Econometrica, Vol. 73, No. 6 (November, 2005), 1849–1892

A SMOOTH MODEL OF DECISION MAKING UNDER AMBIGUITY

BY PETER KLIBANOFF, MASSIMO MARINACCI, AND SUJOY MUKERJI[1]

We propose and characterize 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 von Neumann–Morgenstern 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 beliefs, and ambiguity attitude, a characteristic of the decision maker's tastes. We show that attitudes toward pure risk are characterized by the shape of u, as usual, while attitudes toward ambiguity are characterized by the shape of ϕ. 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). The maxmin expected utility 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 portfolio choice examples are offered.

KEYWORDS: Ambiguity, uncertainty, Knightian uncertainty, ambiguity aversion, uncertainty aversion, Ellsberg paradox, ambiguity attitude.

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; Table I is a stylized description of one of those examples. The table shows four acts, f, g, f', and g', with

[1]We thank N. Al-Najjar, E. Dekel, D. Fudenberg, S. Grant, Y. Halevy, I. Jewitt, 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, M. Machina, and T. Strzalecki for helpful discussions and suggestions. We also thank three referees and the coeditor 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 University of Paris 1 for their hospitality during 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. Tomasz Strzalecki provided excellent research assistance.

TABLE I

ELLSBERG EXAMPLE

	A	B	C
f	10	0	0
g	0	10	0
f'	10	0	10
g'	0	10	10

payoffs contingent on three (mutually exclusive and exhaustive) events, A, B, and C.

Note that P2 implies that if f is preferred to g, then f' is preferred to g'. Consider a situation where a DM, whose preferences over (objective) lotteries are represented by an expected utility functional, "knows" that the probability of event A occurring is 1/3, although 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 that 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 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 where 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 that solves to yield a probability distribution on a set of macroeconomic variables of interest. However, 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 misspecification or "model uncertainty" (e.g., Hansen and Sargent (2000)).

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

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

where f is a real-valued function defined on a state space S (an "act"), u is a von Neumann–Morgenstern (vN–M) utility function, π is a probability measure on S, and 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 set of possible probabilities π over S, and therefore measures the subjective relevance of a particular π as the "right" probability. While u, as usual, characterizes attitude toward pure 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 occurs only when ϕ is linear, a situation that we show is identified with ambiguity neutrality and wherein the preferences are observationally equivalent to those 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 that we call lotteries, i.e., acts measurable with respect to a partition 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, represented 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 assumption states that preferences over second order acts are subjective expected utility (SEU) preferences. The point of defining second order acts and imposing Assumption 2 is to model explicitly the uncertainty about the "right probability" and uncover the DM's subjective beliefs with respect to this uncertainty and attitude to this uncertainty. Indeed, following this assumption, we recover μ and v: the former is a probability measure over possible priors on S that reveal the DM's beliefs, whereas the latter is the vN–M index that summarizes the DM's attitude toward the uncertainty over the "right" prior. Our third assumption connects preferences over second order acts to preferences over acts on S. The assumption 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 assumptions lead to the representation given in (1).

Intuitively, ambiguity averse DMs prefer 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 in the induced distribution of expected utilities, μ_f. This is shown to be equivalent to concavity of ϕ and to the DM being more averse to the subjective uncertainty about priors than he is to the risk in lotteries. In an investment problem, we may think of second order acts as bets on which return distribution is right. It is as if we imagine an ambiguity averse DM to be thinking as follows: "My best guess of the chance that the return distribution is 'π' is 20%. However, this is based on 'softer' information than knowing that the chance of a particular outcome 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 to answer comparative statics questions that involve 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, holding information and risk attitude fixed. Another useful comparative statics exercise is 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 that allows a conceptual/parametric separation of (possibly) ambiguous beliefs and ambiguity attitude, analogous to the distinction usually made between risk and risk attitude. The model and functional representation in the paper allow this separation, 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). A more recent contribution, Ghirardato, Maccheroni, and Marinacci (2004), axiomatizes a model termed α-MEU wherein it is possible in a certain sense to differentiate ambiguity attitude from ambiguity. A more detailed discussion of this model is deferred until Section 5.1. As will be explained in that discussion, the α-MEU model does not, in general, facilitate the first comparative static exercise mentioned above.

