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.marinacci@unito.it

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

September 2002
This Version: April 2004

[1]We thank N. Al-Najjar, E. Dekel, D. Fudenberg, S. Grant, Y. Halevy, B. Lipman, F. Maccheroni, S. Morris, R. Nau, K. Nehring, W. Neilson, K. Roberts, U. Segal, A. Shneyerov, J-M. Tallon, R. Vohra, P. Wakker and especially L. Epstein, I. Gilboa, and M. Machina for helpful discussions and suggestions. We also thank three referees and A. Postlewaite, the co-editor, for offering very useful comments and advice. We also thank a number of seminar and conference audiences. We thank MEDS at Northwestern University, ICER at the University of Torino, and Eurequa at the Universiy of Paris 1 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, ϕ is an increasing transformation, and μ is a subjective probability over the set Π of probability measures π 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 ϕ. 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 Π. 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 it 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 a limiting 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. 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. He calculates the expected utility from f, $\mathbb{E}u\,(f) = u(10) \times 1/3$, but is uncertain about $\mathbb{E}u\,(g)$ beyond knowing that it lies 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 $[u(10) \times 1/3, u(10)]$ depending on how he assigns probability between B and C. 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 assignment is. Hence, his choices. 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 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 above. Among other things, the model could be used to analyze behavior in instances wherein the DM's information is *explicitly* consistent with multiple probabilities on the state space relevant to the decision at hand. One instance is a portfolio investment decision. An investor, in the best circumstances, with access to all publicly available data, will in general be left with a range of return distributions that are plausible. As a second example, think of a monetary policy maker setting policy on the basis of a parametric model which solves to yield a probability distribution on a set of macroeconomic variables of interest. But the probability distribution on variables is conditional on the value of the parameters which, in turn, is uncertain. That might cause the DM to be concerned enough to seek a policy whose performance is more robust to the uncertainty as to which probability applies. Indeed, such a concern is central to the recent literature investigating decision rules robust to model mis-specification or "model uncertainty" (e.g., Hansen and Sargent (2000)).

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 vN-M utility function; π is a probability measure on S; ϕ is a map from reals to reals. There may be subjective uncertainty about what the "right" probability on S is – μ is the DM's subjective prior over Δ, the possible probabilities over S, and therefore measures the subjective relevance of a particular π as the "right" probability. While u, as usual, characterizes attitude toward risk, we show that ambiguity attitude is captured by ϕ. In particular, a concave ϕ characterizes ambiguity aversion, which we define to be an aversion to mean preserving spreads in μ_f, where μ_f is the distribution over expected utility values induced by μ and f. The distribution μ_f represents the uncertainty about ex ante evaluation; it shows the probabilities of different evaluations of the act f. 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, thereby identifying ambiguity with uncertainty/multiplicity with respect to relevant priors and hence, ex ante evaluations. It is worth noting that this preference model does not, in general, impose reduction between μ and the π's in the support of μ. Such reduction only occurs when ϕ 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 idea of modeling ambiguity attitude by relaxing reduction between first and second order probabilities first appeared in Segal (1987) and inspires the analysis in this paper.

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 vN-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 μ and v: the former is a probability measure over possible priors on S revealing the DM's subjective information while the latter is the vN-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 identifies an act f, defined on S, with a second order act that yields for each prior π on S, the certainty equivalent of the lottery induced by f and π. Upon setting $\phi \equiv v \circ u^{-1}$, the three axioms lead to the representation given above. Notice, a concave ϕ implies that v is a concave transform of u. Hence, ambiguity aversion in this framework, defined to be aversion to mean preserving spreads in the induced distribution of expected utilities, turns out to be equivalent to the DM being more risk averse to the subjective uncertainty about priors than he is to the uncertainty in lotteries. Ambiguity neutrality obtains if the DM's attitudes to risk on the two domains of uncertainty are identical.

In the example concerning investment decisions, we may think of second order acts as bets on which return distribution is right. Preferences over second order acts yield a prior on the space of probabilities, which induces a distribution over expected utilities corresponding to each first order act. Ambiguity averse DMs have preference for (first order) acts whose evaluation is more robust to the possible variation in probabilities. In our model that is translated as an aversion to mean preserving spreads of the induced distribution. Since that is equivalent to v being more concave than u (i.e., ϕ is concave), it is as if we imagine the DM in the example to be thinking as follows. "My best guess of the chance that the return distribution is 'π' 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."

Apart from providing (what we think is) a clarifying perspective on ambiguity and ambiguity attitude, this functional representation will be particularly useful in economic modeling in answering comparative statics questions involving ambiguity. Take an economic model where agents' beliefs reflect some ambiguity. Next, without perturbing the information structure, suppose we wanted to ask how the equilibrium would change if the extent of ambiguity aversion were to decrease; e.g., if we were to replace ambiguity aversion with ambiguity neutrality. Another comparative statics exercise might want to hold ambiguity attitudes fixed and ask how the equilibrium is affected if the perceived ambiguity is varied. Working out such comparative statics properly requires a model which allows a conceptual/parametric separation of (possibly) ambiguous beliefs and ambiguity attitude, risk and risk attitude. The model and representation functional in the paper allows that, whereas such a separation is not evident in the pioneering and most popular decision making models that incorporate ambiguity, namely, the maxmin expected utility (MEU) preferences (Gilboa and Schmeidler (1989)) and the Choquet expected utility model of Schmeidler (1989).

To illustrate how the model/representation can facilitate comparative statics we include, in the final section of the paper, a (numerical) analysis of two simple portfolio choice problems. The analysis considers how the choice of an optimal portfolio is affected when risk attitude and, separately, ambiguity attitude parameters are varied. This allows a comparison of the effects of risk attitude with that of ambiguity attitude.

The rest of the paper is organized as follows. Section 2 states our basic assumptions on preferences and derives the representation. Section 3 defines ambiguity attitude and characterizes it in terms of the representation. Section 4 gives a behavioral definition of an ambiguous event and relates this definition to the representation. Section 5 discusses related literature. Finally, Section 6 presents the illustrative portfolio choice problems. 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 σ-algebra of a metric space Ω, and $\mathcal{B}_1$ the Borel σ-algebra of $[0,1]$. Consider the state space $S = \Omega \times [0,1]$, endowed with the product σ-algebra $\Sigma \equiv \mathcal{A} \otimes \mathcal{B}_1$. For the remainder of this paper, all events will be assumed to belong to Σ 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 $[-1,1]$. Given a preference $\succeq$ on the set of Savage acts, $\mathcal{F}$ denotes the set of all bounded Σ-measurable Savage acts; i.e., $f \in \mathcal{F}$ if $\{s \in S : f(s) \succeq x\} \in \Sigma$ for each $x \in \mathcal{C}$, and if there exist $x', x'' \in \mathcal{C}$ such that $x' \succeq f \succeq x''$.

The space $[0,1]$ is introduced simply to model a rich set of lotteries as a set of Savage acts. 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 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 $\omega \in \Omega$.[1]

Given the Lebesgue measure $\lambda : \mathcal{B}_1 \to [0,1]$, 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 $B \in \mathcal{B}_1$. The set of all such probabilities π is denoted by Δ. 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 Δ with the vague topology, that is, the coarsest topology on Δ that makes the following functionals continuous:

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

Throughout the paper we assume Δ to be endowed with the vague topology. Let $\sigma(\Delta)$ be the Borel σ-algebra on Δ generated by the vague topology. Lemma 5 in the Appendix, shows a property of $\sigma(\Delta)$ that we use to guarantee that the integrals in our main representation theorem are well defined.

Since we wish to allow Δ to be another domain of uncertainty for the DM apart from S, we model it explicitly as such. Does the DM regard this domain as uncertain and if so, what are the DM's beliefs? To formally identify this, we look at preferences over *second order acts* which assign consequences to elements of Δ.

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

Let $\succeq^2$ be the DM's preference ordering over $\mathfrak{F}$. The main focus of the model is $\succeq$, a preference relation defined on $\mathcal{F}$ (the set of acts on S). It might be helpful at this point to relate our structure to a more standard Savage-like one. To map our setting to one

[1] 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 and 1997) , and to Anscombe-Aumann style models. 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.

close to Savage, consider a product state space $S \times \Delta$. In a Savage-type theory with this state space the objects of choice, Savage acts, would be all (appropriately measurable) functions from $S \times \Delta$ to an outcome space $\mathcal{C}$. The theory would then take as primitive preferences over Savage acts. In contrast, our theory concerns preferences over only two subsets of Savage acts – those acts that depend either only on S or only on Δ. We do not consider any acts that depend on both, nor do we explicitly consider preferences between these two subsets.

