge in
formulating a general definition of more ambiguous, in keeping with revealed preference
traditions, is that the definition should be preference based but not tied down to par-
ticular parametric preference forms. Following from the question of definition, we wish
to identify what structural properties make one act more ambiguous than another and
how this varies according to the class of preferences one considers.

Two key ideas give us two distinct ways of revealing (via choice behavior) whether
an act is relatively more affected by ambiguity than another act, thereby giving rise to
two (generally) distinct orders of *more ambiguous* on the space of acts. Our definition
of *more ambiguous (I)* says, essentially, that the more ambiguous act is less attractive
to ambiguity averse DMs but not to DMs with preferences neutral to ambiguity. *More
ambiguous (II)* says that act f is more ambiguous than act g if the more ambiguity
averse agent requires more compensation to give up g for f. In other words, the relative
cost of taking on the more ambiguous act, i.e., going from g to f, is costlier for the
more ambiguous type of agent. What lies at the heart of more ambiguous (II) is a
single-crossing notion, suitably strengthened to ensure that transitivity is respected.
The advantage of the first definition is it allows us to identify acts which are separated
purely and solely in terms how much they are affected by ambiguity. An advantage
of the second definition is it allows us to compare acts which are differently affected by
ambiguity, while possibly being different in other dimensions. Note, the common element

---

[1]There is an extensive literature discussing the definition of ambiguous events, e.g., Epstein (1999),
Ghirardato and Marinacci (2002), Nehring (2001) and Klibanoff, Marinacci and Mukerji (2005).

[2]See, e.g., Hansen (2007), Caballero and Krishnamurthy (2008), Epstein and Schneider (2008), Hansen
and Sargent (2010), Uhlig (2010), Boyle, Garlappi, Uppal, and Wang (2011), Collard, Mukerji, Tallon,
and Sheppard (2011), Gollier (2011).

in the definitions: in both instances the order of more ambiguous arises on the back of preferences, more specifically, on a relation on preferences. In the first definition, we compare the choice made by an ambiguity neutral preference with that by a ambiguity averse preference; in the second definition, we compare the choice made by one preference with another which is more ambiguity averse. In this way, the defining properties are universal across preferences. So, fixing a preference class, partially ordered by a more ambiguity averse relation, we may apply these properties to determine whether *that* class deems an act to be more ambiguous than another act.

We also study the case of *events*. Bets on events are acts with binary outcomes and the two notions of more ambiguous acts may be extended, with appropriate qualifications, to very analogous notions of *more ambiguous events*. Conceptually, these notions take forward the literature on definitions of ambiguous events. They are of interest in applications too: for instance, when investigating the effect of ambiguity on contingent contracts, it might be natural to want to compare contingent arrangements across more ambiguous events.

Next, the abstract definitions are implemented to characterize more ambiguous (I) and (II) for some classes of preferences prominent in applications. Two classes of preferences we investigate in particular are, the class of α-*maxmin expected utility* preferences (α-MEU) and *smooth ambiguity* preferences. The α-MEU class, generalize the well known maxmin expected utility preferences due to Gilboa and Schmeidler (1989). For these preferences the decision maker's belief about relevant stochastic environments[3] is represented by a convex, compact set of probabilities on the state space, with acts being evaluated by a weighted average of the maximum and minimum expected utility ranging over the set of probabilities. For smooth ambiguity preferences, decision maker's beliefs about relevant probabilistic environments are represented by a set of probabilities on the state space along with a second-order prior over them.[4]

To get a first idea of the nature and style of the characterizations we obtain we focus in this introductory section on the case of events. Fix a convex, compact set of probabilities Π on a state space $\boldsymbol{S}$ and consider the associated class of α-MEU preferences, α ranging over the interval $[0, 1]$. Given an event $E \subset \boldsymbol{S}$, since Π is compact convex, the set of points $\pi(E) \in [0, 1]$ as π ranges over Π is a closed interval which we denote as $\Pi(E) = \{\pi(E) \mid \pi \in \Pi\} \subset [0, 1]$. We show this preference class considers an event E to be more ambiguous (I) than event E' if and only if $\Pi(E) \supset \Pi(E')$ and the probability intervals are such that they share the same center. Analogous to the centered expansion in the case of α-MEU preferences, for smooth ambiguity the characterizing condition requires that the μ-average of the event probabilities $\pi_i(E)$ is retained (μ being the second order prior) while the $\pi(E')$ are all contained in the convex hull of the $\pi(E)$, $\pi \in \Pi$. Since the more ambiguous (II) notion does away with the requirement of the ambiguity neutral preference, it is intuitive that the characterization of more ambiguous (II) for α-MEU is, essentially, that $\Pi(E)$ is more spread out than $\Pi(E')$, without the requirement of a common center. For smooth ambiguity preferences, the characterization

---

[3] For a discussion of *relevance*, in the sense used here, see Klibanoff, Mukerji, and Seo (2011).

[4] See Section 2.3 for more details on these preference classes, including references.

2

is analogous: two equal μ-measure journeys in the support of μ, one tracking the variation in the probability of E and the other of E', will show the probability of the more ambiguous(II) event E to vary more.

The connection between ordering acts by ambiguity and experiments by information is one of the central themes of this paper and one that we use to formally interpret various characterizations. Consider, again, the case of events. Adopting the language of statistical decision theory, the probability distribution over the *sample space* $\boldsymbol{S}$ is determined by $\pi \in \Pi$, which naturally corresponds to the *parameter space*. In this sense, an event E constitutes a statistical *experiment*, a *statistic* defined on $\boldsymbol{S}$, whose outcome, the occurrence or nonoccurrence of E, may reveal information about the 'true' underlying parameter, π. Intuitively, if $\pi(E)$ does not depend on $\pi \in \Pi$, i.e., $\pi(E) = \pi'(E)$, $\pi, \pi' \in \Pi$, then the event E would be clearly deemed unambiguous by any preference with associated belief in Π. Just as clearly, observing an occurrence or nonoccurrence of E is completely *uninformative* about which distribution $\pi \in \Pi$ actually obtained. These observations equate uninformative with unambiguous in what appears to be a very compelling way, so it is very intuitive that more informative should remain central to the characterization of more ambiguous. Notions of more informative allow us to formally articulate the natural intuition about what is peculiar to the structure of a more ambiguous act: the (probability of) its outcomes are affected more when the probability distribution on the state space is perturbed.[5] There are nuances to the way more ambiguous acts are less robust, depending on the version of more ambiguous and class of preferences under consideration. Notions of more informative are useful in clarifying these nuances.

Some of the characterizations are obtained under a condition imposed on preferences which restricts the nature of associated beliefs. The general condition is *U-comonotonicity*; and, in the special case of events, *event-comonoticity*. A set of probabilities on the state space, Π, is *event-comonotone* for a pair of events E, E' if for all $\pi_1, \pi_2 \in \Pi$, $(\pi_1(E) - \pi_2(E))(\pi_1(E') - \pi_2(E')) \geq 0$. In words, if one of two stochastic environments subjectively thought relevant is better for E then it also better for E'; the events order the relevant stochastic environments in the same way. Evidently, this condition gives a sense in which two events are (stochastically) similar, e.g., a bet on the S&P being less than 11000 at close on a particular day and an analogous bet on the FTSE, but not a bet on a stock market index and a bet on the outcome of a boxing match. The condition is shown to have quite striking implications for the characterizations. For example, the characterizing condition for E being more ambiguous (II) than E', for the class of α-MEU preferences associated with Π and *also* for the class of smooth ambiguity preferences with $\text{supp}(\mu) \subset \Pi$ is essentially, that $\Pi(E)$ is more spread out than $\Pi(E')$. Hence, remarkably, in this case the characterizing conditions for the α-MEU class and the smooth ambiguity class are virtually identical and, with respect to smooth ambiguity preferences, all that matters about the second order prior

---

[5]The state space is an objective construct, as is the mapping describing an act. Hence, how the distribution on outcomes induced by an act and a distribution governing states changes, following a putative change in the governing distribution from π to π', say, is a structural property.

is its support. Furthermore, U-comonotonicity generalizes the result, in a natural way, for the case of acts. Hence, event-comonotonicity (and U-comonotonicity) are important instances in which structural properties distinguishing more ambiguous (II) *do not* vary across a quite wide range of preferences. For an illustration, consider the following example. Let $\boldsymbol{S} = \{r, b, g\}$, and

$$\Pi = \left\{\pi = \lambda P + (1 - \lambda)Q \mid P = (0.2, 0.35, 0.45), P = (0.3, 0.6, 0.1), \lambda \in [0, 1]\right\}.$$

Note, the events $R = \{r\}$ and $B = \{b\}$ are event-comonotone. Further, the interval $\Pi(B)$ is wider than $\Pi(R)$ and without an overlap. Take two classes of preferences, α-MEU and smooth ambiguity, such that the belief associated with α-MEU preferences is the set Π and for the smooth ambiguity preferences, the support of the second order prior μ is a subset of Π. By Proposition 3.5 both classes would deem B as a more ambiguous (II) event than R. Furthermore, the same Proposition shows B is *Blackwell (pairwise) more informative* than R for each dichotomy $\{\pi_1, \pi_2\} \subset \Pi$.

Finally we turn to some applications. As a first application, we use the idea of more ambiguous (I) to identify the ambiguity premium for an act and develop measures of ambiguity based on an approximate formula for the *ambiguity premium* of a small (ambiguous) gamble. Next, we illustrate comparative statics of more ambiguous, (I) and (II). First, we analyze the standard portfolio choice problem with one safe and one uncertain asset and consider the comparative static effect on the optimal weight when the uncertain asset is replaced another which more ambiguous (I). We identify conditions that yield the 'expected' comparative static for the α-MEU case and for the smooth ambiguity case. Secondly, we analyze an optimal saving problem, for α-MEU and smooth ambiguity preferences, in which future income is ambiguous. We explore the impact on savings as future income becomes more ambiguous (II).

The literature on more ambiguous is rather spare. Segal (1987) analyzes preferences over binary acts, e.g., $(x, E; 0, \neg E)$, where you win x if the event E occurs, 0 otherwise. It is assumed that the ambiguity concerning the probability of E in the "ambiguous lottery" $(x, E; 0, \neg E)$ is represented by a probability distribution F^* on $[0, 1]$ governing the probability that E occurs. It is then suggested that to rank "degrees of ambiguity", one should define an order on the set of the distribution functions F^*. Segal considers but rejects the criterion that F^* be riskier than G^* in the sense of Rothschild and Stiglitz (1970) in favour of a more restrictive relation, that F^* crosses G^* only at their common mean from below. Segal writes, referring to an ambiguity averse DM, "one is tempted to assume that if G^* is more ambiguous than F^*, then the value of $(x, E; 0, \neg E)$ under F^* is greater than its value under G^*," but shows that this is not generally true. Segal's counterexample naturally leads one to think of preferences as the starting point for primitive notions of more ambiguous, so it can be seen as an inspiration for the current paper. The analysis in Grant and Quiggin (2005) is also related, but less so. It proceeds in a direction opposite to the one taken in this paper: starting with a primitve notion of a more uncertain act it goes on to characterize corresponding dual notions of *more uncertainty averse* for various preference models. Also, they do not distinguish between ambiguity and risk.

The paper is organized as follows. In Section 2 we first present the definitions of more ambiguous acts and events in pure decision theoretic terms and then introduce particular concepts that are invoked in the characterization results: parametric preference families, the order restrictions on beliefs imposed by comonotonicity ideas, and information orders. Section 3 implements the definitions to characterize more ambiguous events, while Section 4 does the same for more ambiguous acts. Section 5 presents the applications and Section 6 concludes.

2 Decision theoretic considerations

2.1 Preliminaries

Let $\boldsymbol{X}$ be a compact subinterval of $\mathbb{R}$ and $\boldsymbol{L}$ the set of distributions over $\boldsymbol{X}$ with finite supports:

$$\boldsymbol{L} = \left\{ l : \boldsymbol{X} \to [0,1] \mid l(x) \neq 0 \text{ for finitely many } x\text{'s in } \boldsymbol{X} \text{ and } \sum_{x \in \boldsymbol{X}} l(x) = 1 \right\}.$$

Let $\boldsymbol{S}$ be a separable metric space and let $\boldsymbol{\Sigma}$ be an algebra of subsets of $\boldsymbol{S}$. Denote by $\mathcal{F}_0$ the set of all $\boldsymbol{\Sigma}$-measurable finite valued functions from $\boldsymbol{S}$ to $\boldsymbol{L}$. Let $\mathcal{F}$ be a convex subset of $\boldsymbol{L}^{\boldsymbol{S}}$ which includes all constant functions in $\mathcal{F}_0$. In the usual decision theoretic nomenclature, elements of $\boldsymbol{X}$ are (deterministic) *outcomes*, elements of $\boldsymbol{L}$ are *lotteries*, elements of $\boldsymbol{S}$ are *states* and elements of $\boldsymbol{\Sigma}$ are *events*. Elements of $\mathcal{F}$ are *acts* whose state contingent *consequences* are elements of $\boldsymbol{L}$. Hence, given $f \in \mathcal{F}$ and $s \in \boldsymbol{S}$, $f(s)$ is a (finitely supported) probability distribution on $\boldsymbol{X}$ while $f(s)(x)$ denotes the probability of $x \in \boldsymbol{X}$ under $f(s)$. As usual, we may think of an element of $\boldsymbol{L}$ as a *constant act*, i.e., an act with the same consequence in every state. Given an $x \in \boldsymbol{X}$, $\delta_x \in \boldsymbol{L}$ denotes a degenerate lottery such that $\delta_x(x) = 1$. Let $\pi : \boldsymbol{\Sigma} \to [0,1]$ be a countably additive probability. The set of all such probabilities, π, is denoted by Δ. Let $C(\boldsymbol{S})$ be the set of all continuous and bounded real-valued functions on $\boldsymbol{S}$. Using $C(\boldsymbol{S})$ we equip Δ with the vague topology, that is, the coarsest topology on Δ that makes the following functionals continuous:

$$\pi \mapsto \int \psi d\pi \qquad \text{for each } \psi \in C(\boldsymbol{S}) \text{ and } \pi \in \Delta.$$

Let $\mathcal{B}_\Delta$ denote the Borel σ-algebra on Δ generated by the vague topology. Given $\pi \in \Delta$, any act $f \in \mathcal{F}$ induces a corresponding lottery, a probability distribution over outcomes conditional on π. To define this formally, denote by $\mathcal{B}_{\boldsymbol{X}}$ the Borel σ-algebra of $\boldsymbol{X}$ and, for the act f, define the Markov kernel $(\pi, B) \mapsto P_\pi^f(B)$ from $(\Delta, \mathcal{B}_\Delta)$ to $(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}})$ such that[6]

$$P_\pi^f(B) = \int_{\boldsymbol{S}} \sum_{x \in B} f(s)(x) d\pi(s), B \in \mathcal{B}_{\boldsymbol{X}}. \tag{1}$$

---

[6] For a definition of a Markov Kernel see, e.g., Strasser (1985), pg 102, Definition 23.2.

Notation 2.1 To save on notation, we sometimes write $P_\pi^f(x)$ to denote the distribution function *induced* by the act f given a probability π. Specifically, we write $P_\pi^f(x)$ to denote $P_\pi^f((-\infty, x] \cap \boldsymbol{X})$. Note that $x \mapsto P_\pi^f(x)$ is, therefore, well-defined on $\mathbb{R} \supset \boldsymbol{X}$. We will also find the associated inverse distribution functions, or quantile functions, useful. These are defined to be the right continuous functions $Q_\pi^f(p) = \inf\{x \mid P_\pi^f(x) > p\}$, $0 \leq p < 1$, $Q_\pi^f(1) = \inf\{x \mid P_\pi^f(x) \geq 1\}$.

Acts are objects of choice of a decision maker (DM). A binary relation $\succeq$ over $\mathcal{F}$ denotes a DM's preference ordering. Throughout, we will assume a DM's preferences satisfy properties of weak order and monotonicity, defined below.

Axiom 2.1 (Weak order) The preference $\succeq$ is complete and transitive.

Axiom 2.2 (Monotonicity) *(i)* If $x, y \in \boldsymbol{X}$ and $x \geq y$ then $\delta_x \succeq \delta_y$.
(ii) For every $l, l' \in \boldsymbol{L}$, if $l \succ l'$ and $0 \leq \beta < \alpha \leq 1$, then

$$\alpha l + (1 - \alpha)l' \succ \beta l + (1 - \beta)l'.$$

(iii) For every $f, g \in \mathcal{F}$, $f(s) \succeq g(s)$ for all $s \in \boldsymbol{S}$ implies $f \succeq g$.

Note, (i) and (ii) of Axiom 2.2 ensures that preferences over lotteries respect first order stochastic dominance, while (iii) ensures that preferences are state independent.

2.2 Defining more ambiguous

We define ordinal measures of how much the (subjective) evaluation of an act is affected, relative to another act, by (subjectively perceived) ambiguity. The measures are calibrated by explicit reference to individual preferences by comparing how two acts are evaluated by two preferences, one of which is *more ambiguity averse* than the other. Hence, our starting point is a notion of comparative ambiguity aversion. We adopt a notion well entrenched in the literature. Definition 2.1 is, essentially, a restatement of Epstein (1999) and Ghirardato and Marinacci (2002) definitions of comparative uncertainty/ambiguity aversion which were, in turn, a natural adaptation of Yaari (1969) classic formulation of comparative (subjective) risk aversion. Just as the definition of comparative risk aversion requires an a priori definition of a risk-free act, here the analogous role for "ambiguity-free" acts is played by lotteries.

Definition 2.1 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$. Let $\succeq_A, \succeq_B \in \mathcal{P}$. We say $\succeq_B$ is $(\mathcal{P})$-**more ambiguity averse** than $\succeq_A$ if

$$\begin{aligned} f \succeq_B l &\implies f \succeq_A l \\ f \preceq_A l &\implies f \preceq_B l \end{aligned}$$

for all $l \in \boldsymbol{L}$, and for all $f \in \mathcal{F}$.

Remark 2.1 The above definition implies that if two preferences can be ordered in terms of ambiguity aversion then they must rank lotteries in the same way.

As Epstein (1999) notes, to define absolute (rather than comparative) risk aversion, it is necessary to adopt a "normalization" for risk neutrality. The standard normalization is expected value. Analogously, to obtain a notion of absolute ambiguity aversion it is necessary to adopt a normalization for *ambiguity neutrality*. There are two normalizations prominent in the literature. Ghirardato and Marinacci (2002) say a preference is ambiguity neutral if it is a subjective expected utility (SEU) preference. That is, for any $f, g \in \mathcal{F}$, there exists a utility function, $u : \boldsymbol{X} \longrightarrow \mathbb{R}$, and a subjective belief associated with the preference, $\pi \in \Delta$, such that,

$$f \succeq g \Leftrightarrow \int_{\boldsymbol{S}} \left[\sum_{x \in \boldsymbol{X}} u(x) f(s)(x) \right] d\pi(s) \geq \int_{\boldsymbol{S}} \left[\sum_{x \in \boldsymbol{X}} u(x) g(s)(x) \right] d\pi(s).$$

In Epstein (1999), a preference $\succeq$ is ambiguity neutral if it is *probabilistically sophisticated*, that is, a preference that ranks acts or lotteries solely on the basis of their implied probability distributions over outcomes (Machina and Schmeidler (1992)). More precisely, letting $\mathbb{P}$ be the set of all Borel probability measures on $(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}})$, $\succeq$ is probabilistically sophisticated if there exists a function $W : \mathbb{P} \longrightarrow \mathbb{R}$, and an associated belief $\pi \in \Delta$, such that,

$$f \succeq g \Leftrightarrow W\left(P_\pi^f\right) \geq W\left(P_\pi^g\right), f, g \in \mathcal{F}.$$

Although Definition 2.1 says $\mathcal{P}$ is partially ordered by a more ambiguity averse relation, this does not necessarily imply that there exists any distinct pair of preferences in $\mathcal{P}$ which are ordered by the relation.

Definition 2.2 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$. We say $\mathcal{P}$ is **strictly partially ordered** by ($\mathcal{P}$)-more ambiguity averse if for each $\succeq \in \mathcal{P}$ there exists $\succeq' \in \mathcal{P}$, $\succeq \neq \succeq'$, such that $\succeq$ is ($\mathcal{P}$)-more ambiguity averse than $\succeq'$ or $\succeq'$ is ($\mathcal{P}$)-more ambiguity averse than $\succeq$.

The first notion of more ambiguous we offer is in the spirit of the Rothschild and Stiglitz (1970) notion of more risky. We require that an ambiguity neutral decision maker be indifferent between the two acts being compared while the ambiguity averse decision maker disprefers the more ambiguous act. Note, we may use either of the above two normalizations of ambiguity neutrality to obtain a *corresponding* notion of (absolutely) ambiguity averse: an *ambiguity averse* preference is one that is more ambiguity averse than an ambiguity neutral preference.

Definition 2.3 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$ strictly partially ordered by ($\mathcal{P}$)-more ambiguity averse and such that each $\succeq \in \mathcal{P}$ is related to an ambiguity neutral element of $\mathcal{P}$. Given $f, g \in \mathcal{F}$, we say f is a ($\mathcal{P}$)-**more ambiguous (I) act** than g, denoted f ($\mathcal{P}$)-m.a.(I) g, if the following conditions are satisfied:

(i) if $\succeq \in \mathcal{P}$ is ambiguity neutral then $g \sim f$;

(ii) for all $\succeq_A, \succeq_B \in \mathcal{P}$ such that $\succeq_A$ is an ambiguity neutral preference and $\succeq_B$ is $(\mathcal{P})$-more (respectively, less) ambiguity averse than $\succeq_A$ we have $g \succeq_B (\preceq_B) f$.

The notion of an act being more ambiguous than another is calibrated with respect to a reference class $\mathcal{P}$, restricted be a strictly partially ordered preferences. We restrict $\mathcal{P}$ in such a way to discipline its diversity. Recall, for the study of risk (subjective) beliefs are typically assumed to be common across the class of DMs. While $\mathcal{P}$ may well include several ambiguity neutral preferences, incorporating different subjective beliefs and/or risk attitudes, by condition (i) however, each ambiguity neutral preference must deem the acts being compared equivalent thereby restricting the subjective belief associated with the ambiguity neutral preferences included. Furthermore, every preference included in the reference class may be ordered, in terms of the more ambiguity averse relation with respect to *some* ambiguity neutral preference in $\mathcal{P}$.

The requirement in Definition 2.3 that ambiguity neutral agents be indifferent between the acts being compared is very natural but it has two drawbacks. First, we may wish to compare acts with respect to how they are affected by ambiguity, even though they may differ on other dimensions.[7] Second, there are reference classes $\mathcal{P}$ of interest which do not contain ambiguity neutral elements. For example, the set of all α-MEU preferences sharing the same set of priors in the representation functional in general will not include an ambiguity neutral sub-class (see Section 2.3). These considerations lead to our second definition of more ambiguous.

Notation 2.2 Given $y \in \mathbb{R}$, let $(f + \delta_y)$ denote a uniform translation of the contingent distributions on outcomes, that is an act such that,

$$(f + \delta_y)(s)(x + y) = f(s)(x),$$

$s \in \boldsymbol{S}$, $x \in \boldsymbol{X}$. When there is no possibility of confusion, we will sometimes denote the lottery degenerate at $y \in \boldsymbol{X}$ simply by y, in particular we sometimes write $f + y$ to denote $f + \delta_y$.

We propose to translate acts and wish to avoid hitting the bounds of $\boldsymbol{X}$. Let $\boldsymbol{L_J} \subset \boldsymbol{L}$ be the set of all finitely supported lotteries for which each outcome lies in a subinterval $\boldsymbol{J}$ of $\boldsymbol{X}$ with $|\boldsymbol{J}| \leq |\boldsymbol{X}|/3$ and center coinciding with the center of $\boldsymbol{X}$.

Definition 2.4 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$ strictly partially ordered by $(\mathcal{P})$-more ambiguity averse. Given acts f, g with consequences in $\boldsymbol{L_J} \subset \boldsymbol{L}$, we say f is a $(\mathcal{P})$-**more ambiguous (II) act** than g, denoted f $(\mathcal{P})$-m.a.(II) g, if for all $p \in \mathbb{R}$ with $|p| \leq |\boldsymbol{J}|$, $g \succeq_A (f + \delta_p) \Rightarrow g \succeq_B (f + \delta_p)$, whenever $\succeq_B$ is $(\mathcal{P})$-more ambiguity averse than $\succeq_A$.

---

[7]Analogous issues limit the applicability of the Rothschild-Stiglitz notion in risk analyses and led to the development of the notion of *location independent risk* introduced in Jewitt (1989) and analyzed in e.g., Gollier (2001), Chateauneuf, Cohen, and Meilijson (2004).

First, consider the case where $g \sim_A (f + \delta_p) \Rightarrow g \succeq_B (f + \delta_p)$. In this case, the amount p may be interpreted as a "compensating premium"; it measures, behaviorally, A's welfare loss in giving up g for f. Hence, in this case, the defining property for f to be m.a.(II) than g is that the compensating premium good enough for A is not good enough for B, who is more ambiguity averse than A. In general, there might not exist p such that indifference, $g \sim_A (f + \delta_p)$, obtains. If so, suppose $g \succ_A f$, let p be an amount that is not enough to flip A's preference, (i.e., it does not sweeten f enough for A to want to give up g for f) then, given the definition, p certainly won't be enough to flip B's preference, which is more ambiguity averse. As with Definition 2.3 this definition includes the strict partial order condition to discipline the diversity within the reference class $\mathcal{P}$. For every preference in $\mathcal{P}$ there is at least one other preference in $\mathcal{P}$ to which it may be related in terms of the more ambiguity averse relation and preferences, so related, satisfy the condition that the compensating premium is increasing in ambiguity aversion. More abstractly, the definition requires that translations of acts being compared satisfy a single-crossing property:

Definition 2.5 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$. Let $f, g \in \mathcal{F}$. The ordered pair of acts (f,g), satisfies the **single-crossing property for ambiguity** with respect to $\mathcal{P}$, denoted $(f, g) \in SCP(\mathcal{P})$, if for all $\succeq_B (\mathcal{P})$-more ambiguity averse than $\succeq_A$:

(i) $f \succeq_B g \Rightarrow f \succeq_A g$;

(ii) $f \preceq_A g \Rightarrow f \preceq_B g$.