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

The rest of the paper is organized as follows. Section 2 states our main assumptions 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 simple portfolio choice problems. There is a brief concluding section. All proofs, unless otherwise noted in the text, appear in the Appendix.

2. ASSUMPTIONS AND REPRESENTATION

2.1. *Preliminaries*

Let $\mathcal{A}$ be the Borel σ-algebra of a separable metric space Ω and let $\mathcal{B}_1$ be 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 $\mathbb{R}$ that contains 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 $\omega \in \Omega$.[2]

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 (with respect to (w.r.t.) the product topology of S) and bounded real-valued functions on S. Using $C(S)$ we can equip Δ with

[2]Our modeling of lotteries in this way and use of a product state space is similar to the "single-stage" approach in Sarin and Wakker (1992, 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 that 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.

the vague topology, that is, the coarsest topology on Δ that makes the following functionals continuous:

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

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. 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. *Main Assumptions*

Next we describe three assumptions on the preference orderings $\succeq$ and $\succeq^2$. The first assumption 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.

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

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

ASSUMPTION 2—Subjective Expected Utility on Second Order Acts: *There exists a countably additive probability* $\mu: \sigma(\Delta) \to [0, 1]$ *with some* $J \in \sigma(\Delta)$ *such that* $0 < \mu(J) < 1$ *and a continuous, strictly increasing* $v: \mathcal{C} \to \mathbb{R}$, *such that, for all* $\mathfrak{f}, \mathfrak{g} \in \mathfrak{F}$,

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

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

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 that 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.[4] Notice that Π may well be a finite subset of Δ. Finally, the utility function v represents the DM's attitude toward risk generated by payoffs contingent on events in Δ.

[3]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 $u(\cdot)$.

[4]We assume that μ is countably additive to avoid going into technicalities that involve the support. In any case, this is likely to be assumed in any application and, as a matter of theory, Arrow (1971) showed that countably additive probabilities can be derived in an SEU model by adding a monotone continuity axiom. As Arrow (1971, p. 48) remarked, "the assumption of Monotone Continuity seems, I believe correctly, to be the harmless simplification almost inevitable in the formalization of any real-life problem." Furthermore, the foundations for Assumption 2 that we cite in the next paragraph also show how to deliver countable additivity.

Each of the first two assumptions could be replaced by more primitive assumptions on $\succeq$ and $\succeq^2$, respectively, that deliver the expected utility representations. For example, Assumption 1 can be derived along the lines of Grandmont (1972) by identifying lottery acts with their respective distributions. Theorem V.6.1 of Wakker (1989) applied to $\succeq^2$ over second order acts can be used to deliver Assumption 2. This theorem is an axiomatic characterization of continuous subjective expected utility for acts that map from general state spaces to suitably rich consequence spaces. The key axiom in Wakker's treatment is a condition that requires consistency of trade-offs revealed by preferences. These or any other axioms applied to second order acts may not always be easily verifiable, because the payoffs of these acts are contingent on elements of Δ. As observed earlier, in some instances it may not be possible to empirically verify which element in Δ actually obtains. On the other hand, such empirical verification may be relatively simple in experimental settings: in Ellsberg urn experiments, all you would need to do is dump the urn and verify the proportion of balls in it. Verification may also be possible if one has the opportunity to wait and observe a sufficiently long run of data generated by repeated realizations from the $\pi \in \Delta$ that obtains. For instance, a finance professional using a parametric model to evaluate portfolios, where parameters determine the relevant stochastic process on returns, may obtain (econometric) estimates of the true parameter value if there is enough stationarity in the data generating process.