While formally our second order acts may be considered to be a subset of Savage acts, there is a question whether preferences with respect to these acts are observable. The mapping from observable events to events in Δ may not always be evident. When it is not evident we may need something richer than behavioral data, perhaps cognitive data or thought experiments, to help us reveal the DM's beliefs over Δ.

However we would like to suggest that second order acts are not as strange or unfamiliar as they might first appear. Consider any parametric setting, i.e., a finite dimensional parameter space Θ, such that $\Delta = \{\pi_\theta\}_{\theta \in \Theta}$. Second order acts would simply be bets on the value of the parameter. In a parametric portfolio investment example, these could be bets about the parameter values that characterize the asset returns, e.g., means, variances and covariances. Similarly, in model uncertainty applications, second order acts are bets about the values of the relevant parameters in the underlying model. Closer to decision theory, for an Ellsberg urn, second order acts may be viewed as bets on the composition of the urn.

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 an expected utility representation.

Axiom 1 (Expected utility on lotteries) *There exists a unique $u : \mathcal{C} \rightarrow \mathbb{R}$, 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 DM's attitude towards risk generated from the lottery part of the state space.[2] The next axiom is on $\succeq^2$, the preferences over second order acts. These preferences are assumed to have a subjective expected utility representation.

Axiom 2 (Subjective expected utility on 2nd order acts) *There exists a finitely additive probability $\mu : \sigma(\Delta) \rightarrow [0, 1]$, with some $J \in \sigma(\Delta)$ such that $0 < \mu(J) < 1$, and a continuous, strictly increasing $v : \mathcal{C} \rightarrow \mathbb{R}$, such that, for all $\mathfrak{f}, \mathfrak{g} \in \mathfrak{F}$,*

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

Moreover, μ is unique and v is unique up to positive affine transformations.

[2] An alternative approach to deriving risk attitude, suggested to us by Mark Machina, would be to assume appropriate smoothness of our preferences and apply Machina (2004) to identify risk attitudes with preferences over "almost-objective" acts. It may be verified, given our representation, that such preferences are entirely determined by $u(\cdot)$.

We denote by Π the support of μ, that is, the smallest closed (w.r.t. the vague topology) subset of Δ whose complement has measure zero; Π is the subset of Δ the DM *subjectively* considers relevant. Given any $E \subseteq \Pi$, we interpret $\mu(E)$ as the DM's *subjective* assessment of the likelihood that the relevant probability lies in E; hence, μ may be thought of as a "second order probability" over the first order probabilities π. Notice that Π may well be a finite subset of Δ. Finally, the utility function v represents the DM's attitude towards risk generated by payoffs contingent on events in Δ.

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. For example, Axiom 1 can derived as in Grandmont (1972), while Theorem V.6.1 of Wakker (1989) can be used to deliver Axiom 2. Since such developments are by now well known and easily adapted to our setting, we do not do so here.

An act f and a probability π induce a probability distribution π_f on consequences. To define this formally, denote by $\mathcal{B}_c$ the Borel σ-algebra of $\mathcal{C}$, and define $\pi_f : \mathcal{B}_c \to [0,1]$ by $\pi_f(B) = \pi\left(f^{-1}(B)\right)$ for all $B \in \mathcal{B}_c$. The next lemma shows that each distribution π_f can be "replicated" by a suitable lottery act.

Lemma 1 *Given any $f \in \mathcal{F}$ and any $\pi \in \Delta$, there exists a (non-decreasing) lottery act $l_f(\pi) \in \mathcal{L}$ having the same distribution as π_f, i.e., such that $\lambda(l_f(\pi) \in B) = \pi_f(B)$ for all $B \in \mathcal{B}_c$.*

Notation 1 In what follows, δ_x denotes the constant act with consequence $x \in \mathcal{C}$, and $c_f(\pi)$ denotes the certainty equivalent of the lottery act $l_f(\pi)$; i.e., $\delta_{c_f(\pi)} \sim l_f(\pi)$.

Notice that since u is continuous and strictly increasing (Axiom 1), lottery acts have a unique certainty equivalent.

Our final basic axiom requires the preference ordering of primary interest, $\succeq$, to be consistent with Axioms 1 and 2 in a certain way. Since f together with a possible probability π generates a distribution over consequences identical to that generated by $l_f(\pi)$ it is reasonable to assume (for consistency with Axiom 1) that the certainty equivalent of f, given π, be same as the certainty equivalent of $l_f(\pi)$. Thus the certainty equivalent of f depends on which π is the right probability law. It is as if the DM faced a second order act, f^2, yielding $c_f(\pi)$ for each particular π. The axiom says that the DM takes this view and orders act $f \in \mathcal{F}$ identically to second order acts $f^2 \in \mathfrak{F}$.

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 $f^2, g^2 \in \mathfrak{F}$,*

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

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 π in the support of μ; next, compute the expectation (with respect to the measure μ) of the expected utilities obtained in the first stage, each expected utility transformed by the increasing function ϕ.

As will be shown in subsequent analysis, this representation allows a clear decomposition of the DM's tastes and beliefs: u determines risk attitude, ϕ determines ambiguity attitude, and μ determines the subjective belief, including any ambiguity perceived therein by the DM.

Notation 2 *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*

$$V(f) = \int_{\Delta} \phi \left[\int_S u(f(s))\, d\pi \right] d\mu \equiv \mathbb{E}_{\mu} \phi \left(\mathbb{E}_{\pi} u \circ f \right). \tag{1}$$

Given u, the function ϕ 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 $y \in \mathcal{U}$.

P roof. 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,

$$f \succeq g \Longleftrightarrow \int v(c_f(\pi))\, d\mu \geq \int v(c_g(\pi))\, d\mu. \tag{2}$$

Since v and u are strictly increasing, $v(c_f(\pi)) = \phi(u(c_f(\pi)))$ for some strictly increasing ϕ. Since v and u are continuous, so is ϕ. Substituting for $v(c_f(\pi))$ in (2), we get

$$f \succeq g \Longleftrightarrow \int \phi(u(c_f(\pi)))\, d\mu \geq \int \phi(u(c_g(\pi)))\, d\mu. \tag{3}$$

Now, recall,

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

So,

$$u(c_f(\pi)) = \sum_{x \in supp(\pi_f)} u(x) \pi_f(x) = \int_S u(f(s))\, d\pi. \tag{4}$$

Thus, substituting (4) into (3),

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

This proves the representation claim in the Theorem. To see the uniqueness properties of ϕ, 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, $\forall y \in \mathcal{U}$, $\tilde{\phi}(\alpha y + \beta) = \left(v \circ \tilde{u}^{-1}\right)(\alpha y + \beta) = \left(v \circ u^{-1}\right)(y) = \phi(y)$. Finally, v is unique up to positive affine transformations according to Axiom 2, so, fixing u, ϕ is as well. ■

The integrals in (1) are well defined because of Lemma 5 in the Appendix, 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 Appendix A.2 as Theorem 4) for the two orderings $\succeq$ and $\succeq^2$ in which Axioms 1, 2 and 3 are both necessary and sufficient. Theorem 4 also notes explicitly an important point evident in the proof of Theorem 1, that ϕ equals $v \circ u^{-1}$. The functional representation 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 ϕ' is simply the class of ϕ with domain shifted by the positive affine transformation.

We close this subsection by observing that, though in Axiom 1 we assumed expected utility preferences on lotteries, we could relax that assumption by allowing the preferences over lotteries to be Rank Dependent Expected Utility preferences (see Quiggin (1993)) with a suitable probability distortion $\varphi : [0,1] \to [0,1]$. Then the representation of the preferences over acts in $\mathcal{F}$ are as given in the following corollary.