The single-crossing property defines a fundamental comparative static in the sense that it should hold for any comparison of two acts differently affected by ambiguity, irrespective of whatever else may be affecting their evaluation.[8] However, single-crossing is not generally transitive. Transitivity of m.a.(II) relation is ensured by requiring single crossing to continue to be satisfied following arbitrary translations of f. Note, given Monotonicity, if f is m.a.(II) g and $\succeq_B$ is more ambiguity averse than $\succeq_A$, then $g \sim_A (f + \delta_p)$, $g \sim_B (f + \delta_q)$ implies $q \geq p$. [9]

2.2.1 More ambiguous events

As noted in the Introduction, it is of interest to define (comparative) ambiguity of *events*. Preferences for betting on one event rather than another, should reveal (a subjective

---

[8] The analog of Definition 2.5 for risk (with subjective beliefs) allows that the acts differ in aspects other than riskiness (such as different means) but as risk aversion increases, f tends to become less attractive than g due to f having a greater riskiness component. If $\mathcal{P}$ is taken to be SEU preferences with nondecreasing vNM utility and identical belief, π, the condition is equivalent to the distribution functions $P_\pi^f(x)$, $P_\pi^g(x)$, satisfying a single crossing property, see e.g. Gollier (2001), chapter 7. We make use of this fact below (Lemma A.1).

[9] Note, the two definitions of more ambiguous are distinct in that neither relation is strictly weaker than the other. The first definition, requires an ambiguity neutral benchmark, unlike the second. The second definition satisfies a single crossing property. Just as the Rothschild-Stiglitz notion does not generally satisfy single crossing, neither does the relation generated by Definition 2.3.

view) as to how much the event is affected by ambiguity compared to the other event. While the same basic principles applied to the case of acts apply here, there are new considerations to take into account. First, by definition, when we specify two acts we fix their (contingent) payoffs. But specifying two events does not specify their payoffs: *bets* on events are acts, but events themselves are not acts. Second, it seems fundamental to view an event as ambiguous if and only if its complement is ambiguous. It is natural, therefore, to require that if an event is more ambiguous than another the respective complementary events are ranked the same way.

Notation 2.3 If $x, y \in \mathbf{X}$ and $E \in \mathbf{\Sigma}$, xEy denotes the binary act which pays x if the realized state $s \in E$ and y otherwise.

Definition 2.6 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$ strictly partially ordered by ($\mathcal{P}$)-more ambiguity averse and such that each $\succeq \in \mathcal{P}$ is related to an ambiguity neutral element of $\mathcal{P}$. Given events E, $E' \in \mathbf{\Sigma}$,, we say E is a ($\mathcal{P}$)-**more ambiguous (I) event** than E' if: for, all ambiguity neutral $\succeq_A \in \mathcal{P}$,

$$xE'y \sim_A xEy \text{ and } x(\neg E')y \sim_A x(\neg E)y;$$

for all $\succeq_B \in \mathcal{P}$, such that $\succeq_B$ ($\mathcal{P}$)-more ambiguity averse than $\succeq_A$,

$$xE'y \succeq_B xEy \text{ and } x(\neg E')y \succeq_B x(\neg E)y;$$

for all $\succeq_B \in \mathcal{P}$, such that $\succeq_A$ is ($\mathcal{P}$)-more ambiguity averse than $\succeq_B$,

$$xE'y \preceq_B xEy \text{ and } x(\neg E')y \preceq_B x(\neg E)y,$$

where $x, y \in \mathbf{X}$, with $x > y$.

Hence, the act of betting on a more ambiguous event should be m.a.(I) and the same should hold of the complement. As in the case of acts, applying this notion to a class of preferences requires that class to include an ambiguity neutral preference and that such preferences be indifferent between the bets on the two events being compared. The following definition is constructed along the lines of the m.a.(II) definition.

Definition 2.7 Let $\mathcal{P}$ be a class of preferences over $\mathcal{F}$ strictly partially ordered by ($\mathcal{P}$)-more ambiguity averse. Given events $E, E' \in \mathbf{\Sigma}$, we say E is a ($\mathcal{P}$)-**more ambiguous(II) event** than E' if, $\succeq_A, \succeq_B \in \mathcal{P}$, $x, y, p, q \in \mathbf{X}$, with $x > y$,

$$xE'y \succeq_A pEy \Rightarrow xE'y \succeq_B pEy$$

and

$$x(\neg E')y \succeq_A q(\neg E)y \Rightarrow x(\neg E')y \succeq_B q(\neg E)y,$$

whenever $\succeq_B$ is ($\mathcal{P}$)-more ambiguity averse than $\succeq_A$.

For a first intuition, think of a variation of Ellsberg's 2-color, 2-urn example, in which the subject is given imprecise information about the composition of *both* urns, as opposed to the usual example where there is precise information about one urn and no information about the other. Each urn has a total of 100 balls, red and/or black. Let E be the draw of a red ball from the urn I which, the subject is told, has between 30 and 70 red balls and let E' be the draw of a red ball from urn II which is known to have between 40 and 60 red balls. Let p be the stake on E which makes an ambiguity averse agent A indifferent between the bets on E and E'. In this case, the amount $p - x$ may be interpreted as a "compensating stake". For a more ambiguity averse DM, B, the amount $p - x$ (weakly) under-compensates, so B would rather stick with the bet on E'. Like the m.a.(II) notion for acts, here too the fundamental idea is of single-crossing, strengthened to ensure transitivity by requiring the compensation $p - x$ to be monotone in ambiguity aversion. But unlike there, we do not compare preferences, of a less and more ambiguity averse agent, between translations to the *entire* acts. Rather we compare, across two such agents, the effect of a change of stake *on E*, relative to the stake on E' (*and* then, analogously, on the complements), to reveal the perceived comparative ambiguity *about E*. We use *event specific* payoff perturbations, specific to the events being compared.[10]

2.3   Parametric families of preferences considered in characterizations

We will apply the definitions to characterize more ambiguous for two parametric families of preferences, the $\alpha-$maxmin expected utility (α-MEU) family and the smooth ambiguity family. Next, we provide a brief description of these families.

The α-MEU model (Hurwicz (1951), Ghirardato, Maccheroni, and Marinacci (2004), henceforth, GMM)[11] represents preferences over acts in $\mathcal{F}$ according to,

$$V_{\Pi,\alpha,u}(f) = \alpha \min_{\pi \in \Pi} \int_{\boldsymbol{S}} \left[\sum_{x \in \boldsymbol{X}} u(x)f(s)(x) \right] d\pi(s) + (1-\alpha) \max_{\pi \in \Pi} \int_{\boldsymbol{S}} \left[\sum_{x \in \boldsymbol{X}} u(x)f(s)(x) \right] d\pi(s),$$
(2)

where $\alpha \in [0,1]$ is a weight, and $\Pi \subset \Delta$ is a compact, convex set of probability measures on the state space $\boldsymbol{S}$. As usual, $u : \boldsymbol{X} \longrightarrow \mathbb{R}$ is a nondecreasing vN-M utility function, understood to represent risk attitude. The weight α is interpreted to be an index of ambiguity attitude. The set Π is interpreted as the set of probabilities the DM subjectively deems as relevant and is the belief asociated with the preference. Let $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U}$ denote the class of α-MEU preferences where, the set Π is the belief associated with preferences in the class, the ambiguity attitude α ranges over the interval $[0,1]$ and the risk attitude u ranges over a set U. Let $\succeq_A, \succeq_B \in \mathcal{P}$. Then,

---

[10]We will, in contexts where there is no scope of confusion, use the phrase $(\mathcal{P})$-more ambiguous (I) (or, (II)), without appending the qualifier 'acts' or 'events'.

[11]The functional form was first suggested by Hurwicz. GMM axiomatizes a functional form of which the α-MEU form is a special case. However, Eichberger, Grant, Kelsey, and Koshevoy (2011) show that the GMM axiomatization does not provide a complete foundation to the special α-MEU case, in particular when the state space, $\boldsymbol{S}$ is finite. Klibanoff, Mukerji, and Seo (2011) provide an alternative foundation for α-MEU which addresses the problem Eichberger et. al. raise with GMM's axiomatization.

by Proposition 12 in GMM, $\succeq_A$ is ($\mathcal{P}$)-more ambiguity averse than $\succeq_B \Leftrightarrow \alpha_A \geq \alpha_B$, and u_A and u_B are equal up to an affine transformation, where α_A, u_A, and α_B, u_B are associated with $\succeq_A$ and $\succeq_B$, respectively. It is useful to note, given a compact, convex $\Pi \subset \Delta$, $f \in \mathcal{F}$, the kernel P_π^f is *mixture linear* in $\pi \in \Pi$, i.e.,

$$P_{\lambda\pi'+(1-\lambda)\pi''}^f = \lambda P_{\pi'}^f + (1-\lambda)P_{\pi''}^f, \pi', \pi'' \in \Pi \subset \Delta, \lambda \in [0,1].$$

The *smooth ambiguity* model (Klibanoff, Marinacci, and Mukerji (2005), henceforth, KMM)[12] represents preferences over acts according to,

$$V_{\mu,\phi,u}(f) = \int_\Delta \phi \left(\int_S \left[\sum_{x \in \mathbf{X}} u(x)f(s)(x) \right] d\pi(s) \right) d\mu(\pi), \qquad (3)$$

where, $u : \mathbf{X} \longrightarrow \mathbb{R}$ is a nondecreasing vN-M utility function shown to represent risk attitude; $\phi : u(\mathbf{X}) \longrightarrow \mathbb{R}$ is a nondecreasing function which maps (expected) utilities to reals shown to represent ambiguity attitude; $\mu : \mathcal{B}_\Delta \longrightarrow [0,1]$ is a Borel probability measure on Δ. The measure μ is interpreted as representing the DM's belief. The support of μ is taken to be the smallest closed (w.r.t. the vague topology) subset of Δ whose complement has measure zero, i.e., $\text{supp}(\mu) = \bigcap \{D \text{ closed} : \mu(D) = 1, D \subset \Delta\}$. Let $\{(\mu, \phi, u)\}_{\phi \in \Phi(u)}$ denote the class of smooth ambiguity preferences where, the measure μ is the belief associated with the preferences in the class, u is the utility function and the ambiguity attitude function ϕ ranges over some set $\Phi(u)$ of functions $\phi : u(\mathbf{X}) \longrightarrow \mathbb{R}$. Similarly, when the utility function u ranges over a set U, $\{(\mu, \phi, u)\}_{\phi \in \Phi(u), u \in U}$ denotes the class of preferences $\bigcup_{u \in U} \{(\mu, \phi, u)\}_{\phi \in \Phi(u)}$. In the characterizations of more ambiguous to follow, we typically set $U = U_1$, the set of nondecreasing utilities $u : \mathbf{X} \longrightarrow \mathbb{R}$ and $\Phi(u) = \Phi_1(u)$ the set of nondecreasing ambiguity attitudes $\phi : u(\mathbf{X}) \longrightarrow \mathbb{R}$. In this case we abuse notation and write $\{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$. Let $\succeq_A, \succeq_B \in \mathcal{P}\{(\mu, \phi, u)\}_{\phi \in \Phi(u), u \in U}$. Then, by Theorem 2 in KMM, $\succeq_A$ is ($\mathcal{P}$)-more ambiguity averse than $\succeq_B \Leftrightarrow \phi_A = h \circ \phi_B$, where $h : \phi_B(u(\mathbf{X})) \to \mathbb{R}$ is concave, and u_A and u_B are equal up to an affine transformation, where u_A, ϕ_A and u_B, ϕ_B are associated with $\succeq_A$ and $\succeq_B$, respectively.

Given an act, in contrast to α-MEU preferences, smooth ambiguity preferences with beliefs μ naturally induce a joint probability measure on outcomes and possible distributions over states. For each act $f \in \mathcal{F}$, and Borel set $B \in \mathcal{B}_{\mathbf{X}}$, $\pi \to P_\pi^f(B)$ is a $\mathcal{B}_\Delta$ measurable function. The Borel measure μ therefore uniquely[13] defines, for each act $f \in \mathcal{F}$, a probability measure $P^{f,\mu}$ on $(X \times \Delta, \mathcal{B}_{\mathbf{X}} \times \mathcal{B}_\Delta)$ such that for every $C \in \mathcal{B}_\Delta, B \in \mathcal{B}_{\mathbf{X}}$,

$$P^{f,\mu}(B \times C) = \int_C P_\pi^f(B) \, d\mu(\pi). \qquad (4)$$

Recall, the definition of m.a.(I) invokes the existence of an ambiguity neutral element in the relevant preference class. The smooth ambiguity preference (μ, ϕ, u) with ϕ

---

[12] For other preference models with similar representations see Ahn (2008), Ergin and Gul (2009), Nau (2006), Neilson (2010) and Seo (2009).

[13] See, e.g., Meyer (1966), T14, p.15.

affine is an SEU preference. Hence, the class $\{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$ includes an SEU preference. However, for a *given* compact, convex $\Pi \subset \Delta$, the class of α-MEU preferences $\{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, does not in general contain an SEU preference. Rogers and Ryan (2008), however establishes a general condition on the set of beliefs Π that guarantees the existence of an ambiguity neutral preference within the class: the α-MEU preference $(\Pi, 0.5, u)$ is an ambiguity neutral (SEU) preference if Π is *centrally symmetric*.

Definition 2.8 A set $\Pi \subseteq \Delta$ is **centrally symmetric** if there exists $\pi^\star \in \Pi$ (called the center of Π) such that, for any $\pi \in \Delta$, $\pi \in \Pi \Leftrightarrow \pi^\star - (\pi - \pi^\star) \in \Pi$.

As noted in KMM, SEU preferences are the only probabilistically sophisticated preferences within the smooth ambiguity class (so long as preferences over lotteries are expected utility). Marinacci (2002), p.756, shows for α-MEU preferences it is without loss of generality[14] to assume SEU as the benchmark model for ambiguity neutrality; there is no need to consider the more general probabilistically sophisticated preferences. Hence, for the preference classes we characterize SEU is the appropriate benchmark for ambiguity neutrality.

2.4 Event-comonotonicity and U-comonotonicity

Some characterizations, in the sequel, place further structure on the parametric classes of preferences considered by restricting the nature of associated beliefs. In this section we discuss some notions of order on beliefs and introduce two restrictions on preferences which induce a linear order, *event-comonotonicity* and *U-comonotonicity*.

Partial order induced by pairs of events   Quite generally, a set of events $\Sigma' \subset \boldsymbol{\Sigma}$ induces a *partial order* on the set Δ, $\pi \geqslant \pi'$ if $\pi(E) \geq \pi'(E)$ for each $E \in \Sigma'$. We are particularly interested in comparing pairs of events: For the pair of events E and E' from $\boldsymbol{\Sigma}$ we write $\pi' \leqslant_{E,E'} \pi$ if $\pi(E) \geq \pi'(E)$ and $\pi(E') \geq \pi'(E')$.

Notation 2.4 The meet $\bigwedge_{E,E'} \Pi$ denotes the greatest lower bound of the set $\Pi \subset \Delta$, when such a bound exists. That is, $\bigwedge_{E,E'} \Pi$ denotes the largest $\pi' \in \Delta$ such $\pi' \leqslant_{E,E'} \pi$ for all $\pi \in \Pi$, if such a π' exist. Similarly, the join $\bigvee_{E,E'} \Pi$ denotes the least upper bound, when that exists.

In general $(\Delta, \leqslant_{E,E'})$ is not a lattice, to see this let E and E' be disjoint events, there is a $\pi \in \Delta$ which assigns probability 1 to event E and a $\pi' \in \Delta$ which assigns probability 1 to E'. By definition any upper bound to the set $\{\pi, \pi'\} \subset \Delta$ must assign probability 1

---

[14]More precisely, Marinacci shows SEU preferences are the only probabilistically sophisticated preferences within the class of α-MEU preferences defined over acts whose domain includes at least one unambiguous event which is assigned a strictly positive probability by the subjective belief(s) associated with the preferences in the class.

to both events, however since the events are disjoint there is no probability measure in Δ which achieves this, hence there is no least upper bound.[15]

Remark 2.2 If $E, E' \in \mathbf{\Sigma}$, $E \cap E' \neq \emptyset$ and $E' \cup E \neq \mathbf{S}$, $(\Delta, \leqslant_{E,E'})$ is a lattice.

The conditions of Remark 2.2 ensure that $\bigwedge_{E,E'} \Pi, \bigvee_{E,E'} \Pi \in \Delta$ are defined for any compact subset $\Pi \subset \Delta$, but it is clearly not generally necessary that $\bigwedge_{E,E'} \Pi, \bigvee_{E,E'} \Pi \in \Pi$. Both $\arg\max_{\pi \in \Pi} \pi(E)$ and $\arg\max_{\pi \in \Pi} \pi(E')$ are nonempty convex subsets of Π. If these sets have a nonempty intersection, then $\pi \bigvee_{E,E'} \pi'$ exists and is an element of Π. When $\arg\max_{\pi \in \Pi} \pi(E) \cap \arg\max_{\pi \in \Pi} \pi(E') = \emptyset$, it may still be the case that there is some $\pi'' = \pi \bigvee_{E,E'} \pi' \in \Delta, \notin \Pi$ such that $\pi''(E) = \max_{\pi \in \Pi} \pi(E)$ and $\pi''(E') = \max_{\pi \in \Pi} \pi(E')$.

Comonotonicity and linear order   We may think of a bet on event E and a bet on event E' as 'similar' if the events induce the same ordering on the probability measures in Π. A bet that the S&P 500 index exceeds 1500 on January 1 2012 might be regarded as similar to a bet that the S&P 500 exceeds 1550 on February 1 2012 since both are more likely to pay off when π is optimistic about market conditions during the early part of 2012.

Definition 2.9 A set $\Pi \subset \Delta$ is **event-comonotone** for a pair of events $E, E' \in \mathbf{\Sigma}$, if for all $\pi_1, \pi_2 \in \Pi$, $(\pi_1(E) - \pi_2(E))(\pi_1(E') - \pi_2(E')) \geq 0$.

Event comonotonicity for a pair of events $E, E' \in \mathbf{\Sigma}$ imposes, or rather requires, a linear order $\leqq_{E,E'}$ on the set of probability measures $\Pi \subset \Delta$. Regardless of whether the conditions of Remark 2.2 obtain, this defines a lattice $(\Pi, \leqq_{E,E'})$ which given compactness of Π has top and bottom elements $\bigvee_{E,E'} \Pi \in \Pi$ and $\bigwedge_{E,E'} \Pi \in \Pi$ respectively. For α-MEU decision makers, it will become clear, these two, top and bottom, elements of Π contain all the behaviorally relevant information concerning the beliefs about the events E and E'. In general, the requirement of event-comonotonicity with respect to a pair of events restricts preferences over bets on these events by restricting the set of beliefs associated with the preferences. The idea has a natural extension to the case of acts which is the subject of the following definition.

Definition 2.10 A set $\Pi \subset \Delta$ is **U-comonotone** for a collection of acts $A \subset \mathcal{F}$ and class of utilities U if Π can be placed in linear order $\leqq_U$ such that for each $\pi_1, \pi_2 \in \Pi$, $\pi_1 \leqq_U \pi_2$ implies for each $u \in U$

$$\int_{\mathbf{S}} u(f) d\pi_1 \leq \int_{\mathbf{S}} u(f) d\pi_2 \text{ for each act } f \in A. \tag{5}$$

---

[15]Similarly, if the events are mutually exhaustive the probabilities should not add to less than one, however, this is not an issue for us since we do not admit as meaningful the question of whether an event is more ambiguous than its complement.

In the case of acts, as opposed to the case of events, the utility function matters for how the set Π is ordered. That is because bets on events are binary acts with just two outcomes, hence so long as utility indices satisfy monotonicity the choice of a particular utility would not affect the ordering over Π. The relation between event-comonotonicity and U_1-comonotonicity is clarified in the following proposition.

Proposition 2.1 *Let $f, g \in \mathcal{F}$ be acts mapping states to degenerate lotteries over outcomes in $\mathbf{X}$. Let $E_x^f \equiv \{s \in \mathbf{S} : f(s) \leq x\}$, $E_x^g \equiv \{s \in \mathbf{S} : g(s) \leq x\}$ denote the events that the outcome is no greater than $x \in \mathbf{X}$ under acts f and g, respectively. Fix a set $\Pi \subset \Delta$. The following statements are equivalent:*

(i) Π is event-comonotone for each pair of events $(E_x^h, E_{x'}^{h'})$, $h, h' \in \{f, g\}, x, x' \in \mathbf{X}$.

(ii) Π is U_1-comonotone for the pair of acts f, g.

The proposition shows U_1-comonotonicity is equivalent to event-comonotonicity of 'worse-outcome' events under the acts being compared.[16] Evidently, $U-$comonotonicity is a strong condition but it will be a natural and effective analytical tool in economic applications.

2.5 Elements of Statistical Decision Theory (Information Order)

Intuitively, if the distribution P_π^f does not depend on $\pi \in \Pi \subset \Delta$, then the act f would clearly be deemed *unambiguous* by any preference with associated beliefs contained in Π. Just as clearly, observing an outcome resulting from the act f is completely *uninformative* about which distribution $\pi \in \Pi$ actually obtained. These observations equate uninformative with unambiguous in what appears to be a very natural way. It should not, therefore, surprise the reader that we will find concepts from the literature on comparison of experiments (more informative) of direct use in characterizing and interpreting more ambiguous. To this end, in this section we will first associate elements of our decision theoretic set-up with the statistical decision theoretic setup of Wald (1949) used by Blackwell (1953) and then review the concepts of information order we invoke in our characterization results. For each concept reviewed, we discuss how an act deemed more informative by such an order induces distribution on outcomes that may be interpreted as being is more sensitive/less robust to the particular probability generating the states.

As is customary in this theory, we start with the sample space $\mathcal{S}$, a triple[17] consisting of a measurable space, which we will take to be $(\mathbf{S}, \mathbf{\Sigma})$, a parameter space Ω (sometimes

---

[16] Other classes of U, other than non-decreasing utilities may be of interest, for instance, non-decreasing concave utilities. We generally refrain from complicating the paper further by systematically pursuing this line of enquiry, which is left for future research. Though Remark 4.5 is an interesting exception.

[17] This is in accord with Blackwell and Girshick (1954) usage. Blackwell and Girshick also for convenience sometimes refer to the measurable space $(\mathbf{S}, \mathbf{\Sigma})$ itself as the sample space "Though formally the sample space is defined as a triple, we do not always make the distinction between it and the first element of the triple. Thus we speak of an event as a set in the sample space... ." (p.77). Other authors sometimes *define* the measurable set itself to be the sample space.

called–confusingly in this context–the set of states of the world) and a family of probability measures $(P_\omega)_{\omega \in \Omega}$ on (S, Σ). Hence, $\mathcal{S} = ((S, \Sigma), \Omega, (P_\omega)_{\omega \in \Omega})$. We shall be interested in comparing experiments defined on the same sample space $\mathcal{S}$. An *experiment defined on $\mathcal{S}$* is simply a statistic defined on (S, Σ), i.e., a measurable function f from S to some set of possible outcomes of the experiment. We shall limit our attention to statistics which take values in the space of outcomes, i.e., measurable functions $f : S \to X$. The experiment f therefore itself induces a sample space and can be equated with the triple $((X, \mathcal{B}_X), \Omega, (P_\omega^f)_{\omega \in \Omega})$. In order to map this *statistical decision theoretic* framework onto the *decision theoretic* framework of this paper:

1. We equate the first element of the sample space $\mathcal{S}$, (S, Σ), with the state space (S, Σ).

2. We equate Ω to a subset of Δ and the map $\omega \mapsto P_\omega$ is the identity map on Ω. Henceforth, therefore we shall write $\pi \in \Omega$ instead of $\omega \in \Omega$.

3. We equate the statistics f defining experiments with acts $f \in \mathcal{F}$ which have degenerate lotteries, i.e. outcomes, as consequences.

Hence, we associate acts with experiments of the form $((X, \mathcal{B}_X), \Omega, (P_\pi^f)_{\pi \in \Omega})$, with P_π^f defined as in equation (1).

There is, of course, a well-developed theory of what it means for one experiment to be *more informative* than another which it is helpful to briefly review[18]. Blackwell and Girshick (1954) document six equivalent characterizations for one experiment to be more informative than another. In the case of *dichotomies*—that is when the parameter space Ω has only two elements, they furnish a seventh characterization. One of Blackwell's characterizations starts from a general description of a statistical decision problem in which the objective is to make the expected loss resulting from a decision procedure based on the experiment small at each value of the parameter $\pi \in \Omega$. A loss function is a map from the product of some space of actions $\mathcal{A}$ and the parameter space Ω, i.e. $L : \mathcal{A} \times \Omega \to \mathbb{R}$. A decision procedure is a Markov kernel mapping from outcomes (of an experiment) to probability distributions over actions. If for each loss function, for any decision procedure based on experiment g, there is a decision procedure based on f which yields a weakly lower expected loss for each $\pi \in \Omega$, then experiment f is deemed more informative than experiment g. One natural necessary and sufficient condition for this occurs (as shown in the celebrated Blackwell, Sherman, Stein theorem) if one can use the more informative experiment plus randomization devices to construct a garbled experiment equivalent to the less informative experiment.

Definition 2.11 The experiment $((X, \mathcal{B}_X), \Omega, (P_\pi^f)_{\pi \in \Omega})$ is **Blackwell more informative** than the experiment $((X, \mathcal{B}_X), \Omega, (P_\pi^g)_{\pi \in \Omega}))$ on $\Omega \subset \Delta$ (f is Blackwell more informative than g on Ω) if there exists a Markov kernel $(x, B) \mapsto K_x(B)$ from $(X, \mathcal{B}_X)$ to

---

[18]The reader is referred to Blackwell and Girshick (1954), or Ferguson (1967), or Berger (1985) for excellent treatments. Jewitt (2011) gives a detailed discussion of the relationship with Lehmann and Blackwell information.

$(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}})$ such that

$$P_\pi^g(B) = \int_{\boldsymbol{X}} K_x(B) dP_\pi^f(x), B \in \mathcal{B}_{\boldsymbol{X}}, \pi \in \Omega. \tag{6}$$

The experiment f is **pairwise Blackwell more informative** than the experiment g on Ω, if f is Blackwell more informative than g on each dichotomy $\{\pi_1, \pi_2\} \subset \Omega \subset \Delta$.

To set ideas we describe the Markov kernel condition in case of bets on events. Let f be an act describing a unit bet on the event E, i.e.,

$$\left\{ \left(P_\pi^f(\{x=0\}), P_\pi^f(\{x=1\}) \right) \mid \pi \in \Delta \right\} = \{(\pi(E), \pi(\neg E)) \mid \pi \in \Omega\}, \tag{7}$$

and, similarly, let g describe a unit bet of the event E'. If E is Blackwell more informative than E' then on Ω there exists a row stochastic matrix K,

$$K = \begin{bmatrix} b & 1-b \\ c & 1-c \end{bmatrix} \text{ such that, } \begin{pmatrix} \pi(E') \\ \pi(\neg E') \end{pmatrix} = K \begin{pmatrix} \pi(E) \\ \pi(\neg E) \end{pmatrix} \text{ for all } \pi \in \Omega. \tag{8}$$

One easily checks that this implies K is bistochastic, i.e., $c = 1 - b$. Consider, e.g., $\pi^H, \pi^L \in \Delta$ such that $\pi^H(E) > \pi^L(E)$, and $\pi^H(E') > \pi^L(E')$. Hence, if E were pairwise Blackwell more informative than E' on $\{\pi^H, \pi^L\}$ then the likelihood of E is more sensitive than that of E' to whether π^H or π^L obtains.