Even when verifiability is an issue, preference axioms provide a useful conceptual underpinning to choice criteria. For example, economists often apply the subjective expected utility model to a variety of situations characterized by limited verifiability. For instance, consider an investor's portfolio choice problem. The relevant states of the world for a particular stock may include events that take place inside the firm and in the wider market. It cannot be claimed that it is easy to verify, if at all, which of these relevant states actually obtain. Similarly, subjective expected utility is used to model agents in situations of asymmetric information, in a Bayesian game for example, where agents have beliefs about private signals and even beliefs about other players' beliefs. Because signals and beliefs are private, it may not be possible for an observer to verify which of them actually are realized. Nevertheless, in such situations preference axioms are commonly invoked to justify the use of the expected utility criterion and play a useful role in understanding the meaning of this criterion. In much the same way, we believe axioms on second order acts provide a sensible foundation for our assumption of subjective expected utility over second order acts. It is an open question whether an identical representation could be derived using only preferences over acts in $\mathcal{F}$.

Our final main assumption requires the preference ordering of primary interest, $\succeq$, to be consistent with Assumptions 1 and 2 in a certain way. Before stating the assumption, we require a few preliminaries. 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 $\pi_f(B) = \pi(f^{-1}(B))$ 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 (nondecreasing) lottery act $l_f(\pi) \in \mathcal{L}$ that has 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 (Assumption 1), lottery acts have a unique certainty equivalent.

Since f together with a possible probability π generates a distribution over consequences identical to those generated by $l_f(\pi)$, we assume (for consistency with Assumption 1) that the certainty equivalent of f, given π, is the same as the certainty equivalent of $l_f(\pi)$. Thus, as the certainty equivalent of f depends on which π is the right probability law, facing f is like facing a second order act, f^2, yielding $c_f(\pi)$ for each particular π. This motivates the following definition.

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

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

The assumption below says that the DM agrees with the reasoning above and therefore orders acts $f \in \mathcal{F}$ identically to the associated second order acts $f^2 \in \mathfrak{F}$.

ASSUMPTION 3—Consistency with Preferences over Associated Second Order Acts: *Given $f, g \in \mathcal{F}$ and $f^2, g^2 \in \mathfrak{F}$,*

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

2.3. *The Representation*

The preceding three assumptions are basic to our model in that they are all that we invoke to obtain our representation result. Theorem 1 shows that given these assumptions, $\succeq$ is represented by a functional that 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 decomposition of the DM's tastes and beliefs: u determines risk attitude toward acts, ϕ determines ambiguity attitude in the sense that a concave (convex) ϕ implies ambiguity aversion (ambiguity loving), 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 Assumptions* 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*

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

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* $y \in \mathcal{U}$.

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

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

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 (3), we get

$$f \succeq g \iff \int \phi\big(u(c_f(\pi))\big)\, d\mu \geq \int \phi\big(u(c_g(\pi))\big)\, d\mu. \tag{4}$$

Now, recall,

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

So,

$$u(c_f(\pi)) = \sum_{x \in \operatorname{supp}(\pi_f)} u(x)\pi_f(x) = \int_S u(f(s))\, d\pi. \tag{5}$$

Thus, substituting (5) into (4),

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

This proves the representation claim in the theorem. To see the uniqueness properties of ϕ, notice that

$$v(c_f(\pi)) = \phi\big(u(c_f(\pi))\big) \quad \Longleftrightarrow \quad \phi(y) = v(u^{-1}(y)).$$

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

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

Hence, $\forall\, y \in \mathcal{U}$, $\tilde{\phi}(\alpha y + \beta) = (v \circ \tilde{u}^{-1})(\alpha y + \beta) = (v \circ u^{-1})(y) = \phi(y)$. Finally, v is unique up to positive affine transformations according to Assumption 2, so, fixing u, ϕ is as well. *Q.E.D.*

The integrals in (2) 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 Section A.2 as Theorem 4) for the two orderings $\succeq$ and $\succeq^2$ in which Assumptions 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 Assumption 1 imposed 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 preferences over acts in $\mathcal{F}$ is as given in the following corollary.