Corollary 1 *Suppose there exists a continuous and non-decreasing function $\varphi : [0,1] \to [0,1]$ such that, for all $f, g \in \mathcal{L}$, $f \succeq g$ if and only if $\int_{[0,1]} u(f(r))d\varphi(\lambda) \geq \int_{[0,1]} u(g(r))d\varphi(\lambda)$. If $\succeq$ satisfies Axioms 2 and 3, then Eq. (1) of Theorem 1 becomes*

$$V(f) = \int_{\Delta} \phi \left[\int_S u(f(s))\, d\varphi(\pi) \right] d\mu. \tag{5}$$

P roof. It is enough to observe that here (4) becomes $u(c_f(\pi)) = \int_S u(f(s))\, d\varphi(\pi)$. The rest of the proof is identical to that of Theorem 1. ■

Note, the inner integral in (5) is a Choquet integral and that the inner and outer integrals are well defined because of the second part of Lemma 5.

3 Ambiguity Attitude

In this section we first provide a definition of a DM's ambiguity attitude and show that this ambiguity attitude is characterized by properties of ϕ, one of the functions from our representation above. Comparison of ambiguity attitudes across preference relations is dealt with in Section 3.2.

3.1 Characterizing ambiguity attitude

To discuss ambiguity attitude, we first require an additional assumption. In the classical theory, it is commonly implicitly or explicitly assumed or derived 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 weaker assumption suffices for our purposes: the assumption that the two risk attitudes u and v do not vary with Π, the support of an individual's belief μ. Recall that supports are, by definition, closed subsets of Δ.

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 DM. There is a pair of preference relations corresponding to each possible support Π, that is, to each possible state of information he may have about which probabilities π (over S) are relevant to his decision problem. Axioms 4 and 5, which follow, require certain properties of preferences to hold across the different pairs $\{\succeq_\Pi, \succeq^2_\Pi\}_{\Pi \subseteq \Delta}$. We emphasize that these axioms are of a somewhat different nature than the three basic axioms of the previous section. While the basic axioms operate only *within* pairs of preferences $(\succeq_\Pi, \succeq^2_\Pi)$, Axioms 4 and 5 operate across the entire family of pairs of preferences.

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

(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 Π *holds for the risk preferences derived from* $\succeq^2_\Pi$.

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

Notice that there is no restriction on the DM's belief associated with each $\succeq_\Pi$, besides that of having support Π. 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 π induces a distribution π_f on consequences. Each such distribution is naturally associated with a lottery $l_f(\pi) \in \mathcal{L}$, which has a certainty equivalent $c_f(\pi)$. Fixing an act f, the probability μ may then be used to induce a measure μ_f on $\{u(c_f(\pi)) : \pi \in \Pi\}$, the set of expected utility values generated by f corresponding to the different π's in Π (using the utility function u from Axiom 1). When necessary, we denote the belief associated with $\succeq_\Pi$ by μ_Π and the corresponding μ_f by $\mu_{\Pi,f}$. To introduce

μ_f formally, we need the following lemma. Here $\mathcal{B}_u$ denotes the Borel σ-algebra of $\mathcal{U}$, and $u(B) = \{u(x) : x \in B\}$.[3]

Lemma 2 $\mathcal{B}_u = \{u(B) : B \in \mathcal{B}_c\}$.

By Lemma 2, μ_f is defined on $\mathcal{B}_u$.

Definition 3 *Given $f \in \mathcal{F}$, the* ***induced distribution*** $\mu_f : \mathcal{B}_u \to [0,1]$ *is given by:*

$$\mu_f(u(B)) \equiv \mu\left(\left(f^2\right)^{-1}(B)\right) \qquad \text{for each } B \in \mathcal{B}_c.$$

Given an act f, the derived (subjective) probability distribution over expected utilities, μ_f, smoothly aggregates the information the DM has about the relevant π's and how each such π evaluates f, without imposing reduction between μ and the π's. In this framework the induced distribution μ_f represents the DM's subjective uncertainty about the "right" (ex ante) evaluation of an act. The greater the spread in μ_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 μ_f.[4] Just as in the theory of risk aversion, this may be expressed as a preference for getting a sure "average" to getting the act that induces μ_f. To state this formally, we need notation for the mean of μ_f, i.e., for the average expected utility from f.

Notation 3 Let $e\left(\mu_f\right) \equiv \int_{\mathcal{U}} x d\mu_f$. Notice $u^{-1}\left(e\left(\mu_f\right)\right) \in \mathcal{C}$.

Thus $\delta_{u^{-1}\left(e\left(\mu_f\right)\right)}$ is the constant act valued at the average utility of f.

Definition 4 *A DM displays* ***smooth ambiguity aversion at*** (f,Π) *if*

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

where μ has support Π. A DM displays ***smooth ambiguity aversion*** *if she displays smooth ambiguity aversion at (f,Π) for all $f \in \mathcal{F}$ and all closed subsets $\Pi \subseteq \Delta$.*[5]

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 ϕ. The proposition also shows that smooth ambiguity aversion is equivalent to the DM being more risk averse to the uncertainty about the right

[3] Since $\mathcal{U}$ is an interval, $\mathcal{B}_u$ coincides with the restriction on $\mathcal{U}$ of the Borel σ-algebra $\mathcal{B}$ of the real line. The same applies to $\mathcal{B}_1$ and $\mathcal{B}_c$, which are the restictions of $\mathcal{B}$ on $[0,1]$ and $\mathcal{C}$, respectively.

[4] It is important to keep in mind the distinction between μ and μ_f : while μ is a measure on probabilities and does not vary with f, μ_f is a measure on utilities and depends on f.

[5] 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 Π 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.

prior on S than he is to the risk involving lotteries (whose probabilities are objectively known). A result characterizing smooth ambiguity love by convexity of ϕ follows from the same argument. Similarly, smooth ambiguity neutrality is characterized by ϕ 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 DM displays smooth ambiguity aversion.

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

Corollary 2 *Under Axioms 1-4, the following properties are equivalent:*

(i) the DM is smoothly ambiguity neutral;

(ii) ϕ is linear;

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

We claim that this model allows a separation of risk attitude from ambiguity attitude. Proposition 1 shows that ambiguity attitude is represented by $\phi = \nu \circ u^{-1}$. Axiom 1 says that u represents risk attitude on lotteries. Thus we can vary risk attitude on lotteries while holding ambiguity attitude fixed by simultaneously changing ν to leave ϕ unchanged. Similarly, we can vary ambiguity attitude while leaving risk attitude on lotteries unchanged by holding u fixed while changing ν to achieve the desired change in ϕ. So the representation separates risk attitude *over lotteries in* $\mathcal{F}$ from the ambiguity attitude *towards acts in* $\mathcal{F}$. Further, given our three basic axioms, certainty equivalents of acts in $\mathcal{F}$ given a probability π over S are the same as those of the corresponding lotteries. In this sense u also represents risk attitude towards all acts in $\mathcal{F}$ and not just lotteries. (c.f., footnote 2 for an additional justification that u represents risk attitude towards all acts in $\mathcal{F}$.)

Remark 1 An ambiguity averse DM in this model prefers the lottery which pays x with an (objectively determined) probability p (and 0 with probability $1 - p$) to the (second order) act which pays x contingent on an event $E \subseteq \Delta$ to which the DM assigns a subjective prior $\mu(E) = p$. These two options expose the DM to the same uncertainty over payoffs generated in two different ways. Ambiguity aversion is the relative dislike of uncertainty generated by subjective beliefs over probabilities of outcomes compared to uncertainty generated by lotteries. This understanding about what shapes ambiguity attitude in this model also suggests that a DM would evaluate acts by obeying reduction (between μ and π) if π were chosen according to an objective μ. In this case second order acts are just objective lotteries to the DM. Hence v should be equivalent to u implying expected utility.

Our second comparative statics 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 $\mathcal{U}$ over which we require that the DM displays either strict ambiguity aversion, or strict ambiguity love, but not both. What is ruled out is the possibility that the DM'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 DM is ambiguity averse while over others he is ambiguity loving. The statement of the axiom is immediately followed by a proposition which gives an equivalent characterization in terms of ϕ.

Axiom 5 (Consistent ambiguity attitude over some interval) *The DM'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 at all (f, Π) for which supp $(\mu_{\Pi,f})$ 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, Π) for which supp $(\mu_{\Pi,f})$ is a non-singleton subset of K.