If $\Omega \subset \Delta$ is a linearly ordered set, and P_π^f and P_π^g both exhibit monotone likelihood ratio, Lehmann (1988) characterized conditions appropriate to a specific class of loss functions[19] which are simpler to verify than (the first six of) Blackwell and Girshick's (1954) conditions.

Definition 2.12 Let $\leqq$ be a linear order on the set $\Omega \subset \Delta$. The family of probability measures $(P_\pi^f)_{\pi \in \Omega}$ on $\boldsymbol{X}$ satisfies **monotone likelihood ratio** (with respect to $\leqq$) if there is a density p_π^f with respect to a sigma-finite measure λ on $\boldsymbol{X}$ such that, $P_\pi^f(x) = \int_{(-\infty, x]} p_\pi^f(\eta) d\lambda(\eta), x \in \boldsymbol{X}$ with $p_{\pi_1}^f(x_1) p_{\pi_2}^f(x_2) \geq p_{\pi_1}^f(x_2) p_{\pi_2}^f(x_1)$ for all $\pi_1 \leqq \pi_2$ in Ω and $x_1 \leq x_2$ in $\boldsymbol{X}$.

Definition 2.13 Suppose $(P_\pi^f)_{\pi \in \Omega}$ and $(P_\pi^g)_{\pi \in \Omega}$ satisfy monotone likelihood ratio with respect to $\leqq$ on $\Omega \subset \Delta$. Then the experiment f is **Lehmann more informative** than g on Ω if for any $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$, $x \in \boldsymbol{X}$, $\pi_1, \pi_2 \in \Omega$, with $\pi_1 \leqq \pi_2$ there is an $x' \in \boldsymbol{X}$ such that

$$\lambda_1(1 - P_{\pi_1}^f(x')) + \lambda_2 P_{\pi_2}^f(x') \leq \lambda_1(1 - P_{\pi_1}^g(x)) + \lambda_2 P_{\pi_2}^g(x). \tag{9}$$

This has a simple decision theoretic interpretation: given monotone likelihood ratio, the Neymann-Pearson Lemma implies that optimal decisions in simple tests of hypothesis

---

[19]The class is called KR-monotone in Jewitt (2011) after Karlin and Rubin (1956)), similar conditions had already appeared in Blackwell and Girshick (1954). See Jewitt (2011) for an extensive discussion.

$(\pi_1, \pi_2 \in \Omega, \pi_1 < \pi_2, H_0 = \pi_1, H_1 = \pi_2)$ are simple cut-off rules: accept the hypothesis H_0 if the outcome is less than some critical value, reject H_0 in favor of H_1 if the outcome is larger than the critical value and randomize at the critical value. Under the beliefs, λ_1, λ_2, there is some decision rule based on outcomes from act f, which dominates any decision rule based on the outcomes from act g. In this sense, under f (likelihood of) outcomes are more sensitive to whether π_1 or π_2 obtains than they are under g.

Remark 2.3 Lehmann (1988) presented the condition (9), under the stipulation that P_π^f and P_π^g have no atoms and that

$$Q_{\pi_2}^f(P_{\pi_2}^g(x)) \geq Q_{\pi_1}^f(P_{\pi_1}^g(x)), \pi_1 \leqq \pi_2 \in \Pi, \ x \in \mathbf{X}. \tag{10}$$

We give the slightly more general formulation of Definition 2.13 in order to compare better with condition (iii) of Proposition 4.6 below.

Remark 2.4 Suppose $(P_\pi^f)_{\pi \in \Omega}$ and $(P_\pi^g)_{\pi \in \Omega}$ satisfy monotone likelihood ratio with respect to $\leqq$ on $\Omega \subset \Delta$. If experiment f is Lehmann more informative than experiment g on Ω then f is pairwise Blackwell more informative than g on Ω. (*See*, e.g., Jewitt (2011).)

Next, we introduce a notion of garbling distinct from Blackwell garbling. It obtains through interposition of an extra Markov kernel from $(\Omega, \mathcal{B}_\Omega)$ to itself.

Definition 2.14 We say the Markov kernel $(\pi, C) \mapsto K_\pi(C)$ from $(\Omega, \mathcal{B}_\Omega)$ to $(\Omega, \mathcal{B}_\Omega)$ **π-garbles** act f into act g if for all $B \in \mathcal{B}_{\mathbf{X}}$,

$$P_{\pi'}^g(B) = \int_\Delta P_\pi^f(B) dK_{\pi'}(\pi), \pi' \in \Omega. \tag{11}$$

We say g is a **π-garbling** of f if there exists a Markov kernel such that (11) obtains. Given events, $E, E' \in \Sigma$, we say E' is a π-garbling of E if g is a π-garbling of f, where g and f are acts describing unit bets on E' and E, respectively, as in (7).

To illustrate the essence of π-garbling consider a finite $\Omega = \{\pi_1, \ldots, \pi_m\} \subset \Delta$ and two events, $E, E' \subset \Sigma$. In this case, we say E' is a π-*garbling* of E if there exists a row stochastic matrix $[k_{ij}]_{i,j=1,\ldots,m}$, such that

$$\pi_j(E') = \sum_{i=1}^m k_{ij} \pi_i(E). \tag{12}$$

Hence, $\pi_i(E')$, $i = 1, \ldots, m$, are all contained in the convex hull of $\{\pi_i(E)\}_{i=1}^m$; given $\{\pi_1, \ldots, \pi_m\}$, the corresponding event probabilities of E' lie in a more circumscribed set than those for E. In this sense probability of E' is less sensitive than that of E to which element of Ω actually generates the states.

Finally, we briefly discuss *relative entropy* (Kullback and Leibler (1951)), which is a scalar measure preserving information order on dichotomies. Let P and Q be probability

measures on $\boldsymbol{X}$ which are both absolutely continuous with respect to some sigma-finite measure σ. The *relative entropy* or *Kullback-Leibler (K-L) divergence* from P to Q is given by

$$D(P||Q) = -\int_{\boldsymbol{X}} p \log\left(\frac{q}{p}\right) d\sigma$$

where $\frac{dP}{d\sigma} = p$ and $\frac{dQ}{d\sigma} = q$ are the respective Radon-Nikodym derivatives. Hence, if $D(P^f_{\pi'}||P^f_\pi) \geq D(P^g_{\pi'}||P^g_\pi)$ then the two distributions on outcomes corresponding to π and π' induced by act f are further apart than those induced by act g. In this sense, the act f is more sensitive to, or less robust to, the particular probability generating the states, compared to act g.

Remark 2.5 Suppose for each $\pi \in \Omega$, P^f_π and P^g_π are absolutely continuous with respect to the sigma finite measure σ. If the experiment f is pairwise Blackwell more informative than the experiment g on Ω, then for each pair $\pi, \pi' \in \Omega$, $D(P^f_{\pi'}||P^f_\pi) \geq D(P^g_{\pi'}||P^g_\pi)$. Since $x \mapsto -\log(x)$ is convex, this follows immediately from Blackwell and Girshick (1954), Theorem 12.22. Given Remark 2.4, this also demonstrates the link between the notion of Lehmann more informative and measures of K-L divergence. See Remarks 4.1 and 4.2 for the connection between measures of K-L divergence and the π-garbling criterion.

3 Characterizing more ambiguous events

3.1 More ambiguous (I)

α-**MEU preferences** At the outset, it is important to note that, since application of Definition 2.6 requires the existence of an ambiguity neutral element in the preference class, ambiguity neutrality is required for the *full* set of acts $\mathcal{F}$ and not just on bets on the events being compared. Hence, we characterize the definition for a class of preferences corresponding to a belief described by a compact, convex, centrally symmetric $\Pi \subset \Delta$. Given an event $E \in \boldsymbol{\Sigma}$, since Π is compact convex, the set of points $\pi(E) \in [0,1]$ as π ranges over Π is a closed interval which we denote as $\Pi(E) = \{\pi(E) \mid \pi \in \Pi\} = [\min \Pi(E), \max \Pi(E)] \subset [0,1]$. This interval is itself centrally symmetric, by dint of being closed convex and unidimensional, with center $\frac{\min \Pi(E) + \max \Pi(E)}{2}$. It is easy to check, denoting the center of Π as $\pi^\star$, that $\frac{\min \Pi(E) + \max \Pi(E)}{2} = \pi^\star(E)$.

One naturally expects $\Pi(E)$ to expand in some way as the event E is substituted for a more ambiguous one. The following proposition states that $\Pi(E)$ expands while retaining the same center. The characterization may be seen in terms of a π-garbling.

Proposition 3.1 *Let $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, where Π is a compact, convex centrally symmetric subset of Δ with center $\pi^\star$. Consider two events, $E, E' \in \boldsymbol{\Sigma}$. The following are equivalent:*

(i) E is a $(\mathcal{P})$-more ambiguous (I) event than E';

(ii) E' is a π-garbling of E and $\pi^\star(E') = \pi^\star(E)$;

(iii) $\Pi(E') \subset \Pi(E)$ *and* $\pi^\star(E') = \pi^\star(E)$.

Since $\Pi(E')$ is a subset of $\Pi(E)$ and has the same center, it is natural that $\Pi(E')$ has a smaller radius than $\Pi(E)$. Remark 4.1 generalizes this observation for acts when the radius is measured by K-L divergence.

Smooth ambiguity preferences For expositional clarity, we state the analog for smooth ambiguity preferences for the case where μ has finite support. Analogous to the centered expansion in the case of α-MEU preferences, for smooth ambiguity the characterizing condition involves a kind of mean preserving spread of the weighted event probabilities and, once again, may be formally interpreted as a π-garbling.

Proposition 3.2 *Let $\mathcal{P}$ be the class of smooth ambiguity preferences $\{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$, where $\mathrm{supp}(\mu) = \{\pi_i \in \Delta \mid i = 1, ..., m\}$. Consider two events, $E, E' \in \boldsymbol{\Sigma}$. The following are equivalent:*

(i) E is a $(\mathcal{P})$-more ambiguous (I) event than E';

(ii) There exists a row stochastic matrix $[k_{ij}]_{i,j=1,...,m}$ such that

$$\pi_j(E') \;=\; \sum_{i=1}^{m} \pi_i(E) k_{ij} \tag{13}$$

$$\mu_j \;=\; \sum_{i=1}^{m} k_{ij} \mu_i. \tag{14}$$

That is, E' is a π-garbling of E and $\mu_j = \sum_{i=1}^{m} k_{ij} \mu_i$, $j = 1, ..., m$.

The characterization for the smooth ambiguity case is very analogous to the α-MEU case. Condition (14) implies that the μ-average of the event probabilities is the same whether one considers E or E':

$$\sum_{i=1}^{m} \mu_i \pi_i(E') = \sum_{i=1}^{m} \sum_{j=1}^{m} k_{ij} \pi_j(E) \mu_j = \sum_{j=1}^{m} \mu_j \pi_j(E). \tag{15}$$

Condition (13) implies the $\pi_i(E')$ are all contained in the convex hull of the $\pi_i(E)$. Hence, as counterpart to $\Pi(E') \subset \Pi(E)$ in condition (iii) of Proposition 3.1 we have,

$$co\{\pi_1(E'), ..., \pi_m(E')\} \subset co\{\pi_1(E), ..., \pi_m(E)\}.$$

3.2 More Ambiguous (II)

The following notion of a set of points being *doubly star-shaped* is useful for characterizing more ambiguous (II) events.

Definition 3.1 We say $A \subset [0,1]^2$ is **star-shaped** if for $(a_1, a_2), (a_1', a_2') \in A$ and $1 \geq a_1' > a_1 \geq 0$, $\beta a_2 = a_1 \Rightarrow \beta a_2' \geq a_1', \beta \geq 0$; it is **doubly star-shaped** if, in addition, $\beta(1 - a_2) = (1 - a_1) \Rightarrow \beta(1 - a_2') \leq (1 - a_1'), \beta \geq 0$. A function $\zeta : [0,1] \to [0,1]$ is star-shaped (doubly star-shaped) if its graph is star-shaped (doubly star-shaped).

To visualize star shapedness, consider, bottom left and top right corners of a unit square, $[0,1]^2$. Then $A \subset [0,1]^2$ is doubly star-shaped if the slope of the sight line each of the two corners to a point on A increases further away the point is from that corner. It is useful to note that the interval $[a_1, a_1']$ is wider than the interval $[a_2, a_2']$.

α-MEU preferences Since the m.a.(II) notion does away with the requirement of the ambiguity neutral preference, it is intuitive to expect it generalizes the rather intuitive condition $\Pi(E') \subset \Pi(E)$, which itself evidently generalizes the condition of Proposition 3.1, by not requiring the expansion to be centered. The characterization of m.a.(II) for α-MEU requires the doubleton $\{(\min \Pi(E), \min \Pi(E')), (\max \Pi(E), \max \Pi(E'))\}$ to be doubly star-shaped. Evidently, $\Pi(E)$ is more spread out than $\Pi(E')$ though one interval is not necessarily contained in the other. Recall the partial order on events $\leqslant_{E,E'}$, introduced in Section 2.4. When $(\Delta, \leqslant_{E,E'})$ is a lattice, the characterizing condition may be formally interpreted in terms of an information order, with the more ambiguous event being Blackwell more informative.

Proposition 3.3 *Let $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, where Π is a compact, convex subset of Δ. Consider two events, $E, E' \in \mathbf{\Sigma}$. Statements (i) and (ii) are equivalent. If $\bigwedge_{E,E'} \Pi, \bigvee_{E,E'} \Pi \in \Delta$ exist, statement (iii) is equivalent to (i) and (ii).*

(i) E is a ($\mathcal{P}$)-more ambiguous (II) event than E';

(ii) Let $a = (a_1, a_2) = (\min \Pi(E'), \min \Pi(E))$ and $b = (b_1, b_2) = (\max \Pi(E'), \max \Pi(E))$. The set $\{a, b\}$ is doubly star-shaped.

(iii) E is Blackwell more informative than E' for the dichotomy

$$\Omega = \left\{ \bigwedge_{E,E'} \Pi, \bigvee_{E,E'} \Pi \right\} \subset \Delta.$$

Remark 3.1 The conditions (i) and (ii) (of Proposition 3.3) are implied by

$$\Pi(E') \subset \Pi(E). \tag{16}$$

Proposition 3.3 states that if $\bigwedge_{E,E'} \Pi$ and $\bigvee_{E,E'} \Pi$ exist, then E is m.a.(II) event than E' if and only if the partition $(E, \neg E)$ of $\mathbf{S}$ is more informative than the partition $(E', \neg E')$ of $\mathbf{S}$ for the dichotomy $\Omega = \left\{ \bigwedge_{E,E'} \Pi \text{ and } \bigvee_{E,E'} \Pi \right\} \subset \Delta$. Note that the event E plays two roles here—it determines the partition which carries information about which $\pi \in \Delta$ obtains, and it selects the dichotomy $\left\{ \bigwedge_{E,E'} \Pi \text{ and } \bigvee_{E,E'} \Pi \right\}$ which
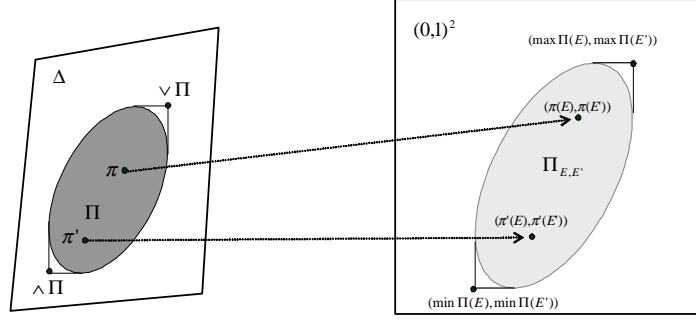
determines the relevant subset of Δ. The fact that the partition $(E, \neg E)$ of $\boldsymbol{S}$ is more informative than the partition $(E', \neg E')$ for the dichotomy $\Omega = \left\{ \bigwedge_{E,E'} \Pi \text{ and } \bigvee_{E,E'} \Pi \right\}$ does not imply that $(E, \neg E)$ is more informative than $(E', \neg E')$ for the dichotomy $\Omega = \left\{ \bigwedge_{\neg E,E'} \Pi \text{ and } \bigvee_{\neg E,E'} \Pi \right\}$.

Smooth ambiguity preferences

Proposition 3.4 *Let $\mathcal{P}$ be the class of smooth ambiguity preferences $\{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$. Consider two events, $E, E' \in \boldsymbol{\Sigma}$. The following are equivalent:*

(i) E is a ($\mathcal{P}$)-more ambiguous (II) event than E';

(ii) there exists a nondecreasing doubly star-shaped function $\zeta : [0,1] \to [0,1]$ such that,

$$\mu(\{\pi \mid \pi(E') \leq q\}) = \mu(\{\pi \mid \pi(E) \leq \zeta(q)\}, q \in [0,1] .$$

Hence, $\pi(E)$ has the same distribution, under μ, as $\zeta(\pi(E'))$. Note, double star-shapedness of ζ means that for each subinterval I of $[0,1]$, $\xi(I)$ is a wider interval than I. Hence, the condition here is analogous to the characterizing condition for α-MEU that $\Pi(E)$ is more spread out than $\Pi(E')$: two equal μ-measure journeys in the support of μ, one tracking the variation in the probability of E and the other of E', will reveal that the probability of more ambiguous event E will vary more.[20]

3.2.1 Adding event comonotonicity: α-MEU and smooth ambiguity

The assumption of event-comonotonicity leads to a particularly striking conclusion: the characterizing conditions for m.a.(II) events for the two classes of preferences collapse to the same condition. Consider α-MEU preferences with beliefs Π, and smooth ambiguity preferences with beliefs μ with support contained in Π. The following proposition asserts

---

[20] A star-shaped ordering of distributions has already been found useful in reliability theory. For non-negative random variables, the distribution F is said to be larger in the star order than the distribution G if $x \mapsto F^{-1}(G(x))$ is a star-shaped function. See e.g. Marshall and Olkin (2007).

22

that for *both* preference classes an event E is m.a.(II) event than E' if and only if E is Blackwell pairwise more informative than E' for all dichotomies from Π. An interesting feature of the condition is that the second order belief, μ, does not matter beyond the determination of its support.

Proposition 3.5 *Let Π be a compact, convex subset of Δ. Suppose Π is event-comonotone for a pair of events $E, E' \in \Sigma$. Let $\mathcal{P}_M(\Pi) = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$; let $\mathcal{P}_S(\Pi) = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$ with $\mathrm{supp}(\mu) = \Pi$. Then the following are equivalent:*

  (i) The set $\Pi_{E,E'} \equiv \{(\pi(E), \pi(E')) \mid \pi \in \Pi\} \subset [0,1]^2$ is doubly star-shaped;

  (ii) E is a $(\mathcal{P}_M(\Pi))$-more ambiguous (II) event than E';

  (iii) E is a $(\mathcal{P}_S(\Pi))$-more ambiguous (II) event than E';

  (iv) E is Blackwell pairwise more informative than E' for each dichotomy $\{\pi_1, \pi_2\} \subset \Pi$.

A first important key to the intuition is that since event-comonotonicity forces $\Pi_{E,E'}$ to be an increasing arc in the unit square the dimension of $\Pi_{E,E'}$ cannot be greater than one. If, in addition, Π is compact convex, this arc is the convex hull of the top and bottom elements of the lattice $(\Pi, \leqslant_{E,E'})$, i.e. the convex hull of the two points, $\left(\bigwedge_{E,E'} \Pi(E'), \bigwedge_{E,E'} \Pi(E) \right)$ and $\left(\bigvee_{E,E'} \Pi(E), \bigvee_{E,E'} \Pi(E') \right)$. Hence, for the case of α-MEU preferences, the characterizing condition of Proposition 3.3 reduces to $\Pi_{E,E'}$ being (doubly) star-shaped. Second, for each pair $\pi_i, \pi_j \in \Pi$ such that $\pi_i \leqslant_{E,E'} \pi_j$ event-comonotonicity allows us to define the intervals $[\pi_i(E'), \pi_j(E')]$ and $[\pi_i(E), \pi_j(E)]$ which, therefore, must have the same μ-measure. Hence, condition (ii) of Proposition 3.4 for smooth ambiguity preferences with with $\mathrm{supp}(\mu) = \Pi$ implies that there is a (doubly) star-shaped function ζ such that $[\pi_i(E), \pi_j(E)] = [\zeta(\pi_i(E')), \zeta(\pi_j(E'))]$.

Remark 3.2 The requirement that Π be convex in Proposition 3.5 is not necessary in the case of smooth ambiguity preferences. Specifically, the equivalence between (i), (iii) and (iv) remains true if the requirement is dropped. Note that there is no presumption in KMM that the support of μ be convex. It is perhaps worth stressing that in applications there are likely to be considerable advantages to dispensing with the requirement (for the same reason that the class of Normal Distributions, although not closed under mixtures, is an important class).

Remark 3.3 Noting that for observation of events monotone likelihood automatically obtains, it is furthermore possible to show that condition (iv) of Proposition 3.5 is equivalent to E being Lehmann more informative than E' on Π (in the sense of Definition 2.13) (see Jewitt (2011)).

4 Characterizing more ambiguous acts

4.1 More ambiguous (I)

4.1.1 Without U-comonotonicity

We begin this section with two closely related sufficient conditions, relating respectively to α-MEU preferences and smooth ambiguity preferences. Both are expressions of the idea that garbling the consequences of an act while preserving its 'balance' makes the act less ambiguous (I). In both cases the notion of garbling condition is that of π-garbling introduced in Section 2.5. The two notions of preserving balance, one to apply to α-MEU preferences and the other for smooth ambiguity preferences, are as follows.

Definition 4.1 Let Π be a compact, convex centrally symmetric subset of Δ with center $\pi^\star$, and let $f \in \mathcal{F}$. We say the Markov kernel $(\pi, C) \mapsto K_\pi(C)$ from $(\Delta, \mathcal{B}_\Delta)$ to itself is (f, Π)-**center preserving** (or, if clear from the context, simply center preserving) if for all Borel sets $B \in \mathcal{B}_X$,

$$P_{\pi^\star}^f(B) = \int_\Pi P_\pi^f(B) dK_{\pi^\star}(\pi). \tag{17}$$

If there is a center-preserving Markov kernel which π-garbles f into g, we say the π-garbling is center preserving. Then (from substituting $\pi^\star$ into (11)) the acts share the same distribution of consequences at the belief over states $\pi = \pi^\star$:

$$P_{\pi^\star}^f(B) = P_{\pi^\star}^g(B), B \in \mathcal{B}_X. \tag{18}$$

The second notion of preserving balance is:

Definition 4.2 Let $\mu : \mathcal{B}_\Delta \longrightarrow [0,1]$ be a Borel probability measure. We say the Markov kernel $(\pi, C) \mapsto K_\pi(C)$ from $(\Delta, \mathcal{B}_\Delta)$ to itself is **measure$-\mu$ preserving** (or, if clear from the context, simply measure preserving) if for all $C \in \mathcal{B}_\Delta$,

$$\mu(C) = \int_\Delta K_\pi(C) d\mu(\pi). \tag{19}$$

If there exists a measure-μ-preserving Markov kernel K which π-garbles f into g, we say the π-garbling is measure-μ preserving. It is useful to note (from integrating both sides of (11)) that then the acts share the same μ-averaged distribution over outcomes:

$$P^{g,\mu}(B \times \Delta) = \int_\Delta P_\pi^g(B) d\mu(\pi) = \int_\Delta P_\pi^f(B) d\mu(\pi) = P^{f,\mu}(B \times \Delta), B \in \mathcal{B}_X. \tag{20}$$

It is immediate from the fact that ambiguity neutral preferences are probabilistically sophisticated that if two acts f and g induce the same marginal distribution of outcomes, then all ambiguity neutral preferences are indifferent between them. Providing the class of preferences under consideration is rich enough, for instance it suffices if $U = U_1$, this is an equivalence which manifests itself in equations (18) and (20).

α-**MEU preferences.** The following proposition shows that the natural generalization of the π-garbling condition of Proposition 3.1 which was a characterization when relating to events also applies *as a sufficient condition* when applied to acts.

Proposition 4.1 *Let* $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, *where* Π *is a compact, convex centrally symmetric subset of* Δ *with center* $\pi^\star$. *Then* f *is a* $(\mathcal{P})-$*more ambiguous (I) act than* g *if there exists a center preserving Markov kernel from* $(\Delta, \mathcal{B}_\Delta)$ *to itself which* π-*garbles* f *into* g.

Remark 4.1 Let Π be a compact, convex centrally symmetric subset of Δ with center $\pi^\star$. If there exists a center preserving Markov kernel from $(\Delta, \mathcal{B}_\Delta)$ to itself which π-garbles f into g, then the radius (measured by K-L divergence from the center) of the set of distributions on outcomes induced by f is greater than the corresponding radius of the set induced by g. That is,

$$\max_{\pi \in \Pi} D(P_\pi^g || P_{\pi*}^g) \leq \max_{\pi \in \Pi} D(P_\pi^f || P_{\pi*}^f).$$

Smooth ambiguity preferences. Similarly, the following proposition states that a measure preserving π-garbling decreases ambiguity (I) for smooth ambiguity preferences.

Proposition 4.2 *Let* $\mathcal{P} = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$. *Then* f *is a* $(\mathcal{P})-$*more ambiguous (I) act than* g *if there is a measure-*μ *preserving Markov kernel from* $(\Delta, \mathcal{B}_\Delta)$ *to itself which* π-*garbles* f *into* g.

Remark 4.2 If there is a measure preserving π-garbling of f into g, then the μ-averaged K-L divergence is less for g than f, that is,

$$\int_{\Delta \times \Delta} D(P_{\pi'}^f || P_\pi^f) d(\mu \times \mu) \geq \int_{\Delta \times \Delta} D(P_{\pi'}^g || P_\pi^g) d(\mu \times \mu).$$

4.1.2 With U-comonotonicity

α-**MEU preferences.** The following proposition establishes that for α-MEU preferences with U_1-comonotonicity and $f, g \in \mathcal{F}$, f m.a.(I) g if and only the bad outcome events under f are m.a.(I) events than the corresponding bad outcome events under g. The result shows under an m.a.(I) act the induced distribution function of outcomes is more sensitive to the probability distribution generating the states: the distribution shifts (downwards) more when the distribution of states changes from π_1 to π_2 with $\pi_1 \leqq_{U_1} \pi_2$.

