# A Smooth Model of Decision Making Under Ambiguity[1]

Peter Klibanoff
MEDS, Kellogg School of Management, Northwestern University
peterk@kellogg.northwestern.edu

Massimo Marinacci
Dip. di Statistica e Matematica Applicata, Università di Torino, and ICER
massimo@econ.unito.it

Sujoy Mukerji
Dept. of Economics, University of Oxford, and ICER
sujoy.mukerji@economics.ox.ac.uk

September 2002
This Version: April 2003

[1]We thank N. Al-Najjar, E. Dekel, D. Fudenberg, S. Grant, Y. Halevy, F. Maccheroni, M. Machina, S. Morris, R. Nau, K. Nehring, K. Roberts, U. Segal, A. Shneyerov, J-M. Tallon, R. Vohra, P. Wakker and especially L. Epstein and I. Gilboa, for helpful discussions and suggestions. We also thank seminar audiences at Chicago, Columbia, Duke, Michigan, the Minneapolis Fed, the Econometric Society N. American Winter Meetings '03, the RUD '02 conference, the Spring '02 Midwest Econ. Theory Conference, Northwestern's Zell Center and CMS-EMS. Mukerji thanks MEDS at Northwestern University and ICER at the University of Torino for their hospitality during the visits when part of the research was completed. Mukerji gratefully acknowledges financial support from the ESRC Research Fellowship Award award R000 27 1065, and Marinacci gratefully acknowledges the financial support of MIUR.

**Abstract**

We propose and axiomatize a model of preferences over acts such that the decision maker prefers act $f$ to act $g$ if and only if $\mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ f) \geq \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ g)$, where $\mathbb{E}$ is the expectation operator, $u$ is a vN-M utility function, $\phi$ is an increasing transformation, and $\mu$ is a subjective probability over the set $\Pi$ of probability measures $\pi$ that the decision maker thinks are relevant given his subjective information. A key feature of our model is that it achieves a separation between ambiguity, identified as a characteristic of the decision maker's subjective information, and ambiguity attitude, a characteristic of the decision maker's tastes. We show that attitudes towards risk are characterized by the shape of $u$, as usual, while attitudes towards ambiguity are characterized by the shape of $\phi$. We also derive $\phi(x) = -\frac{1}{\alpha}e^{-\alpha x}$ as the special case of constant ambiguity aversion. Ambiguity itself is defined behaviorally and is shown to be characterized by properties of the subjective set of measures $\Pi$. This characterization of ambiguity is formally related to the definitions of subjective ambiguity advanced by Epstein-Zhang (2001) and Ghirardato-Marinacci (2002). One advantage of this model is that the well-developed machinery for dealing with risk attitudes can be applied as well to ambiguity attitudes. The model is also distinct from many in the literature on ambiguity in that allows smooth, rather than kinked, indifference curves. This leads to different behavior and improved tractability, while still sharing the main features (e.g., Ellsberg's Paradox, etc.). The Maxmin EU model (e.g., Gilboa and Schmeidler (1989)) with a given set of measures may be seen as an extreme case of our model with infinite ambiguity aversion. Two illustrative applications to portfolio choice are offered.

**JEL Classification Numbers**: D800, D810.

Keywords: Ambiguity, Uncertainty, Knightian Uncertainty, Ambiguity Aversion, Uncertainty Aversion, Ellsberg Paradox.

Corresponding Author: Peter Klibanoff, MEDS Dept., Jacobs Center Rm 548, 2001 Sheridan Road, Evanston, IL 60208. peterk@kellogg.northwestern.edu

# 1 Introduction

Savage's axiom P2, often referred to as "the Sure Thing Principle", states that, if two acts are equal on a given event, then it should not matter (for ranking the acts in terms of preferences) what they are equal to on that event. Even though an "extra-logical principle", in the sense that it is not derived as a logical consequence of more primitive principles, it would find ready acceptance in many circumstances. It has been observed, however, that there is at least one kind of circumstance where a decision maker (DM) might find the principle less persuasive – if the DM were worried by cognitive or informational constraints that leave him uncertain about what odds apply to the payoff relevant events. Ellsberg (1961) presented examples inspired by this observation; the following table is a stylized description of one of those examples. The table shows four acts, $f, g, f'$ and $g'$, with payoffs contingent on three (mutually exclusive and exhaustive) events, $A$, $B$ and $C$.

| | $A$ | $B$ | $C$ |
|---|---|---|---|
| $f$ | 10 | 0 | 0 |
| $g$ | 0 | 10 | 0 |
| $f'$ | 10 | 0 | 10 |
| $g'$ | 0 | 10 | 10 |

Note that P2 implies, if $f$ is preferred to $g$ then $f'$ is preferred to $g'$. Consider a situation where the DM "knows" that the probability of event $A$ occurring is 1/3, though he has no information about how the complementary probability, 2/3, is "divided" between $B$ and $C$. The DM decides to choose $f$ over $g$ *but* $g'$ over $f'$, justifying his choice as follows. Taking into account his perception of the environment, he calculates the expected utility from $f$, $\mathbb{E}u(f) = u(10) \times 1/3$ while he thinks $\mathbb{E}u(g)$ could equal any number in the interval $[u(10) \times 0, u(10) \times 2/3]$; similarly, he calculates $\mathbb{E}u(g') = u(10) \times 2/3$ but realizes ex ante evaluations for $f'$, $\mathbb{E}u(f')$, could be any number in the interval $[u(10) \times 1/3, u(10)]$. He has some aversion to uncertainty about ex ante evaluations: he worries that he may take the "wrong" decision ex ante because he has a relatively vague idea as to what the true probability is and is therefore prone to make a "wrong" evaluation. Hence, his choices. The normative case for the violation of P2 in the circumstance where the DM is bothered by his uncretainty with respect to the relevant priors is, of course, debatable. What (we think) is beyond debate is that, in such a circumstance many DMs would, on deliberation, find it justified to violate P2. This paper presents a model of decision making which can explicitly reflect the circumstance that the DM is (subjectively) uncertain about the priors relevant to his decision. The model allows for the relaxation of P2 specifically and exclusively under such a circumstance, so that behavior, given the uncertainty about ex ante evaluation, may display aversion (or love) for that uncertainty along the lines of the justification discussed in the above example.

The paper first presents a set of assumptions on individual preferences over Savage acts, leading to a functional representation for these preferences. Following that we give results which show how different components of the functional identify elements such as the DM's subjective beliefs and the DM's attitude toward the various uncertainties in the problem and thus the precise links between behavior and these elements. We define behaviorally what it means for a DM's belief about an event to be ambiguous and go on to show that, in our model, this definition is essentially equivalent to the

DM being uncertain about the probability of the event. Thus we are able to identify ambiguity with uncertainty/multiplicity with respect to relevant priors and hence, ex ante evaluations. Correspondingly, we define attitude towards ambiguity as an attitude towards this uncertainty in a way analogous to the standard notion of attitude toward risk. This highlights one advantage of our approach: familiar techniques from the theory of risk attitudes may be used to deal with ambiguity as well, while maintaining a clear distinction between the two.

Preferences axiomatized in this paper are shown to be represented by a functional of the following double expectational form

$$V(f) = \int_{\Delta} \phi \left( \int_{S} u(f)\, d\pi \right) d\mu \equiv \mathbb{E}_{\mu} \phi \left( \mathbb{E}_{\pi} u \circ f \right),$$

where $f$ is a real valued function defined on a state space $S$ (an "act"); $u$ is a von Neumann-Morgenstern utility function; $\pi$ is a probability measure on $S$; $\phi$ is a map from reals to reals. There may be subjective uncertainty about what the "right" probability on $S$ is – $\mu$ is the DM's subjective prior over $\Delta$, the possible probabilities over $S$, and therefore measures the subjective relevance of a particular $\pi$ as the "right" probability. While $u$, as usual, characterizes attitude toward risk, we show that ambiguity attitude is captured by $\phi$. In particular, a concave $\phi$ characterizes ambiguity aversion, which we define to be an aversion to mean preserving spreads in $\mu_f$, where $\mu_f$ is the distribution over expected utility values induced by $\mu$ and $f$. It is worth noting that this preference model does not, in general, impose reduction between $\mu$ and the $\pi$'s in the support of $\mu$. Such reduction only occurs when $\phi$ is linear, a situation that we show is identified with ambiguity neutrality and wherein the preferences are observationally equivalent to that of a subjective expected utility maximizer.

The basic structure of the model and assumptions are as follows. Our focus of interest is the DM's preferences over acts on the state space $S$. This set of acts is assumed to include a special subset of acts which we call lotteries, i.e., acts measurable with respect to a portion of $S$ over which probabilities are assumed to be objectively given (or unanimously agreed upon). We start by assuming preferences over these lotteries are expected utility preferences. From preferences over lotteries, the DM's risk preferences are revealed, identified by v N-M index $u$. We then consider preferences over acts each of whose payoff is contingent on which prior (on $S$) is the "right" probability – we call these acts second order acts. For the moment, to fix ideas, think of these acts as "bets over the right prior". Our second axiom states that preferences over second order acts are subjective expected utility (SEU) preferences. The point of defining second order acts and imposing Axiom 2 is to model explicitly the uncertainty about the "right prior" and uncover the DM's subjective beliefs with respect to this uncertainty and attitude to this uncertainty. Indeed, following this axiom we recover $\mu$ and $v$: the former is a probability measure over possible priors on $S$ revealing the DM's subjective information while the latter is the v N-M index summarizing the DM's risk attitude toward the uncertainty over the "right" prior. Our third axiom connects preferences over second order acts to preferences over acts on $S$. The axiom essentially identifies an act $f$, defined on $S$, with a second order act that yields for each prior $\pi$ on $S$, the certainty equivalent of the lottery induced by $f$ and $\pi$. Upon setting $\phi \equiv v \circ u^{-1}$, the three axioms lead to the representation given above. Notice, a concave $\phi$ implies that $v$ is a concave transform of

$u$. Hence, ambiguity aversion in this framework will be shown to mean that the DM is more risk averse to the subjective uncertainty about priors than he is to the uncertainty in lotteries, suggesting the interpretation that the DM considers the information reflected in the prior $\mu$ to be "less hard", or of a different character, than the information about probabilities contained in the description of lotteries. Relatedly, ambiguity neutrality obtains if the DM's attitudes to risk on the two domains of uncertainty are identical.

It is not hard to think of real world instances wherein a decision maker's information is consistent with multiple probabilities on the state space relevant to the decision at hand. One example is expert opinion. A manager seeking assessments of the probability of success of a venture from a set of experts may find that different experts report different probability numbers. Possibly there would be a further aspect to the manager's information: some prior belief about the credibility of each expert. As a second example, think of a decision maker applying a model to forecast the realization of an observable variable. The model maps a parametric configuration, whose value is not precisely known to the decision maker, to a probability distribution over a set of possible realizations of the observable variable. Data on past realizations of the observable variable would allow the decision maker to update his prior (or, say, estimate a 95% confidence interval) on the possible values of the parameters. Hence, the decision maker will typically be left with a set of probability distributions governing the future realizations of the observable variable, along with some belief over the relevant probabilities. The idea, which echoes the recent literature investigating decision rules robust to model misspecification or "model uncertainty" (Hansen, Sargent, and Tallarini (1999), Hansen, Sargent, Turmuhambetova, and Williams (2001)), finds application in many contexts. Think, for instance, of a monetary policy maker setting policy on the basis of a model which, given parameters, solves to yield a probability distribution on a set of macroeconomic variables of interest. This DM is likely uncertain about the parameters, which involve details of technology, tastes, and political decisions. Or, consider a DM in charge of policy whose impact is contingent on long term realizations of environmental/climatic variables. The DM has access to a model that predicts a probability distribution on the set of climate variables of interest, like precipitation, temperature, given a particular parametrization involving variables like future trajectory of political agreements on emission controls. But, in this case too, the DM is uncertain about the actual values of the parameters and the course of future political decisions.

Notice in the context of the first example, we may think of second order acts as acts which pay off depending on which expert is the best informed, even though the acts are *effectively* bets on which prior is right. Similarly, in the context of the second example, second order acts may be thought of as acts whose payoffs are contingent on a particular parametric configuration being the relevant one. Correspondingly, $\mu$ and $v$ may be regarded as representing preferences for such acts. Ambiguity averse DMs have preference for acts whose performance is more robust to the possible variation in probabilities. In our model, this is equivalent to $v$ being more concave than $u$ ($\phi$ being concave). To see this in the context of our examples, imagine a DM in one of the examples thinking, "My best guess of the chance that expert '$\pi$' is the best informed (or, that politics of emission control will resolve in a particular way) is 20%. However, this is far less informed a guess than knowing that the chance in an objective lottery is 20%. Hence, I would like to behave with more caution, with respect to the former risk." It

is a thought such as this that our model suggests as a way of understanding ambiguity aversion (as we have defined it).

Apart from providing (what we think is) a simple and clarifying perspective on ambiguity and ambiguity attitude, we believe our model, and more specifically our functional representation, will be particularly useful in economic modeling involving comparative statics and ambiguity. Suppose we have an economic model where agents' beliefs reflect some ambiguity and consider, as a benchmark, equilibrium behavior in the model if all agents were neutral to this ambiguity (say, they behaved as SEU maximizers). Next, without perturbing the information structure any further, we want to ask how the equilibrium would change if ambiguity aversion were to replace ambiguity neutrality, both in the case where the aversion is uniform across agents and, more generally, if the ambiguity attitude varied in intensity across the agents. Another comparative statics exercise might hold the taste parameters such as ambiguity attitudes fixed and ask how the equilibrium is affected if the perceived ambiguity varied across agents. The point is, working out such comparative statics properly requires a model which separates an agent's subjective uncertainty from his attitude toward that uncertainty and parametrizes each of these in a tractable way. We argue below that the models allowing for ambiguity attitude available in the literature do not make this separation in a way that is clear and complete enough.

Consider, for instance, the classic model of maxmin expected utility (MEU) preferences (Gilboa and Schmeidler (1989)), which, along with the pioneering Choquet expected utility model of Schmeidler (1989), remains among the most popular models capable of incorporating ambiguity aversion. MEU preferences are represented by the following functional:

$$V(f) = \min_{\pi \in \Pi} \mathbb{E}_{\pi} (u \circ f),$$

where $f$ is an act and $\Pi$ is a set of probability measures. The model does not, in general, impose a separation of information/beliefs and ambiguity attitudes. In general, the set $\Pi$ may not be interpreted as being completely characterized by the decision maker's beliefs. It represents beliefs intertwined with ambiguity attitude in an inseparable way. A common comparative statics exercise that uses this model to reveal the effect of ambiguity aversion proceeds by comparing the behavior corresponding to one set of priors $\Pi$ with the behavior corresponding to a more inclusive set $\Pi'$, i.e., $\Pi \subset \Pi'$, while controlling for risk attitudes by holding $u$ fixed. The larger set of priors could reflect differences in information or perception, differences in ambiguity attitude, or both. This is problematic because we are unable to see if the answers differ because the ambiguity in the environment is changing or because of a change in attitude towards ambiguity. For instance, consider a principal-agent model with moral hazard which is "standard" in every respect except that the players' beliefs about the contractible signal (e.g., output) reflect ambiguity. In such a model it is natural to pose the question, "How does the optimal contract change if the agent becomes more ambiguity averse than the principal?" Suppose we require, as in the standard model, that the comparative statics exercise respect the assumption that the principal and agent have the *same* (prior) information/beliefs about the contractible signal. Such an exercise would be very hard to execute, if not impossible, in the MEU model. The problem is that there are no known results which would tell us how we could alter the set of priors in the representation of the agent's preferences in a way that would incorporate a change in ambiguity attitude while at the same time maintaining that the

agent's priors and the set of priors corresponding to the principal and agent incorporating the same information about the contractible signal.

It may seem that the problem of conflating beliefs and attitudes can be finessed if we take the special case where the set $\Pi$ is interpreted as the set of probabilities the decision maker thinks is relevant, i.e., we interpret it as "pure" information/belief. Then, per force, we fix the decision maker's ambiguity attitude to the extreme case where he only worries about the worst case scenario given his information. This interpretation imposes at least two kinds of limitations on the comparative static exercises. One, we cannot allow for heterogeneity of tastes (specifically, ambiguity attitudes) among agents in the model. Two, we have no way of knowing whether the results of the comparative static exercise are true only for the extreme attitude or extend more generally to intermediate degrees of ambiguity aversion. One advantage of our model is that it does not suffer from these difficulties. For further discussion relating our approach to others in the literature see Section 6.

The rest of the paper is organized as follows. The next section states our basic assumptions on preferences, derives the representation, and presents a definition and characterization of ambiguity attitude. Section 3 defines and characterizes ambiguity in this model. Section 4 focuses on the characterization of comparative ambiguity attitudes. Section 5 contains both formal results and discussion relating the preference-based notion of ambiguity presented in Section 3 to two leading preference-based notions of ambiguity available in the literature. Section 6 contains a general discussion of literature related to our model. Finally, Section 7 presents two illustrative portfolio choice applications of the theory developed here. All proofs, unless otherwise noted in the text, appear in the Appendix.

# 2 Axioms and representation

## 2.1 Preliminaries

Let $\mathcal{A}$ be the Borel $\sigma$-algebra of a metric space $\Omega$, and $\mathcal{B}$ the Borel $\sigma$-algebra of $[0,1)$. Consider the state space $S = \Omega \times [0,1)$, endowed with the product $\sigma$-algebra $\Sigma \equiv \mathcal{A} \otimes \mathcal{B}$. For the remainder of this paper, all events will be assumed to belong to $\Sigma$ unless stated otherwise.

We denote by $f : S \to \mathcal{C}$ a Savage act, where $\mathcal{C}$ is the set of consequences. We assume $\mathcal{C}$ to be an interval in $\mathbb{R}$ containing the interval $[-1,1]$. $\mathcal{F}$ is the set of all simple acts, that is, the set of all $\Sigma$-measurable functions $f : S \to \mathcal{C}$ such that the set $\{x \in \mathbb{R} : f(s) = x$ for some $s \in S\}$ is finite. The space $[0,1)$ is introduced simply to model a rich set of lotteries[1] as a set of Savage acts.[2] An act $l \in \mathcal{F}$ is said to be a lottery if $l$ depends only on $[0,1)$ – i.e., $l(\omega_1, r) = l(\omega_2, r)$ for any $\omega_1, \omega_2 \in \Omega$ and $r \in [0,1)$ – and it is Riemann

[1] By the phrase "a rich set of lotteries" we simply mean that, for any probability $p \in [0,1)$, we may construct an act which yields a consequence with that probability. While this richness is not required in the statement of our axioms or in our representation result, it is invoked later in the paper in Theorems 2 and 3.

[2] Our modelling of lotteries in this way and use of a product state space is similar to the "single-stage" approach in Sarin and Wakker (1992), Sarin and Wakker (1997), and to Anscombe-Aumann style models.

integrable. The set of all such lotteries is $\mathcal{L}$. If $f \in \mathcal{L}$ and $r \in [0,1)$, we sometimes write $f(r)$ meaning $f(\omega, r)$ for any $\omega \in \Omega$.

Given the Lebesgue measure $\lambda : \mathcal{B} \to [0,1]$, let $\pi : \Sigma \to [0,1]$ be a countably additive product probability such that $\pi(A \times B) = \pi(A \times [0,1))\lambda(B)$ for $A \in \mathcal{A}$ and $B \in \mathcal{B}$. The set of all such probabilities $\pi$ is denoted by $\Delta$. Let $C(S)$ be the set of all continuous (w.r.t. the product topology of $S$) and bounded real-valued functions on $S$. Using $C(S)$ we can equip $\Delta$ with the vague topology, that is, the coarsest topology on $\Delta$ that makes the following functionals continuous:

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

Throughout the paper we assume $\Delta$ to be endowed with the vague topology. Let $\sigma(\Delta)$ be the Borel $\sigma$-algebra on $\Delta$ generated by the vague topology. The following lemma shows a property of $\sigma(\Delta)$ that is quite crucial for our purposes, as observed right after Theorem 1 (its routine proof is omitted).

**Lemma 1** *$\sigma(\Delta)$ coincides with the $\sigma$-algebra generated by the real-valued functions on $\Delta$ given by*

$$\pi \mapsto \int \psi d\pi \qquad \pi \in \Delta \text{ and } \psi \in B(\Sigma), \tag{1}$$

*where $B(\Sigma)$ is the set of all bounded and real-valued $\Sigma$-measurable functions.*

Since we wish to allow $\Delta$ to be another domain of uncertainty for the decision maker apart from $S$, we model it explicitly as such. To formally identify the decision maker's subjective uncertainty about this domain, i.e., whether he at all regards this domain as uncertain and if so, what his subjective information and beliefs are, we look at the decision maker's preferences over *second order acts* which assign consequences to elements of this domain. The interpretation of second order acts and their domain is discussed in detail in Remark 2.