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

$$V(f) = \int_{\Delta} \phi\left[\int_S u(f(s))\, d\varphi(\pi)\right] d\mu. \tag{6}$$

PROOF: It is enough to observe that here (5) 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. *Q.E.D.*

Note that while the outer integral is the usual one, the inner integral in (6) is a Choquet integral, and the inner and outer integrals are well defined because 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. Finally, Section 3.3 describes a useful regularity condition on ambiguity attitudes. It also shows that this condition is implied by our other assumptions if ϕ is twice continuously differentiable.

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.

To state this formally in our setting, consider a family $\{\succeq_\Pi, \succeq_\Pi^2\}_{\Pi\subseteq\Delta}$ of pairs of preference relations (over acts and over second order acts, respectively) that characterize each DM. There is a pair of preference relations that correspond to each possible support Π, that is, to each possible state of information the DM may have about which probabilities π (over S) are relevant to his decision problem. Assumptions 4 and 5, which follow here and in Section 3.3, require certain properties of preferences to hold across the different pairs $\{\succeq_\Pi, \succeq_\Pi^2\}_{\Pi\subseteq\Delta}$. We emphasize that these assumptions are of a somewhat different nature than the three assumptions of the previous section. While the earlier assumptions operate only *within* pairs of preferences $(\succeq_\Pi, \succeq_\Pi^2)$, Assumptions 4 and 5 operate across the entire family of pairs of preferences.

For all the definitions, assumptions, and results to come, it would be enough if everywhere that we state something about an entire family $\{\succeq_\Pi, \succeq_\Pi^2\}_{\Pi\subseteq\Delta}$ of pairs of preference relations, we limit our statement to the pair of original preferences $(\succeq, \succeq^2)$ and a pair $\{\succeq_\Pi, \succeq_\Pi^2\}$ with Π that contains exactly two measures that have disjoint support.[5] While we stick with the stronger formulations that use all Π for ease of statement, this observation indicates that many fewer pairs of preference relations need to be considered (two rather than an infinite number) than the stronger versions would lead one to think.

ASSUMPTION 4—Separation of Tastes and Beliefs: *Fix a family of preference relationships* $\{\succeq_\Pi, \succeq^2_\Pi\}_{\Pi\subseteq\Delta}$ *for a given DM.*

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

(ii) *The same invariance with respect to* Π *holds for the risk preferences derived from* $\succeq^2_\Pi$.

Imposing Assumption 4 in addition to the earlier assumptions guarantees that as the support of a DM's subjective belief varies (say, due to conditioning on different information), the DM's attitude toward risk in lotteries, as embodied in u (from Assumption 1), and attitude toward risk on the space Δ, as embodied in v (from Assumption 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 $\succeq_\Pi$ and $\succeq^2_\Pi$ except that of having support Π. Although we do not need to assume it for our results, a natural possibility is that all such beliefs are 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 Assumption 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 $u(B) = \{u(x) : x \in B\}$.[6]

LEMMA 2: *We have* $\mathcal{B}_u = \{u(B) : B \in \mathcal{B}_c\}$.

[5]The significance of two measures with disjoint support is that they allow the entire space of pairs of expected utility values to be generated by varying the act under consideration. This richness is needed for preference over acts to fully pin down convexity properties of ϕ. Such richness comes for free in expected utility theory because acts are not summed over states before the utility function is applied.

[6]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 restrictions of $\mathcal{B}$ on $(0, 1]$ and $\mathcal{C}$, respectively.

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

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

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 about beliefs may affect behavior: ambiguity aversion is an aversion to the subjective 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.[7] 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 $e(\mu_f) \equiv \int_{\mathcal{U}} x \, d\mu_f$. Notice $u^{-1}(e(\mu_f)) \in \mathcal{C}$.

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

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

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

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 supports $\Pi \subseteq \Delta$.

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 subjective uncertainty about the right prior on S than he is to the risk generated by 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.

[7]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.