Proposition 4.3 *Let* $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, *where* Π *is a compact, convex centrally symmetric subset of* Δ *with center* $\pi^\star$. *Suppose* Π *is* U_1-*comonotone for the pair* $f, g \in \mathcal{F}$. *In the case* f *and* g *are acts mapping states into degenerate lotteries over outcomes in* **X**, *the following three conditions are equivalent. In the general case, conditions (i) and (iii). are equivalent.*

(i) *f is a ($\mathcal{P}$)-more ambiguous (I) act than g;*

(ii) *For each $x \in \boldsymbol{X}$, E_x^f, $E_x^g \in \boldsymbol{\Sigma}$, E_x^f is a ($\mathcal{P}$)-more ambiguous (I) event than E_x^g;*

(iii) *The condition (18) holds and for $\pi_1, \pi_2 \in \Pi$, $\pi_1 \leqq_{U_1} \pi_2$, the map $(\alpha, h) \mapsto P_{\alpha\pi_1+(1-\alpha)\pi_2}^h$ is supermodular on $[0,1] \times \{f, g\}$. Specifically for $0 \leq \alpha < \alpha' \leq 1$,*

$$P_{\alpha\pi_1+(1-\alpha)\pi_2}^g - P_{\alpha'\pi_1+(1-\alpha')\pi_2}^g \leq P_{\alpha\pi_1+(1-\alpha)\pi_2}^f - P_{\alpha'\pi_1+(1-\alpha')\pi_2}^f \text{ on } \boldsymbol{X}. \qquad (21)$$

Smooth ambiguity preferences. Given a class of smooth ambiguity preferences $\mathcal{P} = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$, and acts $f, g \in \mathcal{F}$, the fact that μ is common for all preferences within the class means that there is also a consensus on the probability measures $P^{f,\mu}$ and $P^{g,\mu}$ defined on the product space $(\boldsymbol{X} \times \Delta, \mathcal{B}_{\boldsymbol{X}} \times \mathcal{B}_\Delta)$. $P^{f,\mu}$ and $P^{g,\mu}$ have marginal probability measures defined on $(\boldsymbol{X}, \mathcal{B}_{\boldsymbol{X}})$ given by $P^{f,\mu}(B \times \Delta)$ and $P^{g,\mu}(B \times \Delta)$ respectively which, as we have seen in equation (20), represent the beliefs of the ambiguity neutral elements of $\mathcal{P}$ and will be equal if these elements are indifferent between the two acts. $P^{f,\mu}$ and $P^{g,\mu}$ also have marginal probability measures defined on $(\Delta, \mathcal{B}_\Delta)$ given by $P^{f,\mu}(\boldsymbol{X} \times C)$ and $P^{g,\mu}(\boldsymbol{X} \times C)$ respectively, but since by construction $P^{f,\mu}(\boldsymbol{X} \times C)$ and $P^{g,\mu}(\boldsymbol{X} \times C) = \mu(C)$, these are equal also. This means that if f m.a.(I) g, then $P^{f,\mu}$ and $P^{g,\mu}$ have the same marginals. Hence, the ambiguity relation m.a.(I) is determined by properties of the joint probability measures $P^{f,\mu}$ and $P^{g,\mu}$ invariant to the marginals. With U_1-comonotonicity, we induce an order on $\boldsymbol{X} \times \text{supp}(\mu)$ which enables us to express these properties of the joint probability measures in terms of *copulas*.

Notation 4.1 *Denote the collection of lower intervals of $\boldsymbol{X}$ as*

$$\boldsymbol{X}_L = \{\{x \in \boldsymbol{X} \mid x \leq x'\} \mid x' \in \boldsymbol{X}\} \cup \{\{x \in \boldsymbol{X} \mid x < x'\} \mid x' \in \boldsymbol{X}\}.$$

Similarly, for $\Pi \in \mathcal{B}_\Delta$.

$$\Pi_L = \{\{\pi \in \Pi \mid \pi \leqq_U \pi'\} \mid \pi' \in \Pi\} \cup \{\{\pi \in \Pi \mid \pi <_U \pi'\} \mid \pi' \in \Pi\}.$$

Hence, $\boldsymbol{X}_L \times \Pi_L \subset \mathcal{B}_{\boldsymbol{X}} \times \mathcal{B}_\Delta$ is the collection of lower quadrants of $\boldsymbol{X} \times \Pi$.

Proposition 4.4 *Let $\mathcal{P} = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$ with $\text{supp}(\mu) = \Pi$. Suppose Π is U_1-comonotone for the pair $f, g \in \mathcal{F}$. Then, the following are equivalent.*

(i) *f is a ($\mathcal{P}$)-more ambiguous (I) act than g;*

(ii) *The condition (20) holds and*

$$P^{f,\mu} \leq P^{g,\mu} \text{ on } \boldsymbol{X}_L \times \Pi_L. \qquad (22)$$

Remark 4.3 The condition in equation (20) of Proposition 4.3 is sufficient together with (21) for the conclusion of Proposition 4.4: Suppose (21) obtains, then if (20) holds, f ($\mathcal{P}$)-m.a.(I) g for the class $\mathcal{P}$ specified in Proposition 4.4.

To make the connection between condition (22) of Proposition 4.4 with the literature more explicit, we associate with each element in the $\leqq_{U_1}$ ordered set Π a real number: Note that $\pi \leqq_{U_1} (\geqq_{U_1})\pi'$ if and only if $\int_{\boldsymbol{S}} (f+g) \, d\pi \leq (\geq) \int_{\boldsymbol{S}} (f+g) \, d\pi'$ hence, $\pi \mapsto T(\pi)$, with $T(\pi) = \int_{\boldsymbol{S}} (f+g) \, d\pi$ represents the linear order $\leqq_{U_1}$ on Π. Define the distribution function F on $\mathbb{R}$ by $F(a) = \mu(\{\pi \in \Pi \mid \int_{\boldsymbol{S}} (f+g) \, d\pi \leq a\})$. If $a = T(\pi)$, $\pi \in \Pi$, let $F(x|a) = P^f_{T^{-1}(a)}(x)$ denote the conditional distribution (of outcomes given $T(\pi) = a$). We define the joint distribution function of outcomes and $T(\pi)$'s defined on $\mathbb{R}^2$ by

$$F(x,a) = \int_{-\infty}^{a} F(x|\xi) \, dF(\xi) = \int_{\{\pi \in \Pi | \pi \leqq_{U_1} T^{-1}(a)\}} P^f_\pi(x) \, d\mu(\pi).$$

Similarly,

$$G(x,a) = \int_{-\infty}^{a} G(x|\xi) \, dG(\xi) \text{ with } G(x|a) = P^g_{T^{-1}(a)}(x).$$

Since distribution functions are right continuous, conditions (22) and (20) respectively become, after taking liberties with notation in now using π to denote a real number,

$$
\begin{aligned}
F(x,\pi) &\leq G(x,\pi), \ (x,\pi) \in \mathbb{R}^2 \\
F(x) &= F(x,\infty) = G(x,\infty) = G(x), \ x \in \mathbb{R}.
\end{aligned}
$$

Noting also that

$$F(\infty,\pi) = G(\infty,\pi), \ \pi \in \mathbb{R},$$

it is clear that the joint distributions F and G have equal marginals. This means that the condition is a property of the *copulas* [21] of the two joint distributions induced by the pair of acts being compared. In our case, given that $\operatorname{supp}(\mu)$ is linearly ordered by $\leqq_{U_1}$, the copula corresponding to $P^{f,\mu}$ for act f is the function $C^f : [0,1]^2 \to [0,1]$ satisfying

$$C^f(F(x,\infty), F(\infty,\pi)) = F(x,\pi), \tag{23}$$

and similarly for act g. Hence, condition (ii) of the proposition can equivalently be stated as: condition (20) together with

$$C^f \leq C^g \text{ on } [0,1]^2. \tag{24}$$

This condition is discussed in the statistics literature in many places. For instance, Tchen (1980) calls it concordance. This is very natural in our context, it implies for instance that bad news about which probability distribution $\pi \in \Delta$ is operative is more strongly associated with bad news about outcomes—i.e., conditioning on the 'event' $\{\pi' \in \Delta \mid \pi' \leqq_{U_1} \pi\}$ for some given $\pi \in \Pi$ makes the conditional distribution of outcomes worse by first-order stochastic dominance—when the more ambiguous act is taken. We relate the condition to the more informative ordering of Lehmann (1988), the result in the following remark appears in Jewitt (2011).

---

[21] The *copula* C of a random vector (Z_1, Z_2) with cdf $F_{Z_1,Z_2}(z_1,z_2)$ and marginal cdfs $F_{Z_1}(z_1), F_{Z_2}(z_2)$ satisfies $F_{Z_1,Z_2}(z_1,z_2) = C(F_{Z_1}(z_1), F_{Z_2}(z_2))$. By Sklar's theorem (Sklar (1959)), the copula is unique if the marginal distributions are atomless. Otherwise the copula is uniquely defined at points of continuity of the marginal distributions.

Remark 4.4 Suppose $(P_\pi^f)_{\pi \in \Pi}$ and $(P_\pi^g)_{\pi \in \Pi}$ satisfy monotone likelihood ratio with respect to $\leqq_{U_1}$ on $\Pi \subset \Delta$. If experiment f is Lehmann more informative than g on Π, and if μ with support in Π is such that (20) holds, then (22) of condition (ii) Proposition 4.4 holds.

Remark 4.5 We have restricted attention to the class U_1 of monotone utilities, but it may be of interest to extend the analysis to risk averse utilities. Let U_2 be the class of nondecreasing concave utilities and suppose ambiguity is U_2-comonotone for the pair of acts f and g. Then, it can be shown, f is $(\mathcal{P})$-m.a.(I) than g if and only if

$$\int_0^p C^f(\eta, q) d\eta \geq \int_0^p C^g(\eta, q) d\eta \ on \ [0,1]^2, \ \forall p \ s.t. \ 0 \leq p \leq 1.$$

4.2 More ambiguous (II)

The assumption of U-comonotonicity leads to a considerable simplification and much more congenial characterizations than are available in the general case. For completeness we include the characterizations for both α-MEU and smooth ambiguity preferences, without U-comonotonicity in Appendix A.5.

4.2.1 With U-comonotonicity: α-MEU and smooth ambiguity

Single Crossing We begin by characterizing *single-crossing* for α-MEU and smooth ambiguity preferences, this will not only be useful in establishing conditions for m.a.(II), but given the importance of single crossing conditions for comparative statics exercises, it is of significant independent interest. The following condition applies for both smooth and α-MEU preferences.

Condition SCU Suppose Π is U_1-comonotone for the pair $f, g \in \mathcal{F}$. For each $\pi_1 \leqq_U \pi_2$ from Π, there exist $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$ such that for each $x \in \mathbf{X}$,

$$\lambda_1(1 - P_{\pi_1}^f(x)) + \lambda_2 P_{\pi_2}^f(x) \leq \lambda_1(1 - P_{\pi_1}^g(x)) + \lambda_2 P_{\pi_2}^g(x). \tag{25}$$

In the following characterizations, we separate the statements for two classes of preferences since in the case of smooth ambiguity preferences, Π need not be compact convex. Recall, if the ordered pair of acts (f, g), satisfies the single-crossing property for ambiguity with respect to the preference class $\mathcal{P}$ we write $(f, g) \in SCP(\mathcal{P})$.

Proposition 4.5 *Let $\Pi \subset \Delta$ be compact and U_1-comonotone for the pair $f, g \in \mathcal{F}$.*

(a) *Suppose Π is convex. Let $\mathcal{P}_M = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$. The following conditions are equivalent:*

    (i) $(f, g) \in SCP(\mathcal{P}_M)$.

    (ii) Condition SCU holds.

(b) *Suppose* $\Pi = \text{supp}(\mu)$ *is either path-connected or finite. Let* $\mathcal{P}_S = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$.
The following conditions are equivalent:

(i) $(f, g) \in SCP(\mathcal{P}_S)$.

(ii) *Condition SCU holds.*

Condition SCU therefore characterizes both $SCP(\mathcal{P}_M)$ and $SCP(\mathcal{P}_S)$. The defining inequality (25) is strikingly similar to the inequality (9) of Definition 2.13. Recall, Lehmann's condition has a simple interpretation in terms of Neymann-Pearson tests of simple hypotheses. Equation (9) means, essentially, that for any loss function (cost of type I and type II errors), there is some simple cut-off decision rule based on outcomes from act f, which dominates any simple cut-off decision rule based on the outcomes from act g. The stipulation of monotone likelihood ratio implies via the Neymann-Pearson lemma that cut-off rules are optimal. To contrast, condition (25) has the following non-standard statistical decision theoretic interpretation. The DM must guess whether $\pi = \pi_1$ or $\pi = \pi_2$ obtains. If guessing incorrectly she loses £1. Furthermore, the DM is bound to a decision rule of the form: bet that π is small ($\pi = \pi_1$) if the outcome $x' \in \mathbf{X}$ is smaller than some predetermined cut-off x, otherwise bet that $\pi = \pi_2$. The condition asserts that there is a Bayesian DM who prefers the experiment induced by act f rather than act g, given the imposed decision rule and regardless of the cutoff value $x \in \mathbf{X}$. Further clarification of the relationship with Lehmann information is given in the discussion of m.a.(II) below. To set the intuition in a perhaps more straightforward way, note that a sufficient condition (25) for is that $P^g_{\pi_1}$ stochastically dominates $P^f_{\pi_1}$ and $P^f_{\pi_2}$ stochastically dominates $P^g_{\pi_2}$ hence the distribution of outcomes under act f is, in a very strong way, more affected by the change from π_1 to π_2 than is the distribution of outcomes under act g.

Recall, in Proposition 3.5 we obtain a characterizing condition for m.a.(II) given event-comonotonicity that is essentially the same for both classes of preferences, α-MEU and smooth ambiguity. Here we demonstrate something very analogous for acts, given U-comonotonicity. In Propositions 4.6 and 4.7 let $\mathcal{P}_M(\Pi) = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, where Π is a compact, convex subset of Δ; let $\mathcal{P}_S(\Pi) = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$ with $\text{supp}(\mu) = \Pi$. Recall, $\mathbf{L_J}$ consists of lotteries with outcomes in $\mathbf{J}$ (text preceding Definition 2.3).

Proposition 4.6 *Suppose* $f, g \in \mathcal{F}$ *have consequences in* $\mathbf{L_J}$ *and* $\Pi \subset \Delta$ *is* U*-comonotone for the pair* f, g *for* $U = U_1$. *The following conditions are equivalent:*

(i) $f (\mathcal{P}_M(\Pi))$-*m.a.(II)* g.

(ii) $f (\mathcal{P}_S(\Pi))$-*m.a.(II)* g.

(iii) *For each* $\pi_1 \leqq_U \pi_2$ *from* Π, *and* $p \in \mathbb{R}$ *with* $|p| \leq |\mathbf{J}|$, *there exist* $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$ *such that for all* $x \in \mathbf{J}$,

$$\lambda_1(1 - P^{f+p}_{\pi_1}(x)) + \lambda_2 P^{f+p}_{\pi_2}(x) \leq \lambda_1(1 - P^g_{\pi_1}(x)) + \lambda_2 P^g_{\pi_2}(x). \tag{26}$$

Remark 4.6 It is clear from the proof that the equivalence between conditions (i) and (iii) of Proposition 4.6 holds does not require the convexity of Π. It would suffice that Π is connected, or Π finite.

One can easily verify that condition (26) implies condition (9) of Definition 2.13. Hence, it follows that if $(P_\pi^f)_{\pi \in \Pi}$ and $(P_\pi^g)_{\pi \in \Pi}$ satisfy monotone likelihood ratio with respect to $\geqq_{U_1}$ on Π ; if f $(\mathcal{P}_M)$-m.a.(II) g; or f $(\mathcal{P}_S)$-m.a.(II) g, then f is Lehmann more informative than g. The following proposition gives an auxiliary condition under which, act f Lehmann more informative than g on Π, is sufficient for f m.a.(II) g. The auxiliary condition requires for each $\pi \in \Pi$, the differences between quantiles of outcomes under act f are further apart than under act g:

Definition 4.3 For $\pi \in \Delta$, $x \in \boldsymbol{X}$, $x \mapsto P_\pi^f(x)$ is more Bickel-Lehmann dispersed than $x \mapsto P_\pi^g(x)$ if, for all $p_1 < p_2 \in [0,1]$,

$$Q_\pi^g(p_2) - Q_\pi^g(p_1) \leq Q_\pi^f(p_2) - Q_\pi^f(p_1), \tag{27}$$

where $Q_\pi^f(p) = \inf\{x \mid P_\pi^f(x) > p\}$, $Q_\pi^g(p) = \inf\{x \mid P_\pi^g(x) > p\}$.[22] If for each $\pi \in \Pi \subset \Delta$, $x \mapsto P_\pi^f(x)$ is Bickel-Lehmann more dispersed than $x \mapsto P_\pi^g(x)$, then we say act f is **more Bickel-Lehmann dispersed than** g on Π.

Proposition 4.7 *Suppose* $f, g \in \mathcal{F}$ *have consequences in* $\boldsymbol{L_J}$ *and* $\Pi \subset \Delta$ *is* U-*comonotone for the pair* f, g *for* $U = U_1$. *Assume* $(P_\pi^f)_{\pi \in \Pi}$ *and* $(P_\pi^g)_{\pi \in \Pi}$ *are absolutely continuous with respect to Lebesgue measure and satisfy monotone likelihood ratio with respect to* $\leqq_{U_1}$. *Further suppose*

(i) f *is Lehmann more informative than* g *on* Π, *and*

(ii) f *is Bickel-Lehmann more dispersed than* g *on* Π,

then f $(\mathcal{P}_M(\Pi))$-*m.a.(II)* g, *and* $(\mathcal{P}_S(\Pi))$-*m.a.(II)* g.

The first condition of the proposition states that the more ambiguous act has consequences which are more informative (about $\pi \in \Pi$), the second condition states that the more ambiguous act is, in a sense, more highly geared. For example in a simple portfolio problem with a single uncertainty free asset and a single uncertain asset, holding a larger share of the portfolio in terms of the uncertain asset will produce consequences which are *equally* informative as a portfolio with a smaller share of the uncertain asset (providing it is non-zero), this is obvious since there is a one-to-one map between the payoffs, they carry the same information content. On the other hand, the larger portfolio is evidently more highly geared to the uncertainty.

---

[22]Terminology is not uniform in the literature. We use the same terminology as e.g. Landsberger and Meilijson (1994) in reference to Bickel and Lehmann (1976), Doksum (1969), p.1169 defines F to be tail ordered with respect to G if $G^{-1}(F(x)) - x$ is non-decreasing.

5 Applications

5.1 Ambiguity in the small

In this section we apply our concept of more ambiguous, specifically m.a.(I) applied to α-MEU and smooth ambiguity preferences, to derive two approximate measures of ambiguity premia inspired by a famous result in risk theory. Arrow (1965) and Pratt (1964) showed, within the expected utility framework, that under certain conditions, the risk premium of a small gamble is equal to one half the product of the degree of risk aversion as measured by the Arrow-Pratt coefficient of absolute risk aversion and the quantity of risk as measured by variance of the gamble. Arrow and Pratt's objective was to establish the measure of risk aversion, rather than to establish variance as a measure of risk—which seemed as obvious then as now. In our case, since we already have notions of ambiguity aversion in place, we are primarily interested in establishing the small gamble analog of variance for the study of ambiguity. To this end, we develop two measures of ambiguity based on an approximate formula for the *ambiguity premium* of a small (ambiguous) gamble.[23] We identify the ambiguity premium of an act to be the price a DM is willing to pay to swap the act for a lottery comparable to the act in terms of m.a.(I). Why does this make sense? First, that the act and lottery can be compared in terms of m.a.(I) ensures that it is only the differing effect of ambiguity on the two prospects that matters for preferences and secondly, a lottery is completely unaffected by ambiguity. By making the comparison with a lottery, we turn a relative measure, m.a.(I), into an absolute measure.

Consider $\mathcal{P} = \mathcal{P}_M = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, where Π is a compact, convex centrally symmetric subset of Δ with center $\pi^\star$, and $\mathcal{P} = \mathcal{P}_S = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$, and let $l_{\mathcal{P}}^f \in \boldsymbol{L_J}$ be the (unique) lottery such that f $(\mathcal{P})$-m.a.(I) $l_{\mathcal{P}}^f$. Assume $\boldsymbol{S}$ is a finite set,[24] then evidently,

$$\sum_{x \in B} l_{\mathcal{P}_M}^f (x) = P_{\pi^*}^f (B), B \in \mathcal{B}_{\boldsymbol{X}}$$

and

$$\sum_{x \in B} l_{\mathcal{P}_S}^f (x) = \int_\Delta P_\pi^f (B) \, d\mu(\pi), B \in \mathcal{B}_{\boldsymbol{X}}. \tag{28}$$

That is, the ambiguity neutral elements in the respective classes are indifferent between $l_{\mathcal{P}}^f$ and f. The *ambiguity premium* for some $\succeq \in \mathcal{P}$ of an act f is given by $\mathsf{a}^f \in \boldsymbol{X}$, where

$$\left(l_{\mathcal{P}}^f - \mathsf{a}^f \right) \sim f,$$

---

[23] Nau (2006), Izhakian and Benninga (2008), Skiadas (2009) and, Maccheroni, Marinacci, and Ruffino (2010), derive approximations for an *uncertainty premium* in the smooth ambiguity model (under different sets of restrictions). Of these, Maccheroni, Marinacci and Ruffino's result is the most general; they obtain the expression in (30) without the assumption $var_\pi(f)$ (see below) is constant on Π. We are not aware of a primitive definition of *ambiguity premium* in the existing literature nor of uncertainty premia for other preference classes.

[24] This is assumed for simplicity, since lotteries are (also) assumed to have finite supports.

and $\left(l_{\mathcal{P}}^{f} - \mathsf{a}^{f}\right)$ denotes the translated lottery satisfying $\left(l_{\mathcal{P}}^{f} - \mathsf{a}^{f}\right)\left(x - \mathsf{a}^{f}\right) = \left(l_{\mathcal{P}}^{f}\right)(x)$. We say $\left(l_{\mathcal{P}}^{f} - \mathsf{a}^{f}\right)$ is the *lottery equivalent* of act f for preference $\succeq \in \mathcal{P}$. The *certainty equivalent* of act f for $\succeq \in \mathcal{P}$, denoted c^{f}, is the certainty equivalent of the lottery $\left(l_{\mathcal{P}}^{f} - \mathsf{a}^{f}\right)$. Just as in risk theory, the *risk premium* is defined as the difference between the expected value of the lottery $\left(l_{\mathcal{P}}^{f} - \mathsf{a}^{f}\right)$ and the certainty equivalent c^{f}, that is

$$\mathsf{r}^{f} = E[l_{\mathcal{P}}^{f}] - \mathsf{a}^{f} - \mathsf{c}^{f}.$$

The sum $\mathsf{u}^{f} = \mathsf{a}^{f} + \mathsf{r}^{f}$ is therefore the *uncertainty premium.*

The next step in deriving approximations for uncertainty and ambiguity premia for small gambles, is to have a notion of the smallness of a gamble. The variance of the gamble is uniquely defined for the ambiguity neutral members of the class of preferences under consideration and this will be shown to provide a useful metric for approximations. We derive approximations for uncertainty and ambiguity premia for small gambles corresponding to an act f, for the preference classes $\mathcal{P}_S$ and $\mathcal{P}_M$. The key assumptions used for both classes of preferences are (a) a version of Pratt's (1964) condition that the third absolute central moment of the gamble is of smaller order than the second central moment and (b) the variance of outcomes of act f under beliefs $\pi \in \Delta$, denoted $var_{\pi}(f)$ is constant on Π in the $\mathcal{P}_M$ case and constant on the support of μ in the $\mathcal{P}_S$ case.

First, we consider smooth ambiguity preferences. As discussed in KMM, we may write $\phi \equiv v \circ u^{-1}$, where v is the vN-M index in the representation of preferences on second order acts. For the purposes of the present discussion, it is significant that v and u both map from a set denominated in the same units. We use the following notation for the different variances which are needed:

$$
\begin{aligned}
E_{\pi}\left[f\right] &\equiv \int_{\boldsymbol{X}} x \, dP_{\pi}^{f}, \ \ var_{\pi}\left(f\right) \equiv \int_{\boldsymbol{X}} \left(x - E_{\pi}\left[f\right]^{2}\right) dP_{\pi}^{f}, \\
E\left[f\right] &= \int_{\Delta} E_{\pi}\left[f\right] d\mu = \int_{\boldsymbol{X} \times \Delta} x \, dP^{f,\mu}, \ \ var\left(f\right) \equiv \int_{\boldsymbol{X} \times \Delta} \left(x - E\left[f\right]\right)^{2} dP^{f,\mu} \\
var(E_{\pi}\left[f\right]) &= \int_{\Delta} \left(E_{\pi}\left[f\right] - E\left[f\right]\right)^{2} d\mu.
\end{aligned}
$$

Note that a consequence of the law of total probability (the law of total variance), under the assumption that $var_{\pi}(f)$ is constant on the support of μ, is

$$var\left(f\right) = var_{\pi}\left(f\right) + var(E_{\pi}\left[f\right]). \tag{29}$$

This fact is useful in formulating the proposition and in its proof since it implies that if $var\left(f\right)$ is small, then so are $var_{\pi}\left(f\right)$ and $var(E_{\pi}\left[f\right])$, hence we can use $var\left(f\right)$ as a metric for smallness of gambles for all the beliefs π in the support of μ. It is also useful in interpretation, a natural measure of the proportion of the total uncertainty explained by ambiguity is the following *"R-squared"* expression

$$\rho_S \equiv \frac{var(E_{\pi}\left(f\right))}{var_{\pi}\left(f\right) + var(E_{\pi}\left(f\right))}.$$

Using standard big O, little o notation to express orders of smallness, $h(x) = o(k(x))$ if $\lim_{x \to 0} \frac{h(x)}{k(x)} = 0$, in which case h is said to be of smaller order than k. Following KMM (p. 1859), set $v = \phi \circ u : \boldsymbol{X} \longrightarrow \mathbb{R}$. Ambiguity neutrality corresponds to $\phi(x) = x$, that is, $u = v$ (up to a normalization).

Proposition 5.1 *For preferences $\mathcal{P}_S = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$, with u and ϕ strictly increasing three times continuously differentiable functions. Suppose, (a) (Pratt's condition) for each π in the support of μ, the third absolute central moment of P_π^f, is of smaller order than the second central moment, (b) $var_\pi(f)$ is constant on the support of μ. The uncertainty, ambiguity and risk premia satisfy, with $R_v = -v''/v'$ and $R_u = -u''/u'$ evaluated at $E[f]$,*

$$\mathsf{u}^f = \frac{1}{2} R_u var_\pi(f) + \frac{1}{2} R_v var\left(E_\pi[f]\right) + o(var(f)) \tag{30}$$

$$\mathsf{a}^f = \frac{1}{2} \left(R_v - R_u\right) var\left(E_\pi[f]\right) + o(var(f)) \tag{31}$$

$$\mathsf{r}^f = \mathsf{u}^f - \mathsf{a}^f = \frac{1}{2} R_u var(f) + o(var(f)).$$

Hence, the effective total coefficient of uncertainty aversion, i.e. what must be multiplied by one half the total variance of outcomes in order to obtain the uncertainty premium is given by

$$\mathsf{u}^f = \frac{1}{2} var(f) \left(R_u + (R_v - R_u) \frac{var(E_\pi(f))}{var_\pi(f) + var(E_\pi(f))}\right)$$

Note, $R_\phi = -\phi''/\phi' = R_v - R_u$. Evidently, if either $R_v = R_u$ or ρ_S equal zero, there is no ambiguity component of the uncertainty premium. In the first case, it is because the DM is neutral to ambiguity. In the second case, it is because the proportion of total uncertainty caused by ambiguity is zero. This echoes our now familiar insight that for the more ambiguous act the outcomes, $x \in \boldsymbol{X}$, and probabilities, $\pi \in \text{supp}(\mu)$, better explain each other.