**Definition 1** *A **second order act** is any bounded $\sigma(\Delta)$-measurable function $\mathfrak{f} : \Delta \to \mathcal{C}$ that associates an element of $\Delta$ to a consequence. We denote by $\mathfrak{F}$ the set of all second order acts.*

Let $\succeq^2$ be the decision maker's preference ordering over $\mathfrak{F}$. The main focus of the model is $\succeq$, a preference relation defined on $\mathcal{F}$ (the set of simple acts on $S$).

## 2.2 Basic axioms

Next we describe three assumptions on the preference orderings $\succeq$ and $\succeq^2$. The first axiom applies to the preference ordering $\succeq$ when restricted to the domain of lottery acts. Preferences over the lotteries are assumed to have a expected utility representation.

**Axiom 1 (Expected utility on lotteries)** *There exists a unique $u : \mathcal{C} \to \mathbb{R}$, continuous, strictly increasing and normalized so that $u(0) = 0$ and $u(1) = 1$ such that, for all $f, g \in \mathcal{L}$, $f \succeq g$ if and only if $\int_{[0,1)} u(f(r))dr \geq \int_{[0,1)} u(g(r))dr$.*

In the standard way, the utility function, $u$, represents the decision maker's attitude towards risk generated from the lottery part of the state space. The next axiom is on $\succeq^2$, the preferences over second order acts. These preferences are assumed to have a subjective expected utility representation.[3]

**Axiom 2 (Subjective expected utility on 2nd order acts)** *There exists a unique finitely additive probability $\mu : \sigma(\Delta) \to [0,1]$ and a continuous, strictly increasing $v : \mathcal{C} \to \mathbb{R}$, such that*

$$\mathfrak{f} \succeq^2 \mathfrak{g} \Longleftrightarrow \int_\Delta v(\mathfrak{f}(\pi))\, d\mu \geq \int_\Delta v(\mathfrak{g}(\pi))\, d\mu,\ \forall \mathfrak{f}, \mathfrak{g} \in \mathfrak{F}$$

*Further, if there exists a $J \subseteq \Delta$ such that $0 < \mu(J) < 1$, then $v$ is unique up to positive affine transformations.*

We denote by $\Pi$ the support of $\mu$, that is, the smallest closed (w.r.t. the vague topology) subset of $\Delta$ whose complement has measure zero; $\Pi$ is the subset of $\Delta$ the decision maker *subjectively* considers as "relevant". Given any $E \subseteq \Pi$, we interpret $\mu(E)$ as the decision maker's *subjective* assessment of the likelihood that the relevant probability lies in $E$; hence, $\mu$ may be thought of as a "second order probability" over the first order probabilities $\pi$. Notice that $\Pi$ may well be a finite subset of $\Delta$. Finally, the utility function $v$ can be seen as representing the decision maker's attitude towards risks generated from the space of first-order probabilities.

**Notation 1** An act $f$ and a probability $\pi$ induce on consequences a distribution $\pi_f$ defined by $\pi_f(x) = \pi(f^{-1}(x))$ for each $x \in \mathcal{C}$. The support of $\pi_f$ is the range of the act $f$; denote it by $\{x_1, ..., x_n\} \subseteq \mathcal{C}$. We now construct a lottery act $l_f \in L$ inducing a distribution on $\mathcal{C}$ identical to $\pi_f$, as follows:

$$l_f(\omega, r) = \begin{cases} x_1 & if \quad r \in [0, \pi_f(x_1)) \\ x_2 & if \quad r \in [\pi_f(x_1), \pi_f(x_2) + \pi_f(x_1)) \\ \vdots & \vdots \quad \vdots \\ x_n & if \quad r \in [\sum_{i=1}^{n-1} \pi_f(x_i), \sum_{i=1}^{n} \pi_f(x_i) = 1) \end{cases}$$

Finally, let $\delta_x$ denote the constant act with consequence $x \in \mathcal{C}$, and let $c_f(\pi)$ denote the certainty equivalent of the lottery $l_f$, i.e., $\delta_{c_f(\pi)} \sim l_f$.

Continuity of $u$ (Axiom 1) and the assumption that lotteries have finite support guarantee that the certainty equivalent exists for any lottery. We also assumed (in Axiom 1) that $u$ is increasing; this assumption guarantees that the certainty equivalent is unique. Axiom 1 also implies that the order of the enumeration of the support of $\pi_f$ in the construction of $l_f$ does not matter for $\succeq$. Additionally, the fact that each $\pi$ restricted to $\Omega \times \mathcal{B}$ is the Lebesgue measure implies that the only property of the chosen events in the construction of $l_f$ that matters for $\succeq$ is their length.

[3] Each of the first two axioms could be replaced by more primitive assumptions on $\succeq$ and $\succeq^2$, respectively, which deliver the expected utility representations. Since such developments are by now well known and easily adapted to our setting, we do not do so here.

Our final basic axiom requires the preference ordering of primary interest, $\succeq$, to be consistent with Axioms 1 and 2 in a certain way. Essentially, the axiom requires the decision maker to view act $f \in \mathcal{F}$ as equivalent to a second-order act $f^2 \in \mathfrak{F}$ which yields the consequence $c_f(\pi)$ whenever $\pi$ is the "true" probability law. Thus, $f$ is associated with a vector of several certainty equivalents, each of them corresponding to a particular possible probability. Since this vector of certainty equivalents is a second-order act, it is reasonable to ask that its evaluation be consistent with Axiom 2. Since $f$ together with a possible probability $\pi$ generates a distribution over consequences identical to that generated by $l_f$, it seems reasonable to require (for consistency with Axiom 1) that the certainty equivalent of $f$, given $\pi$, be same as the certainty equivalent of $l_f$.

**Definition 2** *Given $f \in \mathcal{F}$, $f^2 \in \mathfrak{F}$ denotes a* ***second order act associated with*** *$f$, defined as follows*

$$f^2(\pi) = c_f(\pi) \qquad \text{for all } \pi \in \Delta.$$

**Axiom 3 (Consistency with preferences over associated 2nd order acts)** *Given $f, g \in \mathcal{F}$ and $f^2, g^2 \in \mathfrak{F}$,*

$$f \succeq g \Longleftrightarrow f^2 \succeq^2 g^2.$$

**Remark 1** Let $\delta_c^2$ denote a second order act such that $\delta_c^2(\pi) = c$, for all $\pi \in \Delta$. Since both $v$ and $u$ are strictly increasing, we have $\delta_c^2 \succeq^2 \delta_{c'}^2$ if and only if $\delta_c \succeq \delta_{c'}$.

**Remark 2** Axioms 1 through 3, taken together, imply that if the DM were to know that the "event $\pi$" in the domain of second order acts has occurred (i.e., $\mu(\pi) = 1$), then he evaluates an act $f \in \mathcal{F}$ using an expected utility functional with probability $\pi$ (on $S$). In other words, the "event $\pi$" is just *some* event such that the DM's conditional preferences given that event is expected utility with probability $\pi$. It is a useful abstraction, facilitating directness and parsimony in the general framework, to describe "the event $\pi$" literally as "the event such that $\pi$ is the relevant/right probability", even though such a description is open to the criticism that such an event may not be objectively verifiable. It is important to recognize, however, that the domain of second order acts does not literally have to be $\Delta$. A set of objectively verifiable, mutually exclusive events, $\mathcal{E} = \{E_\pi\}_{\pi \in \Delta}$, which are *not payoff relevant* with respect to acts in $\mathcal{F}$, would do just as well, provided Axioms 2 and 3 apply to acts defined on $\mathcal{E}$, as explained next. First we need to redefine a second order act as $\mathfrak{f} : \mathcal{E} \to \mathcal{C}$ , i.e., it associates an element of $\mathcal{E}$ to a consequence. Correspondingly, we redefine a second order act associated with $f \in \mathcal{F}$ as, $f^2(E_\pi) = c_f(\pi)$, for all $E_\pi \in \mathcal{E}$. Now we assume that Axioms 2 and 3 to apply to the redefined acts and obtain our representation result as in Theorem 1. As noted earlier, if Axioms 2 and 3 are satisfied by preferences over the redefined acts it will be implied, in particular, that the DM's conditional preferences *given* event $E_\pi$ is expected utility with probability $\pi$. Hence, the notion that "$E_\pi$ is associated with the probability $\pi$" has a meaning that is conceptually clear and something that may be, in principle, verified objectively by observing the DM's preference behavior (conditional on the event $E_\pi$) w.r.t. acts in $\mathcal{F}$. (Note, the description of event $E_\pi$, given that the description uses objective propositions, need not contain a reference to $\pi$.) Given particular contexts, events such as $E_\pi$ will have objectively verifiable identities that go with the context, as suggested by our examples on page 3, and the domain of second order acts will be identified with such

events. For instance, recall the example involving expert advice about the success of a business venture. In that context, an $E_\pi$ could be the contingency described as, "Expert A is the best informed". If the DM were to know that this event is true, he would choose according the A's advice, and his preference behavior would, in principle, reveal (in an objectively verifiable way) the prior $\pi$ applied to acts in $\mathcal{F}$. To summarize, it is not necessary that the domain of second order acts be taken literally as $\Delta$ for our theory to obtain. We write the domain as that to be clear about its relevance to preferences on acts in $\mathcal{F}$, which is purely through its information content about the appropriate prior on $S$.

The above three axioms are basic to our model in that they are all that we invoke to obtain our representation result. Theorem 1 below shows that given these axioms, $\succeq$ is represented by a functional which is an "expected utility over expected utilities". Evaluation of $f \in \mathcal{F}$ proceeds in two stages: first, compute all possible expected utilities of $f$, each expected utility corresponding to a $\pi$ in the support of $\mu$; next, compute the expectation (with respect to the measure $\mu$) of the expected utilities obtained in the first stage, each expected utility transformed by the index $\phi$.

As will be shown in subsequent analysis, this representation allows a clear decomposition of the decision maker's tastes and beliefs: $u$ determines risk attitude, $\phi$ determines ambiguity attitude, and $\mu$ determines the subjective ambiguity of information. The representing functional is also invariant to positive affine transforms of the vN-M utility index that applies to the lotteries. That is, when $u$ is translated by a positive affine transformation to $u'$, the class of associated $\phi'$ is simply the class of $\phi$ with domain shifted by the positive affine transformation.

Formally, let $\mathcal{U}$ denote the range $\{u(x) : x \in \mathcal{C}\}$ of the utility function $u$.

**Theorem 1** *Given Axioms 1, 2 and 3, there exists a continuous and strictly increasing $\phi : \mathcal{U} \to \mathbb{R}$ such that $\succeq$ is represented by the preference functional $V : \mathcal{F} \to \mathbb{R}$ given by*

$$V(f) = \int_\Delta \phi \left[ \int_S u(f(s))\, d\pi \right] d\mu \equiv \mathbb{E}_\mu \phi \left( \mathbb{E}_\pi u \circ f \right) \tag{2}$$

*Furthermore, if there exists $J \subseteq \Delta$ with $0 < \mu(J) < 1$, then, given $u$, the function $\phi$ is unique up to positive affine transformations. Moreover, if $\tilde{u} = \alpha u + \beta$, $\alpha > 0$, then the associated $\tilde{\phi}$ is such that $\tilde{\phi}(\alpha y + \beta) = \phi(y)$, where $y \in \mathcal{U}$.*

The integrals in (2) are well defined because of Lemma 1, which guarantees their existence. Hereafter, when we write a preference relation $\succeq$, we assume that it satisfies the conditions in Theorem 1. This theorem can be viewed as a part of a more comprehensive representation result (reported in the Appendix as Theorem 6) for the two orderings $\succeq$ and $\succeq^2$ in which Axioms 1, 2 and 3 are both necessary and sufficient. Theorem 6 also notes explicitly an important point explained in the proof of Theorem 1, that $\phi$ is to equal $v \circ u^{-1}$.

We close by observing that, though in Axiom 1 we assumed expected utility preferences on lotteries, we could relax that assumption by allowing more general preferences. For example, if preferences over lotteries were Rank Dependent Expected Utility preferences with a suitable probability distortion $\psi : [0,1] \to [0,1]$, then the representation in

Eq. (2) would take the form

$$\int_{\Delta} \phi \left( \int_{S} u\left(f\left(s\right)\right) d\psi\left(\pi\right) \right) d\mu,$$

where the inner integral is in the sense of Choquet. More generally, we conjecture that any suitable preferences over lotteries for which risk attitudes may be usefully defined and for which certainty equivalents exist might be plugged into Axiom 1. Along these lines, the subjective expected utility assumption on preferences for second order acts might also be something that could be relaxed. We suspect that the key features needed would be some version of probabilistic sophistication along with comparability of risk attitudes over second order acts with risk attitudes over lotteries. Paring assumptions down to probabilistic sophistication would, as usual, often restrict the representation to an "implicit" form. We prefer to stick to expected utility in both Axioms principally for two reasons. One, it serves to sharpen the focus of our exercise, and thus enhance clarity. Two, as explained in the introduction, we are strongly motivated by the need to obtain a functional form that would facilitate economically interesting comparative static analyses. An explicit, tractable representation, that would identifiably separate beliefs and attitudes, is absolutely essential for that.

## 2.3 Characterizing ambiguity attitude

In this section we provide a definition of a decision maker's ambiguity attitude and we show that this ambiguity attitude is characterized by properties of $\phi$, one of the functions from our representation above. Comparison of ambiguity attitudes across preference relations is dealt with in Section 4. To discuss ambiguity attitude, we first require an additional assumption. In the classical theory, risk attitude is assumed to be a property of the individual and not of their environment. In particular, it is commonly implicitly or explicitly assumed that a given individual will display the same risk attitude across settings in which she might hold different subjective beliefs. We would like to assume the same. In the context of our theory, this entails the assumption that risk attitudes derived from lotteries and risk attitudes derived from second-order acts are independent of an individual's beliefs. In fact, a slightly weaker assumption suffices for our purposes: the assumption that these two risk attitudes do not vary with the support of an individual's belief. In terms of the axioms and representation of the previous section, this will mean that $u$ and $v$ do not depend on $\Pi$, the support of $\mu$.

To state this formally in our setting, consider a family $\{\succeq_{\Pi}, \succeq^{2}_{\Pi}\}_{\Pi \subseteq \Delta}$ of pairs of preference relations (over acts and over second-order acts respectively) characterizing each decision maker, wherein there is a pair of preference relations corresponding to each possible support $\Pi$, that is, to each possible state of information he may have about which probabilities $\pi$ (over $S$) are relevant to his decision problem. In reading the axiom, recall that supports are, by definition, closed subsets of $\Delta$.

**Axiom 4 (Separation of tastes and beliefs )** *Fix a family of preference relations* $\{\succeq_{\Pi}, \succeq^{2}_{\Pi}\}_{\Pi \subseteq \Delta}$ *for a given decision maker.*

*(i) The restriction of* $\succeq_{\Pi}$ *to lottery acts remains the same for every closed subset* $\Pi \subseteq \Delta$.

*(ii) The same invariance with respect to* $\Pi$ *holds for the risk preferences derived from* $\succeq_{\Pi}^{2}$ .

Imposing Axiom 4 in addition to the earlier axioms guarantees that as the support of a decision maker's subjective belief varies (say, due to conditioning on different information), the decision maker's attitude towards risk in lotteries, as embodied in $u$ (from Axiom 1), and attitude towards risk on the space $\Delta$ of probabilities, as embodied by $v$ (from Axiom 2) remain unchanged. Importantly, this will also mean that the same $\phi$ may be used to represent each $\succeq_{\Pi}$ for a decision maker. To see this, simply recall that $\phi$ may be taken to equal $v \circ u^{-1}$.

Notice that there is no restriction on the decision maker's belief associated with each $\succeq_{\Pi}$, besides that of having support $\Pi$. Though we do not need to assume it for our results, a natural possibility is that all such beliefs be connected via conditioning from some "original" common belief.

We now proceed to develop a formal notion of ambiguity attitude. Recall that an act $f$ together with a probability $\pi$ induces a distribution $\pi_f$ on consequences. Each such distribution is naturally associated with a lottery $l_f \in \mathcal{L}$ that yields each consequence $c$ in the support of $\pi_f$ on a subset of $[0,1)$ with length equal to $\pi_f(c)$. Each such lottery has a certainty equivalent $c_f(\pi)$. By Axiom 2, there is a probability $\mu$ supported by the set of probabilities $\Pi$. Fixing an act $f$, the probability $\mu$ may then be used to induce a measure $\mu_f$ on $\{u(c_f(\pi)) : \pi \in \Pi\}$, the set of expected utility values generated by $f$ corresponding to the different $\pi$'s in $\Pi$ (using the utility function from Axiom 1). This is done in Definition 3, as follows.

**Definition 3** *Given* $f \in \mathcal{F}$, *the* ***induced distribution*** $\mu_f$ *on* $\mathcal{U}$ *is as follows:*

$$\mu_f(u(x)) \equiv \mu\left(\left(f^2\right)^{-1}(x)\right) \qquad \textit{for each } x \in \mathcal{C}.$$

Given an act $f$, the derived (subjective) probability distribution over expected utilities, $\mu_f$, smoothly aggregates the information the decision maker has about the relevant $\pi$'s and how each such $\pi$ evaluates $f$, without imposing reduction between $\mu$ and the $\pi$'s. In this framework the induced distribution $\mu_f$ represents the decision maker's subjective uncertainty about the "right" (ex ante) evaluation of an act. The greater the spread in $\mu_f$, the greater the uncertainty about the ex ante evaluation. In our model it is this uncertainty through which ambiguity of about beliefs may affect behavior: ambiguity aversion is an aversion to the uncertainty about ex ante evaluations. Analogous to risk aversion, aversion to this uncertainty is taken to be the same as disliking a mean preserving spread in $\mu_f$. First we introduce notation for denoting the mean of $\mu_f$.

**Notation 2** Let $e(\mu_f) \equiv \int_{\mathcal{U}} x d\mu_f$. Notice $u^{-1}(e(\mu_f)) \in \mathcal{C}$.

Recalling the notation for constant acts, we see $\delta_{u^{-1}(e(\mu_f))}$ is simply the constant act which yields the same utility as the mean of $\mu_f$. Equivalently, $\mu_{\delta_{u^{-1}(e(\mu_f))}}$ is the degenerate distribution on $e(\mu_f)$. The definition below formally describes the notion that a given decision maker displays ambiguity aversion if each member of the family $\{\succeq_{\Pi}\}_{\Pi \subseteq \Delta}$ shows a preference for an act, $\delta_{u^{-1}(e(\mu_f))}$, that induces the degenerate distribution on the mean of $\mu_f$, $\mu_{\delta_{u^{-1}(e(\mu_f))}}$, over the act $f$ that induces $\mu_f$.

**Definition 4** *A decision maker displays* ***smooth ambiguity aversion*** *if*

$$\delta_{u^{-1}\left(e\left(\mu_f\right)\right)} \succeq_{\Pi} f$$

*for all* $f \in \mathcal{F}$ *and all closed subsets* $\Pi \subseteq \Delta$.[4]

In a similar way, we can define smooth ambiguity love and neutrality. The proposition below shows that smooth ambiguity aversion is characterized in the representing functional by the concavity of $\phi$. The proposition also shows that smooth ambiguity aversion is equivalent to the DM being more risk averse to the uncertainty about the right prior on $S$ than he is to the risk involving lotteries (whose probalities are objectively known). A result characterizing smooth ambiguity love by convexity of $\phi$ (and $u$ being a concave transform of ) follows from the same argument. Similarly, smooth ambiguity neutrality is characterized by $\phi$ linear. It is worth noting that a straightforward adaptation of the proof of the analogous result in risk theory does not suffice here. The reason is that the needed diversity of associated second order acts is not guaranteed in general.

**Proposition 1** *Under Axioms 1-4, the following conditions are equivalent:*

*(i) the function* $\phi : \mathcal{U} \to \mathbb{R}$ *is concave;*

*(ii)* $v$ *is a concave transform of* $u$*;*

*(iii) the decision maker displays smooth ambiguity aversion.*

The proposition has the following corollary (whose simple proof is omitted) which shows that the usual reduction (between $\mu$ and $\pi$) applies whenever ambiguity neutrality holds. In that case we are back to subjective expected utility. An ambiguity neutral decision maker, though informed of the multiplicity of $\pi$'s, is indifferent to the spread in the ex ante evaluation of an act caused by this multiplicity; he only cares about the "expected prior" $\nu$.

**Corollary 1** *Under Axioms 1-4, the following properties are equivalent:*

*(i) the decision maker is smoothly ambiguity neutral;*

*(ii)* $\phi$ *is linear;*

*(iii)* $v$ *is a linear transform of* $u$*;*

*(iv)* $V(f) = \int_S u\left(f\left(s\right)\right) d\nu$*, where* $\nu\left(E\right) = \int_\Delta \pi\left(E\right) d\mu$ *for all* $E \in \Sigma$.