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 Assumptions* 1–4, *the following conditions are equivalent*:

(i) *The function* $\phi : \mathcal{U} \to \mathbb{R}$ *is concave*.
(ii) *Attitude* 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" η.

COROLLARY 2: *Under Assumptions* 1–4, *the following properties are equivalent*:

(i) *The DM is smoothly ambiguity neutral*.
(ii) *The function* ϕ *is linear*.
(iii) *The preference functional* $V(f) = \int_S u(f(s))\, d\eta$, *where* $\eta(E) = \int_\Delta \pi(E)\, d\mu$ *for all* $E \in \Sigma$.

REMARK 1: An ambiguity averse DM in this model prefers the lottery act, $\hat{\ell}$, that pays x with an objective probability p (and 0 with probability $1 - p$) to the second order act, $\hat{f}^2$, that pays x contingent on an event $E \subseteq \Delta$ to which the DM assigns a subjective prior $\mu(E) = p$ (and pays 0 elsewhere). These two options expose the DM to the same uncertainty over payoffs generated in two different ways. Ambiguity aversion is the relative dislike of payoff uncertainty generated by subjective beliefs over probability distributions on S compared to payoff uncertainty generated by lotteries.

Notice that this model does not deal with objective probabilities on Δ; μ is a subjective measure. However, it may be useful to suggest one interpretation of how an "objective μ" might be viewed by the DM. Suppose the DM is informed that each $\pi \in \Delta$ obtains with an objective probability $p(\pi)$, generated, for example, by a randomizing device, such as a roulette wheel. It seems plausible that the DM would then view the second order act $\hat{f}^2$ as equivalent to the lottery act $\hat{\ell}$. The interpretation seems appropriate since $\hat{f}^2$ and $\hat{\ell}$ are identified with identical objective probability distributions over consequences. More generally, in such a case second order acts can be regarded as lotteries, so there is no motivation to distinguish between v and u, i.e., $v = u$. Hence

ϕ is the identity and the DM evaluates acts by expected utility. Needless to say this is only an interpretation. Being strictly formal, lotteries and second order acts are different objects. Hence, it is possible that v could be more concave than u even with objective probabilities. For some recent experimental evidence on this point, see Halevy (2004). These interpretations are limited to the case where the uncertainty about second order acts is wholly objective; it is far from obvious what an appropriate interpretation is if this uncertainty were a combination of objective and subjective.

3.2. *Comparison of Ambiguity Attitudes*

In this section we study differences in ambiguity aversion across DMs. As in the previous section, we identify each DM with a family of preferences $\{\succeq_\Pi, \succeq^2_\Pi\}_{\Pi\subseteq\Delta}$, parametrized by Π. Throughout the section, we assume Assumption 4 holds in addition to the first three assumptions. 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.

DEFINITION 5: Let A and B be two DMs whose families of preferences share the same probability measures μ_Π for each support Π. We say that *A is more ambiguity averse than B* if

$$f \succeq^A_\Pi l \quad\Longrightarrow\quad f \succeq^B_\Pi l \tag{7}$$

for every $f \in \mathcal{F}$, every $l \in \mathcal{L}$, and every support $\Pi \subseteq \Delta$.

The idea behind this definition is that if two DMs, A and B, share the same beliefs but B prefers an uncertain act over a (purely) risky act whenever A does so, then this must be due to B's comparatively lower aversion to ambiguity. Given a lottery l, the set of uncertain acts that B prefers to l is larger than A's preferred set of uncertain acts. Since A and B share the same beliefs, the only factor that could explain the larger set preferred by B is difference in attitudes to the uncertainty. However, differing attitudes toward lotteries (i.e., risk) cannot be the reason. Given that the act f in preference condition (7) may itself be a lottery act and that the condition holds for every $l \in \mathcal{L}$, it essentially follows that A and B must rank lotteries the same way, hence leaving ambiguity attitude as the only factor that may explain the difference in preferences.