We next turn to the class of α-MEU preferences, $\mathcal{P}_M$. With,

$$\overline{\pi} \in \arg\max_{\pi \in \Pi} \int u dP_\pi^f, \underline{\pi} \in \arg\min_{\pi \in \Pi} \int u dP_\pi^f,$$

we may define as above $E_{\pi^*}[f]$, $E_{\alpha\underline{\pi} + (1-\alpha)\bar{\pi}}[f]$, and $var_{\alpha\underline{\pi} + (1-\alpha)\bar{\pi}}(f)$.

Proposition 5.2 *For preferences $\mathcal{P}_M = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, with u a strictly increasing three times continuously differentiable function. Suppose, (a) (Pratt's condition) for each $\pi \in \Pi$, the third absolute central moment of P_π^f, is of smaller order than the second central moment, (b) $var_{\overline{\pi}}(f) = var_{\underline{\pi}}(f)$. The uncertainty, ambiguity and risk premia*

satisfy

$$\mathsf{u}^f = E_{\pi^*}[f] - E_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}[f] + \frac{1}{2}R_u var_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}(f) + o(var_{\pi^*}(f)),$$

$$\mathsf{a}^f = E_{\pi^*}[f] - E_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}[f] - \frac{1}{2}R_u \left(var_{\pi^*}(f) - var_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}(f)\right) + o(var_{\pi^*}(f)),$$

$$\mathsf{r}^f = \frac{1}{2}R_u var_{\pi^*}(f) + o(var_{\pi^*}(f)).$$

Note that condition (b) of Proposition 5.2 relaxes condition (b) of Proposition 5.1. As has already been remarked upon the convexity of Π is a strong condition, indeed apart from degenerate cases, it is incompatible with condition (b) of Proposition 5.1. Note that with $var_{\bar{\pi}}(f) = var_{\underline{\pi}}(f)$, $E_{\bar{\pi}}[f] \neq E_{\underline{\pi}}[f]$, it is easy to check that $var_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}(f)$ is uniquely maximized at $\alpha = 1/2$. Note also since $\alpha\underline{\pi} + (1-\alpha)\bar{\pi}$ is a mixture of distributions, one can view it as being generated by a random selection: α being interpreted as the probability that distribution $\underline{\pi}$ is selected. With this interpretation in place, an application of the law of total variance leads to the decomposition of $var_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}(f)$ into the expected conditional variance $\alpha var_{\underline{\pi}}(f) + (1-\alpha)var_{\bar{\pi}}(f)$ plus the variance of conditional expectations $var(E_{\pi}(f)) = \left(E_{\underline{\pi}}(f) - E_{\bar{\pi}}(f)\right)^2(\alpha - \alpha^2)$. Hence, $var_{\underline{\pi}}(f) = var_{\bar{\pi}}(f)$ implies the difference in variances $var_{\pi^*}(f) - var_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}(f)$ takes the simple form $\left(E_{\underline{\pi}}(f) - E_{\bar{\pi}}(f)\right)^2 (\frac{1}{2} - \alpha)^2$. We may write alternatively, therefore,

$$\mathsf{a}^f = z_\alpha - \frac{1}{2}R_u z_\alpha^2 + o(var_{\pi^*}(f)), \ z_\alpha = \left(\alpha - \frac{1}{2}\right)\left(E_{\bar{\pi}}[f] - E_{\underline{\pi}}[f]\right).$$

or

$$\mathsf{a}^f = E_{\pi^*}[f] - E_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}[f] - \frac{1}{2}R_u\left(E_{\pi^*}[f] - E_{\alpha\underline{\pi}+(1-\alpha)\bar{\pi}}[f]\right)^2 + o(var_{\pi^*}(f)).$$

5.2 Comparative statics of portfolio choice with more ambiguous (I)

A natural test-bed for the applicability of the more ambiguous characterizations is the standard one risky asset one safe asset portfolio problem analyzed by Arrow (1965). In our setting, we modify the risky asset to be one whose return embodies not only risk, but also ambiguity. The safe asset has neither risk nor ambiguity.

Let the act $h \in \mathcal{F}$, with degenerate lotteries as consequences, represent an uncertain asset, and a constant act with a degenerate lottery as consequence, $\mathfrak{f} \in \mathcal{F}$, represent the safe asset. The DM's objective is to select a portfolio share θ for the uncertain asset, in order to maximize the ex ante evaluation of her final wealth position. If initial wealth is w_0, the final wealth is determined by $w_1 = w_0(\theta h + (1 - \theta)\mathfrak{f}) = w_0(\mathfrak{f} + \theta f)$, where $f = h - \mathfrak{f} \in \mathcal{F}$ represents the excess return of the uncertain asset over the safe one. We assume no short-selling and that $w_1 \in \boldsymbol{X}$ for all $0 \leq \theta \leq 1$. Normalizing the utility so that $w_0 = 1, \mathfrak{f} = 0$, the program for α-MEU preferences is

$$\max_{\theta\in[0,1]} \left(\alpha \min_{\pi\in\Pi} \int_{\boldsymbol{X}} u(\theta x) dP_\pi^f(x) + (1-\alpha) \max_{\pi\in\Pi} \int_{\boldsymbol{X}} u(\theta x) dP_\pi^f(x) \right), \tag{32}$$

and for smooth ambiguity preferences it is

$$\max_{\theta \in [0,1]} \int_{\Delta} \phi \left(\int_{\boldsymbol{X}} u(\theta x) dP_{\pi}^f(x) \right) d\mu(\pi). \tag{33}$$

In the former case, the maximization program (32) defines the standard portfolio choice problem identified by $\left((\Pi, \alpha, u) ; \left(P_{\pi}^f \right)_{\pi \in \Pi} \right)$, a tuple of parameters representing the DM and the uncertain asset. Similarly, in the smooth ambiguity case, the standard portfolio choice problem is identified by the parameters $\left((\mu, \phi, u) ; \left(P_{\pi}^f \right)_{\pi \in \Pi} \right)$, where $\Pi = \mathrm{supp}(\mu)$. We suppose u is strictly concave in both cases, and that ϕ is strictly concave in the second. It follows that program (33) is concave in $\theta \in \mathbb{R}$ and strictly so in non-degenerate cases. The presence of the $\max_{\pi \in \Pi}$ operator in program (32) means that concavity is not in general assured. However, U-comonotonicity does imply that the program is concave. In this case, the maxima in the programs (33) and (32) are uniquely attained and denoted respectively $\theta^* \left((\Pi, \alpha, u) ; \left(P_{\pi}^f \right)_{\pi \in \Pi} \right)$ and $\theta^* \left((\mu, \phi, u) ; \left(P_{\pi}^f \right)_{\pi \in \Pi} \right)$.

We do not aim here to mirror the exhaustive study of the portfolio comparative statics problem which has been carried out for risk with expected utility preferences.[25] It is, however, convenient to adopt an assumption on preferences motivated by that literature. It is known for expected utility DMs that a first-order stochastic dominance improvement in the return of the risky asset will never lead to a smaller portfolio share of the risky asset only if the DM's preferences satisfies auxiliary conditions sufficient to imply that in terms of the normalized utility, $(x, \theta) \mapsto u(\theta x)$ is supermodular on $[0, 1] \times \boldsymbol{X}$. It suffices if the DM's utility (not necessarily normalized) has a coefficient of relative risk aversion bounded below unity[26].

Proposition 5.3 *Let* $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, *where* Π *is a compact, convex centrally symmetric subset of* Δ. *Suppose* Π *is* U_1-*comonotone for the acts* f *and* g. *If* $\{(\Pi, \alpha, u)\}_{\alpha \in [0,1]} \subset \mathcal{P}$ *has* u *strictly concave and the normalized utility* $(x, \theta) \mapsto u(\theta x)$ *is supermodular, then for* $\alpha \geq 0.5$, *if act* f *is* $(\mathcal{P})-$*more ambiguous (I) than act* g

$$\theta^* \left((\Pi, \alpha, u) ; \left(P_{\pi}^f \right)_{\pi \in \Pi} \right) \leq \theta^* \left((\Pi, \alpha, u) ; (P_{\pi}^g)_{\pi \in \Pi} \right). \tag{34}$$

To understand this result, note the key simplification induced by U-comonotonicity. Program (32) becomes

$$\max_{\theta \in [0,1]} \int_{\boldsymbol{X}} u(\theta x) dP_{\alpha \underline{\pi} + (1-\alpha)\overline{\pi}}^f \tag{35}$$

and the comparative static question reduces therefore to whether for each given $\alpha \in [0, 1]$ the change in the probability measure $P_{\alpha \underline{\pi} + (1-\alpha)\pi}^f$ induced by replacing f with a less ambiguous act causes the desired portfolio shift, i.e. no change for $\alpha = 0.5$, an increase

---

[25]See, e.g., Gollier (2011) for a study of the comparative statics of more ambiguity averse in the standard portfolio choice problem.

[26]These and other conditions are comprehensively discussed in Sections 4.5, 7.2 of Gollier (2001).

in the optimal θ for $\alpha > 0$ and a decrease for $\alpha < 0$. Using the characterization of m.a.(I) in Proposition 4.3 (i.e. condition (iii)), it is easy to confirm that providing a first-order stochastic dominance improvement in asset returns increases the demand for the risky asset in the Arrow (1965) expected utility portfolio problem, then increased ambiguity reduces the demand for the uncertain asset in the α-MEU portfolio problem.

The next proposition finds sufficient conditions for the comparative static to hold for the case of smooth ambiguity preferences. The key conditions are similar to those invoked for the result in the α-MEU case , however, the proof is more delicate and requires auxiliary assumptions on ϕ, specifically $-\phi''/\phi'$ is nonincreasing.

Proposition 5.4 *Let $\mathcal{P} = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$. Let $\Pi = \mathrm{supp}(\mu)$. Suppose Π is U_1-comonotone for the acts f and g. Suppose $\{(\mu, \phi, u)\}_{\phi \in \Phi_1} \subset \mathcal{P}$ has u strictly concave and the normalized utility $(x, \theta) \mapsto u(\theta x)$ is supermodular, ϕ is concave, $\frac{-\phi''}{\phi'}$ is nonincreasing. If act f is $\mathcal{P}-$more ambiguous (I) than act g and also satisfies condition (iii) of Proposition 4.3 then*

$$\theta^* \left((\mu, \phi, u) ; \left(P_\pi^f \right)_{\pi \in \Pi} \right) \leq \theta^* \left((\Pi, \phi, u) ; (P_\pi^g)_{\pi \in \Pi} \right).$$

Remark 5.1 In contrast to Proposition 4.3 it is *not* required that $\Pi = \mathrm{supp}(\mu)$ is convex.

Remark 5.2 An examination of the proof will show that an alternative to the condition that $\frac{-\phi''}{\phi'}$ is nonincreasing, is $\frac{-\phi''}{\phi'}$ is increasing with $\phi''' \leq 0$. This admits the class of quadratic ϕ.

5.3 Comparative statics of savings with m.a.(II)

To illustrate comparative statics using m.a.(II), we consider the following simple saving problem. The DM lives for two periods, has initial known wealth y_1 and will receive uncertain income Y_2 in period 2 generated by an act $f + a \in \mathcal{F}$. If the DM has α-MEU preferences, they are represented by

$$V_{\alpha, \Pi, u}^f(a) = u(y_1 - a) + \alpha \min_{\pi \in \Pi} \int u(y_2) dP_\pi^{f+a}(y_2) + (1 - \alpha) \max_{\pi \in \Pi} \int u(y_2) dP_\pi^{f+a}(y_2). \quad (36)$$

If the DM has smooth ambiguity preferences, they are represented by

$$V_{\mu, \phi, u}^f(a) = u(y_1 - a) + \phi^{-1} \left(\int \phi \left(\int u dP_\pi^{f+a} \right) d\mu(\pi) \right). \quad (37)$$

The problem is to choose savings[27] $a \in \mathbb{R}$ to maximize (36) or (37). We are interested in investigating the impact on savings of a compensated increase in uncertainty, specifically when f is replaced by g, where $f(\mathcal{P})$-$m.a.(II)$ g while maintaining the DM's standard

---

[27] In this section we ingnore niceties of boundedness of $\boldsymbol{X}$.

of living at the initial level of saving. Hence, we compare the cases, according to which preference obtains, in which g satisfies:

$$a_{\mathcal{M}}^f \in \arg\max_{a \in \mathbb{R}} V_{\mathcal{M}}^f(a), \ V_{\mathcal{M}}^f(a_{\mathcal{M}}^f) = V_{\mathcal{M}}^g(a_{\mathcal{M}}^f), \tag{38}$$

$$a_{\mathcal{S}}^f \in \arg\max_{a \in \mathbb{R}} V_{\mathcal{S}}^f(a), \ V_{\mathcal{S}}^f(a_{\mathcal{S}}^f) = V_{\mathcal{S}}^g(a_{\mathcal{S}}^f). \tag{39}$$

The assumptions in the following proposition will imply that $\arg\max_{a \in \mathbb{R}} V_{\mathcal{M}}^f(a)$ and $\arg\max_{a \in \mathbb{R}} V_{\mathcal{S}}^f(a)$ are uniquely attained.

Proposition 5.5 *Let $\Pi \subset \Delta$ be compact and U_1-comonotone for the pair $f, g \in \mathcal{F}$*

(a) *Let $\mathcal{P}_{\mathcal{M}} = \{(\alpha, \Pi, u)\}_{\phi \in \Phi_1, u \in U_1}$, where Π is convex. Consider a DM with an objective as given in equation (36) and with u strictly concave, CARA. Suppose f is $(\mathcal{P}_{\mathcal{M}})-$more ambiguous (II) than g and satisfies (38). Then $\arg\max_{a \in \mathbb{R}} V_{\mathcal{M}}^f(a) = \arg\max_{a \in \mathbb{R}} V_{\mathcal{M}}^g(a)$.*

(b) *Let $\mathcal{P}_S = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$. $\Pi = \text{supp}(\mu)$. Consider a DM with an objective as given in equation (37), with u strictly concave, CARA. Suppose also, ϕ is concave and satisfies the further condition $-\phi''/\phi'$ is decreasing concave. Suppose f is $(\mathcal{P}_S)-$more ambiguous (II) than g and satisfies (39). Then $\arg\max_{a \in \mathbb{R}} V_{\mathcal{M}}^g(a) \leq \arg\max_{a \in \mathbb{R}} V_{\mathcal{M}}^f(a)$.*

6 Concluding remarks

In closing we discuss, briefly, a couple of questions that appear to follow from the analysis in the paper. In the paper we have discussed a more ambiguous relation on acts and on events, but not on beliefs. It is natural to ask, "how is the optimal portfolio choice affected if the agent's *beliefs* were to become more ambiguous?" Consider the analogous question in a model with a *subjective* expected utility agent, "how is the optimal portfolio choice affected if the agent's *subjective* beliefs were to become more risky?" If we take "subjective beliefs" to mean the agent's (prior) belief on the state space, the question appears to be ill posed since, generally, the state space is not ordered in the way the outcome space is: a same change in prior may cause the distribution induced by an act f to become riskier while causing the distribution induced by another act g to become less risky. One may ask instead, "how is the optimal choice affected if the agent's subjective beliefs were to change such that the probability distribution on payoffs, induced by the uncertain asset (and subjective beliefs), is made riskier?" This question is meaningful, and the answer is the same as the answer to the question as to how optimal portfolio weights change going from one uncertain asset f to a *different* but riskier asset g, holding subjective *beliefs constant*. Consider the comparative static exercise for, say, an α-MEU agent with belief Π, of replacing one asset f with a more ambiguous asset g with corresponding induced distributions $\left(P_{\pi}^f\right)_{\pi \in \Pi}$ and $(P_{\pi}^g)_{\pi \in \Pi}$. As in the SEU case, this

exercise may be reinterpreted as showing the comparative static effect of a change in beliefs, from Π to Π', such that the induced distributions of payoffs of a *given* (uncertain) asset changes from $\left(P_\pi^f\right)_{\pi \in \Pi}$ to $\left(P_\pi^f\right)_{\pi \in \Pi'} = (P_\pi^g)_{\pi \in \Pi}$. That is, the distribution induced by the belief change is the same as that of a more ambiguous asset under unchanged beliefs. So, it is *as if* the belief change has engendered a more ambiguous asset.[28] Such a reinterpretation is one pragmatic response to the question of comparative statics of a "more ambiguous belief". And, under this alternative interpretation, since we are actually thinking of the *same* asset under different beliefs, it is compelling to make the comparison under the assumption of U-comonotonicity.

Our investigation has been static. In dynamic models, the relationship between information and ambiguity is likely to be a crucial element in dynamic decision making. The results in this paper suggest (more) ambiguous acts can be advantageous from the point of learning about the 'true' probability. Avoiding ambiguity in the short term may suppress learning and lead to more exposure to ambiguity in the future. Building models to explore these interactions will be an interesting challenge for the future.

A Appendix

A.1 Proofs of results in Section 2.3

Proof of Remark 2.2. $(\Delta, \leqslant_{E,E'})$ is a lattice if, for any pair $\pi, \pi' \in \Delta$, (a) there is a probability measure in Δ which assigns probability $\max\{\pi(E), \pi'(E)\}$ to event E and probability $\max\{\pi(E'), \pi'(E')\}$ to event E', (b) there is a probability measure in Δ which assigns probability $\min\{\pi(E), \pi'(E)\}$ to event E and probability $\min\{\pi(E'), \pi'(E')\}$ to event E'. We will establish (a), the proof of (b) follows a similar argument.

If $E = E'$ there is nothing to prove, therefore assume without loss of generality that $E'\backslash E \neq \emptyset$. If $E \subsetneq E'$, then all beliefs in Δ attach at least as much mass to E' as to E. Hence, $\max\{\pi(E'), \pi'(E')\} \geq \max\{\pi(E), \pi'(E)\}$ let $\pi'' \in \Delta$ be a probability measure which assigns mass $\max\{\pi(E), \pi'(E)\}$ to some state $s_1 \in E \cap E' = E$ and mass $\max\{\pi(E'), \pi'(E')\} - \max\{\pi(E), \pi'(E)\}$ to some state $s_2 \in E'\backslash E$ any remaining mass is assigned to some $s_3 \in \boldsymbol{S}\backslash(E \cup E')$. Hence, $\pi''(E) = \max\{\pi(E), \pi'(E)\}$, $\pi''(E') = \max\{\pi(E'), \pi'(E')\}$ as required.

If neither event contains the other and $\max\{\pi(E), \pi'(E)\} + \max\{\pi(E'), \pi'(E')\} \leq 1$, let π'' assign mass $m_1 = \max\{\pi(E), \pi'(E)\}$ to $s_1 \in E\backslash E'$, mass $m_2 = \max\{\pi(E'), \pi'(E')\}$ to $s_2 \in E'\backslash E$ and $m_3 = 1 - m_1 - m_2 \geq 0$ to $s_3 \in \boldsymbol{S}\backslash(E \cup E')$ (if $S = E \cup E'$, then $m_3 = 0$).

---

[28]Of course, in general, the same change in belief may well change the induced distribution for another asset in a way that the *initial* distribution corresponded to a characterization for more ambiguous, compared to the distribution following the change. However, there is an instance of a belief change following which induced distributions for *all* events change such that new distributions correspond to a more ambiguous event, for α-MEU preferences (under a given belief). Suppose that Π and Π' are centrally symmetric, share the same center, and there is a Markov Kernel from $(\Pi, \mathcal{B}_\Pi)$ to $(\Pi', \mathcal{B}_{\Pi'})$ with $\Pi' \subset \Pi, \pi' = \int_\Pi \pi dK_{\pi'}(\pi), \pi' \in \Pi'$; from which $\pi'(E) = \int_\Pi \pi(E) dK_{\pi'}(\pi), \pi' \in \Pi'$.

If neither event contains the other and $\max\{\pi(E), \pi'(E)\} + \max\{\pi(E'), \pi'(E')\} > 1$, let π'' assign mass m_1 to $s_1 \in E\backslash E'$, mass m_2 to $s_2 \in E'\backslash E$ and $m_3 = m_1 + m_2 - 1 > 0$ to $s_3 \in E \cap E'$ (recall, it is assumed $E \cap E' \neq \emptyset$) Choosing $m_1, m_2, m_3 \geq 0$ such that $m_1 + m_2 = \max\{\pi(E), \pi'(E)\}, m_2 + m_3 = \max\{\pi(E'), \pi'(E')\}$, which is always possible, establishes $\pi''(E) = \max\{\pi(E), \pi'(E)\}, \pi''(E') = \max\{\pi(E'), \pi'(E')\}$ as required. ∎

Proof of Proposition 2.1. Let $u_x \in U_1$ denote the simple step function, $u_x(x') = 0$ if $x' \leq x$, $u_x(x') = 1$ otherwise. The condition $\int u_x(f(s))d\pi_1(s) \leq \int u_x(f(s))d\pi_2(s)$ becomes $\pi_1(E_x^f) \geq \pi_2(E_x^f)$, similarly for act g. Hence if condition (ii) of the proposition holds, then $\pi_1(E_x^f) \geq \pi_2(E_x^f)$ and $\pi_1(E_x^g) \geq \pi_2(E_x^g)$ for each $\pi_1, \pi_2 \in \Pi$, $x \in \boldsymbol{X}$. This establishes that condition (i) implies (ii).

Conversely, suppose (i) holds, this implies there is a linear order $\leqq$ on Π such that $\pi_1 \leqq \pi_2 \Leftrightarrow \pi_1(E_x^f) \geq \pi_2(E_x^f)$ and $\pi_1(E_x^g) \geq \pi_2(E_x^g)$ for each $x \in \boldsymbol{X}$. A standard stochastic dominance argument completes the argument. Noting that, $\int_{\boldsymbol{S}} \sum_{i=1}^m \beta_i u_{x_i}(f)d\pi_1 = \sum_{i=1}^m \beta_i(1 - \pi_2(E_{x_i}^f))$ it is evident that $\beta_i \geq 0, i = 1, ..., m$, $\pi_2(E_x^f) \leq \pi_1(E_x^f)$, for each $x \in \boldsymbol{X}$ implies $\int_{\boldsymbol{S}} \sum_{i=1}^m \beta_i u_{x_i}(f)d\pi_2 \geq \int_{\boldsymbol{S}} \sum_{i=1}^m \beta_i u_{x_i}(f)d\pi_1$. For any $u \in U_1$, one can construct a sequence of positive linear combinations of the form $\sum_{i=1}^m \beta_i u_{x_i}$ which converges uniformly to u. Similarly for act g. Hence, $\pi_1 \leqq \pi_2$ implies $\int_{\boldsymbol{S}} u(h)d\pi_1 \leq \int_{\boldsymbol{S}} u(h)d\pi_2$, $u \in U_1, h \in \{f, g\}$ as required. ∎

A.2   Proofs of results in Section 3

Proof of Proposition 3.1. Let $u \in U_1$. Since $\Pi(E)$ is a compact interval and $[u(x)\pi(E) + u(y)(1 - \pi(E))]$ is linear in $\pi(E)$, $\min_{\pi \in \Pi}[u(x)\pi(E) + u(y)(1 - \pi(E))]$ and $\max_{\pi \in \Pi}[u(x)\pi(E) + u(y)(1 - \pi(E))]$ are attained at the two extreme points of $\Pi(E)$. For an ambiguity neutral element of the preference class with $\alpha = \frac{1}{2}$, $u \in U_1$, this implies

$$
\begin{aligned}
V_{\Pi, \frac{1}{2}, u}(xEy) &= \frac{1}{2} \min_{\pi \in \Pi}[u(x)\pi(E) + u(y)(1 - \pi(E))] + \frac{1}{2} \max_{\pi \in \Pi}[u(x)\pi(E) + u(y)(1 - \pi(E))] \\
&= u(x)\left(\frac{1}{2}\min\Pi(E) + \frac{1}{2}\max\Pi(E)\right) + u(y)\left(1 - \left(\frac{1}{2}\min\Pi(E) + \frac{1}{2}\max\Pi(E)\right)\right) \\
&= u(x)(\pi^*(E)) + u(y)(1 - \pi^*(E)).
\end{aligned}
$$

Similarly, $V_{\Pi, \frac{1}{2}, u}(xE'y) = u(x)(\pi^*(E')) + u(y)(1 - \pi^*(E'))$. This establishes that if $\pi^*(E) = \pi^*(E')$ all ambiguity elements of the preference class $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$ are indifferent between xEy and $xE'y$. Choosing $x, y \in \boldsymbol{X}$ and $u \in U_1$ such that $u(x) \neq u(y)$ shows the condition also to be necessary.

Using again the fact that $\Pi(E)$ and $\Pi(E')$ are compact intervals, it follows that the condition $\Pi(E) \subset \Pi(E')$ is equivalent to the condition $\min\Pi(E) \leq \min\Pi(E') \leq \max\Pi(E') \leq \max\Pi(E)$. Using this, it is straightforward to show (given $\pi^*(E) = \pi^*(E')$) that all preferences in the class $\mathcal{P}$ which are more ambiguity averse than than the ambiguity neutral element $\left(\Pi, \frac{1}{2}, u\right)$, that is elements of $\mathcal{P}$ with (Π, α, u), $\alpha \geq \frac{1}{2}$, weakly prefer $xE'y$ to xEy.