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

1. *Axiom 5 (i) holds if and only if ϕ linear;*

2. *Axiom 5 (ii) holds if and only if ϕ strictly concave on some open interval $J \subseteq \mathcal{U}$;*

3. *Axiom 5 (iii) holds if and only if ϕ strictly convex on some open interval $K \subseteq \mathcal{U}$.*

The following lemma and remark shows that if ϕ 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 3 *Suppose ϕ is twice continuously differentiable. If ϕ is not linear, then ϕ is either strictly concave or convex over some open interval.*

Remark 2 It follows immediately from Proposition 2 and Lemma 3 that under twice continuous differentiability of ϕ, Axioms 1–4 imply Axiom 5. Note, the conclusion of Lemma 3 may not hold if the hypothesis is weakened to simply ϕ continuous.

3.2 Comparison of ambiguity attitudes

In this section we study differences in ambiguity aversion across DMs. We identify each DM with an entire family of preferences $\{\succeq_\Pi, \succeq^2_\Pi\}_{\Pi\subseteq\Delta}$, parametrized by Π. Throughout the section (with the exception of Proposition 4), we assume Axiom 4 holds in addition to the first three axioms. Hence, ambiguity attitudes do not depend on the support Π of μ.

We begin with our definition of what makes one preference order more ambiguity averse than another. The idea behind it is that if two DMs share the same attitude to risk over acts in $\mathcal{F}$ and the same beliefs but one ranks non-constant acts lower than the other, then this must be due to a relatively greater aversion to ambiguity.

Definition 5 *Let A and B be two DMs whose families of preferences share the same vN-M utility function u and the same probability measures μ_Π for each support Π. We say that DM A* ***is more ambiguity averse than*** *B if*

$$f \succeq^A_\Pi \delta_x \Longrightarrow f \succeq^B_\Pi \delta_x \tag{6}$$

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 DMs are captured by the relative concavity of their functions ϕ,[6] thus showing that the concavity of ϕ plays here the role of the concavity of utility functions in standard risk theory.

Theorem 2 *Let A and B be two DMs whose families of preferences share the same vN-M utility function u and the same probability μ_Π for each support Π. Then, 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 2.

Corollary 3 *Suppose the hypotheses of Theorem 2 hold. If ϕ_A and ϕ_B are twice continuously differentiable, then DM 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

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

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

[6] Or, since u is being held fixed, the relative concavity of their functions v.

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

Remark 3 Corollary 4 connects our definition of smooth ambiguity aversion (Definition 4) to the comparative notion of ambiguity aversion in Definition 5. It shows that they agree, with expected utility taken as the dividing line between ambiguity aversion and ambiguity loving. It can be shown (see Klibanoff, Marinacci, and Mukerji (2003)) 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 6 *Suppose acts f, g, f', g' and $k \in \mathbb{R}$ are such that, for each $s \in S$,*

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

We say that the DM 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 DM views the "ambiguity content" in μ_f and its "translation" $\mu_f + k$ to be the same and so ranking them the same through preferences reveals ambiguity attitude unchanged by the shift in well being. Next we show that constant ambiguity attitudes are characterized by a negative exponential ϕ. It is of some interest to note that the proposition does not assume that ϕ is differentiable.

Proposition 3 *The DM 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 next proposition shows that when ambiguity aversion tends to infinity, our model essentially exhibits a maxmin expected utility behavior à la Gilboa and Schmeidler (1989), where Π is the given set of measures. This result is the analog in our setting of the well-known result in risk theory that extreme risk aversion leads to maxmin behavior.

Notation 4 Set $\operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f) = \sup\{t \in \mathbb{R} : \mu(\{\pi : \mathbb{E}_\pi u(f) < t\}) = 0\}$.

Proposition 4 *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 μ,

(ii) $\lim_n \alpha_n(x) = +\infty$ and $\alpha_n(x) \geq \alpha_{n-1}(x)$ 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,

$$\operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(f) \geq \operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(g),$$

while

$$\operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(f) > \operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(g)$$

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

To make the connection to MEU observe that when Π is finite, then $\operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(f) = \min_{\pi \in \Pi} \mathbb{E}_\pi u(f)$. This also holds under standard topological assumptions, as the next lemma shows.

Lemma 4 *If f is upper semicontinuous (i.e., all preference intervals $\{f \succeq x\}$ are closed), then $\operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(f) = \inf_{\Pi} \mathbb{E}_\pi u(f)$. If, in addition, Π is compact and f is continuous, then $\operatorname{ess}\inf_{\Pi} \mathbb{E}_\pi u(f) = \min_{\pi \in \Pi} \mathbb{E}_\pi u(f)$.*

4 Ambiguity

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. Finally, we briefly comment on the relationship with other notions of ambiguity.

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 $\Omega \times \mathcal{B}_1$. We say an event $E \in \Sigma$ is ambiguous if, analogous to the modal behavior observed in the Ellsberg experiment, betting on E is less desirable than betting on some event B in $\Omega \times \mathcal{B}_1$, and betting on E^c is *also* less desirable than betting on B^c. Similarly, we would also say E is ambiguous if both comparisons were reversed, or if one were indifference and the other were not.

Notation 5 *If $x, y \in \mathcal{C}$ and $A \in \Sigma$, xAy denotes the binary act which pays x if $s \in A$ and y otherwise.*

Definition 7 *An event $E \in \Sigma$ is* **unambiguous** *if, for each event $B \in \Omega \times \mathcal{B}_1$, and for each $x, y \in \mathcal{C}$ 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.*

The next proposition shows a shorter form of the definition that is equivalent to the original given our first three axioms. Though this form lacks immediate identification with the Ellsberg experiment, it helps in understanding 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 $\Omega \times \mathcal{B}_1$.

Proposition 5 *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 $\delta_x \succ \delta_y$,*

$$xEy \sim xBy \Longleftrightarrow yEx \sim yBx. \tag{7}$$

whenever $B \in \Omega \times \mathcal{B}_1$.

Our definition and the analogy with Ellsberg is most compelling when the events in $\Omega \times \mathcal{B}_1$ are themselves unambiguous. Given any particular preference relation, it may be checked using our definition whether this is so. Observe that if $\succeq$ satisfies Axiom 1 then all events in $\Omega \times \mathcal{B}_1$ are indeed unambiguous.[7]

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

Theorem 3 *Assume $\succeq$ satisfies the conditions in Theorem 1. If the event E is ambiguous according to Definition 7, then there exist μ-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 7, then, provided $\succeq$ satisfies Axioms 4 and 5 and is not smoothly ambiguity neutral, there exists a $\gamma \in [0,1]$ such that $\pi(E) = \gamma$, μ-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 Π of μ 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., ϕ linear. Recall that in this case, even if the measures in Π disagree on the probability of an event, the DM *behaves as if* he assigns that event its μ-average probability. Recall that Lemma 3 and Remark 2 showed that under conditions likely to be assumed in any application (twice continuous differentiability of the function ϕ 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.

Epstein and Zhang (2001) and Ghirardato and Marinacci (2002) have proposed behavioral notions of ambiguity meant to apply to a wide range of preferences. In the context of our model, how do their notions compare to the one presented above? It can be shown

[7] Note that the role of $\mathcal{B}_1$ 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 $\Omega \times \mathcal{B}_1$ with the desired alternative set.

that Ghirardato and Marinacci (2002) would identify the same set of ambiguous and unambiguous events as we do while Epstein and Zhang (2001) would yield a somewhat different classification. These results, a discussion of non-constant ambiguity attitude as a source of difference from Epstein and Zhang (2001), and further characterizations and discussion of our definition may be found in Klibanoff, Marinacci, and Mukerji (2003). A result relevant to this discussion also proved in that paper is that, given Axioms 1 through 4, the only departures from expected utility that may arise in this model are also departures from probabilistic sophistication.