[4] This definition is actually stronger than we need for our later results. It is enough that the indicated preference hold for (in addition to the original preference, $\succeq$) some $\succeq_\Pi$ whose set $\Pi$ contains exactly two measures having disjoint support. While we stick with the stronger definition for ease of statement, the observation here indicates that many fewer preference relations need to be considered (two rather than an infinite number) than the stronger version would lead one to think.

Our final axiom imposes a (behavioral) restriction on the preference order so that its ambiguity attitude is "well behaved". This good behavior will be useful in the next section when we discuss ambiguous events and acts. In words, the restriction is that if a preference is not neutral to ambiguity then there exists at least one interval in $\mathcal{U}$ which satisfies the following property. Consider the set of acts for which the set the expected utilities corresponding to each act lies within the given interval. We require that the decision maker displays either strict ambiguity aversion, or strict ambiguity love, but not both, for *all* such acts. What is ruled out (and this is the essence of the restriction) is the possibility that the decision maker's ambiguity attitude flits between ambiguity aversion and ambiguity love, continuously from one point to the next, over the *entire* range of $\mathcal{U}$. Note, it is entirely permissible that there be several intervals, over some of which the decision maker is ambiguity averse while over others he is ambiguity loving. Furthermore, outside a chosen interval, ambiguity attitude may be as inconsistent as desired. The statement of the axiom is immediately followed by a proposition which gives an equivalent characterization in terms of $\phi$.

**Axiom 5 (Consistent ambiguity attitude over some interval)** *The decision maker's family of preferences satisfies at least one of the following three conditions:*

*(i) smooth ambiguity neutrality,*

*(ii) there exists an open interval $J \subseteq \mathcal{U}$ such that smooth ambiguity aversion holds strictly when limited to all $f \in \mathcal{F}$ for which $supp\left(\mu_f\right)$ is a non-singleton subset of $J$,*

*(iii) there exists an open interval $K \subseteq \mathcal{U}$ such that smooth ambiguity love holds strictly when limited to all $f \in \mathcal{F}$ for which $supp\left(\mu_f\right)$ is a non-singleton subset of $K$.*

**Proposition 2** *Under Axioms 1-4, we have:*

1. *Axiom 5 (i) holds if and only if $\phi$ linear;*
2. *Axiom 5 (ii) holds if and only if $\phi$ strictly concave on some open interval $J \subseteq \mathcal{U}$;*
3. *Axiom 5 (iii) holds if and only if $\phi$ strictly convex on some open interval $K \subseteq \mathcal{U}$.*

The following lemma and remark shows that if $\phi$ were twice continuously differentiable, as it is likely to be in any application, then Axiom 5 is actually implied by the other axioms and is not an additional assumption.

**Lemma 2** *Suppose $\phi$ is twice continuously differentiable. If $\phi$ is not linear, then $\phi$ is either strictly concave or convex over some open interval.*

**Remark 3** It follows immediately from Proposition 2 and Lemma 2 that under twice continuous differentiability of $\phi$, Axioms 1–4 imply Axiom 5.

It is perhaps worth noting that there exist functions on $[0, 1]$ that are strictly increasing, continuous and almost everywhere differentiable for which the conclusion of Lemma 2 does not hold. It can be shown that this is the case, for example, of the function presented in Takacs (1978).

# 3 Ambiguity: ambiguous events and ambiguous acts

We have mentioned that an attractive feature of our model is that it allows one to separate ambiguity from ambiguity attitude. In this section we concentrate on the ambiguity part. First, we propose a preference based definition of ambiguity. We then show that this notion of ambiguity has a particularly simple characterization in our model and go on to present results to demonstrate appropriateness of the definition in the context of our model.

What makes an event ambiguous or unambiguous by our definition rests on a test of behavior, with respect to bets on the event, inspired by the Ellsberg 2-color experiment (Ellsberg (1961)). The role corresponding to bets on the draw from the urn with the known mixture of balls is played here by bets on events in $\Omega \times \mathcal{B}$. We dub an event $E \in \Sigma$ as ambiguous if, analogous to the modal behavior observed in the Ellsberg experiment, betting on $E$ is preferred to betting on some event $B$ in $\Omega \times \mathcal{B}$, and betting on $E^c$ is *also* preferred to betting on $B^c$. The proposition following the definition shows an equivalent, shorter, form of the definition which essentially follows from the continuity of $\phi$, which is assured in our model given the first three axioms. This equivalent form, though it lacks immediate intuitive identification with the Ellsberg experiment, adds clarity to our understanding of what makes an event unambiguous: an event is unambiguous if it is possible to calibrate the likelihood of the event with respect to events in $\Omega \times \mathcal{B}$.

**Notation 3** *If $x, y \in X$ and $A \in \Sigma$, $xAy$ denotes the binary act which pays $x$ if $s \in A$ and $y$ otherwise.*

**Definition 5** *An event $E \in \Sigma$ is* **unambiguous** *if, for each event $B \in \Omega \times \mathcal{B}$, and for each $x, y \in X$ such that $\delta_x \succ \delta_y$, either, $[xEy \succ xBy$ and $yEx \prec yBx]$ or, $[xEy \prec xBy$ and $yEx \succ yBx]$ or $[xEy \sim xBy$ and $yEx \sim yBx]$. An event is* **ambiguous** *if it is not unambiguous.*

**Proposition 3** *Assume $\succeq$ satisfies the conditions in Theorem 1. An event $E \in \Sigma$ is unambiguous if and only if for each $x$ and $y$ with $\delta_x \succ \delta_y$,*

$$xEy \sim xBy \Longleftrightarrow yEx \sim yBx. \tag{3}$$

*whenever $B \in \Omega \times \mathcal{B}$.*

If $\succeq$ satisfies Axioms 1 through 3 then all events in $\Omega \times \mathcal{B}$ are unambiguous. Given any particular preference relation, it of course may be checked using our definition whether an event in $\Omega \times \mathcal{B}$ is unambiguous. We also observe here that the role of $\mathcal{B}$ in our definition may be played equally well by some other rich set of events over which preferences display a likelihood relation representable by a probability measure. Furthermore, the product structure of our state space also does not play an essential role in formulating such a definition. In general, replace $\Omega \times \mathcal{B}$ with the desired alternative set.

The next theorem relates ambiguity of an event to event probabilities in our representation.

**Theorem 2** *Assume $\succeq$ satisfies the conditions in Theorem 1. If the event $E$ is ambiguous according to Definition 5, then there exist $\mu$-non-null sets $\Pi' \subseteq \Pi$ and $\Pi'' \subseteq \Pi$ and*

*$\gamma \in (0,1)$, such that $\pi(E) < \gamma$ for all $\pi \in \Pi'$ and $\pi(E) > \gamma$ for all $\pi \in \Pi''$. If the event $E$ is unambiguous according to Definition 5, then, provided $\succeq$ satisfies Axioms 4 and 5 and is not smoothly ambiguity neutral, there exists a $\gamma \in [0,1]$ such that $\pi(E) = \gamma$, $\mu$-a.e.*

Thus, in our model, if there is agreement about an event's probability then that event is unambiguous. Furthermore, if $\succeq$ has *some* range over which it is either strictly smoothly ambiguity averse or strictly smoothly ambiguity loving then disagreement about an event's probability implies that the event is ambiguous. When the support $\Pi$ of $\mu$ is finite, the meaning of disagreement about an event's probability in the theorem above simplifies to: there exist $\pi, \pi' \in \Pi$ such that $\pi(E) \neq \pi'(E)$.

To understand why conditions are needed for one direction of the theorem think of the case of ambiguity neutrality i.e., $\phi$ linear. Recall that in this case, even if the measures in $\Pi$ disagree on the probability of an event, the decision maker *behaves as if* he assigns that event its $\mu$-average probability. Recall that Lemma 2 and Remark 3 showed that under conditions likely to be assumed in any application (twice continuous differentiability of the function $\phi$ and Axiom 4) ambiguity neutrality is the *only* case where there will fail to be a range of strict ambiguity aversion (or love) and so the only case where disagreement about an event's probability will not imply that the event is ambiguous.

We say that a collection of events $\mathcal{A} \subseteq \Sigma$ is a $\lambda$-system if (i) $S \in \mathcal{A}$; (ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$; and (iii) $A_1, A_2 \in \mathcal{A}$ and $A_1 \cap A_2 = \varnothing \Rightarrow A_1 \cup A_2 \in \mathcal{A}$. This could be called a finite $\lambda$-system since (iii) requires only closure under finite disjoint unions rather than the closure under countable disjoint unions required in e.g., Billingsley (1986), p. 36. It has been widely argued in the literature (e.g., Epstein and Zhang (2001), Ghirardato and Marinacci (2002), Zhang (2002)) that the maximal collection of unambiguous events forms a $\lambda$-system and not necessarily an algebra. Hence we include the following corollary to Theorem 2 which observes that in our model, given our definition of ambiguous events, the collection of all unambiguous events does form a $\lambda$-system.

**Corollary 2** *Suppose $\succeq$ satisfies Axioms 1 through 5. Let $\Lambda \subseteq \Sigma$ be the collection of all unambiguous events in $\Sigma$. Then $\Lambda$ is a (finite) $\lambda$-system.*

Next, we identify an unambiguous act as an act which is measurable with respect to $\Lambda$, the collection of all unambiguous events in $\Sigma$. Notice, if an act $f$ is unambiguous then the induced distribution $\mu_f$ is degenerate.

**Definition 6** *An act $f \in \mathcal{F}$ is an **unambiguous act** if it is measurable with respect to $\Lambda$. Let $\mathcal{H}$ be the set of all unambiguous acts.*

To gain further insight into our definition of an ambiguous event and its appropriateness in the context of our model, we introduce the following concept combining the standard notion of a qualitative probability with an additional necessary condition for the existence of a probability and likelihood revealed through bets on events. (Conditions $(i)$ through $(iv)$ in the definition define a qualitative probability, condition $(v)$ is also necessary for the existence of a representing probability and condition $(vi)$ connects this qualitative probability with $\succeq$ over bets.)

**Definition 7** *We call a binary relation $\succeq_q$ on a $\lambda$-system $\mathcal{A} \subseteq \Sigma$ a* ***qualitative probability relation for*** $\succeq$ *if it is (i) complete and transitive, (ii) $S \succ_q \emptyset$ , (iii) $A \succeq_q \emptyset$ for $A \in \mathcal{A}$, (iv) for $A, B, C \in \mathcal{A}$, if $A \cap C = B \cap C = \emptyset$ then*

$$A \succ_q B \iff A \cup C \succ_q B \cup C$$

*(v) for $A, B \in \mathcal{A}$,*

$$A \succ_q B \iff B^c \succ_q A^c$$

*and (vi) for $A, B \in \mathcal{A}$, $A \succeq_q B$ if there exist consequences $x, y$ with $\delta_x \succ \delta_y$ such that*

$$xAy \succeq xBy.$$

$\succeq$ *is said to have* ***probabilistic beliefs*** *if there exists a qualitative probability relation for $\succeq$ on all of $\Sigma$.*

If $\mathcal{A}$ is an algebra, conditions $(i)$-$(iv)$ imply $(v)$ (see e.g., Kreps (1988), p. 118). However, as observed by Zhang (1999), this is not true if $\mathcal{A}$ is merely a $\lambda$-system. Since $(v)$ is a necessary condition for the existence of a probability representing $\succeq_q$ on $\mathcal{A}$, it makes sense to include it here. In any event, the result below on the existence of a qualitative probability relation for $\succeq$ is false without condition $(v)$.

The next result uses this preference based notion of qualitative probability relation to give an alternative characterization of ambiguous events in our setting.

**Corollary 3** *Assume $\succeq$ satisfies Axioms 1 through 5. Fix an event $E \subseteq \Sigma$. Then $E$ is unambiguous if and only if there exists a qualitative probability relation for $\succeq$ on some $\lambda$-system that is a superset of $\{E, E^c, \Omega \times \mathcal{B}\}$.*

**Remark 4** The "if" part of the claim in Corollary 3 does not depend on $\succeq$ satisfying Axioms 1 through 5.

The existence of a qualitative probability relation for $\succeq$ on a $\lambda$-system containing $\{E, E^c, \Omega \times \mathcal{B}\}$ intuitively means that the event $E$ can be compared in a consistent way with the rich set of events in the Borel $\sigma$-algebra $\mathcal{B}$.

The most important class of preferences exhibiting probabilistic beliefs are the probabilistically sophisticated preferences of Machina and Schmeidler (1992). Besides probabilistic beliefs, they also require some additional conditions which we do not consider because they are superfluous for our purposes. Given Axioms 1 through 5, the only departure from expected utility that may arise in our model is one due to ambiguity sensitive behavior, behavior that is not probabilistically sophisticated, as formally detailed in the following corollary.

**Corollary 4** *Assume $\succeq$ satisfies Axioms 1 through 4. Consider the following four properties:*

*(i) $\succeq$ has probabilistic beliefs*

*(ii) $\succeq$ is probabilistically sophisticated*

*(iii)* $\succeq$ *has a subjective expected utility representation*

*(iv) For each event $E$, there exists a $\gamma \in [0,1]$ such that $\pi(E) = \gamma$, $\mu$-a.e.*

*Then,*

$$(iv) \Longrightarrow (iii) \Longrightarrow (ii) \Longrightarrow (i).$$

*Moreover, all four properties are equivalent whenever $\succeq$ satisfies Axiom 5(ii) or (iii), while the first three are equivalent (and true) whenever $\succeq$ satisfies Axiom 5(i).*

When $\phi$ is linear ($\succeq$ satisfies Axiom 5(i)), recall that reduction applies and $\succeq$ has a SEU representation with the subjective prior on $E$ equal to $\int \pi(E)\, d\mu$. Hence, if $\phi$ is linear $\succeq$ is SEU even if all the $\pi$'s do not agree. This is meaningful since all it says is that if the decision maker is ambiguity neutral his behavior is indistinguishable from an SEU maximizer even if the information is ambiguous.

# 4 Comparison of ambiguity attitudes

In this section we study differences in ambiguity aversion across decision makers. For each decision maker we consider the entire family of preferences $\{\succeq_\Pi, \succeq^2_\Pi\}_{\Pi \subseteq \Delta}$, parametrized by the state of information given by $\Pi$. Throughout the section (with the exception of Proposition 5), we assume Axiom 4 holds, in addition to the first three axioms. Hence, ambiguity attitudes do not depend on the support $\Pi$ of $\mu$, that is, on the particular information available to the decision makers.

We begin with our definition of what makes one preference order more ambiguity averse than another. The idea behind it is that if two decision makers are such that they share the same attitude to risk and the same beliefs and yet one of the two is more averse to acts with uncertain consequences, then the comparatively "extra" aversion to uncertainty is due to a relatively greater aversion to ambiguity.

**Definition 8** *Let $A$ and $B$ be two decision makers whose families of preferences share the same vN-M utility function $u$ and share the same probability measures $\mu_\Pi$ for each support $\Pi$. We say that decision maker $A$* ***is more ambiguity averse than*** *$B$ if*

$$f \succeq^A_\Pi \delta_x \Longrightarrow f \succeq^B_\Pi \delta_x \tag{4}$$

*for every $f \in \mathcal{F}$, every $x \in \mathcal{C}$, and every closed subset $\Pi \subseteq \Delta$.*

We can now state our comparative result, which shows that differences in ambiguity aversion across decision makers are captured by the relative concavity of their functions $\phi$, thus showing that the concavity of $\phi$ plays here the role of the concavity of utility functions in standard risk theory.

**Theorem 3** *Let $A$ and $B$ be two decision makers whose families of preferences share the same vN-M utility function $u$ and the same probability $\mu_\Pi$ for each support $\Pi$. Then, decision maker $A$ is more ambiguity averse than $B$ if and only if*

$$\phi_A = h \circ \phi_B$$

*for some strictly increasing and concave $h : \phi_B(\mathcal{U}) \to \mathbb{R}$.*

Using results from standard risk theory we get the following corollary as an immediate consequence of Theorem 3.

**Corollary 5** *Suppose the hypotheses of Theorem 3 hold. If $\phi_A$ and $\phi_B$ are twice continuously differentiable, then decision maker $A$ is more ambiguity averse than $B$ if and only if, for every $x \in \mathcal{U}$,*

$$-\frac{\phi_A''(x)}{\phi_A'(x)} \geq -\frac{\phi_B''(x)}{\phi_B'(x)}.$$

Analogous to risk theory, we will call the ratio

$$\lambda(x) = -\frac{\phi''(x)}{\phi'(x)}$$

the **coefficient of ambiguity aversion** at $x \in \mathcal{U}$.

**Corollary 6** *Under Axioms 1-4, $\phi$ is concave if and only if the decision maker is more ambiguity averse than an expected utility decision maker, that is, a decision maker all of whose associated preferences $\succeq_\Pi$ are expected utility.*

**Remark 5** Corollary 6 connects our definition of smooth ambiguity aversion (Definition 4) to the comparative notion of ambiguity aversion in Definition 8. It shows that they agree, with expected utility taken as the dividing line between ambiguity aversion and ambiguity loving. Corollary 4, in Section 3, shows that if Axiom 5 holds nothing would change if we were to take probabilistic sophistication, rather than expected utility, as the benchmark.

We close the section by considering the two important special cases of constant and extreme ambiguity attitudes. We begin by defining a behavioral notion of constant ambiguity attitude.

**Definition 9** *Suppose acts $f$, $g$, $f'$, $g'$ and $k \in \mathbb{R}$ are such that, for each $\omega \in \Omega$ and $r \in [0,1)$,*

$$\begin{aligned} u(f'(\omega,r)) &= u(f(\omega,r)) + k \\ u(g'(\omega,r)) &= u(g(\omega,r)) + k. \end{aligned}$$

*We say that the decision maker displays **constant ambiguity attitude** if, for each closed subset $\Pi \subseteq \Delta$,*

$$f \succeq_\Pi g \Longleftrightarrow f' \succeq_\Pi g'.$$

To see the spirit of the definition notice, by bumping up utility (not the raw payoffs) in each state by a constant amount we achieve a uniform shift in the induced distribution over ex-ante evaluations, i.e.,

$$\mu_{f'} = \mu_f + k \text{ and } \mu_{g'} = \mu_g + k$$

The intuition of constant ambiguity attitude is that the decision maker views the "ambiguity content" in $\mu_f$ and its "translation" $\mu_f + k$ to be the same. Next we show that constant ambiguity attitudes are characterized by a negative exponential $\phi$. It is of some interest to note that the proposition does not assume that $\phi$ is differentiable.

**Proposition 4** *The decision maker displays constant ambiguity attitude if and only if there exists an $\alpha \neq 0$ such that, for all $x \in \mathcal{U}$, either $\phi(x) = x$ or $\phi(x) = -\frac{1}{\alpha}e^{-\alpha x}$, up to positive affine transformations.*

We now turn to extreme ambiguity attitudes. The Maxmin expected utility model (e.g., Gilboa and Schmeidler (1989)) with a given set of measures $\Pi$, may be viewed as an extreme case of our model as the ambiguity aversion tends to infinity, as shown in the following proposition. It may similarly be shown that Maxmax expected utility with a given set of measures $\Pi$, may be viewed as the case where ambiguity aversion tends to negative infinity (extreme ambiguity love).

**Proposition 5** *Let $\succeq$ be an ordering on $\mathcal{F}$ satisfying Axioms 1, 2 and 3. Then there exists a sequence $\{\succeq_n\}_{n=1}^{\infty}$ of orderings on $\mathcal{F}$ satisfying Axioms 1, 2 and 3, with $\succeq_1 = \succeq$, such that*

*(i) all $\succeq_n$ share the same vN-M utility function $u$ and the same measure $\mu$,*

*(ii) $\lim_n \lambda_n(x) = +\infty$ and $\lambda_n(x) \geq \lambda_{n-1}(x)$ for all $n \geq 1$ and all $x \in \mathcal{U}$.*

*Moreover, given any $f$ and $g$ in $\mathcal{F}$ if it holds, eventually, that $f \succeq_n g$, then,*

$$\min_{\pi \in \Pi} \mathbb{E}_\pi u(f) \geq \min_{\pi \in \Pi} \mathbb{E}_\pi u(g),$$

*while*

$$\min_{\pi \in \Pi} \mathbb{E}_\pi u(f) > \min_{\pi \in \Pi} \mathbb{E}_\pi u(g)$$

*implies that, eventually, $f \succ_n g$.*

# 5 Relating to other notions of ambiguity

In this section, we compare, in the context of our model, our definition of ambiguity with the definitions proposed in Epstein and Zhang (2001) and Ghirardato and Marinacci (2002). Throughout this section we assume that $\Omega$ is finite and that $\Sigma = 2^\Omega \otimes \mathcal{B}$.