We can now state our comparative result, which shows that differences in ambiguity aversion across DMs who share the same belief μ are completely characterized by the relative concavity of their functions ϕ. Significantly, the result shows that Definition 5 *implies* that ambiguity aversion is comparable across two DMs *only if their risk attitudes coincide*.[8] Note that the relative concavity of ϕ plays here a role analogous to the role of the relative concavity of utility functions in standard risk theory.

THEOREM 2: *Let A and B be two DMs whose families of preferences share the same probability μ_Π for each support Π. Then A is more ambiguity averse than B if and only if they share the same (normalized) vN–M utility function u and*

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

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 hypothesis of Theorem 2 holds. If ϕ_A and ϕ_B are twice continuously differentiable, then A is more ambiguity averse than B if and only if they share the same (normalized) vN–M utility function u and, 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 $x \in \mathcal{U}$.

COROLLARY 4: *A DM's preferences are represented using a concave ϕ if and only if he is more ambiguity averse than some expected utility DM (i.e., a DM all of whose associated preferences $\succeq_\Pi$ are expected utility).*

REMARK 2: 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 also be shown (see Klibanoff, Marinacci, and Mukerji (2003)) that if our Assumption 5 (presented in the next section) holds, nothing would change if we were to take probabilistic sophistication, rather than expected utility, as the benchmark.

[8]This feature, that comparison of ambiguity attitudes is restricted to DMs with the same risk attitude, is shared by the α-MEU type models. See, for instance, Proposition 6 in Ghirardato, Maccheroni, and Marinacci (2004). It is also present in the approaches to comparative ambiguity aversion developed in Epstein (1999) and Ghirardato and Marinacci (2002).

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

DEFINITION 6: We say that the DM displays constant ambiguity attitude if, for each support $\Pi \subseteq \Delta$,

$$f \succeq_{\Pi} g \quad \Longleftrightarrow \quad f' \succeq_{\Pi} g'$$

whenever acts f, g, f', and g' are such that, for some $k \in \mathbb{R}$ and for each $s \in S$,

$$\begin{aligned} u(f'(s)) &= u(f(s)) + k, \\ u(g'(s)) &= u(g(s)) + k. \end{aligned} \tag{8}$$

Recall that under Assumption 1 (or a Grandmont (1972) style axiomatization thereof) u may be recovered from preferences on lotteries only up to normalization. However, this is enough for the definition to be meaningful, since any quadruple of acts satisfies the conditions in the definition with u if and only if it satisfies such conditions with any renormalization $\tilde{u} = \alpha u + \beta$ for any $\alpha > 0$ and $\beta \in \mathbb{R}$ (the particular values of k relating the acts may change, but this is immaterial).

To see the spirit of the definition, notice that 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.,

$$\mu_{f'}(z+k) = \mu_f(z) \quad \text{and} \quad \mu_{g'}(z+k) = \mu_g(z).$$

The intuition of constant ambiguity attitude is that the DM views the "ambiguity content" in μ_f and its "translation" $\mu_{f'}$ 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 an exponential ϕ. It is of some interest to note that the proposition does not assume that ϕ is differentiable.

PROPOSITION 2: *The DM displays constant ambiguity attitude if and only if either $\phi(x) = x$ for all $x \in \mathcal{U}$ or there exists an $\alpha \neq 0$ such that $\phi(x) = -\frac{1}{\alpha}e^{-\alpha x}$ for all $x \in \mathcal{U}$, up to positive affine transformations.*

We now turn to extreme ambiguity attitudes. The next proposition shows that when ambiguity aversion is taken 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.[9]

[9]Tomasz Strzalecki helped us improve this result from that in an earlier version and provided a key step in its proof. Observe that when $\mathcal{C}$ is bounded, then by the Dini theorem, condition (iii) can be weakened to $\lim_n \alpha_n(x) = +\infty$ for all $x \in \mathcal{U}$.