To see this, suppose $u(x) > u(y)$, $\alpha > \frac{1}{2}$, then

$$V_{\Pi,\alpha,u}(xEy) = \alpha \left[u(x)\min \Pi(E) + u(y)(1 - \min \Pi(E))\right]$$
$$+ (1-\alpha)\left[u(x)\max \Pi(E) + u(y)(1 - \max \Pi(E)))\right]$$
$$= \left[u(x)\max \Pi(E) + u(y)(1 - \max \Pi(E)))\right]$$
$$- \alpha\left([u(x)\max \Pi(E) + u(y)(1 - \max \Pi(E)))] - [u(x)\min \Pi(E) + u(y)(1 - \min \Pi(E))]\right)$$
$$= \left[u(x)\max \Pi(E) + u(y)(1 - \max \Pi(E)))\right] - \alpha\left([(u(x) - u(y))(\max \Pi(E) - \min \Pi(E))]\right).$$

Hence,

$$V_{\Pi,\alpha,u}(xEy) - V_{\Pi,\frac{1}{2},u}(xEy) = (\frac{1}{2} - \alpha)\left([(u(x) - u(y))(\max \Pi(E) - \min \Pi(E))]\right).$$

Similarly,

$$V_{\Pi,\alpha,u}(xE'y) - V_{\Pi,\frac{1}{2},u}(xE'y) = (\frac{1}{2} - \alpha)\left([(u(x) - u(y))\left(\max \Pi(E') - \min \Pi(E')\right)]\right).$$

Since, $V_{\Pi,\frac{1}{2},u}(xEy) = V_{\Pi,\frac{1}{2},u}((xE'y)$, it follows that $V_{\Pi,\alpha,u}(xEy) < V_{\Pi,\alpha,u}(xE'y)$ if and only if

$$\left((\max \Pi(E) - \min \Pi(E)) < \left(\max \Pi(E') - \min \Pi(E')\right)\right).$$

If $u(x) < u(y)$, the proof proceeds in the same way, $u(x) = u(y)$ is trivial. Likewise all preferences with $\alpha < \frac{1}{2}$ weakly prefer xEy to $xE'y$. This establishes the equivalence of conditions (i) and (iii) of the proposition.

The equivalence with (ii) is a simple application of Theorem 108 in Hardy, Littlewood and Polya (1952). Specifically, the conditions $\min \Pi(E) \leq \min \Pi(E') \leq \max \Pi(E') \leq \max \Pi(E)$, $\pi^*(E) = \pi^*(E')$ are equivalent to the statement that the vector $(\min \Pi(E), \max \Pi(E))$ *majorizes* the vector $(\min \Pi(E'), \max \Pi(E'))$, hence there exists a bistochastic matrix $B = \begin{bmatrix} b & 1-b \\ 1-b & b \end{bmatrix}$ such that $(\min \Pi(E'), \max \Pi(E')) = B(\min \Pi(E), \max \Pi(E))$. It follows that for any $\pi_\gamma \in \Pi$, $\pi_\gamma(E') = \gamma \min \Pi(E') + (1-\gamma)\max \Pi(E'), 0 \leq \gamma \leq 1$. $\pi_\gamma(E') = \gamma' \min \Pi(E) + (1-\gamma')\max \Pi(E)$ with $\gamma' = b\gamma + (1-b)(1-\gamma)$. Hence, $\pi_\gamma(E') = \int \pi(E)dK_\gamma(\pi)$, with K_γ placing mass $b\gamma + (1-b)(1-\gamma)$ on π such that $\pi(E) = \min \Pi(E)$ and the remaining mass on π such that $\pi(E) = \max \Pi(E)$. ∎

Proof of Proposition 3.2. Let $u \in U_1$, $x, y \in \mathbf{X}$. Setting $a = u(y)$, $b = u(x) - u(y)$, we can write

$$V_{\mu,\phi,u}(xEy) = \sum_{i=1}^{m} \phi\left(a + b\pi_i(E)\right)\mu_i, V_{\mu,\phi,u}(xE'y) = \sum_{i=1}^{m} \phi\left(a + b\pi_i(E')\right)\mu_i.$$

Hence, (i) implies that

$$\sum_{i=1}^{m} \phi\left(a + b\pi_i(E)\right)\mu_i \leq \sum_{i=1}^{m} \phi\left(a + b\pi_i(E')\right)\mu_i,$$

for all concave nondecreasing $\phi : u(\boldsymbol{L}) \to \mathbb{R}$. The inequality is required to hold with equality when ϕ is affine, corresponding to the case of ambiguity neutral preferences. If $b = 0$, there is nothing to prove, hence suppose $b > 0$ ($b < 0$ leads to an equivalent argument). Let $A = \{z \in \mathbb{R} \mid a + bz \in u(\boldsymbol{L})\} \neq \varnothing$. The condition can be stated equivalently as:

$$\sum \pi_i(E)\mu_i = \sum \pi_i(E')\mu_i$$
$$\sum \phi\left(\pi_i(E)\right)\mu_i \leq \sum \phi\left(\pi_i(E')\right)\mu_i$$

for all nondecreasing concave $\phi : A \to \mathbb{R}$. We wish to show that this is equivalent to condition (ii). That (ii) implies the condition is a direct implication of Jensen's inequality. The converse follows from Blackwell (1953)'s extension of Hardy, Littlewood, and Pólya (1929) (see *Theorem* in Sherman (1951)). To apply this theorem, we need to show that the condition implies

$$\sum \phi\left(\pi_i(E)\right)\mu_i \geq \sum \phi\left(\pi_i(E')\right)\mu_i$$

for all concave $\phi : A' \to \mathbb{R}$, for some interval $A' \subset A$ which contains each $\pi_i(E), \pi_i(E')$, $i = 1, ..., m$. Denoting the largest of the $\pi_i(E), \pi_i(E')$, $i = 1, ..., m$, by a, let $A' = A \cap (-\infty, a]$. Any concave function $\phi : A \to \mathbb{R}$ can be decomposed into the sum of a concave function which is nondecreasing on A' and an affine support function at a. This establishes the required condition and a direct application of the theorem (taking into account it is expressed in terms of convex, rather than concave functions) establishes the equivalence between (i) and (ii). ∎

Proof of Proposition 3.3. Let $x, y, p \in \boldsymbol{X}$, $x > y$, $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha=[0,1], u\in U_1}$. Applying the first part of Definition (2.7), we require for E $(\mathcal{P})$-m.a.(II) E' that $1 \geq \alpha' > \alpha \geq 0$, and

$$\alpha\min_{\pi\in\Pi}(\pi(E)u(x) + (1 - \pi(E))u(y)) + (1 - \alpha)\max_{\pi\in\Pi}(\pi(E)u(x) + (1 - \pi(E))u(y))$$
$$= \alpha\min_{\pi\in\Pi}(\pi(E')u(p) + (1 - \pi(E'))u(y)) + (1 - \alpha)\max_{\pi\in\Pi}(\pi(E')u(p) + (1 - \pi(E'))u(y))$$

implies

$$\alpha'\min_{\pi\in\Pi}(\pi(E)u(x) + (1 - \pi(E))u(y)) + (1 - \alpha')\max_{\pi\in\Pi}(\pi(E)u(x) + (1 - \pi(E))u(y))$$
$$\leq \alpha'\min_{\pi\in\Pi}(\pi(E')u(p) + (1 - \pi(E'))u(y)) + (1 - \alpha')\max_{\pi\in\Pi}(\pi(E')u(p) + (1 - \pi(E'))u(y)).$$

Equivalently, $x > y$, $u \in U_1$, $0 \leq \alpha \leq 1$, and

$$\left(\alpha\min\Pi(E) + (1 - \alpha)\max\Pi(E)\right)(u(x) - u(y)) \tag{40}$$
$$= \left(\alpha\min\Pi(E') + (1 - \alpha)\max\Pi(E')\right)(u(p) - u(y))$$

implies

$$\left(\max\Pi(E) - \min\Pi(E)\right)(u(x) - u(y)) \geq \left(\max\Pi(E') - \min\Pi(E')\right)(u(x') - u(y)). \tag{41}$$

Evidently, for each choice of $x > y$, the condition (40) can be met by some choice of $u \in U_1$ and $p \in \boldsymbol{X}$, which satisfies $u(x) > u(y)$. Hence, solving (42) for $(u(p) - u(y)) / (u(x) - u(y))$ and substituting into (41) we require, for all $0 \leq \alpha \leq 1$,

$$\frac{(\max \Pi(E) - \min \Pi(E))}{(\alpha \min \Pi(E) + (1 - \alpha) \max \Pi(E))} \geq \frac{(\max \Pi(E') - \min \Pi(E'))}{(\alpha \min \Pi(E') + (1 - \alpha) \max \Pi(E'))}. \quad (42)$$

This is easily seen to be equivalent to

$$\frac{\min \Pi(E')}{\max \Pi(E')} \geq \frac{\min \Pi(E)}{\max \Pi(E)}. \quad (43)$$

The second part of definition (2.7) simply repeats the first part except it applies to the complementary events, hence (43) becomes

$$\frac{\min\limits_{\pi \in \Pi} (1 - \pi(E'))}{\max\limits_{\pi \in \Pi} (1 - \pi(E'))} \geq \frac{\min\limits_{\pi \in \Pi} (1 - \pi(E))}{\max\limits_{\pi \in \Pi} (1 - \pi(E))},$$

i.e.

$$\frac{1 - \max \Pi(E')}{1 - \min \Pi(E')} \geq \frac{1 - \max \Pi(E)}{1 - \min \Pi(E)}.$$

This establishes the equivalence of (i) and (ii).

In order to complete the proof, we apply theorem 12.2.2 of Blackwell and Girshick (1954) applied to the dichotomy $\Omega = \{\bigvee_{E,E'} \Pi, \bigwedge_{E,E'}\}$. The information available to the decision maker depends on which of two alternative experiments she chooses to undertake. The first experiment consists of observing whether event E occurs or not. The second experiment is to observe whether E' occurs or not. Criterion *4.* of Blackwell and Girshick's theorem states that the first experiment is more informative than the second if it has a riskier likelihood ratio: specifically, for all convex $\upsilon : \mathbb{R} \to \mathbb{R}$,

$$\upsilon \left(\frac{\bigvee_{E,E'} \Pi(E)}{\bigwedge_{E,E'} \Pi(E)} \right) \bigwedge_{E,E'} \Pi(E) + \upsilon \left(\frac{1 - \bigvee_{E,E'} \Pi(E)}{1 - \bigwedge_{E,E'} \Pi(E)} \right) \left(1 - \bigwedge_{E,E'} \Pi(E) \right)$$

$$\geq \upsilon \left(\frac{\bigvee_{E,E'} \Pi(E')}{\bigwedge_{E,E'} \Pi(E')} \right) \bigwedge_{E,E'} \Pi(E') + \upsilon \left(\frac{1 - \bigvee_{E,E'} \Pi(E')}{1 - \bigwedge_{E,E'} \Pi(E')} \right) \left(1 - \bigwedge_{E,E'} \Pi(E') \right).$$

Since, when υ affine the inequality holds with equality and

$$\frac{\bigvee_{E,E'} \Pi(E)}{\bigwedge_{E,E'} \Pi(E)} \geq 1 \geq \frac{1 - \bigvee_{E,E'} \Pi(E)}{1 - \bigwedge_{E,E'} \Pi(E)}, \frac{\bigvee_{E,E'} \Pi(E')}{\bigwedge_{E,E'} \Pi(E')} \geq 1 \geq \frac{1 - \bigvee_{E,E'} \Pi(E')}{1 - \bigwedge_{E,E'} \Pi(E')}$$

trivially, this condition becomes equivalent to the following condition (a simple mean preserving spread)

$$\frac{\bigvee_{E,E'} \Pi(E)}{\bigwedge_{E,E'} \Pi(E)} \geq \frac{\bigvee_{E,E'} \Pi(E')}{\bigwedge_{E,E'} \Pi(E')}, \frac{1 - \bigvee_{E,E'} \Pi(E')}{1 - \bigwedge_{E,E'} \Pi(E')} \geq \frac{1 - \bigvee_{E,E'} \Pi(E)}{1 - \bigwedge_{E,E'} \Pi(E)}.$$

$$\frac{\bigvee_{E,E'}\Pi(E)}{\bigwedge_{E,E'}\Pi(E)} \geq \frac{\bigvee_{E,E'}\Pi(E')}{\bigwedge_{E,E'}\Pi(E')}, \frac{1-\bigvee_{E,E'}\Pi(E')}{1-\bigwedge_{E,E'}\Pi(E')} \geq \frac{1-\bigvee_{E,E'}\Pi(E)}{1-\bigwedge_{E,E'}\Pi(E)}.$$

Equivalently, by the existence of $\bigvee_{E,E'}$ and $\bigwedge_{E,E'}$,

$$\frac{\max\Pi(E)}{\min\Pi(E)} \geq \frac{\max\Pi(E')}{\min\Pi(E')}, \frac{1-\max\Pi(E')}{1-\min\Pi(E')} \geq \frac{1-\max\Pi(E)}{1-\min\Pi(E)}.$$

∎

The following Lemma is rather well known. The sufficiency part is implicit in e.g. Karlin and Novikoff (1963), or see e.g. Gollier (2001, Chapter 4) for a more explicit discussion.

Lemma A.1 *Let F and G be distribution functions with supports in an interval $\boldsymbol{I} \subset \mathbb{R}$. The following two conditions are equivalent.*

(a) $\int_{\boldsymbol{I}} \upsilon_A dF \geq \int_{\boldsymbol{I}} \upsilon_A dG \Rightarrow \int_{\boldsymbol{I}} \upsilon_B dF \geq \int_{\boldsymbol{I}} \upsilon_B dG$ for all integrable nondecreasing functions υ_A, $\upsilon_B : \boldsymbol{I} \to \mathbb{R}$ with υ_A more concave than υ_B (υ_A is a continuous concave transformation of υ_B).

(b) Single crossing. $\mathbb{R}$ can be partitioned into two intervals (one of which may be null), $\mathbb{R} = \boldsymbol{I}_1 \cup \boldsymbol{I}_2$, $\boldsymbol{I}_1 < \boldsymbol{I}_2$ such that $F \leq G$ on $\boldsymbol{I}_2$, $F \geq G$ on $\boldsymbol{I}_2$.

Proof. Let the random variable X have cdf F and Y have cdf G, denote the cdf of $\upsilon_A(X)$, by F_A and $\upsilon_A(Y)$ by G_A. Denote $F_A - G_A = H_A$. If (a) holds then, equivalently, $\int_{\upsilon_A(\boldsymbol{I})} v dH_A(v) \geq 0 \Rightarrow \int_{\upsilon_A(\boldsymbol{I})} \varphi(v) dH_A(v) \geq 0$ whenever φ is nondecreasing concave. Integration by parts gives the implication $\int_{\upsilon_A(\boldsymbol{I})} H_A(v) dv \leq 0 \Rightarrow \int_{\upsilon_A(\boldsymbol{I})} H_A(v) d\varphi(v) \geq 0$ and since φ is absolutely continuous, we may write $\int_{\upsilon_A(\boldsymbol{I})} H_A(v) dv \leq 0 \Rightarrow \int_{\upsilon_A(\boldsymbol{I})} H_A(v) \varphi'(v) dv \leq 0$ for some nonincreasing $\varphi' \geq 0$. Suppose H_A has a single sign change at v', from negative to positive, then $\int_{\upsilon_A(\boldsymbol{I})} H_A(v) (\varphi'(v') - \varphi'(v)) dv \geq 0$, so $\varphi'(v') \int_{\upsilon_A(\boldsymbol{I})} H_A(v) dv \geq \int_{\upsilon_A(\boldsymbol{I})} H_A(v) \varphi'(v) dv$. Evidently, if $\varphi'(v') > 0$, then $\int_{\upsilon_A(\boldsymbol{I})} H_A(v) dv < 0 \Rightarrow \int_{\upsilon_A(\boldsymbol{I})} H_A(v) \varphi'(v) dv < 0$. If $\varphi'(v') = 0$, then $\varphi' = 0$ on the interval where $H_A(v)$ is positive, evidently $\int_{\upsilon_A(\boldsymbol{I})} H_A(v) \varphi'(v) dv < 0$ unless $\varphi' = 0$ on the set where F differs from G. Necessity: if there are $x_1 < x_2 \in \boldsymbol{I}$ with $F(x_1) > G(x_1)$ and $F(x_2) < G(x_2)$ then with $\upsilon_A(x)$ defined to equal 0 on $x < x_1$, 1 on $x_1 \leq x < x_2$, $1+B$ on $x \geq x_2$, with $B = \frac{F(x_1)-G(x_1)}{G(x_2)-F(x_2)}$ one verifies that $\int_{\boldsymbol{I}} \upsilon_A dF = \int_{\boldsymbol{I}} \upsilon_A dG$. However, with $\varphi(v) = \min\{v, 1\}$, $\int_{\boldsymbol{I}} \varphi(\upsilon_A) dF = 1 - F(x_1) < \int_{\boldsymbol{I}} \varphi(\upsilon_A) dG = 1 - G(x_1)$, this contradicts (a). ∎

Lemma A.2 *Let F and G be distribution functions with supports in $\mathbb{R}_+$, and inverses $F^{-1} : (0,1) \to \mathbb{R}_+$, $F^{-1}(\xi) = \inf\{\eta \in \mathbb{R}_+ \mid F(\eta) > \xi\}$ and $G^{-1} : (0,1) \to \mathbb{R}_+$, $G^{-1}(\xi) = \inf\{\eta \in \mathbb{R}_+ \mid G(\eta) > \xi\}$. Let $\mathcal{C}$ the set of nondecreasing functions $\phi : \mathbb{R}_+ \to \mathbb{R}$. The following conditions are equivalent:*

(a) $\int \upsilon_A(\alpha\eta) dF(\eta) \geq \int \upsilon_A(\beta\eta) dG(\eta) \Rightarrow \int \upsilon_B(\alpha\eta) dF(\eta) \geq \int \upsilon_B(\beta\eta) dG(\eta)$, υ_A, $\upsilon_B \in \mathcal{C}$ with υ_A more concave than υ_B (υ_A is a concave transformation of υ_B), for all $\alpha, \beta > 0$.

(b) $G(\eta/\beta) - F(\eta/\alpha)$ has at most a single sign change which if one occurs is from negative to positive, for all $\alpha, \beta > 0$.

(c) $\eta \mapsto F^{-1}(G(\eta))$ is star-shaped $(F^{-1}(G(\lambda\eta)) \leq \lambda F^{-1}(G(\eta))$ for all $\lambda \in [0,1]$ and $\eta \in \mathbb{R}_+)$.

Proof. Condition (a) can be written after a change of variables as

$$\int \upsilon_A(\eta) dF(\eta/\alpha) \geq \int \upsilon_A(\eta) dG(\eta/\beta) \Rightarrow \int \upsilon_B(\eta) dF(\eta/\alpha) \geq \int \upsilon_B(\eta)(\eta) dG(\eta/\beta).$$

Hence, by Lemma A.1 the difference between the cdfs $G(\eta/\beta) - F(\eta/\alpha)$ must satisfy (b). The equivalence with (c), together with other equivalent conditions is asserted in Proposition C.11. in Marshall and Olkin (2007). ∎

Proof of Proposition 3.4. Definition 2.7 requires for $x > y \in \mathbf{X}$, $p \in \mathbf{X}$, that $xEy \succeq_A pEy \Rightarrow xEy \succeq_B pEy$ whenever the preference $\succeq_B$ is more ambiguity averse than the preference $\succeq_A$. For smooth ambiguity preferences this can be written in the notation introduced in section 3.2 as

$$\int \phi_A \left(\pi(E') \left(u(x) - u(y) \right) + u(y) \right) d\mu(\pi) \geq \int \phi_A \left(\pi(E) \left(u(p) - u(y) \right) + u(y) \right) d\mu(\pi) \tag{44}$$

implies

$$\int \phi_B \left(\pi(E') \left(u(x) - u(y) \right) + u(y) \right) d\mu(\pi) \geq \int \phi_B \left(\pi(E) \left(u(p) - u(y) \right) + u(y) \right) d\mu(\pi). \tag{45}$$

whenever $u \in U_1$, $\phi_A : u(\mathbf{L}) \rightarrow \mathbb{R}$, $\phi_B : u(\mathbf{L}) \rightarrow \mathbb{R}$, and ϕ_B is more concave than ϕ_A. Fixing some $u \in U_1$, with $u(x) - u(y) = \alpha > 0$ (if $u(x) = u(y)$ the implication is trivial), and noting that $u(p) - u(y) \leq 0$ satisfies the implication trivially by monotonicity, let $\beta = u(p) - u(y) > 0$. Hence, with $v_A(z) = \phi_A \left(z + u(y) \right), v_B(z) = \phi_B \left(z + u(y) \right)$ we require

$$\int \upsilon_A \left(\alpha\pi(E') \right) d\mu(\pi) \geq \int \upsilon_A \left(\beta\pi(E) \right) d\mu(\pi) \tag{46}$$

implies

$$\int \upsilon_B \left(\alpha\pi(E') \right) d\mu(\pi) \geq \int \upsilon_B \left(\beta\pi(E) \right) d\mu(\pi) \tag{47}$$

whenever v_B is more concave than v_A. Applying Lemma A.2 establishes that with

$$F(x) = \mu(\{\pi \mid \pi(E') \leq x\}), \ G(x) = \mu(\{\pi \mid \pi(E') \leq x\}), q \in [0,1],$$

$q \mapsto F^{-1}(G(q))$ is star-shaped. To complete the proof one repeats the exercise replacing E and E' by their respective complementary events. This establishes that $q \mapsto 1 - F^{-1}(G(1-q))$ is also star-shaped. Hence, $q \mapsto F^{-1}(G(q)) = \zeta(q)$ is doubly star-shaped.

Let Z be a random variable with distrubution function G. By star-shapedness, G is strictly increasing on its support, it follows that $G(Z) = W$ is uniformly distributed

on $[0, 1]$. Hence, $\Pr[\zeta(Z) \leq q] = \Pr[F^{-1}(G(Z)) \leq q] = \Pr[F^{-1}(W) \leq q] = F(q)$. These observations equate to condition (ii) of the proposition. ∎

Proof of Proposition 3.5. (i) $\Leftrightarrow$ (ii). This is a consequence of convexity and dimensionality. $\Pi-$comonotonicity implies that $\Pi_{E,E'} = \{(\pi(E), \pi(E')) \mid \pi \in \Pi\} \subset [0,1]^2$ is a nondecreasing arc in the unit square, therefore not space filling. Given that Π is closed convex, this arc must be the convex hull of two points. This set is doubly star-shaped if the set comprising the two extreme points is doubly star-shaped. Hence, the result follows immediately from the equivalence of conditions (i) and (ii) of Proposition 3.3.

(i) $\Leftrightarrow$ (iii). Condition (ii) and the fact that $\Pi_{E,E'}$ is the convex hull of two points imply the representation $\pi(E') = a + b\pi(E)$, for some real $a, b \geq 0, a + b \leq 1$ whenever $(\pi(E), \pi(E') \in \Pi$. Evidently, if $b = 0$, E' is a (completely) unambiguous event and the implication holds, therefore let $b > 0$. Hence, for each $q \in [0,1]$, $\mu(\{\pi \in \Pi \mid \pi(E) \leq q\}) = \mu(\{\pi \in \Pi \mid \pi(E') \leq \frac{q-a}{b}\})$. Since $q \to \frac{q-a}{b}$ is doubly star-shaped, Proposition 3.4 applies.

(i) $\Leftrightarrow$ (iv). Any closed compact subset $\Pi' \subset \Pi$ maps into a closed compact subset $\Pi'_{E,E'} \subset \Pi_{E,E'}$ which given (i) is evidently doubly star-shaped. $\bigvee_{E,E'} \Pi', \bigwedge_{E,E'} \Pi'$ both exist in Π' by comonotonicity and compactness of Π'. Moreover, $\Pi'_{E,E'}$ is the convex hull of $(\bigvee_{E,E'} \Pi'(E), \bigvee_{E,E'} \Pi'(E'))$ and $(\bigvee_{E,E'} \Pi'(E), \bigvee_{E,E'} \Pi'(E'))$. Given these observations, that E is Blackwell pairwise more informative than E' for the dichotomy $(\bigwedge_{E,E'} \Pi', \bigvee_{E,E'} \Pi')$ follows from the equivalence of (ii) and (iii) in proposition 3.3. The proof is completed by observing that for $\pi_1, \pi_2 \in \Pi$, we can without loss of generality, take $\pi_1 \leqq_{E,E'} \pi_2$, choosing Π' equal to the convex hull of $\{\pi_1, \pi_2\}$ evidently implies $(\bigvee_{E,E'} \Pi', \bigwedge_{E,E'} \Pi') = (\pi_1, \pi_2)$. Hence, the conclusion. ∎

A.3 Proofs of results in Section 4

Proof of Proposition 4.1. From the preference representation (2),

$$
\begin{aligned}
V_{\Pi,\alpha,u}(f) &= \alpha \min_{\pi \in \Pi} \int_S u(f) d\pi + (1 - \alpha) \max_{\pi \in \Pi} \int_S u(f) d\pi \\
&= \alpha \min_{\pi \in \Pi} \int_{\boldsymbol{X}} u dP_{\pi}^f + (1 - \alpha) \max_{\pi \in \Pi} \int_{\boldsymbol{X}} u dP_{\pi}^f.
\end{aligned}
$$

At $\alpha = 1/2$,

$$
V_{\Pi,\alpha,u}(f) = \int_S u(f) d \left(0.5 \underline{\pi}^u + 0.5 \overline{\pi}^u \right)
$$

where $\underline{\pi}^u$ and $\overline{\pi}^u$ respectively minimize and maximize the expected utility over Π. It follows from Π is centrally symmetric, that for any $u \in U_1$, $0.5 \underline{\pi}^u + 0.5 \overline{\pi}^u = \pi^\star$, so we have

$$
V_{\Pi,0.5,u}(f) = \int_S u(f) d\pi^\star = \int_{\boldsymbol{X}} u dP_{\pi^*}^f.
$$

45

Similarly,

$$V_{\Pi,0.5,u}(g) = \int_S u(g)d\pi^\star = \int_{\mathbf{X}} u dP^g_{\pi^*}.$$

Suppose that g is a center preserving π-garbling of f, then $P^f_{\pi^*} = P^g_{\pi^*}$ (equation (18)). Hence, $V_{\Pi,0.5,u}(f) = V_{\Pi,0.5,u}(g)$ for all the risk neutral elements of $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, i.e. part (a) of Definition 2.3 holds. To establish part (b) it will suffice to show that if g is a center preserving π-garbling of f,

$$\max_{\pi \in \Pi} \int_{\mathbf{X}} u dP^f_\pi - \min_{\pi \in \Pi} \int_{\mathbf{X}} u dP^f_\pi \geq \max_{\pi \in \Pi} \int_{\mathbf{X}} u dP^g_\pi - \min_{\pi \in \Pi} \int_{\mathbf{X}} u dP^g_\pi \qquad (48)$$

since this implies f is preferred (dispreferred) to g when $\alpha > 0.5$ ($\alpha < 0.5$). Evidently, since the Markov kernel $(\pi, C) \mapsto K_\pi(C)$ from $(\Omega, \mathcal{B}_\Omega)$ to $(\Omega, \mathcal{B}_\Omega)$ postulated in (11) 'averages' over Π rather than maximizes,

$$\max_{\pi \in \Pi} \int_{\mathbf{X}} u dP^f_\pi \geq \int_\Delta \left[\int_{\mathbf{X}} u dP^f_\pi \right] dK_{\pi'}(\pi), \pi' \in \Pi. \qquad (49)$$