We first relate our definition to the behavioral notion of ambiguity developed in Epstein and Zhang (2001). Their notion of ambiguity was designed to apply to a wide variety of models of preferences.

**Definition 10** *(Epstein-Zhang (2001)) An event $T$ is **unambiguous** if: (a) for all disjoint subevents $A$, $B$ of $T^c$, acts $h$, and outcomes $x^*, x, z, z'$,*

$$\begin{pmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \in B \\ h(s) & \text{if} & s \in T^c \setminus (A \cup B) \\ z & \text{if} & s \in T \end{pmatrix} \succsim \begin{pmatrix} x & \text{if} & s \in A \\ x^* & \text{if} & s \in B \\ h(s) & \text{if} & s \in T^c \setminus (A \cup B) \\ z & \text{if} & s \in T \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \in B \\ h(s) & \text{if} & s \in T^c \setminus (A \cup B) \\ z' & \text{if} & s \in T \end{pmatrix} \succsim \begin{pmatrix} x & \text{if} & s \in A \\ x^* & \text{if} & s \in B \\ h(s) & \text{if} & s \in T^c \setminus (A \cup B) \\ z' & \text{if} & s \in T \end{pmatrix};$$

*and (b) the condition obtained if $T$ is everywhere replaced by $T^c$ in (a) is also satisfied. Otherwise, $T$ is **ambiguous** .*

This definition works by looking for conditional likelihood reversals over events in the complement of the event being tested for ambiguity. Our definition instead looks for likelihood reversals involving the event being tested for ambiguity and events in $\mathcal{B}$. Note that such a strategy was not available to Epstein and Zhang since they wished their definition to have power in environments that might lack an appropriately rich set of unambiguous events.

Is our definition of ambiguity identical to that of Epstein and Zhang (2001) when applied to our model of preferences? The answer is no. How do they differ? It is of interest to note at the outset, given the many discussions comparing notions of ambiguity due to Epstein and Zhang (2001) and Ghirardato and Marinacci (2002), the reason for the difference cannot be that our definition confounds ambiguity with probabilistically sophisticated departures from expected utility. As we have shown in the previous subsection, whenever our definition identifies an event as ambiguous, behavior is not probabilistically sophisticated. Nevertheless, one might still expect, since we can take advantage of the rich structure of our model, that any difference would lie in the direction of our definition classifying more events as ambiguous than Epstein and Zhang. This is part of the story: in Example 1 we show that there may be some events that are ambiguous according to our definition that are not according to Epstein and Zhang.

**Example 1** Let $\Omega = \{\omega_1, \ldots, \omega_n\}$. The measure $\mu$ assigns probability 1/2 to both $\pi_0$ and $\pi_1$, where $\pi_0$ and $\pi_1$ yield marginals on $\Omega$ of

$$\pi_0(\omega_1) = \lambda_1, \ \pi_0(\omega_j) = \frac{1-\lambda_1}{n-1}, \ j \neq 1$$

and

$$\pi_1(\omega_1) = \lambda_2, \ \pi_1(\omega_j) = \frac{1-\lambda_2}{n-1}, \ j \neq 1,$$

respectively, with $\lambda_1 < \lambda_2$. The utility function is $u(x) = x$ (risk neutrality). The function $\phi$ is $\phi(x) = -e^{-\alpha x}$ with $\alpha > 0$.

Since $\pi_0(\omega_1) = \lambda_1 < \lambda_2 = \pi_1(\omega_1)$ and $\phi$ is strictly concave, Theorem 2 implies that the event $\omega_1 \times [0,1)$ is ambiguous according to the definition in this paper. We now demonstrate that it is unambiguous according to the definition of Epstein and Zhang (2001) (Definition 10). To this end, consider $T = \omega_1 \times [0,1)$ in their definition. Notice that all events in $T^c$ are assigned the same relative weights under $\pi_0$ as under $\pi_1$. Specifically, for any $E \subseteq T^c$, $\pi_0(E) = \frac{1-\lambda_1}{1-\lambda_2}\pi_1(E)$. Under the specified preferences, the top pair of acts in Definition 10 is evaluated according to

$$-\frac{1}{2}e^{-\alpha(\lambda_1 z + \pi_0(A)x^* + \pi_0(B)x + \int_{s\in T^c\setminus(A\cup B)} h(s)\pi_0(s))} - \frac{1}{2}e^{-\alpha(\lambda_2 z + \pi_1(A)x^* + \pi_1(B)x + \int_{s\in T^c\setminus(A\cup B)} h(s)\pi_1(s))}, \tag{5}$$

and

$$-\frac{1}{2}e^{-\alpha(\lambda_1 z + \pi_0(A)x + \pi_0(B)x^* + \int_{s\in T^c\setminus(A\cup B)} h(s)\pi_0(s))} - \frac{1}{2}e^{-\alpha(\lambda_2 z + \pi_1(A)x + \pi_1(B)x^* + \int_{s\in T^c\setminus(A\cup B)} h(s)\pi_1(s))}, \tag{6}$$

respectively. Substituting $\pi_0(E) = \frac{1-\lambda_1}{1-\lambda_2}\pi_1(E)$ and simplifying yields (5) $\geq$ (6) if and only if

$$\pi_1(A)x^* + \pi_1(B)x \geq \pi_1(A)x + \pi_1(B)x^*. \tag{7}$$

The evaluation of the lower pair of acts in Definition 10 differs only by substituting $z'$ for $z$. It is not hard to show that preference in the lower pair is also determined by (7). Therefore condition (a) of the definition is satisfied. Setting $T = (\omega_1 \times [0,1))^c$ and noting that, again, all events in $T^c$ will be assigned the same relative weights under any $\pi$ leads to the conclusion that condition (b) of the definition holds as well. Therefore, $\omega_1 \times [0,1)$ is unambiguous according to Definition 10. ■

The intuition behind this example is that Epstein and Zhang (2001) ambiguity requires a reversal in the relative likelihoods of two events lying in the complement of the candidate ambiguous event. The preferences in the example have the property that the ambiguity on the candidate event affects the likelihoods of all events in its complement in the same way – thus relative likelihoods in the complement are unchanged. In the formal result below, we show that such situations are in some sense rare in that by perturbing the beliefs of the decision maker even slightly one may have the ambiguity generate the differences in relative likelihoods needed for Definition 10.

This is not the whole story however. We show in Example 2 that the Epstein and Zhang definition may, more surprisingly, classify as ambiguous some events that are unambiguous according to our definition. This may appear somewhat strange, given that if $E$ is unambiguous in our framework (with $\succeq$ satisfying Axioms 1 through 5) then preferences restricted to acts measurable with respect to the $\lambda$-system generated by $\{\emptyset, E, E^c, \Omega \times \mathcal{B}, S\}$ are probabilistically sophisticated (and, in fact, are expected utility).

**Example 2** Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. The measure $\mu$ assigns probability 1/2 to both $\pi_0$ and $\pi_1$, where $\pi_0$ and $\pi_1$ yield marginals on $\Omega$ of

$$\pi_0(\omega_1) = 0.5, \pi_0(\omega_2) = 0.22, \pi_0(\omega_3) = 0.18, \pi_0(\omega_4) = 0.1,$$

and

$$\pi_1(\omega_1) = 0.5, \pi_1(\omega_2) = 0.02, \pi_1(\omega_3) = 0.18, \pi_1(\omega_4) = 0.3,$$

respectively. The utility function is $u(x) = x$ (risk neutrality). The function $\phi$ is $\phi(x) = ln(x)$.

Since $\pi_0(\omega_1) = \pi_1(\omega_1) = 0.5$, the event $\omega_1 \times [0,1)$ is unambiguous according to the definition in this paper. We now demonstrate that it is ambiguous according to the definition of Epstein and Zhang (2001) (Definition 10). In their definition, let $T = \omega_1 \times [0,1)$, $A = \omega_3 \times [0,1)$, $B = \omega_4 \times [0,1)$, $x^* = 1, x = 0, h(s) = 0, z = 0$, and $z' = 100$. Calculation shows that for the top pair of acts in their definition the act on the left is strictly preferred to the one on the right, since $ln(0.18) > 0.5\ln(0.1) + 0.5\ln(0.3)$. However, for the bottom pair of acts the preference is reversed, as $ln(50.18) < 0.5\ln(50.1) + 0.5\ln(50.3)$. ■

What is going on in this example? The function $\phi(x) = ln(x)$ does not reflect a constant ambiguity attitude. In particular, as the acts under consideration get better and better in terms of expected utilities, the ambiguity aversion of such a decision maker diminishes. This is in close analogy to risk theory, since $ln(x)$ displays constant relative risk aversion but diminishing absolute risk aversion when used as a utility function. In our example, $A$ is unambiguous, while $B$ is ambiguous but has a higher average probability than $A$. When $z = 0$ the expected utilities of the acts are relatively low under all the

measures and the portion of the $ln(x)$ function that is quite ambiguity averse is the relevant one. Here $A$ gets favored over $B$. When considering $z' = 100$ however, the acts get considerably higher expected utilities and a portion of the $ln(x)$ function that is much less concave (ambiguity averse) is relevant. In the second case, ambiguity aversion has diminished enough that the decision maker is now willing to favor the ambiguous-but-higher-average-probability event, $B$ over the unambiguous one, $A$. Examples of this kind may be constructed quite generally when ambiguity aversion is not constant. In such a case, changing payoffs on *any* event $E$, no matter what its ambiguity status, may lead to a conditional likelihood reversal between two events in $E^c$ *if at least one of the two events in* $E^c$ *is ambiguous*. This occurs because changing payoffs may change ambiguity attitude thus possibly affecting the decision maker's ranking of events. If all events in $E^c$ are unambiguous, changing ambiguity attitude cannot affect the ranking of these events and thus a conditional likelihood reversal would not occur in this case.

This example suggests that when a rich set of events like $\mathcal{B}$ over which the decision maker has a probability is available, our approach allows one to distinguish between reversals due to the ambiguity of the event being tested and those due to changing ambiguity attitude.

As the next theorem shows, the differences identified in the above two examples are in some sense the only ones separating the two definitions of ambiguity in our setting. The second example is dealt with by the assumption of constant ambiguity aversion. The first example is dealt with by showing that perturbing the beliefs of the decision maker by as small amount as one wishes can eliminate this type of disagreement. We should note up-front that the statement in part (b) of the theorem below (concerning differences in the direction of the first example) does not go as far as one might hope. In particular, the perturbation argument we develop works for events in $2^{\Omega} \times [0,1)$ rather than general events in $\Sigma$, assumes that $\mu$ has a finite support and, even though the perturbations required are arbitrarily small, we have not been able to rule out that they might change the ambiguity classification of some compound events (i.e., events not of the form $\omega \times [0,1)$). This contrasts with part (a) of the theorem (concerning differences in the direction of the second example), which applies to all events and whose proof does not make use of the finite support assumption. Part (a) shows quite strongly that differences in the direction of the second example indeed stem from non-constant ambiguity attitude.

**Definition 11** *An event $E \subseteq \Omega$ is **null** if, for all $x, y, z \in \mathcal{C}$,*

$$\begin{pmatrix} x \text{ if } s \in E \times [0,1) \\ y \text{ if } s \in E^c \times [0,1) \end{pmatrix} \sim \begin{pmatrix} z \text{ if } s \in E \times [0,1) \\ y \text{ if } s \in E^c \times [0,1) \end{pmatrix}.$$

**Proposition 6** *An event $E \subseteq \Omega$ is **null** if and only if $\pi(E \times [0,1)) = 0$, $\mu$-a.e.*

**Theorem 4** *Assume that $\Omega$ contains at least three non-null states and that $\phi(x) = -e^{-\alpha x}$, $\alpha > 0$.*

*(a) If an event $E \in \Sigma$ is unambiguous according to Definition 5, then it is unambiguous according to the definition in Epstein-Zhang (2001).*

*(b) If an event $E \times [0,1)$ is ambiguous according to Definition 5, then there exists a sequence of perturbations of $\Pi$, denoted $\Pi(\epsilon_n)$, with $\lim_{n\to\infty} \Pi(\epsilon_n) = \Pi$ and $\lim_{n\to\infty} \epsilon_n = 0$, such that for all $n$, given $\Pi(\epsilon_n)$:*

*(i) The event $E \times [0,1)$ is ambiguous according to the definition in Epstein-Zhang (2001), and,*

*(ii) for each $\omega \in \Omega$, $\omega \times [0,1)$ is ambiguous according to Definition 5 if and only if $\omega \times [0,1)$ was ambiguous according to Definition 5 given $\Pi$.*

An analogous result is true with $\alpha < 0$ and $\phi(x) = e^{-\alpha x}$ (constant ambiguity love).

Next we relate our behavioral definition of ambiguity to the behavioral definition proposed by Ghirardato and Marinacci (2002). The definition they propose applies to a very general class of preferences which includes the preferences axiomatized in the present paper. In their definition, they invoke a notion of a benchmark preference for $\succeq$ which is any SEU preference that is less ambiguity averse than $\succeq$ (in the sense of Definition 8).

**Definition 12** *(Ghirardato and Marinacci (2002)) For a given benchmark preference $\geqslant$ for $\succeq$ the set of $\geqslant$ −**unambiguous acts**, denoted $\mathcal{H}^{gm}_{\geqslant}$, is the largest subset of $\mathcal{F}$ satisfying the following two conditions:*

*(A) For every $f \in \mathcal{H}^{gm}_{\geqslant}$ and every $x \in \mathcal{C}$,*

$$\delta_x \geqslant f \Longleftrightarrow \delta_x \succeq f \text{ and } \delta_x \leqslant f \Longleftrightarrow \delta_x \preceq f;$$

*(B) For every $f \in \mathcal{H}^{gm}_{\geqslant}$ and every $g \in \mathcal{F}$, if $\{g^{-1}(x) : x \in \mathcal{C}\} \subseteq \{f^{-1}(x) : x \in \mathcal{C}\}$, then $g \in \mathcal{H}^{gm}_{\geqslant}$.*

Theorem 4 in the working paper version of Ghirardato and Marinacci (2002),[5] shows that $\mathcal{H}^{gm}_{\geqslant}$ is actually independent of the particular benchmark $\geqslant$ and hence we may drop the subscript and let $\mathcal{H}^{gm}$ denote the set of acts identified as unambiguous in the Ghirardato-Marinacci definition. Correspondingly, they define the set of unambiguous events $\Lambda^{gm}$ as the collection of *all* the 'upper pre-image' sets of the unambiguous acts $f \in \mathcal{H}^{gm}$. (An upper pre-image set of an act $f : S \to \mathcal{C}$ is a set of the form $\left\{s : \delta_{f(s)} \succeq \delta_x\right\}$ for $x \in \mathcal{C}$.)

In the following theorem we show if $\succeq$ satisfies all our axioms, the collection of unambiguous events ($\Lambda$) and the collection of unambiguous acts ($\mathcal{H}$) identified by the definitions in this paper (Definitions 5 and 6, respectively) are identical to the collections identified by the corresponding definitions in Ghirardato and Marinacci (2002).

**Theorem 5** *Suppose $\succeq$ satisfies Axioms 1 through 5. Then, $\mathcal{H} = \mathcal{H}^{gm}$ and $\Lambda = \Lambda^{gm}$.*

While Ghirardato and Marinacci (2002) does provide a simple characterization of ambiguous acts and events for case of biseparable preference they do not have a corresponding result for general preferences. In particular, the class of preferences axiomatized in the present paper is not contained in the biseparable class. Hence, the formal relationship described above between our definitions and the Ghirardato-Marinacci definitions of ambiguous acts and events in the context of $\succeq$, does not follow from the characterization obtained in Ghirardato and Marinacci (2002). What explains the above result that the two definitions of ambiguous acts, ours and Ghirardato-Marinacci's, coincide in our

[5] For convenience, we reproduce the relevant section of the theorem in the Appendix as Theorem 7. It appears just before the statement of the proof of our Theorem 5.

model? The explanation lies in two key properties shared by the definitions. One is that according to both definitions, if two acts are measurable with respect to the same collection of events, then if one of the acts is unambiguous so is the other; the actual payoffs do not matter only the partition generated by the acts. This property is, for instance, not shared by the Epstein-Zhang definition: we saw in Example 2 simply changing payoffs (without changing the measurability of the act) can affect the ambiguity status of the underlying events. The other key property is that the set of unambiguous events, according to Ghirardato-Marinacci, is a partition such that the ranking between acts measurable with respect to that partition is the same, whether given by the benchmark preference relation or whether given by the ambiguity sensitive preference relation: behavior with respect to these events is unaffected by ambiguity attitude. In our model there is a natural benchmark corresponding to any given preference, namely the SEU relation obtained by making reduction hold (i.e., Axiom 5 (i) ) without changing any other aspect of the given preference. Our definition of ambiguity identifies those events as unambiguous for which there is no disagreement about probabilities. In our model, ranking between acts measurable with respect to such events according to the given preference must coincide with that corresponding to the benchmark SEU preference.

# 6 More on related literature

A key idea in the present paper, relaxing reduction between first and second order probabilities to accommodate ambiguity sensitive preferences, owes its inspiration to the research reported in Segal (1987). That paper presented a model of decision making under uncertainty which assumes a unique second order probability over a set of given first order probabilities, but relaxes reduction and weights the possible first order probabilities non-linearly. Using examples, Segal observed that such a model would be flexible enough to accommodate both Allais and Ellsberg type behavior. While Segal (1987) did not axiomatize a general preference order incorporating ambiguity aversion, Segal (1990) developed the key idea of relaxing reduction further in the context of choice under risk and obtained a novel axiomatization of the Anticipated Utility model.

Another paper that invokes this same idea of relaxing reduction to model ambiguity aversion is Nau (2001), (revised and expanded in Nau (2002)). The paper presents an axiomatic model of partially separable preferences where the decision maker may satisfy the independence axiom selectively within partitions of the state space whose elements have "similar degrees of uncertainty". The axiomatization makes no attempt to uniquely separate beliefs from state-dependent cardinal utilities. Indeed, a major and innovative contribution of the paper(s) is to present an intuitive notion of ambiguity aversion in a state dependent utility framework; this goes beyond what is in the previous literature, which does not allow state dependent utility. Our paper, on the other hand, axiomatizes and investigates a model of ambiguity aversion staying very much within the state independent utility framework, separating beliefs and ambiguity attitude from state independent utilities, while using Segal's idea of relaxing reduction. The notion of ambiguity aversion characterized in our paper is not as general as Nau's in that it is not applicable to a state dependent framework. Also, as we show in our formal results, our notion of ambiguity is closely related to the principal notions in the existing literature. A related point

which underscores the more modest scope of our analysis, compared to Nau's, is that in our framework the only departure from SEU is ambiguity aversion/preference, whereas in Nau's framework departure from SEU can occur in various ways. Ergin and Gul (2002), a work contemporaneous to ours, considers a preference framework very analogous to Nau's and obtains a representation which, at least in a special case, is essentially the same as obtained in the present paper. Just as Nau's framework has two possible partitions of the state space with the DM being (possibly) differently risk averse on one partition as compared to the other, Ergin and Gul's framework is also a product space, each component of which they call an "issue". Their key axiom permits "issue preference", that is, agents may not be indifferent among gambles that yield the same probability distribution if they depend on different issues. In other words, the DM is permitted to have different risk attitudes on different components of the product space. Unlike Nau, Ergin and Gul do not allow for state-dependence. However, an important respect in which their analyses is more general than either Nau's or ours is that they do not invoke expected utility on each component but only require probabilistic sophistication. They do raise a matter of interpretation, in particular the verifiability of the states underlying second order acts, that might differentiate our work from theirs. This partly motivated our discussion of verifiability in Remark 2.

The seminal work of Kreps and Porteus (1978) is not concerned with ambiguity, but is quite related to our modeling approach in that the representation we derive has a two-stage recursive form with expected utility at each stage. This is similar to a simplified, two-stage, with only last stage consumption version of their functional. Moreover, as they relate the convexity/concavity of their first stage evaluation function to attitude towards the timing of the resolution of risk, we are able to relate convexity/concavity of our first stage evaluation function with attitude towards ambiguity. In both cases, neutrality leads reduction to hold and preferences to be expected utility. The analogy is not perfect, however. Aside from the obvious fact that Kreps and Porteus (1978) works with purely objective temporal lotteries while we are in a more subjective setting, there is at least one important structural difference in the objects over which preferences are formed. While it is possible to associate a natural two-stage lottery with each act in our setting (imagine first choosing a probability over the state space according to a subjective distribution and then imagine that probability determining the state and thus the realized outcome according to the act) only a restricted subset of two-stage lotteries will be so associated. This is true both because there is only one distribution over probabilities over the state space and, more importantly, because the second stage lotteries induced by an act as the probabilities over the state space vary will be strongly related due to the underlying mapping of states to outcomes. For this reason it does not appear fruitful for us to try to directly adapt existing work on recursive preferences over two-stage lotteries to our setting. For example, Grant, Kajii, and Polak (2000) examine the degree to which the assumption of expected utility at each stage may be relaxed for recursive preferences over two-stage lotteries. They find some striking results. For instance, if the decision-maker always prefers later resolution of risk, if both first and second stage preferences satisfy the rank dependent model, and if second stage preferences display risk aversion then second stage preferences must in fact be expected utility. Reasoning by analogy and drawing on some of our results, this suggests that in our model, if the decision maker is both ambiguity averse and risk averse, then our assumption that preferences

over lotteries are expected utility may be relaxed to rank dependence while leaving the representation unchanged (i.e., still implying that preferences over lotteries are expected utility). However, we believe that this is not true as explain at the end of Section 2.2. We conjecture that the Grant, Kajii, and Polak (2000) results fail to translate to our setting precisely because of the restricted set of two-stage lotteries implicit in our framework.