5 Related literature

5.1 MEU and related models

Schmeidler (1989) was seminal in formalizing a decision theoretic model of ambiguity. It introduced the Choquet expected utility (CEU) model, which models uncertainty with nonadditive measures, with respect to which one takes the Choquet integral of the utility function. The MEU model of Gilboa and Schmeidler (1989) suggests that a DM entertains a set of priors, and computes the minimal expected utility for each act, where the prior ranges on this set. In general, the two models are distinct, but for a convex nonadditive measure (taking the set of priors to be the core of this measure) the two models give the same decision rule. The CEU and MEU models have been influential and have been applied in a variety of economic settings. Many applications of CEU use convex nonadditive measures, so they can be viewed as using either CEU or MEU. But observers have criticized the MEU/CEU model with the question, "Why evaluate acts by their minimal expected utility? Isn't this too extreme?" One could argue that it is not as extreme as it might first appear: the minimum is taken over a set of priors, but this need not be the set of priors that is literally deemed possible by the DM. However this argument undermines the attractive cognitive interpretation of the set of priors as the ambiguous information the DM has. For instance, take two DMs who share the same information, i.e., they both think a certain set of priors is possible. One is less cautious than the other, however. Suppose the first evaluates an action by the minimum expected utility over the literal set of priors while the other uses the expected utility at the 25^{th} percentile rather than the minimum. The MEU sets of priors representing the two DMs' preferences would be different and thus at least one must differ from the literal set of priors. In contrast to the CEU/MEU model, the present paper offers a model that allows for a set of priors that may be interpreted literally without necessarily implying the maxmin criterion.

Next we consider the relationship with a generalization of the maxmin functional to the α-maxmin EU model (α-MEU):

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

As in the MEU model, Π still might not be the literal set of priors, although there is more flexibility with α-MEU in capturing the ambiguity attitude (parameterized by α) of the DM. If one does interpret the Π literally, the model shares with MEU the limitation that it does not smoothly aggregate how the act performs under each possible π but only looks at the extremal performance values (the best and the worst).

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 Π, $\mathbb{E}_\pi (u \circ f) > \mathbb{E}_\pi (u \circ g)$. The α-maxmin rule must rank the acts equally, while our model would not.

Finally, we remark that it may be helpful to think of part of the difference between the model in this paper and models such as CEU, MEU and α-MEU as analogous to that between models of first and second order risk aversion (Segal and Spivak (1990), Loomes and Segal (1994)). Models such as MEU and α-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."

5.2 Models which relax reduction

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, 1990). The former 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 ambiguity aversion is not defined per se, Theorem 4.2 in that paper, which gives conditions (on the weighting function on the probabilities) under which a (binary) "nonambiguous lottery is preferred to an ambiguous one," appears to conceptualize aversion to ambiguity as an aversion to spreads in the second order probability. In our model the second order probability is μ. For general acts, aversion to spreads in μ and aversion to spreads in μ_f are distinct. Recall from Section 3 that we define ambiguity aversion as aversion to spreads in μ_f. 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.

Neilson (1993) uses lack of reduction to axiomatize a model of ambiguity attitude with a functional form identical to ours. This work, of which we were unaware while writing this paper, also contains the idea of using an Arrow-Pratt type index to measure ambiguity aversion. The axiomatic set-up differs from ours and the nature of ambiguity (as opposed to ambiguity attitude) is not explored. Another paper that relaxes reduction is Nau (2003) (a revised and expanded version of Nau (2001)). The paper presents an axiomatic model of partially separable preferences where the DM 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 utilities. Section 5 of that paper discusses, without axiomatization, a functional form like ours with separate first and second order probabilities as a special case of the state-dependent utility form. A major contribution of the paper is to present an intuitive notion of ambiguity aversion in a state dependent utility framework.

Ergin and Gul (2002) 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 this 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 a product state space. Their key axiom permits the DM to have different risk attitudes on different ordinates of the product space. A significant feature of Ergin and Gul's model is that it allows probabilistically sophisticated non-expected utility preference conditional on each ordinate. Unlike Nau, Ergin and Gul do not allow for state-dependence.

An important difference between our paper and Ergin and Gul is the domain over which preferences are defined. Ergin and Gul denote their product space $\Omega_a \times \Omega_b$. The objects of choice in their theory are the *full* set of Savage acts mapping $\Omega_a \times \Omega_b$ to an outcome space. How does this relate to our structure? First, observe that it is *not* the case that $\Omega_a \times \Omega_b$ corresponds to $S \times \Delta$; rather it corresponds to $[0, 1] \times \Delta$ in our model. We derive a recursive representation of preferences over acts on S that is completely determined by preferences over acts that depend only on Δ and acts that depend only on $[0, 1]$, while Ergin and Gul derive a recursive representation of preferences over acts on $\Omega_a \times \Omega_b$ that is completely determined by preferences over acts that depend only on Ω_b and acts that depend only on Ω_a. This difference in the domain of acts over which a recursive representation is derived has strong implications for the modeling of ambiguity. Specifically, if the domain is $\Omega_a \times \Omega_b$ as in Ergin and Gul, for any preferences either (1) preference is globally probabilistically sophisticated and *all* events are unambiguous; or (2) *all* non-null events that do not depend exclusively on either Ω_a or Ω_b alone are ambiguous (in the sense of our definition in Section 4). Thus, if ambiguity is present in their model, its scope is determined entirely by the exogenous structure of the state space. In contrast, in our model, the events in the Ω part of S may display a wide variety of patterns of ambiguity/unambiguity. The DM's preferences reveal which events are ambiguous and which are not, offering flexibility in modeling ambiguity and (partially) endogenizing its domain.

The seminal work of Kreps and Porteus (1978) is not concerned with ambiguity, or indeed with subjective probabilities, but is related to our modeling approach in that the representation we derive has a two-stage recursive form with expected utility at each stage. Halevy and Feltkamp (2001) try to rationalize ambiguity aversion by assuming that a DM 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. Chew and Sagi (2003) presents a model with endogenously defined "domains" within which the DM has the same risk attitude but across which they do not. Their approach involves domain-specific applications of the independence axiom leading to "domain recursive" preferences.

6 Portfolio Choice Examples

In this section we consider two examples of simple portfolio choice problems. The examples are intended both as a 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 Ω contains two elements, ω_1

and ω_2. The measure μ assigns probability 1/2 to both π_1 and π_2, which yield marginals on Ω of

$$\pi_1(\omega_1) = \frac{1}{4}, \pi_1(\omega_2) = \frac{3}{4} \text{ and } \pi_2(\omega_1) = \frac{3}{4}, \pi_2(\omega_2) = \frac{1}{4},$$

respectively. The function u is given by $u(x) = \begin{cases} \frac{1-e^{-ax}}{1-e^{-a}} & if \quad a > 0 \\ x & if \quad a = 0 \end{cases}$. This utility function displays constant absolute risk aversion with a as the coefficient of absolute risk aversion and is normalized so that $u(0) = 0$ and $u(1) = 1$.[8] The function ϕ is given by $\phi(x) = \begin{cases} \frac{1-e^{-\alpha x}}{1-e^{-\alpha}} & if \quad \alpha > 0 \\ x & if \quad \alpha = 0 \end{cases}$. This function may be said to display *constant ambiguity aversion* with this terminology justified by Proposition 3 in Section 3. α 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 and 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 α 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 (α) 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 (α) leads more to be invested in the safe asset. Table 2 gives a numerical illustration of this effect when risk aversion is fixed at $a = 2$.

[8]It is important that the normalization of u is maintained as risk aversion is varied. If it were not, then a corresponding shift in ϕ would be required to compensate for any change in normalization of u in order to maintain ambiguity aversion fixed. This follows from our representation theorem, but is worth pointing out again here. We thank Klaus Nehring for suggesting this.

ambiguity aversion (α)	amount allocated to safe asset
0	0.132699
0.02	0.133009
2	0.162948
20	0.368953
100	0.637172
200	0.678876

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) 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 (a), ambiguity aversion (α)) and beliefs (μ) 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 safe asset, a risky asset and an ambiguous asset**)

Here we consider the allocation problem where the risky (but unambiguous) asset underlying l is available in addition to the ambiguous and safe assets of the previous example. As risk aversion increases, holding ambiguity aversion fixed, the agent will want to diversify into both the safe asset and the ambiguous asset f (since it is not perfectly correlated with l), trading-off expected return against risk. In particular, the ratio of holdings of f to l increases. On the other hand, as ambiguity aversion increases, holding risk aversion fixed, the ambiguity about the payoff from f drives the agent away from it as f becomes a less effective diversifier and less valuable. Here the ratio of holdings of f to l decreases. Thus risk aversion and ambiguity aversion work in opposite directions in terms of the composition of the risky part of the agent's portfolio. Tables 3 and 4 give numerical illustrations of these effects.[9],[10]

Ambiguity aversion (α)	Amount allocated to l	Amount allocated to f
0	0.628076	0.867301
0.02	0.628076	0.867098
2	0.628076	0.847394
20	0.628076	0.697172
100	0.628076	0.41994
200	0.628076	0.338991

Table 3. Optimal amount out of $1 allocated to the risky (l) and ambiguous (f) assets as ambiguity aversion increases, holding risk aversion at $a = 2$.