Maximizing over the right hand side of (49) establishes $\max_{\pi \in \Pi} \int_{\mathbf{X}} u dP^f_\pi \geq \max_{\pi' \in \Pi} \int_\Delta \left[\int_{\mathbf{X}} u dP^f_\pi \right] dK_{\pi'}(\pi)$. The π-garbling condition (11) and T.16, p.16 of Meyer (1966) implies $\int_\Delta \left[\int_{\mathbf{X}} u dP^f_\pi \right] dK_{\pi'}(\pi) = \int_{\mathbf{X}} u dP^g_{\pi'}$. Hence, $\max_{\pi \in \Pi} \int_{\mathbf{X}} u dP^f_\pi \geq \max_{\pi' \in \Pi} \int_{\mathbf{X}} u dP^g_{\pi'}$. The same argument applied to min rather than max establishes $\min_{\pi \in \Pi} \int_\Delta u dP^f_\pi \leq \min_{\pi \in \Pi} \int_\Delta u dP^g_\pi$, hence (48) is established as required. ∎

Proof of Proposition 4.2. If act $g \in \mathcal{F}$ is a measure preserving $\pi-$garbling of $f \in \mathcal{F}$, then using (19) we have

$$
\begin{aligned}
V(f) &= \int_\Delta \phi \left(\int_{\mathbf{X}} u dP^f_\pi \right) d\mu \\
&= \int_\Delta \phi \left(\int_{\mathbf{X}} u dP^f_\pi \right) d \int_\Delta K_{\pi'}(\pi) d\mu(\pi') \\
&= \int_\Delta \left[\int_\Delta \phi \left(\int_{\mathbf{X}} u dP^f_\pi \right) dK_{\pi'}(\pi) \right] d\mu(\pi')
\end{aligned}
$$

(the second equality is justified e.g. by T.16, p. 16 of Meyer (1966)). Hence by Jensen's inequality and T.16 of Meyer (1966) again, it follows from (11) that

$$
\begin{aligned}
V(f) &\leq \int_\Delta \phi \left(\int_\Delta \int_{\mathbf{X}} u dP^f_\pi dK_{\pi'}(\pi) \right) d\mu(\pi') \\
&= \int_\Delta \phi \left(\int_{\mathbf{X}} u dP^g_\pi \right) d\mu.
\end{aligned}
$$

This shows all ambiguity averse preferences in $\mathcal{P}$ weakly prefer g to f. By the same argument all ambiguity seeking preferences weakly prefer f to g and all ambiguity neutral are indifferent. ∎

Proof of Remark 4.1. If there is a center preserving π-garbling of f into g, $P_\pi^g = \int_\Delta P_{\pi'}^f dK_\pi(\pi')$, $P_{\pi*}^f = P_{\pi*}^g$. Hence, using Jensen's inequality, for $\pi \in \Pi$

$$
\begin{aligned}
D(P_\pi^g || P_{\pi*}^g) &= D(\int_\Delta P_{\pi'}^f dK_\pi(\pi'') || P_{\pi*}^f) \\
&\leq \int_\Delta D(P_{\pi'}^f || P_{\pi*}^f) dK_\pi(\pi') \leq \max_{\pi \in \Pi} D(P_\pi^f || P_{\pi*}^f).
\end{aligned}
$$

Hence,

$$
\max_{\pi \in \Pi} D(P_\pi^g || P_{\pi*}^g) \leq \max_{\pi \in \Pi} D(P_\pi^f || P_{\pi*}^f).
$$

∎

Proof of Remark 4.2. Since D is convex, if there is a measure preserving π-garbling of f into g, we have

$$
\begin{aligned}
\int_\Delta D(P_{\pi'}^f || P_\pi^f) d\mu(\pi) &= \int_\Delta D(P_{\pi'}^f || P_\pi^f) d \int K_{\pi''}(\pi) d\mu(\pi'') \\
&\geq \int_\Delta D(P_{\pi'}^f || \int P_\pi^f dK_{\pi''}(\pi)) d\mu(\pi'') \\
&= \int_\Delta D(P_{\pi'}^f || P_{\pi''}^g) d\mu(\pi'').
\end{aligned}
$$

Repeating the argument gives

$$
\int_{\Delta \times \Delta} D(P_{\pi'}^f || P_\pi^f) d\mu \times \mu \geq \int_{\Delta \times \Delta} D(P_{\pi'}^g || P_\pi^g) d\mu \times \mu.
$$

That is, the μ-averaged K-L divergence is less for g than f. ∎

Proof of Proposition 4.3.

Recall, by convexity of Π, P_π^f, P_π^g are mixture linear (see Section 2). U_1-comonotonicity for the acts $f, g \in \mathcal{F}$ means that for each $x \in \mathbf{X}$, $\int u dP_\pi^g$ and $\int u dP_\pi^f$ are both non-decreasing on Π in the linear order $\leq_{U_1}$. Hence, since Π is compact, there exist top and bottom elements of Π, denoted respectively $\overline{\pi}$ and $\underline{\pi}$ such for all $u \in U_1$, $\pi \in \Pi$, $\int u dP_{\underline{\pi}}^f \leq \int u dP_\pi^f \leq \int u dP_{\overline{\pi}}^f$ and $\int u dP_{\underline{\pi}}^g \leq \int u dP_\pi^g \leq \int u dP_{\overline{\pi}}^g$. Hence, $\lambda : \Pi \to [0,1]$, defined by $\pi =_{U_1} \lambda(\pi)\overline{\pi} + (1 - \lambda(\pi))\underline{\pi}$, represents $\leq_{U_1}$ and $P_\pi^f = \lambda(\pi)P_{\overline{\pi}}^f + (1 - \lambda(\pi))P_{\underline{\pi}}^f$, $f \in \mathcal{F}$.

$(i) \iff (iii)$. If (i) holds, then for $u \in U_1$,

$$
\alpha \int u dP_{\underline{\pi}}^f + (1 - \alpha) \int u dP_{\overline{\pi}}^f \geq (\leq)\alpha \int u dP_{\underline{\pi}}^g + (1 - \alpha) \int u dP_{\overline{\pi}}^g \tag{50}
$$

whenever $\alpha \geq (\leq)0.5$. Choosing $u_x \in U_1$ equal to the unit step function at $x \in \mathbf{X}$: $u_x(x') = 0$ for $x' \leq x$, $u_x(x') = 1$ for $x' > x$ gives $\alpha P_{\underline{\pi}}^f(x) + (1 - \alpha)P_{\overline{\pi}}^f(x) \leq (\geq)$ $\alpha P_{\underline{\pi}}^g(x) + (1 - \alpha)P_{\overline{\pi}}^g(x)$ whenever $\alpha \geq (\leq)$ 0.5. Since Π is centrally symmetric, choosing

47

$\alpha = 0.5$ gives $P_{\pi^*}^f(x) = P_{\pi^*}^g(x)$. Choosing $\alpha \neq 0.5$ gives $P_{\underline{\pi}}^f(x) - P_{\overline{\pi}}^f(x) \geq P_{\underline{\pi}}^g(x) - P_{\overline{\pi}}^g(x)$. Since, $P_{\pi_1}^f(x) - P_{\pi_2}^f(x) = (\lambda(\pi_1) - \lambda(\pi_2))\left(P_{\overline{\pi}}^f(x) - P_{\underline{\pi}}^f(x)\right)$, $P_{\pi_1}^g(x) - P_{\pi_2}^g(x) = (\lambda(\pi_1) - \lambda(\pi_2))\left(P_{\overline{\pi}}^g(x) - P_{\underline{\pi}}^g(x)\right)$ it follows that for $\pi_1 \leqq_{U_1} \pi_2$ $P_{\pi_1}^f(x) - P_{\pi_2}^f(x) \geq P_{\pi_1}^g(x) - P_{\pi_2}^g(x)$. Repeating the argument for each $x \in \boldsymbol{X}$, establishes $(i) \Rightarrow (iii)$. The reverse implication follows from standard stochastic dominance arguments. Condition (iii) implies that for all $u \in U_1$, $\pi_1 \leqq_{U_1} \pi_2$, $\int u dP_{\pi^*}^f = \int u dP_{\pi^*}^g$, $\int u d\left(P_{\pi_1}^f - P_{\pi_2}^f\right) \leq \int u d\left(P_{\pi_1}^g - P_{\pi_2}^g\right)$, hence (50) holds, hence (i).

$(ii) \Longleftrightarrow (iii)$. Given that Π is compact convex and centrally symmetric, condition (ii) of Proposition 3.1 is clearly implied by condition (iii) of Proposition 4.3: set $\pi_1 = \underline{\pi}$, $\pi_2 = \overline{\pi}$ and observe $\Pi(E_x^f) = co\{\underline{\pi}(E_x^f), \overline{\pi}(E_x^f)\}$, $\Pi(E_x^g) = co\{\underline{\pi}(E_x^g), \overline{\pi}(E_x^g)\}$. Conversely, given that Π is compact convex and centrally symmetric, condition (ii) of Proposition 3.1 implies condition (iii) of Proposition 4.3 by the fact that $\pi =_{U_1} \lambda(\pi)\overline{\pi} + (1 - \lambda(\pi))\underline{\pi}$, represents $\leqq_{U_1}$. If $f, g \in \mathcal{F}$ map states into degenerate lotteries on $\boldsymbol{X}$, the equivalence follows immediately since then $P_\pi^f(x) = \pi(E_x^f)$, $x \in \boldsymbol{X}$. ∎

Lemma A.3 *Let* $\mathcal{P} = \{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U}$, *$\mu$ has support* Π*. Suppose that* Π *is U-comonotone for acts f and g. The following conditions are equivalent (we denote the extension of U to functions defined on $\boldsymbol{X} \times \Pi$ and constant in the π coordinate also by U)*

(i) f $(\mathcal{P})$-m.a.(I) g.

(ii)

$$\int_{\boldsymbol{X} \times C} u dP^{f,\mu} \leq \int_{\boldsymbol{X} \times C} u dP^{g,\mu}, \tag{51}$$

for all $u \in U$, and $C \in \Pi_L$.

Proof of Lemma A.3. Suppose (i) holds, i.e. for all nondecreasing concave (convex) $\phi : u(\boldsymbol{X}) \to \mathbb{R}$ we have

$$\int_\Pi \int_{\boldsymbol{X}} u dP_\pi^f d\mu = \int_\Pi \int_{\boldsymbol{X}} u dP_\pi^g d\mu \tag{52}$$

$$\int_\Pi \phi\left(\int_{\boldsymbol{X}} u dP_\pi^f\right) d\mu \leq (\geq) \int_\Pi \phi\left(\int_{\boldsymbol{X}} u dP_\pi^g\right) d\mu. \tag{53}$$

Rewriting (52) (KMM, Corollary 2) and using a generalization of Hardy, Littlewood and Polya's 'angles' theorem (1929, p.152) theorem (see e.g. Chong (1974), Theorem 2.5 and Corollary 1.8), this is equivalent to: for all $t \in \mathbb{R}$,

$$\int_{\boldsymbol{X} \times \Pi} u dP^{f,\mu} = \int_{\boldsymbol{X} \times \Pi} u dP^{g,\mu}$$

$$\int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^f - t\right)^- d\mu \leq \int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^g - t\right)^- d\mu,$$

where $(\xi)^- = \min\{\xi, 0\}$. Given U-comonotonicity, $\int u dP_\pi^f$ is $\leqq_U$-nondecreasing in π for all $u \in U$. Therefore, for any constant t, $\pi \longmapsto \int u dP_\pi^f - t$ has at most a single sign change as $\pi \leqq_U$-increases over Π, similarly for $g \in \mathcal{F}$. There exist, therefore, 'intervals' $I_t^f \subset \Pi$, which are either closed $I_t^f = \{\pi \in \Pi \mid \pi \leqq_U \pi_f\}$ or open $I_t^f = \{\pi \in \Pi \mid \pi < \pi_f\}$, for some $\pi_f \in \Pi$, such that

$$
\int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^f - t \right)^- d\mu \;=\; \int_{I_t^f} \left(\int_{\boldsymbol{X}} u dP_\pi^f - t \right) d\mu = \int_{I_t^f} \int_{\boldsymbol{X}} u dP_\pi^f - t\mu(I_t^f)
$$
$$
= \int_{\boldsymbol{X} \times I_t^f} u dP^{f,\mu} - t\mu(I_t^f),
$$

similarly for g

$$
\int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^g - t \right)^- d\mu = \int_{\boldsymbol{X} \times I_t^g} u dP^{g,\mu} - t\mu(I_t^g).
$$

If $\int_{\boldsymbol{X} \times I} u dP^{f,\mu} \leq \int_{\boldsymbol{X} \times I} u dP^{g,\mu}$ for each $I \in \Pi_L$, then evidently, for each $t \in \mathbb{R}$,

$$
\int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^f - t \right)^- d\mu \;\leq\; \int_{\boldsymbol{X} \times I_t^g} u dP^{f,\mu} - t\mu(I_t^g)
$$
$$
\leq\; \int_{\boldsymbol{X} \times I_t^g} u dP^{g,\mu} - t\mu(I_t^g)
$$
$$
=\; \int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^g - t \right)^- d\mu.
$$

This establishes sufficiency. For necessity, suppose that contrary to condition (ii) of the lemma there exist $u \in U$, $C \in \Pi_L$ such that

$$
\int_{\boldsymbol{X} \times C} u dP^{f,\mu} > \int_{\boldsymbol{X} \times C} u dP^{g,\mu}. \tag{54}
$$

Setting $\pi' = \sup C$ and $t = \int_{\boldsymbol{X}} u dP_{\pi'}^f$, (54) evidently implies

$$
\int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^f - t \right)^- d\mu(\pi) \;=\; \int_C \left(\int_{\boldsymbol{X}} u dP_\pi^f - t \right) d\mu(\pi)
$$
$$
=\; \int_{\boldsymbol{X} \times C} u dP^{f,\mu} - t\mu(J)
$$
$$
>\; \int_{\boldsymbol{X} \times C} u dP^{g,\mu} - t\mu(J)
$$
$$
\geq\; \int_\Pi \left(\int_{\boldsymbol{X}} u dP_\pi^g - t \right)^- d\mu(\pi).
$$

∎

Proof of Proposition 4.4. Apply Lemma A.3. The proof is a standard stochastic dominance argument. Choosing $u \in U_1$ to be simple step functions establishes necessity.

Approximating each $u \in U_1$ uniformly by a sequence of positive linear combinations of simple step functions establishes sufficiency. ∎

Proof of Proposition 4.5. $(f, g) \in SCP(\mathcal{P}_S) \Leftrightarrow$ Condition SCU. Equation (25) of SCU is equivalent to the existence of $(\lambda_1, \lambda_2) \gneqq 0$ such that

$$\lambda_1 \left[P_{\pi_1}^g(x) - P_{\pi_1}^f(x) \right] \leq \lambda_2 \left[P_{\pi_2}^g(x) - P_{\pi_2}^f(x) \right]. \tag{55}$$

Equivalently, with $(y_1(x), y_2(x)) = \left(P_{\pi_1}^g(x) - P_{\pi_1}^f(x), P_{\pi_2}^g(x) - P_{\pi_2}^f(x) \right)$, there is a hyperplane which separates $Y = \{ (y_1(x), y_2(x)) \mid x \in \mathbb{R} \} \subset \mathbb{R}^2$ from the orthant $\mathbb{O} = \{ (y_1, y_2) \in \mathbb{R}^2 \mid y_1 > 0, y_2 < 0 \}$. Absent such a hyperplane, by the separating hyperplane theorem and Caratheodory's theorem, there is a convex combination of at most three points in Y which lies in $\mathbb{O}$. Hence, denoting these points x_1, x_2, x_3 with weights $\gamma_1, \gamma_2, \gamma_3$ and denoting by $u_{x_i} \in U_1$ the step function taking values 1 on $x \geq x_i$, 0 elsewhere, $u = \sum_i \gamma_i u_i \in U_1$, we have

$$\int_{\boldsymbol{X}} u dP_{\pi_1}^f \quad > \quad \int_{\boldsymbol{X}} u dP_{\pi_1}^g = a \tag{56}$$

$$\int_{\boldsymbol{X}} u dP_{\pi_2}^f \quad < \quad \int_{\boldsymbol{X}} u dP_{\pi_2}^g = b. \tag{57}$$

If Π is path connected, it follows from U_1-monotonicity that there is an open interval containing π_1 for which $\int_{\boldsymbol{X}} u dP_{\pi_1}^g \leq a$ but not $\int_{\boldsymbol{X}} u dP_{\pi_1}^f \leq a$. Since $\pi \mapsto \int u dP_\pi^f, \pi \mapsto \int u dP_\pi^g$ are continuous functionals and μ is a strictly positive measure on Π (recall, it is assumed that $\text{supp}(\mu) = \Pi$),

$$\mu \left(\left\{ \pi \in \Pi \mid \int_{\boldsymbol{X}} u dP_{\pi_1}^f \leq a \right\} \right) \quad < \quad \mu \left(\left\{ \pi \in \Pi \mid \int_{\boldsymbol{X}} u dP_{\pi_1}^g \leq a \right\} \right)$$

$$\mu \left(\left\{ \pi \in \Pi \mid \int_{\boldsymbol{X}} u dP_{\pi_1}^f \leq b \right\} \right) \quad > \quad \mu \left(\left\{ \pi \in \Pi \mid \int_{\boldsymbol{X}} u dP_{\pi_1}^g \leq b \right\} \right).$$

The same conclusion evidently obtains if Π is finite. These inequalities express a single crossing condition on distribution functions of expected utilities. Hence, applying Lemma A.1, there are nondecreasing ϕ_1, ϕ_2 with ϕ_2 a concave transformation of ϕ_1 such that

$$\int_\Delta \phi_1 \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) d\mu \quad > \quad \int_\Delta \phi_1 \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) d\mu \tag{58}$$

$$\int_\Delta \phi_2 \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) d\mu \quad < \quad \int_\Delta \phi_2 \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) d\mu. \tag{59}$$

This contradicts $SCP(\mathcal{P}_S)$, hence Condition SCU (25) is necessary for $(f, g) \in SCP(\mathcal{P}_S)$.

Sufficiency follows similarly, if Condition SCU holds, there is a hyperplane which separates Y from $\mathbb{O}$. It follows that the system of inequalities (56) and (57) does not obtain for any $u \in U_1$ of the form $u = \sum_i^n \gamma_i u_i$, with $\gamma_i > 0$ and u_i a nondecreasing step

50

function, $i = 1, ..., n$. Standard approximation arguments implies the system does not hold for any $u \in U_1$. Applying Lemma A.1 again, the system of inequalities (58) and (59) never obtain. Hence, $(f, g) \in SCP(\mathcal{P}_S)$ holds.

$(f, g) \in SCP(\mathcal{P}_M) \Leftrightarrow$ Condition SCU. Since Π is compact and linearly ordered, there exist top and bottom elements, respectively $\overline{\pi}, \underline{\pi} \in \Pi$ such for all $u \in U_1$, $\pi \in \Pi$, $\int u dP_{\underline{\pi}}^f \leq \int u dP_{\pi}^f \leq \int u dP_{\overline{\pi}}^f$ and $\int u dP_{\underline{\pi}}^g \leq \int u dP_{\pi}^g \leq \int u dP_{\overline{\pi}}^g$. Hence, if (ii) holds, then for each $u \in U_1$, with $\pi_1 = \underline{\pi}$, $\pi_2 = \overline{\pi}$.

$$
\begin{aligned}
&\alpha \int_{\boldsymbol{X}} u dP_{\pi_1}^f + (1 - \alpha) \int_{\boldsymbol{X}} u dP_{\pi_2}^f - \left(\alpha \int_{\boldsymbol{X}} u dP_{\pi_1}^g + (1 - \alpha) \int_{\boldsymbol{X}} u dP_{\pi_2}^g \right) \\
&= \alpha \left(\int_{\boldsymbol{X}} u dP_{\pi_1}^f(x) - \int_{\boldsymbol{X}} u dP_{\pi_1}^g \right) + (1 - \alpha) \left(\int_{\boldsymbol{X}} u dP_{\pi_2}^f(x) - \int_{\boldsymbol{X}} u dP_{\pi_2}^g \right) \\
&= \alpha A + (1 - \alpha) B
\end{aligned}
$$

has (at most) single sign change from negative to positive as α increases from 0 to 1. This rules out the configuration $A < 0$, $B > 0$ but all others are admissible. Choosing u to be step functions and arbitrary convex combinations of step functions requires, therefore as for the $SCP(\mathcal{P}_S)$ case, that Y be separated from $\mathbb{O}$. This establishes $(f, g) \in SCP(\mathcal{P}_M) \Leftrightarrow$ equation (25) with $\pi_1 = \underline{\pi}$, $\pi_2 = \overline{\pi}$. Using the convexity of Π, and the fact that P_{π}^f, P_{π}^g are mixture linear, the equivalence is seen to extend to all $\pi_1 \leqq_{U_1} \pi_2 \in \Pi$. Hence, $(f, g) \in SCP(\mathcal{P}_M) \Leftrightarrow$ Condition SCU. ∎

Proof of Proposition 4.6. Sufficiency of the condition follows immediately from the fact that for each fixed $p \in \mathbb{R}$, with $|p| \leq |\boldsymbol{J}|$, condition (iii) of the Proposition implies Condition SCU holds with f replaced by $f + p$. Application of Proposition 4.5 implies $(f + p, g) \in SCP(\mathcal{P}_S)$ and $(f + p, g) \in SCP(\mathcal{P}_M)$, which is the desired result. Necessity is equally immediate, if condition (iii) fails to hold then there exists some $p \in \mathbb{R}$, with $|p| \leq |\boldsymbol{J}|$, such that $(f + p, g) \notin SCP(\mathcal{P}_S)$ and $(f + p, g) \notin SCP(\mathcal{P}_M)$. ∎

Proof of Proposition 4.7. To establish the conditions (i) and (i) suffice for f $(\mathcal{P}_S)$-m.a.(II) g. it suffices (by Lemma A.1) to establish that for each $p \in \mathbb{R}$, with $|p| \leq |\boldsymbol{J}|$, the map $\pi \mapsto \int_{\boldsymbol{X}} u dP_{\pi}^{f+p} - \int_{\boldsymbol{X}} u dP_{\pi}^g$ has at most a single sign change as π increases in the $\leqq_{U_1}$ order, which, if one occurs, is from negative to positive. For act g we may express the expected utility $\int_X u dP_{\pi}^g$ in terms of the quantile function as

$$
\int_{\boldsymbol{X}} u dP_{\pi}^g = \int_{\boldsymbol{J}} u dP_{\pi}^g = \int_0^1 u(Q_{\pi}^g(\xi)) d\xi, \pi \in \Pi.
$$

Using the fact that $x \mapsto P_{\pi}^f(x)$ is absolutely continuous with respect to Lebesgue measure λ, we express expected utility for act f as

$$
\int_{\boldsymbol{J}} u dP_{\pi}^{f+p} = \int_{\boldsymbol{J}} u p_{\pi}^{f+p} d\lambda, \pi \in \Pi, \tag{60}
$$

and the following further change of variables is permissible

$$\int_0^1 u(Q_\pi^g(\xi))d\xi = \int_{\boldsymbol{J}} u(Q_\pi^g(P_\pi^{f+p}(\zeta)))p_\pi^{f+p}(\zeta)d\lambda(\zeta).$$

Hence, we seek conditions under which the map $\pi \mapsto \int_{\boldsymbol{J}} \left[u(\zeta) - u(Q_\pi^g(P_\pi^{f+p}(\zeta)))\right] p_\pi^{f+p}(\zeta)d\lambda(\zeta)$ has a single change of sign in the manner prescribed above. Specifically, we require for $\pi_1 \leqq_{U_1} \pi_2$

$$\int_{\boldsymbol{J}} \left[u(\zeta) - u(Q_{\pi_1}^g(P_{\pi_1}^{f+p}(\zeta)))\right] p_{\pi_1}^{f+p}(\zeta)d\lambda(\zeta) \geq 0 \Rightarrow$$
$$\int_{\boldsymbol{J}} \left[u(\zeta) - u(Q_{\pi_2}^g(P_{\pi_2}^{f+p}(\zeta)))\right] p_{\pi_2}^{f+p}(\zeta)d\lambda(\zeta) \geq 0. \tag{61}$$

This is established this by establishing in turn the implication

$$\int_{\boldsymbol{J}} \left[u(\zeta) - u(Q_{\pi_1}^g(P_{\pi_1}^{f+p}(\zeta)))\right] p_{\pi_1}^{f+p}(\zeta)d\lambda(\zeta) \geq 0 \Rightarrow \int_{\boldsymbol{J}} \left[u(\zeta) - u(Q_{\pi_1}^g(P_{\pi_1}^{f+p}(\zeta)))\right] p_{\pi_2}^{f+p}(\zeta)d\lambda(\zeta) \geq 0, \tag{62}$$

and the inequality

$$u(Q_{\pi_1}^g(P_{\pi_1}^{f+p}(\zeta))) \geq u(Q_{\pi_2}^g(P_{\pi_2}^{f+p}(\zeta))), \; \zeta \in \boldsymbol{J}. \tag{63}$$

Implication (62) is established by noting: 1. f is Bickel-Lehmann more dispersed than g, this implies that $\zeta \mapsto \left[u(\zeta) - u(Q_{\pi_1}^{g+q}(P_{\pi_1}^{f+p}(\zeta)))\right]$ has at most a single sign change, if one occurs it is from negative to positive. 2. P_π^{f+p} has monotone likelihood ratio. Application of Karlin (1968)'s Theorem 3.1 gives the result (or see Jewitt (1987), Athey (2002)). Inequality (63) follows from the fact that $u \in U_1$ is nondecreasing, and that act f is assumed Lehmann more informative than act g, hence by (10) $Q_{\pi_1}^g(P_{\pi_1}^{f+p}(\zeta)) \geq Q_{\pi_2}^{g+}(P_{\pi_2}^{f+p}(\zeta))$, $\zeta \in \boldsymbol{X}$. This establishes the claim that conditions (i) and (ii) imply f $(\mathcal{P}_S)$-m.a.(II) g. That f $(\mathcal{P}_M)$-m.a.(II) g follows immediately from the equivalence established in Proposition 4.6. ■

A.4   Proofs of results in Section 5

Proof of Proposition 5.1.   The ambiguity premium and total uncertainty premium corresponding to an act f, denoted a^f and u^f, respectively, is defined implicitly as follows:

$$\int_\Delta \phi\left(\int_{\boldsymbol{X}} u dP_\pi^f\right) d\mu = \phi\left(\int_{\boldsymbol{X}} u\left(x - \mathsf{a}^f\right) d\int_\Delta P_\pi^f d\mu\right) \tag{64}$$
$$\int_\Delta \phi\left(\int_{\boldsymbol{X}} u dP_\pi^f\right) d\mu = \phi\left(u\left(E[f] - \mathsf{u}^f\right)\right) \tag{65}$$