Another paper modeling ambiguity attitude by indirectly relaxing the reduction assumption is Halevy and Feltkamp (2001). They try to rationalize ambiguity aversion by assuming that a decision maker mistakenly views his choice of an action as determining payoffs for two positively related replications of the same environment, rather than simply for a single environment. If he is risk averse and has expected utility preferences over a single instance then this "bundling" of problems may lead to Ellsberg type behavior. Nehring (2001) defines a class of preferences, termed "utility sophisticated", whose identifying spirit is that the members of this class do not show Allais type departures from expected utility though they do allow for Ellsberg type behavior. The preferences modeled in the present paper do satisfy this spirit though it remains to be formally verified whether the preferences actually satisfy utility sophistication as defined by Nehring.

In the introduction, we mentioned that the MEU model of preferences may be criticized for conflating perception of ambiguity (or more generally, beliefs) with attitude towards ambiguity (tastes). In particular we pointed out two shortcomings in assuming that the set of priors $\Pi$ reflected beliefs only: (1) homogeneity of ambiguity attitude would be imposed and (2) that ambiguity attitude would be extreme. The two problems may seemingly be addressed by a generalization of the maxmin functional to the so called $\alpha$-maxmin form($\alpha$-MEU):[6]

$$\hat{V}(f) = \alpha \max_{\pi \in \Pi} \mathbb{E}_\pi (u \circ f) + (1-\alpha) \min_{\pi \in \Pi} \mathbb{E}_\pi (u \circ f),$$

where the ambiguity attitude may be varied parametrically by varying $\alpha$. However, there is another limitation of the maxmin model that is not resolved by the $\alpha$-maxmin generalization and from which our model does not suffer. The limitation is, once again, particular to the case where we interpret $\Pi$ simply as the set of probabilities that the decision maker thinks is relevant. This interpretation of the $\alpha$-maxmin functional implies that the preferences do not take into account any information that the decision maker may have about plausibility of each $\pi$ in $\Pi$. The functional does not smoothly aggregate the information concerning how the act performs under each possible $\pi$ but only looks at the extremal performance values (the worst under maxmin, the best and the worst under $\alpha$-maxmin). Taking $\Pi$ to be the set of probabilities subjectively thought as possible, the choice rule generated by the functional looks almost pathological and even, in a way, contrary to the general intuition following from Ellsberg type behavior. For instance, take two acts $f$ and $g$ which share the same extremal valuations (i.e., $\max_{\pi \in \Pi} \mathbb{E}_\pi (u \circ f) = \max_{\pi \in \Pi} \mathbb{E}_\pi (u \circ g)$ and $\min_{\pi \in \Pi} \mathbb{E}_\pi (u \circ f) = \min_{\pi \in \Pi} \mathbb{E}_\pi (u \circ g)$) but for "almost all" probabilities in $\Pi$, $\mathbb{E}_\pi (u \circ f) > \mathbb{E}_\pi (u \circ g)$. The choice rule would rank the acts equally. Relatedly, take two acts $f$ and $g$ which share the same extremal valuations but suppose valuations for $f$ are far less spread out than $g$'s. Again, $f$ and $g$ would be ranked equally, which is somewhat contrary to the general Ellsberg intuition that ambiguity averse decision makers have preference for acts whose performance is more stable,

[6] This form has recently been axiomatized in Ghirardato, Maccheroni, and Marinacci (2002). A similar functional was proposed previously in Jaffray (1995).

more robust to the possible variation in probabilities. The following example serves to illustrate the point.

**Example 3** Consider two acts $f$ and $g$ whose utility payoffs in states 1, 2, 3 and 4 are as given in the following table:

| | State 1 | State 2 | State 3 | State 4 |
|---|---|---|---|---|
| $u \circ f$ | 1 | 2 | 3 | 4 |
| $u \circ g$ | 1 | 1 | 4 | 4 |

Figure 1, depicts the two acts.

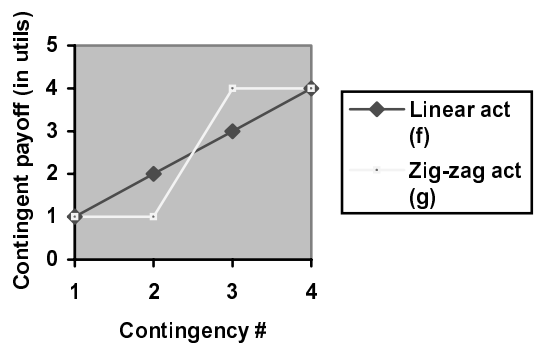


Figure 1: Contingent utility payoffs of $f$ and $g$

Suppose a decision maker thinks the four probabilities, $\pi_1$, $\pi_2$, $\pi_3$, $\pi_4$, described in the table below, are possible/relevant.

| | State 1 | State 2 | State 3 | State 4 |
|---|---|---|---|---|
| $\pi_1$ | .7 | .1 | .1 | .1 |
| $\pi_2$ | .1 | .7 | .1 | .1 |
| $\pi_3$ | .1 | .1 | .7 | .1 |
| $\pi_4$ | .1 | .1 | .1 | .7 |

The points to note about the possible $\pi_i$'s are that each favors one state, is uniform across the other states, and every state is favored by some $\pi_i$. The "linear" act $f$ is more robust to variation in probabilities than is the "zigzag" act $g$. The latter, which concentrates its payoffs on states 3 and 4, does very well under $\pi_3$, $\pi_4$ but rather poorly under $\pi_1$, $\pi_2$. Notice, though, the best and worst expected utility evaluations are the same for the two acts, hence any $\alpha$-maxmin functional will evaluate the two acts identically. The functional is insensitive to the non-robustness of $g$ relative to $f$. The insensitivity arises because of the "non-smooth" aggregation of the information that results from exclusive concentration on extremal values. The point is made more vividly if one considers a prior on the $\pi_i$'s, say $\mu$, as an example of a decision maker's information about the relevance of each $\pi_i$, and use the prior to aggregate the information smoothly by looking at the induced distribution on the expected utility values generated by each act. Given $f$ and

$\mu$, $\mu_f$ is generated in the following way: $\mu_f\left(E_{\pi_i}\left(u \circ f\right)\right) = \mu\left(\pi_i\right)$. Figures 2 and 3 show the distributions corresponding to $\mu$ being the uniform prior. As is evident, $\mu_g$ is a mean preserving spread of $\mu_f$. ■

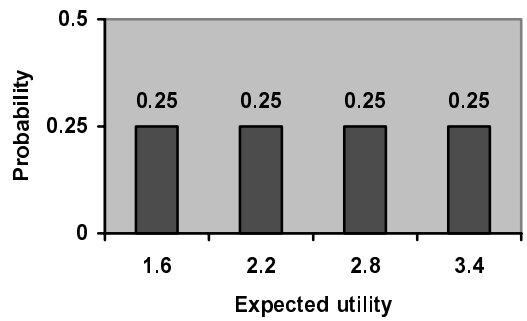


Figure 2: $\mu_f$ : Probability distribution on expected utility generated by $f$

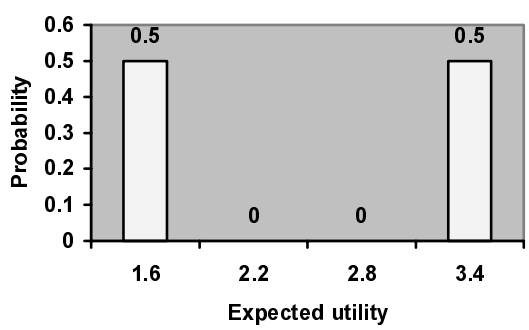


Figure 3: $\mu_g$ : Probability distribution on expected utility generated by $g$.

Finally, we remark that it may be helpful to think of the part of the difference between the model in this paper and models such as MEU and $\alpha$-MEU as analogous to that between first and second order risk aversion (Segal and Spivak (1990), Loomes and Segal (1994)). Models such as MEU and $\alpha$-MEU display ambiguity averse behavior when the corresponding indifference curves in the utility space are kinked (behavior which may be dubbed first-order ambiguity aversion). The model in this paper focuses on incorporating ambiguity aversion even when the indifference curves are not kinked ("second order ambiguity aversion"), thus the moniker "a smooth theory."[7]

[7] Of course, kinks are not ruled out in our model and may be included if desired – they are simply not needed to generate behavior reflecting ambiguity.

# 7 Portfolio Choice Examples

In this section we consider two examples of simple portfolio choice problems. The examples are intended both as a simple, concrete illustration of our framework and as suggestive of the potential of our approach in applications. We focus, in particular, on comparative statics in risk attitude and in ambiguity attitude.

The environment for the examples is as follows. The space $\Omega$ contains two elements, $\omega_1$ and $\omega_2$. The measure $\mu$ assigns probability 1/2 to both $\pi_0$ and $\pi_1$, where $\pi_0$ and $\pi_1$ yield marginals on $\Omega$ of

$$\pi_0(\omega_1) = \frac{1}{4}, \pi_0(\omega_2) = \frac{3}{4} \text{ and } \pi_1(\omega_1) = \frac{3}{4}, \pi_1(\omega_2) = \frac{1}{4},$$

respectively. The function $u$ is given by $u(x) = \begin{cases} -e^{-ax} & if \quad a > 0 \\ x & if \quad a = 0 \end{cases}$. This utility function displays constant absolute risk aversion with $a$ as the coefficient of absolute risk aversion. The function $\phi$ is given by $\phi(x) = \begin{cases} -e^{-\alpha x} & if \quad \alpha > 0 \\ x & if \quad \alpha = 0 \end{cases}$. This function may be said to display *constant ambiguity aversion* with this terminology justified by Proposition 4 of the previous section. $\alpha$ is thus the *coefficient of ambiguity aversion.*

Table 1 illustrates the acts that will appear in our examples. Each of these acts is meant to represent the gross payoff (in dollars) per dollar invested in a particular asset as a function of the state of the world.

| | $\omega_1 \times [0, \frac{1}{2})$ | $\omega_1 \times [\frac{1}{2}, 1)$ | $\omega_2 \times [0, \frac{1}{2})$ | $\omega_2 \times [\frac{1}{2}, 1)$ |
|---|---|---|---|---|
| $f$ | 2 | 2 | 1 | 1 |
| $l$ | 3 | 1 | 3 | 1 |
| $\delta_{1.15}$ | 1.15 | 1.15 | 1.15 | 1.15 |

Table 1. Gross \$ payoff per \$ invested for each of three assets.

Observe that $f$ is an example of an ambiguous act, as its payoff depends on the ambiguous events $\omega_1 \times [0, 1)$ and $\omega_2 \times [0, 1)$. $l$ is an example of an unambiguous, but risky, act (it is also a lottery). $\delta_{1.15}$ is an example of a constant act, involving neither risk nor ambiguity. Thinking of these in terms of assets and asset returns, $f$ reflects a 100% return when the state of the world $s \in \omega_1 \times [0, 1)$ and 0% otherwise; $l$ reflects a return of 200% with probability 1/2 and a return of 0% with probability 1/2; and $\delta_{1.15}$ reflects a sure return of 15%.

**Example 4**: (**Allocating \$1 between a safe asset & an ambiguous asset**)

The classic simple example of a static portfolio choice problem is the decision of how to allocate wealth between a safe asset and a risky asset. As is well known, an increase in risk aversion (here an increase in $a$) leads more wealth to be invested in the safe asset. Here, the asset underlying $f$ is not only risky but is also ambiguous. $\delta_{1.15}$ is the safe asset. By varying $\alpha$ and $a$ in this example, we can vary the ambiguity aversion and risk aversion of the agent respectively. What are the comparative statics results in this framework? Just as with a risky asset, holding ambiguity aversion ($\alpha$) fixed, an increase in risk aversion ($a$) leads more to be invested in the safe asset. Furthermore, holding risk

aversion ($a$) fixed, an increase in ambiguity aversion ($\alpha$) leads more to be invested in the safe asset. Table 2 gives a numerical illustration of this effect when risk aversion is fixed at $a = 2$.

| ambiguity aversion ($\alpha$) | amount allocated to safe asset |
|---|---|
| 0 | 0.1327 |
| 0.02 | 0.132973 |
| 2 | 0.158941 |
| 20 | 0.344943 |
| 100 | 0.620079 |
| 200 | 0.675267 |

Table 2. Optimal amount out of \$1 allocated to the safe asset as ambiguity aversion varies holding risk aversion at $a = 2$.

In this example, ambiguity aversion and risk aversion work in the same direction. If we view the ambiguous asset as a proxy for equities, this example suggests that if observed portfolio allocations between equities and safe assets are rationalized by risk aversion only – ignoring ambiguity aversion – then levels of risk aversion may be overestimated. Thus ambiguity aversion may play a role in helping to explain the equity premium puzzle. A number of previous papers have noted this possible role for ambiguity aversion, including Chen and Epstein (2002), Epstein and Wang (1994). Also, work including Hansen, Sargent, and Tallarini (1999) and Hansen, Sargent, Turmuhambetova, and Williams (2001) has suggested that model uncertainty plays a similar role in reinforcing risk. While the cited papers are complete, dynamic models and we present merely a very simple, static example, one reason to think that our approach may be useful here is the particularly clean separation between tastes (risk aversion, ambiguity aversion) and beliefs ($\mu$) it provides, which allows one to be confident in doing comparative statics that the intended feature is all that is being varied. ■

Our second example will show that ambiguity aversion and risk aversion do not always reinforce each other. In particular, there can be a trade-off between risk and ambiguity.

**Example 5** : (**Allocating \$1 between a risky asset & an ambiguous asset**)

Here we consider the allocation problem where the assets available are the risky (but unambiguous) asset underlying $l$ and the ambiguous asset underlying $f$. If the agent is both ambiguity neutral ($\alpha = 0$) and risk neutral ($a = 0$), the optimal allocation is to invest as much as possible in asset $l$, since its return would first-order stochastically dominate that of $f$. However, as risk aversion increases, holding ambiguity aversion fixed, the agent will want to diversify into the asset $f$ since it is not perfectly correlated with $l$, trading-off expected return against risk. On the other hand, as ambiguity aversion increases, holding risk aversion fixed, the ambiguity about the payoff from $f$ drives the agent towards $l$, as $f$ becomes a less effective diversifier and less valuable. Thus risk aversion and ambiguity aversion work in opposite directions here. Tables 3 and 4 give

numerical illustrations of these effects.

| Coefficient of ambiguity aversion ($\alpha$) | Optimal amount allocated to asset $l$ |
|---|---|
| 0 | 0.47577 |
| 0.02 | 0.475807 |
| 2 | 0.479518 |
| 20 | 0.511136 |
| 100 | 0.599734 |
| 200 | 0.635793 |

Table 3. Optimal amount out of \$1 allocated to the risky asset as ambiguity aversion increases holding risk aversion at $a = 2$

| Coefficient of risk aversion ($a$) | Optimal amount allocated to asset $l$ |
|---|---|
| 0 | $+\infty$ (1) |
| 0.02 | 21.2026 (1) |
| 2 | 0.47577 |
| 20 | 0.344886 |
| 100 | 0.335644 |
| 200 | 0.334489 |

Table 4. Optimal amount out of \$1 allocated to the risky asset as risk aversion increases holding ambiguity aversion at $\alpha = 0$.

In this case, if such behavior is examined ignoring ambiguity aversion, the agent's level of risk aversion will typically be underestimated. This suggests that ambiguity may play a role in explaining the underdiversification puzzle – the finding that the portfolios of risky assets that individuals hold are not diversified as much as plausible levels of risk aversion say they should be. One example of the underdiversification puzzle is home-bias, where the assets that are not sufficiently diversified into are those of companies geographically removed from the investor. If one hypothesizes that investors are ambiguity averse and perceive more ambiguity with increased distance then this could generate home-bias. Generation of underdiversification in the context of a model uncertainty framework appears in Uppal and Wang (2002). Epstein and Miao (2001) generates home-bias in a heterogeneous agent dynamic multiple priors setting. ■

# A Appendix: Proofs and Related Material

## A.1 Theorem 1

By Axiom 3, $f \succeq g \Leftrightarrow f^2 \succeq^2 g^2$. By Axiom 2, $f^2 \succeq^2 g^2 \Leftrightarrow \int v(c_f(\pi))\, d\mu \geq \int v(c_g(\pi))\, d\mu$. Hence,

$$f \succeq g \Longleftrightarrow \int v(c_f(\pi))\, d\mu \geq \int v(c_g(\pi))\, d\mu. \tag{8}$$

Given Remark 1, $v$ and $u$ must represent the same ordinal preferences on $\mathcal{C}$. Therefore, $v(c_f(\pi)) = \phi(u(c_f(\pi)))$ for some strictly increasing $\phi$. Since $v$ and $u$ are continuous, so is $\phi$. Substituting for $v(c_f(\pi))$ in (8), we get

$$f \succeq g \Longleftrightarrow \int \phi(u(c_f(\pi)))\, d\mu \geq \int \phi(u(c_g(\pi)))\, d\mu. \tag{9}$$

Now, recall,

$$\delta_{c_f(\pi)} \sim l_f \Longleftrightarrow u(c_f(\pi)) = \int_{[0,1)} u(l_f(r))\, dr.$$

So,

$$u(c_f(\pi)) = \sum_{x \in supp(\pi_f)} u(x)\pi_f(x) = \int_S u(f(s))\, d\pi. \tag{10}$$

Thus, substituting (10) into (9),

$$f \succeq g \Longleftrightarrow \int \phi\left(\int_S u(f(s))\, d\pi\right) d\mu \geq \int \phi\left(\int_S u(g(s))\, d\pi\right) d\mu$$

This proves the representation claim in the Theorem. To see the uniqueness properties of $\phi$, notice that

$$v(c_f(\pi)) = \phi(u(c_f(\pi))) \Leftrightarrow \phi(y) = v\left(u^{-1}(y)\right).$$

Let $\tilde{u} = \alpha u + \beta$ and let $y \in \mathcal{U}$. Then,

$$\begin{aligned}
\left(\tilde{u}^{-1}\right)(\alpha y + \beta) &= \{x : \tilde{u}(x) = \alpha y + \beta\} \\
&= \{x : \alpha u(x) + \beta = \alpha y + \beta\} \\
&= \{x : u(x) = y\} \\
&= u^{-1}(y).
\end{aligned}$$

Hence, $\forall y \in \mathcal{U}$, $\tilde{\phi}(\alpha y + \beta) = (v \circ \tilde{u}^{-1})(\alpha y + \beta) = (v \circ u^{-1})(y) = \phi(y)$. Finally, if there exists $J \subseteq \Delta$ with $0 < \mu(J) < 1$, then $v$ is unique up to positive affine transformations according to Axiom 2, so, fixing $u$, $\phi$ is as well. ■

We close this section by stating the result mentioned right after Theorem 1.

**Theorem 6** *Let $\succeq$ and $\succeq^2$ be two binary relations on $\mathcal{F}$ and $\mathfrak{F}$, respectively. The following statements, (i) and (ii), are equivalent:*

*i. Axioms 1, 2 and 3 hold.*

*ii. There exists a continuous, strictly increasing $\phi : \mathcal{U} \to \mathbb{R}$, a unique finitely additive probability $\mu : \sigma(\Delta) \to [0,1]$, continuous and strictly increasing utility functions $v : \mathcal{C} \to \mathbb{R}$ and $u : \mathcal{C} \to \mathbb{R}$, such that*

*(a) $\phi = v \circ u^{-1}$*

*(b)* $\succeq^2$ *is represented by the preference functional* $V^2 : \mathfrak{F} \to \mathbb{R}$ *given by*

$$V^2(\mathfrak{f}) = v(\mathfrak{f})\, d\mu$$

*(c)* $\succeq$ *is represented by the preference functional* $V : \mathcal{F} \to \mathbb{R}$ *given by*

$$V(f) = \int_{\mathcal{U}} \phi(x)\, d\mu_f = \int \phi \left[ \int_S u(f)\, d\pi \right] d\mu \equiv \mathbb{E}_\mu \phi\left(\mathbb{E}_\pi u \circ f\right).$$

*If there exists* $J \subseteq \Delta$ *with* $0 < \mu(J) < 1$*, then* $v$ *is unique up to positive affine transformations. Moreover,* $u$ *is unique up to positive affine transformations and if* $\tilde{u} = \alpha u + \beta$*,* $\alpha > 0$*, then the associated* $\tilde{\phi}$ *is such that* $\tilde{\phi}(\alpha y + \beta) = \phi(y)$*, where* $y \in \mathcal{U}$*.*

## A.2 Results of Subsection 2.3

We begin with a useful lemma (see Theorems 88 and 91 in Hardy, Littlewood, and Polya (1952)).