Risk aversion (a)	Amount allocated to l	Amount allocated to f
0.02	62.8076	31.7994
1	1.25615	1.57176
2	0.628076	0.847394
20	0.0628076	0.0867301
100	0.0125615	0.017346

Table 4. Optimal amount out of $1 allocated to the risky (l) and ambiguous (f) assets as risk aversion increases holding ambiguity aversion at $\alpha = 2$.

In this case, if such behavior is examined ignoring ambiguity aversion, not only will the amount allocated to the safe asset seem to indicate higher risk aversion, as in the previous example, but an examination of the mix of risky assets (ratio of holdings of f to l) would indicate a lower level of risk aversion than the agent possesses. 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 (2003). Epstein and Miao (2003) generates home-bias in a heterogeneous agent dynamic multiple priors setting. See also Schroder and Skiadas (2003) for a related general framework.■

7 Conclusion

In conclusion, we summarize the main contributions of the paper:

[9] It is worth noting that the direction of these numerical comparative statics on the ratio of holdings of f to l continue to hold when the safe asset is elimated from this example, when f rather than l has the higher average payoff (switch the 3's and 2's), or when f and l yield the same average payoff (replace the 3's with 2's).

[10] The numbers in a row may add to more than 1 due to short sales of the safe asset.

1. It offers a model that allows for a set of priors to be present in a decision problem without necessarily implying the maxmin criterion. In doing so it generalizes MEU to a class of less extreme decision rules, and, at the same time, allows a three way separation of ambiguity, ambiguity attitudes and risk attitudes.

2. The paper also shows how familiar techniques from the literature on risk and risk attitude may be used to analyze ambiguity and ambiguity attitude.

3. The paper provides a simple behavioral definition of an ambiguous event. It shows that such events are identified in an easy and natural way within the model.

4. It offers a model that is smooth. Rather than the minimum operator, that generates kinks, here the model would "normally" allow for smooth operators that are much easier to use in economic applications.

A Appendix: Proofs and Related Material

A.1 Preliminaries

Denote by $\int \psi d\varphi(\pi)$ the standard Choquet integral, i.e.,

$$\int \psi d\varphi(\pi) = \int_0^{+\infty} \varphi(\pi(\psi \geq t))\, dt + \int_{-\infty}^{0} (1 - \varphi(\pi(\psi \geq t)))\, dt$$

w.r.t. the set function $\varphi(\pi) : \Sigma \to [0,1]$ induced by a continuous and non-decreasing function $\varphi : [0,1] \to \mathbb{R}$.

Lemma 5 $\sigma(\Delta)$ *coincides with the σ-algebra generated by the real-valued functions on Δ given by*

$$\pi \mapsto \int \psi d\pi \qquad \pi \in \Delta \text{ and } \psi \in B(\Sigma), \tag{8}$$

where $B(\Sigma)$ is the set of all bounded and real-valued Σ-measurable functions. In particular, given a continuous and non-decreasing function $\varphi : [0,1] \to \mathbb{R}$, the map

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

is $\sigma(\Delta)$-measurable.

P roof. The case $\varphi(x) = x$ is a routine exercise, and we omit the proof. Suppose first that $\varphi : [0,1] \to \mathbb{R}$ is monotone convex and continuous. Set $\varphi_0 = \varphi - \varphi(0)$, so that φ_0 is a monotone convex and continuous function with $\varphi_0(0) = 0$. For all $\psi \in B(\Sigma)$, set

$$\int \psi d\varphi(\pi) \equiv (\sup \psi - \inf \psi)\, \varphi(0) + \int \psi d\varphi_0(\pi). \tag{9}$$

We begin by showing that the map $\psi \to \int \psi d\varphi_0(\pi)$ is $\sigma(\Delta)$-measurable. Let $D = \{x_n\}_{n=1}^{\infty}$ be a countable dense subset of $(0,1)$. Since φ_0 is convex and continuous on $(0,1)$, its subdifferential $\partial \varphi_0(x_n) \subseteq \mathbb{R}$ is non-empty for each $n \geq 1$. Let $\alpha_n \in \partial \varphi_0(x_n)$

and set $\chi_n(x) = \varphi_0(x_n) + \alpha_n(x - x_n)$ for each $x \in [0,1]$. Since both φ_0 and all χ_n are continuous, we have $\chi_n(x) \leq \varphi_0(x)$ for each $x \in [0,1]$; moreover, $\varphi_0(x) = \max_n \chi_n(x)$ for each $x \in D$. We want to show that $(\varphi_0 > t) = \bigcup_{n=1}^{\infty}(\chi_n > t)$ for all $t \in \mathbb{R}$. Fix $t \in \mathbb{R}$, and suppose $\widetilde{x}$ is such that $\varphi_0(\widetilde{x}) > t$. Since φ_0 is continuous, there exists a neighborhood V of $\widetilde{x}$ such that $\varphi_0(x) > t$ for each $x \in V$. Hence, there exists $x_n \in D \cap V$ such that $x_n \leq \widetilde{x}$ and $\chi_n(x_n) = \varphi_0(x_n) > t$. As χ_n is monotonic, $\chi_n(\widetilde{x}) \geq \chi_n(x_n) > t$, and so $\widetilde{x} \in (\chi_n > t)$. Since t was arbitrary, this shows that $(\varphi_0 > t) \subseteq \bigcup_{n=1}^{\infty}(\chi_n > t)$ for all $t \in \mathbb{R}$. Conversely, suppose $\widetilde{x} \in \bigcup_{n=1}^{\infty}(\chi_n > t)$. Then, there is some $n \geq 1$ such that $\chi_n(\widetilde{x}) > t$. As $\varphi_0 \geq \chi_n$, we have $\varphi_0(\widetilde{x}) \geq \chi_n(\widetilde{x}) > t$, and so $\bigcup_{n=1}^{\infty}(\chi_n > t) \subseteq (\varphi_0 > t)$. We conclude that $\bigcup_{n=1}^{\infty}(\chi_n > t) = (\varphi_0 > t)$.

On the other hand, for each χ_n define $\int \psi d\chi_n(\pi)$ as in (9). Since the map $\pi \to \int \psi d\pi$ is $\sigma(\Delta)$-measurable, it is easily seen that the map $\pi \to \int \psi d\chi_n(\pi)$ as well is $\sigma(\Delta)$-measurable. Hence, the equality $\bigcup_{n=1}^{\infty}(\chi_n > t) = (\varphi_0 > t)$ shows that also the map $\psi \to \int \psi d\varphi_0(\pi)$ is $\sigma(\Delta)$-measurable. In turn, this implies that the map $\psi \to \int \psi d\varphi(\pi)$ is $\sigma(\Delta)$-measurable.

Suppose now that φ is any continuous and monotone function on $[0,1]$. By the Weierstrass Approximation Theorem, there is a sequence of polynomials $P_n : [0,1] \to \mathbb{R}$ uniformly converging to φ. Given $P_n(x) = \sum_{i=0}^{k} a_i x^i$, for each $x \in [0,1]$ set $P_n^+(x) = \sum_{i=0}^{k}(a_i \vee 0)x^i$ and $P_n^+(x) = \sum_{i=0}^{k} -(a_i \wedge 0)x^i$. Then, both P_n^+ and P_n^- are monotone, convex and continuous functions on $[0,1]$. For each n we have $\int \psi dP_n(\pi) = \int \psi dP_n^+(\pi) - \int \psi dP_n^-(\pi)$ – where the integrals are defined as in (9) – and so, by what has been proved before, each map $\pi \to \int \psi dP_n(\pi)$ is $\sigma(\Delta)$-measurable.