Consider the l.h.s. of equation (65). Under the assumption (Pratt (1964)) that the absolute third central moments of P_π^f are of smaller order than $var_\pi[f]$ for each $\pi \in \Pi$,

$$\int_\Delta \phi \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) d\mu = \int_\Delta v \left(u^{-1} \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) \right) d\mu$$

$$= \int v \left(E_\pi[f] - \frac{1}{2} R_u(E_\pi[f])) var_\pi[f] + o(var_\pi[f]) \right) d\mu(\pi),$$

where the second equality is the Arrow-Pratt approximation for certainty equivalents. Note by the law of total variance and assumption that $var_\pi[f]$ is constant on $supp\mu$, if $var[f]$ is small, then $var_\pi[f]$ is smaller, hence we may substitute $o(var[f])$ for $o(var_\pi[f])$, similarly for $var(E_\pi[f])$. Given R_u has a continuous second derivative, by Taylor's theorem, $R_u(E_\pi[f]) = R_u(E[f]) + R_u'(E[f])(E_\pi[f] - E[f]) + o(E_\pi[f]) - E[f])$. It follows that

$$var \left(E_\pi[f] - \frac{1}{2} R_u(E_\pi[f])) var_\pi[f] + o(var_\pi[f]) \right)$$

$$= var(E_\pi[f]) \left(1 + R_u'(E[f]) var_\pi[f] \right) var(E_\pi[f]) + o(E[var_\pi[f]])$$

$$= var(E_\pi[f]) + o(var[f]).$$

Using these facts in applying the Arrow-Pratt approximation again now gives

$$\int_\Delta \phi \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) d\mu = v \left(E[f] - \frac{1}{2} R_u(E[f])) var_\pi[f] - \frac{1}{2} R_v(E[f])) var(E_\pi[f]) + o(var[f]) \right).$$

From (64) similar arguments give

$$\int_\Delta \phi \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) d\mu = v \left(E[f] - \mathsf{u}^f - \frac{1}{2} R_u(E[f])) var[f] + o(var[f]) \right).$$

Trivially, $\phi \left(u \left(E[f] - \mathsf{u}^f \right) \right) = v(E[f] - \mathsf{u}^f)$. Hence, equating terms and using (29), one obtains

$$\mathsf{u}^f = \frac{1}{2} R_u var_\pi(f) + \frac{1}{2} R_v var(E_\pi[f]) + o(var(f))$$

$$\mathsf{a}^f = \frac{1}{2} (R_v - R_u) var(E_\pi[f]) + o(var(f))$$

$$\mathsf{r}^f = \mathsf{u}^f - \mathsf{a}^f = \frac{1}{2} R_u var(f) + o(var(f)),$$

where R_v and R_u are evaluated at $E[f]$. ∎

Proof of Proposition 5.2. Since, $\alpha \min_{\pi \in \Pi} \int u dP_\pi^f + (1 - \alpha) \max_{\pi \in \Pi} \int u dP_\pi^f = \int u dP_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}^f$, the ambiguity and uncertainty premia are defined respectively by

$$\int u dP_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}^f = \int u(x - \mathsf{a}^f) dP_{\pi^\star}^f(x),$$

$$\int u dP_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}^f = u(E_{\pi^\star}[f] - \mathsf{c}^f).$$

In terms of the Arrow-Pratt certainty equivalents, since $var_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}(f) \leq var_{0.5\underline{\pi}+0.5\overline{\pi}}(f) = var_{\pi^\star}(f)$,

$$E_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}[f] - \frac{1}{2}var_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}(f) = E_{\pi^\star}[f] - \mathsf{a}^f - \frac{1}{2}var_{\pi^\star}(f) + o\left(var_{\pi^\star}(f)\right),$$

$$E_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}[f] - \frac{1}{2}var_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}(f) = E_{\pi^\star}[f] - \mathsf{c}^f + o\left(var_{\pi^\star}(f)\right)$$

Hence,

$$\mathsf{a}^f = E_{\pi^\star}[f] - E_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}[f] - \frac{1}{2}\left(var_{\pi^\star}(f) - var_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}(f)\right) + o\left(var_{\pi^\star}(f)\right),$$

$$\mathsf{c}^f = E_{\pi^\star}[f] - E_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}[f] + \frac{1}{2}var_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}(f) + o\left(var_{\pi^\star}(f)\right).$$

∎

Proof of Proposition 5.3. U_1-comonotonicity implies the objective function may be written as

$$\alpha \min_{\pi\in\Pi}\int_{\boldsymbol{X}}u(\theta x)dP^f_\pi(x) + (1-\alpha)\max_{\pi\in\Pi}\int_{\boldsymbol{X}}u(\theta x)dP^f_\pi(x) = \int_{\boldsymbol{X}}u(\theta x)dP^f_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}$$

$$= \int_0^1 u(\theta Q^f_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}(p))dp.$$

If $(\theta, x) \mapsto u(\theta x)$ is supermodular on the lattice $[0,1] \times \boldsymbol{X}$ (with the partial order (θ, x) larger than (θ', x') if $\theta \geq \theta'$ and $x \geq x'$), then $(\theta, P) \mapsto \int_{\boldsymbol{X}} u(\theta x)dP(x)$ is easily seen to be supermodular on the lattice $[0,1]\times\boldsymbol{L}$ (with (θ, P) larger than (θ', P') if $\theta \geq \theta'$ and P first-order stochastically dominates P'). Condition (iii) implies that for $\alpha > \frac{1}{2}$, $P^g_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}$ first order stochastically dominates $P^f_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}$, conversely for $\alpha < \frac{1}{2}$, $P^f_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}$ first order stochastically dominates $P^g_{\alpha\underline{\pi}+(1-\alpha)\overline{\pi}}$. The result follows immediately. ∎

Remark A.1 *In light of Proposition 4.4, the following equivalence is a restatement of Tchen (1980) (also, Epstein and Tanny (1980)). Let $\mathrm{supp}(\mu)$ be U_1-comonotone for the pair $f, g \in F$. f m.a.(II) g if and only if for all supermodular functions $\nu : \boldsymbol{X} \times \Delta \to \mathbb{R}$,*

$$\int_{\boldsymbol{X}\times\Delta} \nu dP^{f,\mu} \geq \int_{\boldsymbol{X}\times\Delta} \nu dP^{g,\mu}.$$

Proof of Proposition 5.4. At the portfolio share $\theta^* = \theta^*\left((\mu, \phi, u); (P^g_\pi)_{\pi\in\Pi}\right) \geq 0$, the following first-order condition holds

$$\int_{\Delta\times\boldsymbol{X}} \phi'\left(\int_{\boldsymbol{X}} u dP^g_\pi\right) u'\left(\theta^* x\right) x dP^{g,\mu} = 0. \tag{66}$$

54

It suffices for the result that (66) implies

$$\int_{\Delta \times \boldsymbol{X}} \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) u' \left(\theta^* x \right) x dP^{f,\mu} \leq 0,$$

since, by concavity the θ satisfying the first-order condition for f must be greater than θ^*. Since $u'(w + \theta^* x) x$ is nondecreasing in x and $\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right)$ is nonincreasing in $\pi \in \Pi$, by concavity and U_1-monotonicity, it follows from $P^{f,\mu}$ more ambiguous (I) than $P^{g,\mu}$ (Remark (A.1)) that (66) implies

$$\int_{\Delta \times \boldsymbol{X}} \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) u' \left(\theta^* x \right) x dP^{f,\mu} \leq 0.$$

Hence, it suffices to establish the implication

$$\int_{\Delta \times \boldsymbol{X}} \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) u' \left(\theta^* x \right) x dP^{f,\mu} \leq 0$$

$$\Rightarrow \int_{\Delta \times \boldsymbol{X}} \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) u' \left(\theta^* x \right) x dP^{f,\mu} \leq 0.$$

This will be achieved by showing that there exists $\lambda \geq 0$ such that

$$\int_{\Delta \times \boldsymbol{X}} \left(\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) - \lambda \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) \right) u' \left(\theta^* x \right) x dP^{f,\mu} \leq 0. \tag{67}$$

To this end, choose $\lambda = \lambda^* > 0$ so that

$$\int_{\Delta \times \boldsymbol{X}} \left(\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) - \lambda^* \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) \right) dP^{f,\mu} = 0. \tag{68}$$

It follows from the assumptions that ϕ' is convex, hence from f m.a.(I) g it follows that

$$\int_\Delta \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) d\mu \geq \int_\Delta \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) d\mu,$$

therefore $\lambda^* \geq 1$. Hence, since ϕ' is decreasing, $\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) \geq \lambda^* \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right)$ implies $\int_{\boldsymbol{X}} u dP_\pi^f \leq \int_{\boldsymbol{X}} u dP_\pi^g$. Let $\zeta(\pi) = \int_{\boldsymbol{X}} u dP_\pi^f - \int_{\boldsymbol{X}} u dP_\pi^g$, ζ is a nondecreasing function by assumption. If ϕ' is logconvex, then $\frac{\phi'(\eta - \zeta)}{\phi'(\eta)}$ is nondecreasing in η for all $\zeta < 0$ and, since ϕ' is decreasing, is nondecreasing in ζ. It follows that, for $\pi' \geqq \pi$,

$$\frac{\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f - \zeta(\pi) \right)}{\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right)} \leq \frac{\phi' \left(\int_{\boldsymbol{X}} u dP_{\pi'}^f - \zeta(\pi) \right)}{\phi' \left(\int_{\boldsymbol{X}} u dP_{\pi'}^f \right)} \leq \frac{\phi' \left(\int_{\boldsymbol{X}} u dP_{\pi'}^f - \zeta(\pi') \right)}{\phi' \left(\int_{\boldsymbol{X}} u dP_{\pi'}^f \right)}.$$

Hence,

$$\left[\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) - \lambda^* \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) \right] = \left[\frac{\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right)}{\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f - \zeta(\pi) \right)} - \lambda^* \right] \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right)$$

55

has at most a single sign change which is from positive to negative if one occurs. The rest of the proof is standard. Since $\pi \mapsto \int_{\boldsymbol{X}} u'(\theta^* x)\, x dP_\pi^f$ is nondecreasing, there exists a $k \in \mathbb{R}$ such that

$$\int_\Delta \left(\phi' \left(\int_{\boldsymbol{X}} u dP_\pi^f \right) - \lambda^* \phi' \left(\int_{\boldsymbol{X}} u dP_\pi^g \right) \right) \left(\int_{\boldsymbol{X}} u'(w + \theta^* x)\, x dP_\pi^f - k \right) d\mu \leq 0.$$

Using (68) this is easily seen to imply (67) as required. ∎

Proof of Proposition 5.5. For the case of α-MEU preferences, one obtains, by the envelope theorem, the first order condition

$$u'(y_1 - a) = \alpha \int u'(y_2 + a) dP_{\underline{\pi}}^f(y_2) + (1 - \alpha) \int u'(y_2 + a) dP_{\overline{\pi}}^f(y_2)), \qquad (69)$$

this condition uniquely determines the optimum given u is strictly concave. For the CARA case, compensation

$$\alpha \int u dP_{\underline{\pi}}^f + (1 - \alpha) \int u dP_{\overline{\pi}}^f = \alpha \int u dP_{\underline{\pi}}^g + (1 - \alpha) \int u dP_{\overline{\pi}}^g \qquad (70)$$

means the first order condition is unchanged

$$\alpha \int u' dP_{\underline{\pi}}^f + (1 - \alpha) \int u' dP_{\overline{\pi}}^f = \alpha \int u' dP_{\underline{\pi}}^g + (1 - \alpha) \int u' dP_{\overline{\pi}}^g.$$

As noted in Gierlinger and Gollier (2008), Hardy, Littlewood, and Pólya (1952)'s generalization of Minkowski's inequality, that the generalized mean $v \mapsto \mathfrak{M}_\varphi(v) = \varphi^{-1} \left(\int \varphi(v(\pi)) d\mu(\pi) \right)$ is concave if $x \mapsto -\varphi'(x)/\varphi''(x)$ is a concave function. Hence, with $v_1(\pi) = \int u(y_2 + s_1) dP_\pi^f(y_2)$, $v_2(\pi) = \int u(y_2 + s_2) dP_\pi^f(y_2)$, $v_\lambda(\pi) = \int u(y_2 + \lambda s_1 + (1 - \lambda)s_2) dP_\pi^f(y_2)$ we have $\lambda \mathfrak{M}_\phi(v_1) + (1 - \lambda)\mathfrak{M}_\phi(v_2) \leq \mathfrak{M}_\phi(\lambda v_1 + (1 - \lambda)v_2)$. It follows that if u is strictly concave, $a \mapsto V_{\mathcal{S}}^f(a)$ is strictly concave. The first-order conditions are (having set the optimal saving $a = 0$ without loss of generality) are

$$u'(y_1) = \frac{\int \phi' \left(\int u dP_\pi^f \right) \left(\int u' dP_\pi^f \right) d\mu}{\phi' \left(\phi^{-1} \left(\int \phi \left(\int u dP_\pi^f \right) d\mu \right) \right)}.$$

We wish to show that for an m.a.(II) compensated increase in ambiguity

$$u'(y_1) \leq \frac{\int \phi' \left(\int u dP_\pi^g \right) \left(\int u' dP_\pi^g \right) d\mu}{\phi' \left(\phi^{-1} \left(\int \phi \left(\int u dP_\pi^g \right) d\mu \right) \right)}$$

since, a strict inequality will require a reduction in saving to restore the first order constraint. That is, we wish to establish

$$\int \phi(U_f) d\mu = \int \phi(U_g) d\mu \Rightarrow$$
$$\int \phi'(U_f) U_f' d\mu \geq \int \phi'(U_g) U_g' d\mu. \qquad (71)$$

where $U_f = \int u dP_\pi^f$, $U_f' = \int u' dP_\pi^f$ and $U_g = \int u dP_\pi^g$, $U_g' = \int u' dP_\pi^g$. Using CARA, this becomes

$$\int \phi\left(U_f\right) d\mu = \int \phi\left(U_g\right) d\mu \Rightarrow$$

$$\int \phi'\left(U_f\right) U_f d\mu \leq \int \phi'\left(U_g\right) U_g d\mu.$$

Note that the function $\widetilde{\phi}(U) = \phi'(U)U$ is a nondecreasing concave transformation of ϕ on $\mathbb{R}_-$. Increasing is immediate, the concavity part can be seen from the fact that the ratio of derivatives $\frac{\phi''(U)U + \phi'(U)}{\phi'(U)} = \frac{-\phi''(U)}{\phi'(U)}(-U) + 1$ is the product of two positive decreasing function and therefore decreasing. The result now follows since a compensated increase in m.a.(II) satisfying the equality condition in (71) implies that $\int \widetilde{\phi}\left(U_f\right) d\mu \leq \int \widetilde{\phi}\left(U_g\right) d\mu$ for any $\widetilde{\phi}$ which is a concave transformation of ϕ. ∎

A.5 Characterizing more ambiguous (II) acts without $U-$comonotonicity

A.5.1 $\alpha\text{-}MEU$ and smooth ambiguity

As in Proposition 2.1 we begin by restricting attention to acts f and g having degenerate lotteries as consequences, $E_x^g = \{s \in \mathbf{S} \mid g(s) \leq x\}$ and $E_x^f = \{s \in \mathbf{S} \mid f(s) \leq x\}$. For any m-vector $x = (x_i, ..., x_m) \in \mathbf{X}^m, i = 1, ..., m$, the set $\{(\pi(E_{x_1+\rho}^f), ..., \pi(E_{x_m+\rho}^f)) \mid \pi \in \Pi\}$ is a closed convex subset of $[0,1]^m$ which since $E_{x_i-\infty}^f = \emptyset$ and $E_{x_i+\infty}^f = \mathbf{S}$ expands from $\{(0, ..., 0)\}$ and then contracts to $\{(1, ..., 1)\}$ as ρ traverses the real line from $-\infty$ to ∞. For any given $\beta \in [0,1]^m$, $\beta \cdot x = \tau \in [0,1]$ defines a hyperplane. Denote $cl_\beta[Z]$ as the union of all such hyperplanes which have a non-null intersection with $Z \subset [0,1]^m$.

Proposition A.1 *Let $\mathcal{P} = \{(\Pi, \alpha, u)\}_{\alpha \in [0,1], u \in U_1}$, where Π is a compact, convex subset of Δ. In the case that $f, g \in \mathcal{F}$ are acts mapping states into degenerate lotteries over outcomes in $\mathbf{X}$, the following conditions are equivalent.*

(i) Act f is $(\mathcal{P})-$more ambiguous (II) than act g;

(ii) There is no $m \in \mathbb{N}, x \in \mathbf{X}^m, \beta \in [0,1]^m, \rho \in \mathbb{R}$ such that

$$cl_\beta[\{(\pi(E_{x_1+\rho}^f), ..., \pi(E_{x_m+\rho}^f)) \mid \pi \in \Pi\} \subsetneqq cl_\beta[\{(\pi(E_{x_1}^g), ..., \pi(E_{x_m}^g)) \mid \pi \in \Pi\}$$

however, there do exist $m \in \mathbb{N}, x \in \mathbf{X}^m, \beta \in [0,1]^m, \rho \in \mathbb{R}$ such that

$$cl_\beta[\{(\pi(E_{x_1}^g), ..., \pi(E_{x_m}^g)) \mid \pi \in \Pi\} \subsetneqq cl_\beta[\{(\pi(E_{x_1+\rho}^f), ..., \pi(E_{x_m+\rho}^f)) \mid \pi \in \Pi\}.$$

Although this proposition gives a reasonably intuitive characterization of increased ambiguity in terms of a collection of possible event probabilities becoming "enlarged", it will be hard to apply in specific circumstances

Proposition A.2 *Let $\mathcal{P}$ be the class of smooth ambiguity preferences $\{(\mu, \phi, u)\}_{\phi \in \Phi_1, u \in U_1}$, the following conditions are equivalent.*

(i) *The act f is $(\mathcal{P})-$more ambiguous (II) than act g.*

(ii) *For each $u \in U_1$, $\xi_1, \xi_2, \rho \in \mathbb{R}, \xi_1 < \xi_2$,*

$$\mu\left(\left\{\pi \in \Delta \mid \int_X u(g(s) + \delta_\rho)d\pi \leq \xi_1\right\}\right) \quad \geq \quad (>)\mu\left(\left\{\pi \in \Delta \mid \int_X u(f(s) + \delta_\rho)d\pi \leq \xi_1\right\}\right) \ implies$$

$$\mu\left(\left\{\pi \in \Delta \mid \int_X u(g(s) + \delta_\rho)d\pi \leq \xi_2\right\}\right) \quad \geq \quad (>)\mu\left(\left\{\pi \in \Delta \mid \int_X u(f(s) + \delta_\rho)d\pi \leq \xi_2\right\}\right).$$

References

AHN, D. (2008): "Ambiguity without a state space," *The Review of Economic Studies*, 75(1), 3.

ARROW, K. (1965): *Aspects of the theory of risk-bearing.* Yrjö Jahnssonin Säätiö.

ATHEY, S. (2002): "Monotone comparative statics under uncertainty," *The Quarterly Journal of Economics*, 117(1), 187.

BERGER, J. (1985): *Statistical decision theory and Bayesian analysis.* Springer.

BICKEL, P., AND E. LEHMANN (1976): "Descriptive statistics for nonparametric models. III. Dispersion," *The Annals of Statistics*, 4(6), 1139–1158.

BLACKWELL, D. (1953): "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, 24(2), 265–272.

BLACKWELL, D., AND M. GIRSHICK (1954): *Theory of games and statistical decisions.* John Wiley & Sons.

BOYLE, P., L. GARLAPPI, R. UPPAL, AND T. WANG (2011): "Keynes Meets Markowitz: The Trade-Off Between Familiarity and Diversification," *Management Science*, p. mnsc.1110.1349.

CABALLERO, R., AND A. KRISHNAMURTHY (2008): "Collective risk management in a flight to quality episode," *The Journal of Finance*, 63(5), 2195–2230.

CHATEAUNEUF, A., M. COHEN, AND I. MEILIJSON (2004): "Four notions of mean-preserving increase in risk, risk attitudes and applications to the rank-dependent expected utility model," *Journal of Mathematical Economics*, 40(5), 547–571.

CHONG, K. M. (1974): "Some extensions of a theorem of Hardy, Littlewood and Pólya and their applications," *Canadian Journal of Mathematics*, 26(6), 1321–1340.

COLLARD, F., S. MUKERJI, J.-M. TALLON, AND K. SHEPPARD (2011): "Ambiguity and the Historical Equity Premium," Working Paper, Department of Economics 550, Oxford University, Oxford, U.K.

DOKSUM, K. (1969): "Starshaped Transformations and the Power of Rank Tests," *The Annals of Mathematical Statistics*, 40(4), 1167–1176.

EICHBERGER, J., S. GRANT, D. KELSEY, AND G. KOSHEVOY (2011): "The [alpha]-MEU model: A comment," *Journal of Economic Theory*.

ELLSBERG, D. (1961): "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics*, 75, 643–669.

EPSTEIN, L., AND M. SCHNEIDER (2008): "Ambiguity, information quality, and asset pricing," *The Journal of Finance*, 63(1), 197–228.

EPSTEIN, L., AND S. TANNY (1980): "Increasing generalized correlation: a definition and some economic consequences," *The Canadian Journal of Economics*, 13(1), 16–34.

EPSTEIN, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579–608.

ERGIN, H., AND F. GUL (2009): "A subjective theory of compound lotteries," *Journal of Economic Theory*, 144(3), 899–929.

FERGUSON, T. (1967): *Mathematical statistics: A decision theoretic approach*, vol. 7. Academic Press New York.

GHIRARDATO, P., F. MACCHERONI, AND M. MARINACCI (2004): "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118, 133–173.

GHIRARDATO, P., AND M. MARINACCI (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251–289.

GIERLINGER, J., AND C. GOLLIER (2008): "Socially efficient discounting under ambiguity aversion," Working paper 561, IDEI.

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

GOLLIER, C. (2001): *The economics of risk and time*. The MIT Press, MA.

——— (2011): "Portfolio choices and asset prices: The comparative statics of ambiguity aversion," *The Review of Economic Studies*.

GRANT, S., AND J. QUIGGIN (2005): "Increasing uncertainty: a definition," *Mathematical Social Sciences*, 49(2), 117–141.

HANSEN, L., AND T. SARGENT (2010): "Fragile beliefs and the price of uncertainty," *Quantitative Economics*, 1(1), 129–162.

HANSEN, L. P. (2007): "Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk," *American Economic Review*, 97(2), 1–30.

HARDY, G., J. LITTLEWOOD, AND G. PÓLYA (1952): *Inequalities*. Cambridge Univ. Press, Cambridge, 2nd edn.

HARDY, G. H., J. E. LITTLEWOOD, AND G. PÓLYA (1929): "Some simple inequalities satisfied by convex functions," *Mess. of Math.*, 58, 145–152.

HURWICZ, L. (1951): "Optimality criteria for decision making under ignorance," *Cowles commission papers*, 370.

IZHAKIAN, Y., AND S. BENNINGA (2008): "The uncertainty premium in an ambiguous economy," mimeo., Tel-Aviv University.

JEWITT, I. (1987): "Risk aversion and the choice between risky prospects: the preservation of comparative statics results," *The Review of Economic Studies*, 54(1), 73.

——— (1989): "Choosing between risky prospects: the characterization of comparative statics results, and location independent risk," *Management Science*, 35(1), 60–70.

JEWITT, I. (2011): "Information in decision and agency problems," mimeo.

KARLIN, S. (1968): *Total positivity*. Stanford Univ Pr.

KARLIN, S., AND A. NOVIKOFF (1963): "Generalized convex inequalities.," *Pacific Journal of Mathematics*, 13(4), 1251–1279.

KARLIN, S., AND H. RUBIN (1956): "The theory of decision procedures for distributions with monotone likelihood ratio," *The Annals of Mathematical Statistics*, 27(2), 272–299.

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), 1849–1892.

KLIBANOFF, P., S. MUKERJI, AND SEO (2011): "Relevance and Symmetry," Working Paper Series 539, Department of Economics, Oxford University.

KNIGHT, F. (1921): *Risk, uncertainty and profit*. Houghton Miffin.

KULLBACK, S., AND R. LEIBLER (1951): "On information and sufficiency," *The Annals of Mathematical Statistics*, 22(1), 79–86.

LANDSBERGER, M., AND I. MEILIJSON (1994): "The generating process and an extension of Jewitt's location independent risk concept," *Management Science*, 40(5), 662–669.

LEHMANN, E. (1988): "Comparing location experiments," *The Annals of Statistics*, 16(2), 521–533.

MACCHERONI, F., M. MARINACCI, AND D. RUFFINO (2010): "Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis," Working paper, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

MACHINA, M., AND D. SCHMEIDLER (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780.

MARINACCI, M. (2002): "Probabilistic sophistication and multiple priors," *Econometrica*, 70(2), 755–764.

MARSHALL, A., AND I. OLKIN (2007): *Life distributions*. Springer, New York.

MEYER, P. A. (1966): *Probability and Potentials*. Blaisdell Publishing Company, MA.

NAU, R. (2006): "Uncertainty aversion with second-order utilities and probabilities," *Management Science*, 52(1), 136.

NEHRING, K. (2001): "Ambiguity in the Presence of Probabilistic Beliefs," Discussion paper, University of California, Davis.

NEILSON, W. (2010): "A simplified axiomatic approach to ambiguity aversion," *Journal of Risk and Uncertainty*, pp. 1–12.

PRATT, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.

ROGERS, A., AND M. J. RYAN (2008): "Lower Probabilities of Centrally Symmetric Sets," mimeo., University of Auckland.

ROTHSCHILD, M., AND J. STIGLITZ (1970): "Increasing risk: I. A definition," *Journal of Economic theory*, 2(3), 225–243.

SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57(3), 571–587.

SEGAL, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202.

SEO, K. (2009): "Ambiguity and Second-Order Belief," *Econometrica*, 77(5), 1575–1605.

SHERMAN, S. (1951): "On a theorem of Hardy, Littlewood, Polya, and Blackwell," *Proceedings of the National Academy of Sciences of the United States of America*, 37(12), 826.

SKIADAS, C. (2009): "Smooth ambiguity toward small risks and continuous-time recursive utility," mimeo., Northwestern University.

SKLAR, A. (1959): "Fonctions de répartition à n dimensions et leurs marges," *Publ. Inst. Statist. Univ. Paris*, 8, 229–231.

STRASSER, H. (1985): *Mathematical theory of statistics: statistical experiments and asymptotic decision theory*. Walter de Gruyter.

TCHEN, A. (1980): "Inequalities for distributions with given marginals," *The Annals of Probability*, 8(4), 814–827.

UHLIG, H. (2010): "A model of a systemic bank run," *Journal of Monetary Economics*, 57(1), 78–96.

WALD, A. (1949): "Statistical decision functions," *The Annals of Mathematical Statistics*, 20(2), 165–205.

YAARI, M. (1969): "Some remarks on measures of risk aversion and on their uses," *Journal of Economic theory*, 1(3), 315–329.