**Lemma 3** *Let* $\phi : A \subseteq \mathbb{R} \to \mathbb{R}$ *be a continuous function defined on a convex set* $A$*. Then,* $\phi$ *is concave (strictly concave) if and only if there exists* $\lambda \in (0,1)$ *such that, for all* $x, y \in A$ *with* $x \neq y$*,*

$$\phi(\lambda x + (1-\lambda) y) \geq (>)\, \lambda \phi(x) + (1-\lambda)\phi(y). \tag{11}$$

**Proof of Proposition 1.** (i) implies (iii): By the Jensen inequality, $\phi\left(\int x d\mu_f\right) \geq \int \phi(x)\, d\mu_f$. Thus, $\phi\left(e\left(\mu_f\right)\right) \geq \int \phi(x)\, d\mu_f$, which in turn implies $\delta_{u^{-1}\left(e\left(\mu_f\right)\right)} \succeq f$ by Theorem 1.

(iii) implies (i): Suppose $\Pi$ consists of two mutually singular probability measures $\pi'$ and $\pi''$, i.e., there is some event $E$ with $\pi'(E) = 1$ and $\pi''(E) = 0$. Given any $x, y \in \mathcal{U}$ let $a = u^{-1}(x)$ and $b = u^{-1}(y)$. Hence, $a, b \in \mathcal{C}$ and so $f \equiv aEb \in \mathcal{F}$. Then, $u(c_f(\pi')) = u(a) = x$ and $u(c_f(\pi'')) = u(b) = y$. Since, by definition, $\mu_\Pi$ has full support on $\Pi$, there is $\lambda \in (0,1)$ such that $\mu(\pi') = \lambda$ and $\mu(\pi'') = 1 - \lambda$. Thus, $\mu_{\Pi,f}(x) = \lambda$ and $\mu_{\Pi,f}(y) = 1 - \lambda$. By (iii) and the representation,

$$\phi(\lambda x + (1-\lambda) y) \geq \lambda \phi(x) + (1-\lambda)\phi(y). \tag{12}$$

So, there exists $\lambda \in (0,1)$ such that, given any $x, y \in \mathcal{U}$, Equation (12) holds. By Lemma 3, $\phi$ is concave. Finally, by Axiom 4, $\phi$ is independent of the choice of $\Pi$ above

(i) is equivalent to (ii): Follows from the fact that $\phi = v \circ u^{-1}$and thus $v = \phi \circ u$ up to a positive affine transformation. ■

**Proof of Proposition 2.** To prove (1.), apply Proposition 1 and its analogue for smooth ambiguity love and note that $\phi$ both concave and convex is equivalent to $\phi$ linear. Now turn to the proof of (2.). $\phi$ strictly concave on an open interval $J \subseteq \mathcal{U}$ implies $\phi\left(\int x d\mu_f\right) > \int \phi(x)\, d\mu_f$ for all $\mu_f$ with non-singleton $supp\left(\mu_f\right) \subseteq J$ by the strict

version of Jensen's inequality. Thus, $\phi\left(e\left(\mu_f\right)\right) > \int \phi(x)\, d\mu_f$, which in turn implies $\delta_{u^{-1}(e(\mu_\Pi))} \succ f$ for all $f \in \mathcal{F}$ with non-singleton $supp\left(\mu_f\right) \subseteq J$ by Theorem 1. The reverse direction follows directly from the argument in the proof of Proposition 1 that smooth ambiguity aversion implies concavity of $\phi$ with the weak inequalities replaced with strict and attention limited to the restriction of $\phi$ to $J$. Part (3.) follows exactly as (2.) with concavity replaced by convexity and inequalities reversed. ■

**Proof of Lemma 2.** Suppose $\phi : \mathcal{U} \to \mathbb{R}$ is twice continuously differentiable and is not linear. There exists $x_0 \in \mathcal{U}$ such that $\phi''(x_0) \neq 0$. For, suppose *per contra* that $\phi''(x) = 0$ for all $x \in \mathcal{U}$. Then $\phi'(x) = k \in \mathbb{R}$ for all $x \in \mathcal{U}$. Hence, $\phi(x) = kx + c$ for some $k, c \in \mathbb{R}$, a contradiction. We conclude that there is $x_0 \in \mathcal{U}$ such that $\phi''(x_0) \neq 0$. Since $\phi''$ is continuous, there exists an interval $(\alpha, \beta) \subseteq \mathcal{U}$, with $x_0 \in [\alpha, \beta]$, such that $\phi''(x)\, \phi''(x_0) > 0$ for all $x \in (\alpha, \beta)$, which implies the desired conclusion. ■

## A.3 Proposition 3

Suppose (3) holds. Let $E$ be such that $xEy \succ xBy$. By Theorem 1, $V(xBy) = \phi(u(x)\beta + u(y)(1-\beta))$, where $\beta = \pi(B)$ for all $\pi \in \Pi$. Since $\phi(u(y)) \leq V(xEy) \leq \phi(u(x))$, by the continuity of $\phi$ there is $\beta^* \geq \beta$ such that

$$\phi(u(x)\beta^* + u(y)(1-\beta^*)) = V(xEy). \tag{13}$$

Since $\lambda$ is non-atomic, there is $\Omega \times \mathcal{B} \ni B^* \supseteq B$ such that $\pi(B^*) = \beta^*$ for all $\pi \in \Pi$. Hence, by (13) and by Theorem 1, $xEy \sim xB^*y$. By (3), this implies that $yEx \sim yB^*x$. As $\phi$ is strictly increasing, $\phi(u(x)(1-\beta^*) + u(y)\beta^*) < \phi(u(x)(1-\beta) + u(y)\beta)$, and so, by Theorem 1, $yB^*x \prec yBx$. Hence, $yEx \prec yBx$, and we conclude that

$$xEy \succ xBy \Longrightarrow yEx \prec yBx.$$

A similar argument proves the converse implication, and so

$$xEy \succ xBy \Longleftrightarrow yEx \prec yBx.$$

Finally, again a similar argument shows that

$$xEy \prec xBy \Longleftrightarrow yEx \succ yBx,$$

as desired. This completes the proof as the "only if" part is trivial. ■

## A.4 Theorem 2

Let $I$ be an index set for $\Pi$, i.e., $\Pi = \{\pi_i : i \in I\}$. By assumption, $\succeq$ satisfies the conditions in Theorem 1 and so the representation there applies. Fix an event $E$. Suppose that $E$ is ambiguous. This means that there exists an event $B \in \Omega \times \mathcal{B}$ and $x, y \in X$ with $\delta_x \succ \delta_y$ such that either $[xEy \succ xBy$ and $yEx \succeq yBx]$ or $[xEy \prec xBy$ and $yEx \preceq yBx]$ or $[xEy \sim xBy$ and $yEx \nsim yBx]$. Let $\beta$ denote $\pi_i(B) (= \pi_j(B)$ for all $j \in I)$. If $\pi_i(E)$ were equal to some fixed $\alpha \in [0,1]$ for $\mu$-almost-all $i$, then, by the representation, for all $w, z \in X$,

$$wEz \succeq wBz \Longleftrightarrow \alpha u(w) + (1-\alpha)\, u(z) \geq \beta u(w) + (1-\beta)\, u(z).$$

However, this makes it impossible for $E$ to be ambiguous. Therefore $\pi_i(E)$ must vary with $i$. Specifically, if $\gamma = \int_I \pi_i(E)\, d\mu$, then there exist $\mu$-non-null sets $I' \subseteq I$ and $I'' \subseteq I$ such that $\pi_i(E) < \gamma$ for $i \in I'$ and $\pi_i(E) > \gamma$ for $i \in I''$ and the first claim in the theorem is proved.

Next, suppose that $\succeq$ are not smoothly ambiguity neutral, Axioms 4 and 5 hold and $E$ is unambiguous. Proposition 2 implies that $\phi$ is strictly concave (or strictly convex) on a non-empty open interval $(u_1, u_2) \subseteq \mathcal{U}$. Fix $k, l \in \mathcal{U}$ such that $u_1 < k < l < u_2$. Let $\gamma = \int_I \pi_i(E)\, d\mu$. One can think of $\gamma$ as the decision maker's "expected" probability of the event $E$. According to our representation of preferences the following is true:

$$\begin{aligned}
V\left(u^{-1}(l)\,\Omega \times [0,\gamma)\, u^{-1}(k)\right) &= \phi(\gamma l + (1-\gamma)\, k), \\
V\left(u^{-1}(l)\, E u^{-1}(k)\right) &= \int_I \phi\left(\pi_i(E) l + (1-\pi_i(E))\, k\right) d\mu, \\
V\left(u^{-1}(k)\,\Omega \times [0,\gamma)\, u^{-1}(l)\right) &= \phi(\gamma k + (1-\gamma)\, l), \\
V\left(u^{-1}(k)\, E u^{-1}(l)\right) &= \int_I \phi(\pi_i(E) k + (1-\pi_i(E))\, l) d\mu.
\end{aligned}$$

Since $\phi$ is strictly concave (the strictly convex case follows similarly) on the interval $[k, l]$, Jensen's inequality (and the definition of $\gamma$) implies that

$$V\left(u^{-1}(l)\, E u^{-1}(k)\right) \leq V\left(u^{-1}(l)\,\Omega \times [0,\gamma)\, u^{-1}(k)\right)$$

and

$$V\left(u^{-1}(k)\, E u^{-1}(l)\right) \leq V\left(u^{-1}(k)\,\Omega \times [0,\gamma)\, u^{-1}(l)\right)$$

with both inequalities strict if it is not the case that $\pi_i(E)$ takes on the same value everywhere (specifically, $\pi_i(E) = \gamma$ for $\mu$-almost-all $i$). Suppose that both inequalities are indeed strict. This says that

$$u^{-1}(l)\, E u^{-1}(k) \prec u^{-1}(l)\,\Omega \times [0,\gamma)\, u^{-1}(k)$$

and

$$u^{-1}(k)\, E u^{-1}(l) \prec u^{-1}(k)\,\Omega \times [0,\gamma)\, u^{-1}(l)$$

implying that $E$ is ambiguous, a contradiction. Therefore it must be that $\pi_i(E) = \gamma$ for $\mu$-almost-all $i$ and the second claim in the theorem is proved. ■

## A.5 Corollary 2

Let $I$ be an index set for $\Pi$, i.e., $\Pi = \{\pi_i : i \in I\}$. Let $\succeq$ satisfy Axioms 1 through 4. First consider the case where $\succeq$ also satisfies Axiom 5(i). In this case, by Corollary 1, preferences are SEU and hence $\Lambda$ coincides with $\Sigma$, a $\sigma$-algebra and therefore, a $\lambda$-system. Next consider the case where either Axiom 5(ii) or Axiom 5(iii) holds and let $E, F \in \Sigma$ *s.t.* $E \cap F = \varnothing$. By Theorem 2, $E$ unambiguous implies there exists a $\gamma \in [0,1]$ such that $\pi_i(E) = \gamma$ for $\mu$ almost all $i$. Also, $\pi_i(E^c) = 1 - \gamma$ for $\mu$ almost all $i$. Similarly, there exists a $\gamma' \in [0,1]$ such that $\pi_i(F) = \gamma'$ and $\pi_i(F^c) = 1 - \gamma'$and for $\mu$ almost all $i$. Since $\pi_i$ is a probability measure and $E \cap F = \varnothing$, it follows that $\pi_i(E \cup F) = \gamma + \gamma'$ for $\mu$ almost all $i$. Hence, by Theorem 2, $E \cup F$ is unambiguous. Finally, note that it follows directly from Definition 5 that $S$ is unambiguous and that if $E$ is unambiguous then so is $E^c$. ■

## A.6 Corollary 3

Let $I$ be an index set for $\Pi$, i.e., $\Pi = \{\pi_i : i \in I\}$. "If part": Fix an event $E$. Suppose that $\succeq_q$ is a qualitative probability relation for $\succeq$ on a $\lambda$-system $\mathcal{A} \supseteq \mathcal{E} = \{E, E^c, \Omega \times \mathcal{B}\}$. By properties (i), (v), and (vi) in the definition of a qualitative probability relation, $E$ is unambiguous.

"Only if part": Let $\Lambda$ denote the set of all unambiguous events in $\Sigma$. Suppose that $E$, and therefore all elements of $\mathcal{E}$, is unambiguous. By Corollary 2, $\Lambda$ is a $\lambda$-system. By our hypothesis, $\mathcal{E} \subseteq \Lambda$. Let $\succeq$ satisfy Axioms 1 through 4. First, suppose $\succeq$ also satisfies either Axiom 5(ii) or Axiom 5(iii). By Theorem 2 for every $A \in \Lambda$ there exists a $\gamma(A) \in [0,1]$ such that $\pi_i(A) = \gamma(A)$ for $\mu$ almost all $i$. So, a probability $\gamma$ representing the likelihood relation on $\lambda$-system $\Lambda \supseteq \mathcal{E}$ exists and therefore a qualitative probability relation for $\succeq$ exists on the same $\lambda$-system. Next, suppose $\succeq$ satisfies Axiom 5(i) (in addition to Axioms 1 through 4). In this case, by Corollary 1, preferences are SEU and hence a probability $\gamma$ representing the likelihood relation on $\lambda$-system $\Lambda \supseteq \mathcal{E}$ exists and therefore a qualitative probability relation for $\succeq$ exists on the same $\lambda$-system. Therefore $E$ unambiguous implies the existence of a qualitative probability relation for $\succeq$ on a $\lambda$-system containing $\mathcal{E}$. ■

## A.7 Corollary 4

By the representation $(iv) \Longrightarrow (iii)$. The implications $(iii) \Longrightarrow (ii) \Longrightarrow (i)$ are obvious. By Theorem 2, $(i) \Longrightarrow (iv)$ when $\succeq$ satisfies Axiom 5(ii) or (iii). To see this, observe that if $\succeq_q$ is a qualitative probability relation for $\succeq$ on $\Sigma$, then by Corollary 3 all events $E \in \Sigma$ are unambiguous. By Theorem 2, provided $\succeq$ satisfies Axiom 5(ii) or (iii), this implies agreement about each event's probability. As to the case where $\succeq$ satisfies Axiom 5(i), recall that $\succeq$ has an SEU representation if $\succeq$ is smoothly ambiguity neutral. ■

## A.8 Theorem 3

The "if" part follows easily from the Jensen inequality. As to the "only if", set $h(x) = \left(\phi_A \circ \phi_B^{-1}\right)(x)$ for all $x \in \mathcal{U}$. The function $h$ is clearly strictly increasing. Moreover, since $\left(\phi_A^{-1} \circ \phi_A\right)(x) = x = \left(\phi_B^{-1} \circ \phi_B\right)(x)$ for all $x \in \mathcal{U}$, we have $\phi_A = h \circ \phi_B$. We want to show that $h$ is concave if and only if $A$ is more ambiguity averse than $B$.

By Definition 8, $\int \phi_A d\mu_f \geq \phi_A(u(x))$ implies $\int \phi_B d\mu_f \geq \phi_B(u(x))$ for all $f \in \mathcal{F}$ and $x \in \mathcal{C}$. Since $\mathcal{U}$ is an interval, given any $f \in \mathcal{F}$ there exists $x_f \in \mathcal{C}$ such that $\int \phi_A d\mu_f = \phi_A(u(x_f))$. Hence, $f \sim_1 \delta_{x_f}$ and so, by (4), $\int \phi_B d\mu_f \geq \phi_B(u(x_f))$. In turn this implies that, for all $f \in \mathcal{F}$,

$$\phi_B^{-1}\left(\int \phi_B d\mu_f\right) \geq \phi_A^{-1}\left(\int \phi_A d\mu_f\right)$$

and so

$$h\left(\int \phi_B d\mu_f\right) \geq \int \phi_A d\mu_f = \int (h \circ \phi_B)\, d\mu_f. \tag{14}$$

Let $\phi_B(x), \phi_B(y) \in \phi_B(\mathcal{U})$. By proceeding as in the proof of Proposition 1, there is a set $\Pi$, an act $f$ and a $\lambda \in (0,1)$ such that $\mu_{\Pi,f}(x) = \lambda$ and $\mu_{\Pi,f}(y) = 1 - \lambda$. Hence, Eq.

(14) reduces to

$$h\left(\lambda\phi\left(x\right)+\left(1-\lambda\right)\phi\left(y\right)\right)\geq\lambda h\left(\phi\left(x\right)\right)+\left(1-\lambda\right)h\left(\phi\left(y\right)\right).$$

Since $\phi_B\left(\mathcal{U}\right)$ is an interval, by Lemma 3 we conclude that $h$ is concave. ■

## A.9 Corollary 6

Construct a family of expected utility preferences $\{\succeq_\Pi^{eu}\}_{\Pi\subseteq[0,1]}$ as follows: Fix $u$ and $\mu_\Pi$ so that they match those for $\succeq_\Pi$ and take $\phi^{eu}$ to be the identity. Suppose $\phi$ is concave. Then $\phi$ is an increasing, concave transformation of $\phi^{eu}$. By Theorem 3, $\succeq$ is more ambiguity averse than $\succeq^{eu}$. In the other direction, suppose $\succeq$ is more ambiguity averse than some $\succeq^{eu}$. Then by Theorem 3, $\phi$ is an increasing, concave transformation of $\phi^{eu}$. Since $\phi^{eu}$ must be linear (as Axiom 4 implies $\phi$ is the same for each $\Pi\subseteq\Delta$), this implies $\phi$ is concave. ■

## A.10 Proposition 4

W.l.o.g., assume that $\mathcal{U}=[0,1]$. Let $k\in(0,1)$ and set $\mathcal{U}_k=[0,1-k]$. Let $\mathcal{C}_k\subseteq\mathcal{C}$ be such that $u\left(\mathcal{C}_k\right)=\mathcal{U}_k$ and consider

$$\mathcal{F}^k=\left\{f\in\mathcal{F}:f\left(s\right)\in\mathcal{C}_k\text{ for each }s\in S\right\}.$$

Define $\succeq_\Pi^k$ on $\mathcal{F}^k$ as follows: $f\succeq_\Pi^k g$ if and only if

$$\int\phi_k\left(\int u\left(f\left(s\right)\right)d\pi\right)d\mu_\Pi\geq\int\phi_k\left(\int u\left(g\left(s\right)\right)d\pi\right)d\mu_\Pi,$$

where $\phi_k\left(x\right)=\phi\left(x+k\right)$ for each $x\in\mathcal{U}_k$. We have:

$$\begin{aligned}
f\quad\succeq\quad{}^k_\Pi\delta_x &\Longleftrightarrow\int\phi_k\left(\int u\left(f\left(s\right)\right)d\pi\right)d\mu_\Pi\geq\phi_k\left(u\left(x\right)\right)\\
&\Longleftrightarrow\int\phi\left(\int\left(u\left(f\left(s\right)\right)+k\right)d\pi\right)d\mu_\Pi\geq\phi\left(u\left(x\right)+k\right)\\
&\Longleftrightarrow\int\phi\left(\int u\left(f'\left(s\right)\right)d\pi\right)d\mu_\Pi\geq\phi\left(u\left(x\right)+k\right)\\
&\Longleftrightarrow f'\succeq_\Pi\delta_{u^{-1}(u(x)+k)}\Longleftrightarrow f\succeq_\Pi\delta_x,
\end{aligned}$$

where the last equivalence follows from Definition 9. Hence, $\succeq^k$ is as ambiguity averse as $\succeq$, when restricted to $\mathcal{F}^k$. By Theorem 3, there exist $a\left(k\right)>0$ and $b\left(k\right)\in\mathbb{R}$ such that, for all $x\in[0,1-k]$,

$$\phi\left(x+k\right)=\phi_k\left(x\right)=a\left(k\right)\phi\left(x\right)+b\left(k\right).\tag{15}$$

Since $k$ was arbitrary, we conclude that the functional equation (15) holds for all $k\in(0,1)$ and all $x\in(0,1)$ such that $x+k\leq1$. This is a variation of Cauchy's functional equation (see p. 150 of Aczel (1966)), and its only strictly increasing solutions are (up to positive affine transformations) $\phi\left(x\right)=x$ or $\phi\left(x\right)=-\frac{1}{\alpha}e^{-\alpha x}$, $\alpha\neq0$ . ■

## A.11 Proposition 5

The proposition is a consequence of the following lemma, which we prove for the sake of completeness.