As $\lim_n \int \psi dP_n(\pi) = \int \psi d\varphi(\pi)$, we conclude that also the map $\pi \to \int \psi dP_n(\pi)$ is $\sigma(\Delta)$-measurable, as desired. ■

The lottery act $l_f(\pi)$ of Lemma 1 is constructed as follows: set

$$F(x) = \pi_f(\mathrm{supp}(\pi_f) \cap (-\infty, x]),$$

and define its generalized inverse $F^{-1} : [0,1] \to \mathbb{R}$ by

$$F^{-1}(r) = \inf\{x \in \mathrm{supp}(\pi_f) : F(x) \geq r\} \qquad \text{for each } r \in [0,1].$$

Then, $l_f(\pi)$ is given by $l_f(\pi)(\omega, r) = F^{-1}(r)$ for each $r \in [0,1]$ and each $\omega \in \Omega$.

For example, consider a simple act f, that is, an act taking on a finite number of values. In this case, $\mathrm{supp}(\pi_f)$ is finite, say $\mathrm{supp}(\pi_f) = \{x_1, ..., x_n\}$, with $x_1 < \cdots < x_n$. It is easily seen that here $l_f(\pi)$ is given by:

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

for each $r \in [0,1]$ and each $\omega \in \Omega$.

Proof of Lemma 1. Since f is bounded, $\mathrm{supp}(\pi_f)$ is a compact subset of $\mathbb{R}$. As $F^{-1}(r) \leq x$ if and only if $F(x) \geq r$, we have

$$\lambda\left(F^{-1}(r) \leq x\right) = \lambda([0, F(x)]) = F(x) = \pi_f(\mathrm{supp}(\pi_f) \cap (-\infty, x]).$$

In turn, this implies $\lambda(F^{-1} \in B) = \pi_f(B)$ for all Borel subsets $B \subseteq \text{supp}(\pi_f) \subseteq \mathcal{C}$. We conclude that the desired lottery act $l_f(\pi)$ is given by $l_f(\omega, r) = F^{-1}(r)$ for each $r \in [0,1]$ and each $\omega \in \Omega$ (notice that F^{-1} is non-decreasing and so it is Riemann integrable). ■

A.2 Representation Theorem

We state the more comprehensive representation result mentioned right after Theorem 1 in which the axioms are both necessary and sufficient.

Theorem 4 *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*

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

Moreover, v and u are 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.3 Results on Ambiguity Attitude

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

Lemma 6 *Let $\phi : A \subseteq \mathbb{R} \to \mathbb{R}$ be a continuous function defined on a convex set A. Then, ϕ is concave (strictly concave) if and only if there exists $\lambda \in (0,1)$ such that, for all $x, y \in A$ with $x \neq y$,*

$$\phi(\lambda x + (1-\lambda) y) \geq (>)\, \lambda \phi(x) + (1-\lambda)\phi(y). \tag{10}$$

Proof of Lemma 2. First notice that, since both $\mathcal{C}$ and $\mathcal{U}$ are Borel subsets of $\mathbb{R}$, both $\mathcal{B}_c$ and $\mathcal{B}_u$ coincide with the restrictions of the Borel σ-algebra of $\mathbb{R}$ on $\mathcal{C}$ and $\mathcal{U}$, respectively. Since u is injective and Borel measurable, each set $u(B)$, with $B \in \mathcal{B}_c$, belongs to $\mathcal{B}_u$ (see, e.g., Corollary 15.2 of Kechris (1995)). Hence, $\{u(B) : B \in \mathcal{B}_c\} \subseteq \mathcal{B}_u$.

On the other hand, let $B \in \mathcal{B}_u$. Since u is Borel measurable, $u^{-1}(B) \in \mathcal{B}_c$. Hence, $B = u(u^{-1}(B)) \in \{u(B) : B \in \mathcal{B}_c\}$, as desired. ■

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

(iii) implies (i): Suppose Π consists of two mutually singular probability measures π' and π'', 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, μ_Π has full support on Π, there is $\lambda \in (0,1)$ such that $\mu_\Pi(\pi') = \lambda$ and $\mu_\Pi(\pi'') = 1 - \lambda$. Thus, $\mu_{\Pi,f}(x) = \lambda$ and $\mu_{\Pi,f}(y) = 1 - \lambda$. By (iii) and the representation,

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

So, there exists $\lambda \in (0,1)$ such that, given any $x, y \in \mathcal{U}$, Equation (11) holds. By Lemma 6, ϕ is concave. Finally, by Axiom 4, ϕ is independent of the choice of Π 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 ϕ both concave and convex is equivalent to ϕ linear. Now turn to the proof of (2.). ϕ strictly concave on an open interval $J \subseteq \mathcal{U}$ implies $\phi\left(\int x d\mu_{\Pi,f}\right) > \int \phi(x)\, d\mu_{\Pi,f}$ for all $\mu_{\Pi,f}$ with non-singleton *supp* $\left(\mu_{\Pi,f}\right) \subseteq J$ by the strict version of Jensen's inequality. Thus, $\phi\left(e\left(\mu_{\Pi,f}\right)\right) > \int \phi(x)\, d\mu_{\Pi,f}$, which in turn implies $\delta_{u^{-1}(e(\mu_\Pi))} \succ_\Pi f$ for all (f, Π) with non-singleton *supp* $\left(\mu_{\Pi,f}\right) \subseteq J$ by Theorem 1. The reverse direction follows directly from the argument in the proof of Proposition 1 that smooth ambiguity aversion implies concavity of ϕ with the weak inequalities replaced with strict and attention limited to $x, y \in J \subseteq \mathcal{U}$. Part (3.) follows exactly as (2.) with concavity replaced by convexity, inequalities reversed and $x, y \in K \subseteq \mathcal{U}$. ■

Proof of Lemma 3. 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 ϕ'' 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.4 Theorem 2

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 5, $\int \phi_A d\mu_f \geq \phi_A(u(x))$ 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 (6), $\int \phi_B d\mu_f \geq \phi_B(u(x_f))$. 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

$$h\left(\int \phi_B d\mu_f\right) \geq \int \phi_A d\mu_f = \int (h \circ \phi_B)\, d\mu_f. \tag{12}$$

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 Π, 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. (12) reduces to

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

Since $\phi_B(\mathcal{U})$ is an interval, by Lemma 6 we conclude that h is concave. ■

A.5 Corollary 4

Construct a family of expected utility preferences $\{\succeq_\Pi^{eu}\}_{\Pi \subseteq [0,1]}$ as follows: Fix u and μ_Π so that they match those for $\succeq_\Pi$ and take ϕ^{eu} to be the identity. Suppose ϕ is concave. Then ϕ is an increasing, concave transformation of ϕ^{eu}. By Theorem 2, $\succeq$ is more ambiguity averse than $\succeq^{eu}$. In the other direction, suppose $\succeq$ is more ambiguity averse than some $\succeq^{eu}$. Then by Theorem 2, ϕ is an increasing, concave transformation of ϕ^{eu}. Since ϕ^{eu} must be linear (as Axiom 4 implies ϕ is the same for each $\Pi \subseteq \Delta$), this implies ϕ is concave. ■

A.6 Proposition 3

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 $u(\mathcal{C}_k) = \mathcal{U}_k$ and consider

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

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

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

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

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

where the last equivalence follows from Definition 6. Hence, $\succeq^k$ is as ambiguity averse as $\succeq$, when restricted to $\mathcal{F}^k$. By Theorem 2, there exist $a(k) > 0$ and $b(k) \in \mathbb{R}$ such that, for all $x \in [0, 1-k]$,

$$\phi(x+k) = \phi_k(x) = a(k)\,\phi(x) + b(k). \tag{13}$$

Since k was arbitrary, we conclude that the functional equation (13) holds for all $k \in (0,1)$ and all $x \in (0,1)$ such that $x + k \leq 1$. 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) $\phi(x) = x$ or $\phi(x) = -\frac{1}{\alpha}e^{-\alpha x}$, $\alpha \neq 0$. ■

A.7 Lemma 4

Suppose f is upper semicontinuous. As u is continuous, then $u(f)$ as well is upper semicontinuous. Then, the map $\pi \to \mathbb{E}_\pi u(f)$ is upper semicontinuous in the vague topology of Δ (see, e.g., Thm 14.5 in Aliprantis and Border (1999)). Clearly, $\mu(\mathbb{E}_\pi u(f) \geq \operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f)) = 1$. As the map $\pi \to \mathbb{E}_\pi u(f)$ is upper semicontinuous, the set $(\mathbb{E}_\pi u(f) \geq \operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f))$ is closed in S, and so by the definition of support we have $\Pi \subseteq (\mathbb{E}_\pi u(f) \geq \operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f))$. Hence, $\inf_\Pi \mathbb{E}_\pi u(f) \geq \operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f)$. On the other hand, $\mu(\mathbb{E}_\pi u(f) < \inf_\Pi \mathbb{E}_\pi u(f)) \leq \mu(\Pi^c) = 0$, and so, by the definition of $\operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f)$, $\inf_\Pi \mathbb{E}_\pi u(f) \leq \operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f)$. We conclude that $\inf_\Pi \mathbb{E}_\pi u(f) = \operatorname{ess\,inf}_\Pi \mathbb{E}_\pi u(f)$, as desired.