**Lemma 4** *Let $\Delta(X)$ be the set of all finitely supported probabilities defined on a consequence space $X$. Let $\lambda_n$ be the Arrow-Pratt coefficient of $u_n$. Let $\succsim$ be a vN-M ordering on $\Delta(X)$. There exists a sequence $\{\succsim_n\}_{n=1}^{\infty}$ of vN-M orderings on $\Delta(X)$, with $\succsim_1 = \succsim$, such that*

*i all $\succsim_n$ agree on $X$,*

*ii $\lim_n \lambda_n(x) = +\infty$ and $\lambda_n(x) \geq \lambda_{n-1}(x)$ for all $n \geq 1$ and all $x \in X$,*

*Given any $p$ and $q$ in $\Delta(X)$, if it holds, eventually, that $p \succsim_n q$, then*

$$\min_{x \in supp(p)} u(x) \geq \min_{x \in supp(q)} u(x), \tag{16}$$

*while*

$$\min_{x \in supp(p)} u(x) > \min_{x \in supp(q)} u(x) \tag{17}$$

*implies that, eventually, $p \succ_n q$.*

**Proof.** Let $\succsim_n$ be the vN-M ordering whose vN-M utility index is given by $u_1(x) = u(x)$, and $u_n(x) = -u(x)^{-n}$ for all $n > 1$. Since $u_n$ is obtained from $u$ through the increasing transformation $-x^{-n}$, point (i) is satisfied. On the other hand, $\lambda_n(x) = n\lambda(x)$, and so point (ii) as well is satisfied.

We begin by showing that

$$\lim_n \left( -\int -u(x)^{-n} dp \right)^{-\frac{1}{n}} = \min_{x \in supp(p)} u(x). \tag{18}$$

To prove it, first observe that, given any bounded $\mathcal{X}$-measurable function $\psi : X \to \mathbb{R}$, we have (see, e.g., Lemma 12.1 in Aliprantis and Border (1999)):

$$\lim_n \left( \int (\psi(x))^n dp \right)^{\frac{1}{n}} = \max_{x \in supp(p)} \psi(x). \tag{19}$$

As,

$$\lim_n \left( -\int -u(x)^{-n} dp \right)^{-\frac{1}{n}} = \left( \min_{x \in supp(p)} u(x) \right) \lim_n \left( \int \left( \frac{u(x)}{\min_{x \in supp(p)} u(x)} \right)^{-n} dp \right)^{-\frac{1}{n}},$$

and

$$\lim_n \left( \int \left( \frac{u(x)}{\min_{x \in supp(p)} u(x)} \right)^{-n} dp \right)^{-\frac{1}{n}} = \lim_n \frac{1}{\left( \int \left( \frac{\min_{x \in supp(p)} u(x)}{u(x)} \right)^n dp \right)^{\frac{1}{n}}} = 1,$$

we conclude that Eq. (18) holds.

On the other hand, since $-x^{-\frac{1}{n}}$ is an increasing transformation, it is easy to check that, given any $p$ and $q$ in $\Delta(X)$, we have $\int -u(x)^{-n}\,dp \geq \int -u(x)^{-n}\,dq$ if and only if

$$\left(-\int -u(x)^{-n}\,dp\right)^{-\frac{1}{n}} \geq \left(-\int -u(x)^{-n}\,dq\right)^{-\frac{1}{n}}. \tag{20}$$

Hence, $p \succsim_n q$ implies Eq. (20), and so if eventually it holds $p \succsim_n q$, we have

$$\lim_n \left(-\int -u(x)^{-n}\,dp\right)^{-\frac{1}{n}} \geq \lim_n \left(-\int -u(x)^{-n}\,dq\right)^{-\frac{1}{n}}.$$

By Eq. (18), we finally conclude that Eq. (16).

To complete the proof, assume that Eq. (17) holds. By Eq. (18), there exists $n_0$ large enough so that, for all $n \geq n_0$,

$$\begin{aligned}
\left|\left(-\int -u(x)^{-n}\,dp\right)^{-\frac{1}{n}} - \min_{x\in supp(p)} u(x)\right| &< \frac{\min_{x\in supp(p)} u(x) - \min_{x\in supp(q)} u(x)}{2}, \\
\left|\left(-\int -u(x)^{-n}\,dq\right)^{-\frac{1}{n}} - \min_{x\in supp(q)} u(x)\right| &< \frac{\min_{x\in supp(p)} u(x) - \min_{x\in supp(q)} u(x)}{2}.
\end{aligned}$$

Hence, for all $n \geq n_0$ it holds

$$\left(-\int -u(x)^{-n}\,dp\right)^{-\frac{1}{n}} > \left(-\int -u(x)^{-n}\,dq\right)^{-\frac{1}{n}},$$

which in turn implies $p \succ_n q$. ■

## A.12 Results of Section 5

For the following proofs we need a piece of notation. As in Section 5 we assumed $\Omega$ to be finite, we can use $[0,1]$ as an index set for $\Delta$, that is, $\Delta = \{\pi_i : i \in [0,1]\}$. In this case, $\mu$ can be viewed as defined on the Borel $\sigma$-algebra of $[0,1]$, with support $I$.

### A.12.1 Proposition 6

Suppose $\pi_i(E \times [0,1)) = 0$, for $\mu$ almost all $i$. By Theorem 1,

$$f = \begin{pmatrix} x \text{ if } s \in E \times [0,1) \\ y \text{ if } s \in E^c \times [0,1) \end{pmatrix} \sim \begin{pmatrix} z \text{ if } s \in E \times [0,1) \\ y \text{ if } s \in E^c \times [0,1) \end{pmatrix} = g,$$

iff

$$\begin{aligned}
& \int_I \phi\left[\sum_{\omega\in E}\int_{[0,1)} u(x)\,\pi_i(\{\omega\times[0,1)\})\,dr+\sum_{\omega\in E^c}\int_{[0,1)} u(y)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
= & \int_I \phi\left[\sum_{\omega\in E}\int_{[0,1)} u(z)\,\pi_i(\{\omega\times[0,1)\})\,dr+\sum_{\omega\in E^c}\int_{[0,1)} u(y)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
\Leftrightarrow & \int_I \phi\left[\sum_{\omega\in E^c}\int_{[0,1)} u(y)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
= & \int_I \phi\left[\sum_{\omega\in E^c}\int_{[0,1)} u(y)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu.
\end{aligned}$$

Therefore $E$ is null. To prove the other direction, suppose $E$ is null. To generate a contradiction assume $\pi_i(E\times[0,1))>0$ for $i\in H$ where $\mu(H)=c>0$. Fix $x$ and $z$ such that $\delta_x\succ\delta_z$. Set $y=0$. By Theorem 1,

$$f=\begin{pmatrix} x \text{ if } s\in E\times[0,1) \\ y \text{ if } s\in E^c\times[0,1)\end{pmatrix}\sim\begin{pmatrix} z \text{ if } s\in E\times[0,1) \\ y \text{ if } s\in E^c\times[0,1)\end{pmatrix}=g,$$

iff

$$\begin{aligned}
& \int_I \phi\left[\sum_{\omega\in E}\int_{[0,1)} u(x)\,\pi_i(\{\omega\times[0,1)\})\,dr+\sum_{\omega\in E^c}\int_{[0,1)} u(0)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
= & \int_I \phi\left[\sum_{\omega\in E}\int_{[0,1)} u(z)\,\pi_i(\{\omega\times[0,1)\})\,dr+\sum_{\omega\in E^c}\int_{[0,1)} u(0)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
\Leftrightarrow & \int_I \phi\left[\sum_{\omega\in E}\int_{\{r\in[0,1)|(\omega,r)\in E\}} u(x)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
= & \int_I \phi\left[\sum_{\omega\in E}\int_{\{r\in[0,1)|(\omega,r)\in E\}} u(z)\,\pi_i(\{\omega\times[0,1)\})\,dr\right]d\mu \\
\Leftrightarrow & \ u(x)=u(z),
\end{aligned}$$

a contradiction. ■

### A.12.2 Theorem 4

**(a)** Fix any event $E$ that is unambiguous according to Definition 5. Suppose,

$$\begin{pmatrix} x^* & if & s\in A \\ x & if & s\in B \\ h(s) & if & s\in C\equiv E^c\backslash(A\cup B) \\ z & if & s\in E\end{pmatrix}\succeq\begin{pmatrix} x & if & s\in A \\ x^* & if & s\in B \\ h(s) & if & s\in C\equiv E^c\backslash(A\cup B) \\ z & if & s\in E\end{pmatrix}.$$

Then according to our representation,

$$\begin{aligned}
& \int_I -e^{-\alpha\left[u(z)\pi_i(E)+u(x^*)\pi_i(A)+u(x)\pi_i(B)+\int_C u(h(s))d\pi_i\right]}d\mu \qquad (21) \\
\geq & \int_I -e^{-\alpha\left[u(z)\pi_i(E)+u(x)\pi_i(A)+u(x^*)\pi_i(B)+\int_C u(h(s))d\pi_i\right]}d\mu.
\end{aligned}$$

Since $E$ is unambiguous according to Definition 5 and $\phi$ is strictly concave, Theorem 2 implies $\pi_i(E)$ is almost everywhere constant in $i$ and hence the $u(z)$ term can be taken outside the integral on both sides and canceled. This can clearly be done, in exactly the same way, if $z$ is replaced by some $z'$ in the acts above. Therefore for any $z' \in \mathcal{C}$,

$$\begin{pmatrix} x^* & if & s \in A \\ x & if & s \in B \\ h(s) & if & s \in C \equiv E^c \backslash (A \cup B) \\ z' & if & s \in E \end{pmatrix} \succeq \begin{pmatrix} x & if & s \in A \\ x^* & if & s \in B \\ h(s) & if & s \in C \equiv E^c \backslash (A \cup B) \\ z' & if & s \in E \end{pmatrix}.$$

Since, under Definition 5, $E$ unambiguous implies that $E^c$ is unambiguous, a similar preference implication can be derived replacing $E$ by $E^c$. Thus $E$ is unambiguous according to the definition in Epstein and Zhang (2001).

**(b)**[8] Fix any $E \subseteq \Omega$ such that $E \times [0,1)$ is ambiguous according to definition 5. The strategy for showing that this event is ambiguous according to Epstein and Zhang (2001) after a suitable perturbation of the measures in $\Pi$ is as follows: We eventually select a particular pair of acts for which we will show that there is the type of conditional likelihood reversal required by Epstein and Zhang (2001). Using our representation, we are able to find equations (22, 23) that are necessary for no reversal to occur. It turns out that there are essentially only two ways these two equations can hold simultaneously given that $\pi_i(E \times [0,1))$ is not constant. One way is for the relative probabilities of the events in the complement of $E \times [0,1)$ to remain unchanged, as happened in Example 1. A second, related way is that although the relative probabilities of the events in the complement do change, they do so in a way which happens to exactly "cancel out" on level sets of $\pi_i(E \times [0,1))$. We construct the perturbation of $\Pi$ precisely to prevent these two circumstances from happening (and also to maintain the ambiguity/unambiguity of all events of the form $\omega \times [0,1)$ as required by part $b(ii)$ of the theorem). We then are able to conclude that the needed reversal in preference does occur and $E \times [0,1)$ is indeed ambiguous according to Epstein and Zhang (2001).

We begin by constructing the needed perturbation to prevent both the "cancelling out" on level sets problem (by making sure all such sets are singletons) and the constancy of relative probabilities problem (by direct perturbation). Since $E \times [0,1)$ is ambiguous and unambiguous events are closed under disjoint unions, there must exist an $\omega_1 \in E$ such that $\omega_1 \times [0,1)$ is ambiguous. Fix such an $\omega_1$. Choose $\omega_2$, $\omega_3 \in \Omega$ as follows. By Theorem 2, there exist $i, j \in I$ such that $\pi_i(E \times [0,1)) \neq \pi_j(E \times [0,1))$. Given such $i$ and $j$, since probabilities sum to 1, there must be an $\omega' \in \Omega \backslash E$ such that $\omega' \times [0,1)$ is ambiguous. Set $\omega_2$ equal to some such $\omega'$. Choose $\omega_3$ to be any element of $\Omega \backslash (E \cup \{\omega_2\})$ such that $\omega_3 \times [0,1)$ is non-null. Without loss of generality we may assume that such a $\omega_3$ exists.[9]

Given $\Pi$, construct $\Pi(\epsilon_n)$ as follows:

[8] We are grateful to Rakesh Vohra for his help in constructing this proof. (Full responsibility lies with the authors for any mistakes.)

[9] If no such $\omega_3$ exists, since there were assumed to be at least three $\omega \in \Omega$ such that $\omega \times [0,1)$ is non-null, it must be that at least two of these lie in $E$. If $E \times [0,1)$ above is replaced by $(\Omega \backslash E) \times [0,1)$, then $(\Omega \backslash E) \times [0,1)$ will be ambiguous according to definition 5 and $\omega_2, \omega_3 \in E$ with the desired properties will exist. The proof may then be carried through for $(\Omega \backslash E) \times [0,1)$. This will then imply the result for $E \times [0,1)$ since both definitions of ambiguous event are closed under complementation.

*Step 1*: Recall that $I \subseteq [0,1]$. Divide $I$ into sets of the form

$$I_\kappa \equiv \{i \in I \mid \pi_i (E \times [0,1)) = \kappa\}.$$

Let $J'$ denote the collection of sets $I_\kappa$ including exactly each such set that is not a singleton, has $\kappa \neq 0$ and $\kappa \neq 1$, and has $\pi_i (\omega_2 \times [0,1)) > 0$ for all $i \in I_\kappa$. Let $J = \{i \mid i \in I_\kappa \text{ for some } I_\kappa \in J'\}$. Let $H'$ denote the collection of sets $I_\kappa$ including exactly those that are not in $J'$, and have $\kappa \neq 0$. Let $H = \{i \mid i \in I_\kappa \text{ for some } I_\kappa \in H'\}$. Choose $\epsilon_n > 0$ no larger than the minimum of {half the minimum distance between the $\kappa$ values associated with the $I_\kappa$ sets, half of $\min_{i \in J} (1 - \pi_i (E \times [0,1)))$, half of $\min_{i \in J} \pi_i (\omega_2 \times [0,1))$, half of $\min_{i \in H} (1 - \pi_i (\omega_2 \times [0,1)))$, half of $\min_{i \in H} \pi_i (E \times [0,1))$, half of

$$\max_{i,j \in I} |\pi_i (\omega_2 \times [0,1)) - \pi_j (\omega_2 \times [0,1))|,$$

and half of

$$\min_{i,j \in I, \omega \in E \text{ s.t. } \pi_i(\omega \times [0,1)) \neq \pi_j(\omega \times [0,1))} |\pi_i (\omega \times [0,1)) - \pi_j (\omega \times [0,1))|\}.$$

If there are $\omega \in E$ such that $\omega \times [0,1)$ is unambiguous, then transfer a total mass of $\epsilon_n$ to $\omega_1$ taken evenly from these unambiguous states for each $i$.

For each $i \in J$, let

$$\pi_i^1 (\epsilon_n) (\omega_1 \times [0,1)) = \pi_i (\omega_1 \times [0,1)) + i\epsilon_n$$

and

$$\pi_i^1 (\epsilon_n) (\omega_2 \times [0,1)) = \pi_i (\omega_2 \times [0,1)) - i\epsilon_n.$$

Note that the way we have chosen $\epsilon_n$ and the fact that $I \subseteq [0,1]$ ensures that $E \times [0,1)$, $\omega_1 \times [0,1)$ and $\omega_2 \times [0,1)$ remain ambiguous according to definition 5 and that no new "bunchings" of $\pi_i (E \times [0,1))$ have been created.

*Step 2*: For each $i \in H$, let (if there was mass transferred to $\omega_1$ from the unambiguous states in $E$ in the previous step)

$$\pi_i^1 (\epsilon_n) (\omega_1 \times [0,1)) = \pi_i (\omega_1 \times [0,1)) - i\epsilon_n,$$

otherwise

subtract the $i\epsilon_n$ from the ambiguous states in $E$ in any way that doesn't push any below 0

and

$$\pi_i^1 (\epsilon_n) (\omega_2 \times [0,1)) = \pi_i (\omega_2 \times [0,1)) + i\epsilon_n.$$

Note that the way we have chosen $\epsilon_n$ and the fact that $I \subseteq [0,1]$ ensures that $E \times [0,1)$, $\omega_1 \times [0,1)$ and $\omega_2 \times [0,1)$ and any other ambiguous $\omega \in E$ remain ambiguous according to definition 5 and that no new "bunchings" of $\pi_i (E \times [0,1))$ have been created.

*Step 3:* For all points and for all $i \in I$ at which $\pi_i^1 (\epsilon_n)$ has yet to be defined, let

$$\pi_i^1 (\epsilon_n) (.) = \pi_i (.).$$

The end result of Steps 1 through 3 is a set of probabilities $\Pi^1(\epsilon_n) \equiv \{\pi_i^1(\epsilon_n)\}_{i\in I}$ that has the same set of ambiguous events of the form $\omega \times [0,1)$ as $\Pi$ but, except (possibly) where $\pi_i^1(\epsilon_n)(E \times [0,1)) = 0$, $\pi_i^1(\epsilon_n)(E \times [0,1))$ takes on distinct values for each $i \in I$ and $\pi_i^1(\epsilon_n)(\omega_2 \times [0,1)) > 0$ for all $i \in I$.

*Step 4*: This step should only be undertaken if the ratio $\frac{\pi_i^1(\epsilon_n)(\omega_3\times[0,1))}{\pi_i^1(\epsilon_n)(\omega_2\times[0,1))}$ is constant across all $i \in I$ such that $\pi_i^1(\epsilon_n)(E \times [0,1)) > 0$. Notice that since $\omega_2 \times [0,1)$ is ambiguous, this can only occur if $\omega_3 \times [0,1)$ is ambiguous as well. Select an $m \in I$ such that

$$\pi_m^1(\epsilon_n)(E \times [0,1)) > 0.$$

Set

$$\pi_m^2(\epsilon_n)(\omega_3 \times [0,1)) = \pi_m^1(\epsilon_n)(\omega_3 \times [0,1)) + \frac{m}{2}\epsilon_n$$

and

$$\pi_m^2(\epsilon_n)(\omega_2 \times [0,1)) = \pi_m^1(\epsilon_n)(\omega_2 \times [0,1)) - \frac{m}{2}\epsilon_n.$$

*Step 5*: For all points and for all $i \in I$ at which $\pi_i^2(\epsilon_n)$ has yet to be defined, let

$$\pi_i^2(\epsilon_n)(.) = \pi_i^1(.).$$

Then define $\Pi(\epsilon_n) = \Pi^2(\epsilon_n)$. Notice, because of Step 4, that the ratio $\frac{\pi_i(\epsilon_n)(\omega_3\times[0,1))}{\pi_i(\epsilon_n)(\omega_2\times[0,1))}$ cannot be constant across all $i \in I$ such that $\pi_i(\epsilon_n)(E \times [0,1)) > 0$. This fact will be used in the proof below.