Finally, if f is continuous and Π is compact, then the map $\pi \to \mathbb{E}_\pi u(f)$ is continuous on a compact set, and so by the Weierstrass Theorem it attains a minimum on Π.

A.8 Proposition 4

The proof is based on the following result.[11]

Lemma 7 *Given any bounded $\sigma(\Delta)$-measurable function $\psi : \Delta \to \mathbb{R}$, we have*

$$\lim_{n\to+\infty} -\frac{1}{n}\log\int_\Delta e^{-n\psi}d\mu = \operatorname{ess\,inf}\psi.$$

Let $\succsim_n$ be the orderings on $\mathcal{F}$ sharing the same function u and the same measure μ, and with $v_1 = v$, and $v_n = -e^{-nv}$ for each $n > 1$. Hence, $\phi_n = -e^{-n\phi}$ for each $n > 1$. Notice that each ϕ_n is obtained from ϕ through the increasing transformation $-e^{-nx}$. Moreover, $\alpha_n = \alpha + n\phi'$, and so point (ii) is satisfied.

Given $f, g \in \mathcal{F}$, set $F(\pi) = \int u(f)\,d\pi$ and $G(\pi) = \int u(g)\,d\pi$ for each $\pi \in \Delta$. Since $u(f)$ and $u(g)$ are simple functions, by Lemma 5 both F and G are bounded $\sigma(\Delta)$-measurable functions. Suppose that, for some $n_0 > 1$, $f \succeq_n g$ for each $n > n_0$. By Theorem 1,

$$\int_\Delta \phi_n(F(\pi))\,d\mu \geq \int_\Delta \phi_n(G(\pi))\,d\mu \qquad \text{for each } n > n_0.$$

[11] See Maccheroni, Marinacci, and Rustichini (2004) for the proof, which is a variation on a result of Donsker-Varadhan (see, e.g., Prop. 1.4.1 of Dupuis and Ellis (1997)).

Hence,

$$\int_{\Delta} -e^{-n\phi(F(\pi))} d\mu \geq \int_{\Delta} -e^{-n\phi(G(\pi))} d\mu \qquad \text{for each } n > n_0,$$

which implies

$$-\frac{1}{n} \log \int_{\Delta} e^{-n\phi(F(\pi))} d\mu \geq -\frac{1}{n} \log \int_{\Delta} e^{-n\phi(G(\pi))} d\mu \qquad \text{for each } n > n_0.$$

Since ϕ is continuous and strictly increasing, both $\phi(F)$ and $\phi(G)$ are bounded $\sigma(\Delta)$-measurable functions. By the Claim, we then have $\operatorname{ess\,inf} \phi(F) \geq \operatorname{ess\,inf} \phi(G)$. As ϕ is strictly increasing, we conclude that $\operatorname{ess\,inf} F \geq \operatorname{ess\,inf} G$, as desired.

To complete the proof, suppose $\operatorname{ess\,inf} F > \operatorname{ess\,inf} G$, so that $\operatorname{ess\,inf} \phi(F) > \operatorname{ess\,inf} \phi(G)$. By the Claim, there exists n_0 large enough so that, for all $n \geq n_0$,

$$\begin{aligned}
\left| -\frac{1}{n} \log \int_{\Delta} e^{-n\phi(F(\pi))} d\mu - \operatorname{ess\,inf} \phi(F) \right| &< \frac{\operatorname{ess\,inf} \phi(F) - \operatorname{ess\,inf} \phi(G)}{2}, \\
\left| -\frac{1}{n} \log \int_{\Delta} e^{-n\phi(G(\pi))} d\mu - \operatorname{ess\,inf} \phi(G) \right| &< \frac{\operatorname{ess\,inf} \phi(F) - \operatorname{ess\,inf} \phi(G)}{2}.
\end{aligned}$$

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

$$-\frac{1}{n} \log \int_{\Delta} e^{-n\phi(F(\pi))} d\mu > -\frac{1}{n} \log \int_{\Delta} e^{-n\phi(G(\pi))} d\mu,$$

which in turn implies $f \succ_n g$. ■

A.9 Proposition 5

Suppose (7) 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 ϕ there is $\beta^* \geq \beta$ such that

$$\phi(u(x)\beta^* + u(y)(1-\beta^*)) = V(xEy). \tag{14}$$

Since λ is non-atomic, there is $\Omega \times \mathcal{B}_1 \ni B^* \supseteq B$ such that $\pi(B^*) = \beta^*$ for all $\pi \in \Pi$. Hence, by (14) and by Theorem 1, $xEy \sim xB^*y$. By (7), this implies that $yEx \sim yB^*x$. As ϕ 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.10 Theorem 3

Let I be an index set for Π, 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 $B \in \Omega \times \mathcal{B}_1$ and $x, y \in \mathcal{C}$ 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 β 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 μ-almost-all i, then, by the representation, for all $w, z \in \mathcal{C}$,

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

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 μ-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 ϕ 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 γ as the DM'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 ϕ is strictly concave (the strictly convex case follows similarly) on the interval $[k, l]$, Jensen's inequality (and the definition of γ) 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 μ-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 μ-almost-all i and the second claim in the theorem is proved. ■

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.

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

Chew, H. S., and J. S. Sagi (2003): "Small Worlds: Modeling Attitudes Towards Sources of Uncertainty," mimeo, http://faculty.haas.berkeley.edu/sagi/03SmallWorlds.pdf.

Dupuis, P., and R. S. Ellis (1997): *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York.

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

Epstein, L. G., and J. Miao (2003): "A Two-Person Dynamic Equilibrium under Ambiguity," *Journal of Economic Dynamics and Control*, 27, 1253–1288.

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

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

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

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.

Grandmont, J.-M. (1972): "Continuity Properties of a Von Neumann-Morgenstern Utility," *Journal of Economic Theory*, 4(1), 45–57.

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

Hansen, L., and T. Sargent (2000): "Wanting Robustness in Macroeconomics," http://home.uchicago.edu/ lhansen/.

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

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

Kechris, A. S. (1995): *Classical Descriptive Set Theory.* Springer-Verlag, New York.

Klibanoff, P., M. Marinacci, and S. Mukerji (2003): “A Smooth Model of Decision Making under Ambguity,” Discussion Paper Working paper 113, Department of Economics, Oxford University, http://www.econ.ox.ac.uk/Research/wp/pdf/paper113.pdf.

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.

Maccheroni, F., M. Marinacci, and A. Rustichini (2004): “Variational Representation of Preferences under Ambiguity,” mimeo.

Machina, M. (2004): “Almost-Objective Uncertainty,” *Economic Theory*, forthcoming.

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

——— (2003): “Uncertainty Aversion with Second-Order Probabilities and Utilities,” mimeo, http://faculty.fuqua.duke.edu/

Neilson, W. S. (1993): “Ambiguity Aversion: An Axiomatic Approach Using Second Order Probabilities,” mimeo.

Quiggin, J. (1993): *Generalized Expected Utility Theory: The Rank-Dependent Expected Utility Model.* Kluwer-Nijhoff, Amsterdam.

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.

Schroder, M., and C. Skiadas (2003): “Optimal Lifetime Consumption-Portfolio Strategies under Trading Constraints and Generalized Recursive Preferences,” *Stochastic Processes and their Applications*, 108, 155–202.

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.

UPPAL, R., AND T. WANG (2003): "Model Misspecification and Under Diversification," *Journal of Finance*, 58, 2465–2486.

WAKKER, P. (1989): *Additive representations of preferences, a new foundation of decision analysis.* Kluwer (Academic Publishers), Dordrecht.