Above, $\epsilon_n$ was allowed to be real number strictly between zero and an upper bound that was the minimum of several terms. To form a sequence, select any sequence of numbers $\{\epsilon_n\}_{n=1}^{\infty}$ in this range such that the $\lim_{n\to\infty}\epsilon_n = 0$. Observe, that given our construction, $\lim_{n\to\infty}\Pi(\epsilon_n) = \Pi$ (specifically, $\lim_{n\to\infty}\pi_i(\epsilon_n)(E) = \pi_i(E)$ for all $i \in I$ and all $E \in \Sigma$). Now we proceed to show that, given any such $\Pi(\epsilon_n)$ ,the event $E \times [0,1)$ is ambiguous according to the definition in Epstein and Zhang (2001). Consider the pair of acts depicted below (where $E, \omega_2$ and $\omega_3$ are as in the construction of $\Pi(\epsilon_n)$ and $C \equiv (E \times [0,1))^c \setminus ((\omega_3 \times [0,t)) \cup (\omega_2 \times [t,1))))$.

$$\begin{pmatrix} 1 & if & s \in \omega_3 \times [0,t) \\ 0 & if & s \in \omega_2 \times [t,1) \\ 0 & if & s \in C \\ 0 & if & s \in E \times [0,1) \end{pmatrix}, \begin{pmatrix} 0 & if & s \in \omega_3 \times [0,t) \\ 1 & if & s \in \omega_2 \times [t,1) \\ 0 & if & s \in C \\ 0 & if & s \in E \times [0,1) \end{pmatrix}.$$

Given our representation of preferences, the decision maker is indifferent between these two acts iff

$$\sum_{i=1}^{k} -e^{-\alpha t \pi_i(\epsilon_n)(\omega_3\times[0,1))}\mu(i) = \sum_{i=1}^{k} -e^{-\alpha(1-t)\pi_i(\epsilon_n)(\omega_2\times[0,1))}\mu(i). \tag{22}$$

Let $t^*$ be the unique $t$ that solves equation (22). Such a $t^*$exists since the left-hand side is continuously increasing in $t$ and is equal to $-1$ at $t = 0$ while the right-hand side is

continuously decreasing in $t$ and is equal to $-1$ at $t=1$. If $E\times[0,1)$ were unambiguous according to Epstein and Zhang (2001), it would be the case that

$$\begin{pmatrix} 1 & if & s\in\omega_3\times[0,t^*) \\ 0 & if & s\in\omega_2\times[t^*,1) \\ 0 & if & s\in C \\ u^{-1}(c) & if & s\in E\times[0,1) \end{pmatrix} \sim \begin{pmatrix} 0 & if & s\in\omega_3\times[0,t^*) \\ 1 & if & s\in\omega_2\times[t^*,1) \\ 0 & if & s\in C \\ u^{-1}(c) & if & s\in E\times[0,1) \end{pmatrix}$$

for all $c\in\mathcal{U}$. In terms of our representation, this means

$$\begin{aligned} & \sum_{i=1}^{k} -e^{-\alpha[t^*\pi_i(\epsilon_n)(\omega_3\times[0,1))+c\pi_i(\epsilon_n)(E\times[0,1))]}\mu(i) \\ = & \sum_{i=1}^{k} -e^{-\alpha[(1-t^*)\pi_i(\epsilon_n)(\omega_2\times[0,1))+c\pi_i(\epsilon_n)(E\times[0,1))]}\mu(i) \text{ for all } c\in\mathcal{U}. \end{aligned} \tag{23}$$

Taking the derivative of both sides of equation (23) with respect to $c$ and evaluating at $c=0$,

$$\begin{aligned} & \sum_{i=1}^{k} -e^{-\alpha t^*\pi_i(\epsilon_n)(\omega_3\times[0,1))}\mu(i)\,\pi_i(\epsilon_n)(E\times[0,1)) \\ = & \sum_{i=1}^{k} -e^{-\alpha(1-t^*)\pi_i(\epsilon_n)(\omega_2\times[0,1))}\mu(i)\,\pi_i(\epsilon_n)(E\times[0,1)). \end{aligned} \tag{24}$$

In fact, since equation (23) is an identity in $c$, we can differentiate both sides as many times as we wish while maintaining equality. Calculating this out, we see

$$\begin{aligned} & \sum_{i=1}^{k} -e^{-\alpha t^*\pi_i(\epsilon_n)(\omega_3\times[0,1))}\mu(i)\,[\pi_i(\epsilon_n)(E\times[0,1))]^m \\ = & \sum_{i=1}^{k} -e^{-\alpha(1-t^*)\pi_i(\epsilon_n)(\omega_2\times[0,1))}\mu(i)\,[\pi_i(\epsilon_n)(E\times[0,1))]^m \end{aligned} \tag{25}$$

for all $m=0,1,2,\ldots$ . Since by construction of $\{\pi_i(\epsilon_n)\}_{i=1}^{k}$, either $\pi_i(\epsilon_n)(E\times[0,1))=0$, in which case those terms in the system of equations (25) get no weight, or, for the remaining $i$'s we can strictly order the $\pi_i(\epsilon_n)$'s by the weight they give to $E\times[0,1)$. Let $i^*$ be the index of the $\pi_i(\epsilon_n)$ that gives the largest such weight. Divide both sides of the each equation in the system of equations above by $[\pi_{i^*}(\epsilon_n)(E\times[0,1))]^m$. We can then rewrite the system as, for all $m=0,1,2,\ldots$,

$$\begin{aligned} & -e^{-\alpha t^*\pi_{i^*}(\epsilon_n)(\omega_3\times[0,1))}\mu(i^*) \\ & +\sum_{i\neq i^*}\frac{-e^{-\alpha t^*\pi_i(\epsilon_n)(\omega_3\times[0,1))}\mu(i)\,[\pi_i(\epsilon_n)(E\times[0,1))]^m}{[\pi_{i^*}(\epsilon_n)(E\times[0,1))]^m} \\ = & -e^{-\alpha(1-t^*)\pi_{i^*}(\epsilon_n)(\omega_2\times[0,1))}\mu(i^*) \\ & +\sum_{i\neq i^*}\frac{-e^{-\alpha(1-t^*)\pi_i(\epsilon_n)(\omega_2\times[0,1))}\mu(i)\,[\pi_i(\epsilon_n)(E\times[0,1))]^m}{[\pi_{i^*}(\epsilon_n)(E\times[0,1))]^m}. \end{aligned} \tag{26}$$

A necessary condition for the above system of equations to hold is that

$$-e^{-\alpha t^* \pi_{i^*}(\epsilon_n)(\omega_3 \times [0,1))} \mu(i^*) = -e^{-\alpha(1-t^*)\pi_{i^*}(\epsilon_n)(\omega_2 \times [0,1))} \mu(i^*). \tag{27}$$

To see this, observe that for any $\delta > 0$ there exists an $M$ such that for all $m > M$,

$$\begin{aligned} \delta > \quad & \sum_{i \neq i^*} \frac{-e^{-\alpha t^* \pi_i(\epsilon_n)(\omega_3 \times [0,1))} \mu(i) \left[\pi_i(\epsilon_n)(E \times [0,1))\right]^m}{\pi^n_{i^*}(\epsilon_n)(E \times [0,1))} \\ & - \sum_{i \neq i^*} \frac{-e^{-\alpha(1-t^*)\pi_i(\epsilon_n)(\omega_2 \times [0,1))} \mu(i) \left[\pi^n_i(\epsilon_n)(E \times [0,1))\right]^m}{\pi^n_{i^*}(\epsilon_n)(E \times [0,1))} > -\delta. \end{aligned}$$

This is true because $0 \leq \pi_i(\epsilon_n)(E \times [0,1)) < \pi_{i^*}(\epsilon_n)(\omega_1 \times [0,1)) \leq 1$ and all the other terms are bounded.

Given equation (27), we can cancel the $i^*$ terms from both sides of the equations in the system (25). This gives a new system of equations. For this new system find the $i$ such that $\pi_i(\epsilon_n)(E \times [0,1))$ gives the largest weight and repeat the above steps to show that

$$-e^{-\alpha t^* \pi_i(\epsilon_n)(\omega_3 \times [0,1))} \mu(i) = -e^{-\alpha(1-t^*)\pi_i(\epsilon_n)(\omega_2 \times [0,1))} \mu(i),$$

for that $i$. Canceling and repeating $k-2$ more times (or until the largest remaining $\pi_i(\epsilon_n)(E \times [0,1)) = 0$), we find

$$-e^{-\alpha t^* \pi_i(\epsilon_n)(\omega_3 \times [0,1))} = -e^{-\alpha(1-t^*)\pi_i(\epsilon_n)(\omega_2 \times [0,1))},$$

for all $i \in I$ such that $\pi_i(\epsilon_n)(\omega_1 \times [0,1)) > 0$. This is only possible if

$$\frac{\pi_i(\omega_3 \times [0,1))}{\pi_i(\omega_2 \times [0,1))} = \frac{1-t^*}{t^*}$$

for all $i \in I$ such that $\pi_i(\epsilon_n)(\omega_1 \times [0,1)) > 0$. As noticed above in the construction of $\Pi(\epsilon_n)$, this cannot be true. Therefore we have a contradiction and it cannot be that $E \times [0,1)$ unambiguous according to the definition in Epstein and Zhang (2001). ■

### A.12.3 Theorem 5

We first report a result from the working paper version of Ghirardato and Marinacci (2002), where it is stated as Theorem 4.

**Theorem 7** *Let $\succeq$ be an ambiguity averse monotonic preference relation. For all $\geqslant_1$, $\geqslant_2$ belonging to the set of benchmark preferences corresponding to $\succeq$, we have*

$$\mathcal{H}^{gm}_{\geqslant_1} = \mathcal{H}^{gm}_{\geqslant_2} \equiv \mathcal{H}^{gm} \text{ and } \Lambda^{gm}_{\geqslant_1} = \Lambda^{gm}_{\geqslant_2} \equiv \Lambda^{gm}.$$

We can now prove Theorem 5.

**Proof of Theorem 5.** Let $\geqslant$, the benchmark, be the special case of $\succeq$ where $\phi$ is linear. That is, $\geqslant$ is represented by $V_{\geqslant}(f) = \sum_\omega \left[\int_{[0,1)} u(f(\omega, r))\, dr\right] \nu(\omega)$, where $\nu(\omega) = \int_I \pi_i(\{\omega \times [0,1)\})\, d\mu$ for all $\omega \in \Omega$.

Suppose $\phi$ is linear (i.e., Axiom 5(i) holds). Then $\mathcal{H}^{gm} = \mathcal{F}$. On the other hand, by Corollaries 3 and 4, $\phi$ linear means that all events are unambiguous. Hence, $\mathcal{H} = \mathcal{F}$ as well. For the rest of the proof we suppose $\phi$ is strictly concave or strictly convex over some open interval $J \subseteq \mathcal{U}$ (i.e., Axiom 5(ii) or (iii) holds)

First we prove that $\mathcal{H} \subseteq \mathcal{H}^{gm}$. By Theorem 2, $\pi_i = \pi_j \equiv \bar{\pi}$ on $\Lambda$. Hence, we have, for $f \in \mathcal{H}$,

$$V(f) = \phi\left(\sum_{\omega}\left[\int_{[0,1)} u\left(f\left(\omega, r\right)\right) dr\right] \bar{\pi}\left(\omega\right)\right). \tag{28}$$

Since $\phi$ is strictly increasing, the way $\succeq$ ranks acts belonging to $\mathcal{H}$ is equivalently given by the functional,

$$\sum_{\omega}\left[\int_{[0,1)} u\left(f\left(\omega, r\right)\right) dr\right] \bar{\pi}\left(\omega\right).$$

Notice, $V_{\geqslant}(f)$ as well reduces to the functional (28). Hence, $\succeq$ and $\geqslant$ agree on $\mathcal{H}$. Since $\mathcal{H}$ contains all the constant acts this proves that part (A) of Definition 12 holds for $\mathcal{H}$. On the other hand, it is immediate to see that part (B) holds as well. Thus, since $\mathcal{H}_{\geqslant}$ is the largest subset satisfying (A) and (B), it follows that $\mathcal{H} \subseteq \mathcal{H}^{gm}_{\geqslant}$. Furthermore, by Theorem 7, $\mathcal{H}^{gm}_{\geqslant} = \mathcal{H}^{gm}$.

As to the converse inclusion, we consider the case of $\phi$ strictly concave on $J$ (the strictly convex case being similar). We take an act $f \notin \mathcal{H}$ and go on to show that $f \notin \mathcal{H}^{gm}_{\geqslant}$ either. It is enough to consider acts $f$ such that $supp\left(\mu_f\right) \subseteq J$. To see why this is the case, assume $f$ is such that $supp\left(\mu_f\right) \nsubseteq J$. Let $\{x_1, ..., x_n\}$ be the range of the finite-valued act $f$. Since $u$ is strictly increasing on the interval $\mathcal{C}$, it is differentiable on $\mathcal{C}$, except on at most a countable subset $M$ of $\mathcal{C}$. The function $u$ is therefore locally Lipschitz on $\mathcal{C} - M$. Since $J$ is an open interval and $u$ is strictly increasing and continuous, $u^{-1}\left(J\right)$ is an open interval. Hence, $u^{-1}\left(J\right) \cap \left(\mathcal{C} - M\right) \neq \emptyset$, and so there exists $c \in u^{-1}\left(J\right)$ at which $u$ is locally Lipschitz. Let $\left(c - \varepsilon, c + \varepsilon\right)$ be a neighborhood of $c$ over which $u$ is locally Lipschitz. Since $J$ is an open interval, by taking $\varepsilon$ small enough, we can assume that $\left[u\left(c\right) - \varepsilon, u\left(c\right) + \varepsilon\right] \subseteq u^{-1}\left(J\right)$. It is easy to check that there exist $\alpha \neq 0$ and $\beta \in \mathbb{R}$ such that $\left|\alpha x_i + \beta - c\right| \leq \varepsilon$ for each $i = 1, ...n$. Hence, by the local Lipschitz property, $\left|u\left(\alpha x_i + \beta\right) - u\left(c\right)\right| \leq \varepsilon$, and so

$$\left|\int u\left(\alpha f + \beta\right) d\pi - u\left(c\right)\right| \leq \varepsilon$$

for all probabilities $\pi \in \Delta$. Hence,

$$\begin{aligned} supp\left(\mu_{\alpha f + \beta}\right) &\subseteq \left[\inf_{\pi \in \Delta} \int u\left(\alpha f + \beta\right) d\pi, \sup_{\pi \in \Delta} \int u\left(\alpha f + \beta\right) d\pi\right] \\ &\subseteq \left[u\left(c\right) - \varepsilon, u\left(c\right) + \varepsilon\right] \subseteq J. \end{aligned}$$

Set $g = \alpha f + \beta$. As $g\left(s\right) = g\left(s'\right)$ if and only if $f\left(s\right) = f\left(s'\right)$ for all $s, s' \in S$, we have $\{f^{-1}(x) : x \in \mathcal{C}\} = \{g^{-1}(x) : x \in \mathcal{C}\}$. Hence, by (B) of Definition 12, $f \in \mathcal{H}^{gm}_{\geqslant}$ if and only if $g \in \mathcal{H}^{gm}_{\geqslant}$ and by Definition 6, $f \in \mathcal{H}$ if and only if $g \in \mathcal{H}$. All this shows that in what follows we can assume that $supp\left(\mu_f\right) \subseteq J$.

As in Definition 4, take the constant act $\delta_{u^{-1}(e(\mu_f))}$. As $f \notin \mathcal{H}$, either $\delta_{u^{-1}(e(\mu_f))} \succ f$ or $\delta_{u^{-1}(e(\mu_f))} \sim f$. Suppose, first, the preference is strict. Since for the benchmark preference $\geqslant$, $\delta_{u^{-1}(e(\mu_f))}$ and $f$ are indifferent, part (A) of Definition 12 is violated for $f$.

Next, suppose $\delta_{u^{-1}(e(\mu_f))} \sim f$. By the strict concavity of $\phi$ (on $J \supseteq supp(\mu_f)$) this implies that $\mu_f$ has a singleton support. For, if $\mu_f$ did not have a singleton support then by strict concavity of $\phi$,

$$V(f) = \int \phi(x)\, d\mu_f < \phi\left(\int x d\mu_f\right) = \phi(e(\mu_f)) = V\left(\delta_{u^{-1}(e(\mu_f))}\right), \tag{29}$$

and so $\delta_{u^{-1}(e(\mu_f))} \succ f$. Hence it remains to consider the case $f \notin \mathcal{H}$ with $\mu_f$ having singleton support. We show, next, that in this case the condition (B) of Definition 12 is violated for $f$.

Since $f \notin \mathcal{H}$ there exists an $x \in \mathcal{C}$ such that the event $f^{-1}(x)$ is ambiguous. Call that event $A(x)$. Define $g : S \longrightarrow \mathcal{C}$ as follows,

$$\begin{aligned} g(s) &= f(s) \text{ if } s \notin A(x), \\ g(s) &= x' \text{ if } s \in A(x), \end{aligned}$$

with $x' \notin \{f(s) : s \in S\}$ (recall that all our acts are finite valued). Since $A(x)$ is ambiguous there exist $\mu$-non-null sets $I' \subseteq I$ and $I'' \subseteq I$ such that for any $i \in I'$ and $j \in I''$, $\pi_i(A(x)) \neq \pi_j(A(x))$. Since, $\mu_f$ has singleton support, for $\mu$-almost-all $i, j$, $E_{\pi_i}(u \circ f) = E_{\pi_j}(u \circ f)$. Hence, for $\mu$-almost-all $i \in I'$ and $j \in I''$, $E_{\pi_i}(u \circ f) = E_{\pi_j}(u \circ f)$. For all such $i, j$ pairs,

$$\begin{aligned} E_{\pi_i} u(g) - E_{\pi_j} u(g) &= E_{\pi_i} u(f) - \pi_i(A(x))[u(x) - u(x')] \\ &\quad - E_{\pi_j} u(f) + \pi_j(A(x))[u(x) - u(x')] \\ &= [u(x) - u(x')][\pi_j(A(x)) - \pi_i(A(x))] \\ &\neq 0. \end{aligned}$$

Therefore, $\mu_g$ does not have a singleton support since it is not true that $E_{\pi_i}(u \circ f) = E_{\pi_j}(u \circ f)$ for $\mu$-almost-all $i, j$. By strict concavity of $\phi$ (see Eq. 29) we then conclude that $\delta_{u^{-1}(e(\mu_g))} \succ g$. This violates part (B) of the Definition 12.

Hence, summing up, we may conclude $f \notin \mathcal{H}^{gm}_{\geqslant}$, and so, again by Theorem 7, $f \notin \mathcal{H}^{gm}$. This completes the proof that $\mathcal{H} = \mathcal{H}^{gm}$.

Since all acts are finite-valued, the $\lambda$-system $\Lambda$ can be viewed as the set of all pre-image sets of the acts in $\mathcal{H}$. As $\Lambda^{gm}$ is defined as the collection of all pre-image sets of acts in $\mathcal{H}^{gm}$, it follows that $\Lambda = \Lambda^{gm}$. ■

# References


Aczel, J. (1966): *Lectures on Functional Equations and their Applications.* Academic Press, New York and London.

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

Billingsley, P. (1986): *Probability and Measure*. John Wiley, New York, second edn.

Chen, Z., and L. G. Epstein (2002): "Ambiguity, Risk and Asset Returns in Continuous Time," *Econometrica*, 70(4), 1403–1443.

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

Epstein, L., and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62(3), 283–322.

Epstein, L., and J. Zhang (2001): "Subjective probabilities on subjectively unambiguous events," *Econometrica*, 69, 265–306.

Epstein, L. G., and J. Miao (2001): "A Two-Person Dynamic Equilibrium under Ambiguity," mimeo., University of Rochester.

Ergin, H., and F. Gul (2002): "A Subjective Theory of Compound Lotteries," Discussion paper, Princeton University.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2002): "Ambiguity from the Differential Viewpoint," Discussion paper, ICER.

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

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

Grant, S., A. Kajii, and B. Polak (2000): "Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models," *Econometrica*, 68(2), 425–434.

Halevy, Y., and V. Feltkamp (2001): "A Bayesian Approach to Uncertainty Aversion," Discussion paper, University of British Columbia.

Hansen, L. P., T. J. Sargent, and T. D. Tallarini, Jr. (1999): "Robust Permanent Income and Pricing," *The Review of Economic Studies*, 66.

Hansen, L. P., T. J. Sargent, G. Turmuhambetova, and N. Williams (2001): "Robustness and Uncertainty Aversion," Discussion paper, University of Chicago.

Hardy, G., J. E. Littlewood, and G. Polya (1952): *Inequalities*. Cambridge University Press, Cambridge and New York.

Jaffray, J.-Y. (1995): "Dynamic Decision Making With Belief Functions," in *Advances in the Dempster-Shafer Theory of Evidence*, ed. by e. a. Yeager.

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

Kreps, D., and E. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

Loomes, G., and U. Segal (1994): "Observing Orders of Risk Aversion," *Journal of Risk and Uncertainty*, 9, 239–256.

Machina, M., and D. Schmeidler (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780.

Nau, R. (2001): "Uncertainty Aversion with Second-Order Probabilities and Utilities," in *Second International Symposium on Imprecise Probabilities and Their Application*, Ithaca, N.Y.

——— (2002): "Uncertainty Aversion with Second-Order Probabilities and Utilities," mimeo., Duke University.

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

Sarin, R., and P. Wakker (1992): "A simple axiomatization of nonadditive expected utility," *Econometrica*, 60(6), 1255–1272.

——— (1997): "A Single-Stage Approach to Anscombe and Aumann's Expected Utility," *Review of Economic Studies*, 64(3), 399–409.

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.

——— (1990): "Two-Stage Lotteries Without the Reduction Axiom," *Econometrica*, 58, 349–77.

Segal, U., and A. Spivak (1990): "First Order versus Second Order Risk Aversion," *Journal of Economic Theory*, 51, 111–125.

Takacs, L. (1978): "An Increasing Continuous Singular Function," *American Mathematical Monthly*, 85, 35–37.

Uppal, R., and T. Wang (2002): "Model Misspecification and Under-Diversification," Discussion paper, University of British Columbia.

Zhang, J. (1999): "Qualitative Probabilities on $\lambda$-Systems," *Mathematical Social Sciences*, 38, 11–20.

——— (2002): "Subjective Ambiguity, Expected Utility and Choquet Expected Utility," *Economic Theory*, 20, 159–181.