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

PROPOSITION 3: *Let A_n be any sequence of DMs such that*:
(i) *all A_n share the same measures μ_{Π}*;
(ii) *for all n, A_{n+1} is more ambiguity averse than A_n*;
(iii) $\lim_n(\inf_{x \in \mathcal{U}} \alpha_n(x)) = +\infty$;
(iv) *each ϕ_{A_n} is everywhere twice continuously differentiable*.
Given any f and g in $\mathcal{F}$, if $f \succsim_{\Pi}^{A_n} g$ for all n sufficiently large, then

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

Moreover,

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

implies that, *for all n large enough*, $f \succ_{\Pi}^{A_n} g$.

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 3: *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*, *$\Pi$ is compact and f is continuous*, *then* $\operatorname{ess\,inf}_{\Pi} \mathbb{E}_{\pi} u(f) = \min_{\pi \in \Pi} \mathbb{E}_{\pi} u(f)$.

We have claimed that our representation of preferences over acts in $\mathcal{F}$ allows a separation of beliefs and tastes. We now make explicit the sense in which this is so. In the representation, subjective beliefs (or information) are captured by μ, including the ambiguity in the beliefs, as the results and discussion in Section 4 explain. With respect to tastes, three conceptually distinct attitudes are present in our model: risk attitude on acts in $\mathcal{F}$, risk attitude on second order acts in $\mathfrak{F}$, and ambiguity attitude on acts in $\mathcal{F}$. To what extent are the representations of these attitudes separately embodied in u, v, and ϕ, respectively? It is clear from Assumption 1 that u represents risk attitude on lotteries (a subset of acts in $\mathcal{F}$). However, given our three main assumptions, 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 for all acts in $\mathcal{F}$ and not simply lotteries. (See footnote 3 for an additional justification that u represents risk attitude toward acts in $\mathcal{F}$.) Assumption 2 guarantees that v represents risk attitude on second order acts in $\mathfrak{F}$. Proposition 1 shows that ϕ represents absolute ambiguity attitude on acts in $\mathcal{F}$; i.e., the shape of ϕ completely determines whether the DM is ambiguity averse, ambiguity loving, or ambiguity neutral. Theorem 2 shows the sense in which ϕ represents comparative ambiguity attitude; the comparison of ambiguity attitudes of two DMs who share the same information, μ, is completely determined by the comparative concavity of the respective ϕ's; however, the ambiguity attitudes of two DMs are not comparable if they do not share the same risk attitude, as represented by u. To claim a complete separation of these three attitudes and beliefs, it would have to be possible to meaningfully compare properties of preferences with representations involving (arbitrarily) different specifications of μ, ϕ, u, and v. However, our separation is not complete in this sense because of two limitations, which are explained next.

First, it is clear that a three-way separation of taste "parameters" ϕ, v, and u is not possible, because they are related by the equation $\phi = v \circ u^{-1}$. Nevertheless, it is equally clear that any two of these three parameters ϕ, v, and u may be specified independently when using the model. For example, any ϕ may be combined with any u (and then v will be $\phi \circ u$). We are primarily interested in behavior toward acts in $\mathcal{F}$; hence, the fact that we are able to specify ϕ and u (which, unlike v, represent attitudes toward acts in $\mathcal{F}$) independently is arguably "good enough." For instance, much of the impact of the well-known preference model due to Epstein and Zin (1989) rests on the fact that they, unlike standard models, allow a separation of intertemporal substitution from intratemporal risk aversion. However, just as here, there are three attitudinal aspects of preference in their model: willingness to intertemporally substitute, intratemporal risk aversion, and preference over the timing of the resolution of risk. Moreover, only two aspects may be specified separately, and once this is done, the third is constrained. Nonetheless, it is customarily accepted that in their model there is an effective separation of the two attitudes of primary interest—intertemporal substitution and intratemporal risk aversion—and that this separation is good enough for economic modeling where one is less often interested in the third attitude.

A second limitation arises due to the fact that even though we may specify ϕ, u, and μ independently of each other, our results limit the extent to which we may infer ambiguity attitude on the basis of ϕ when comparing preferences across these specifications. Comparing preferences of two DMs, we can infer (using Proposition 1) whether each DM is ambiguity averse or ambiguity loving or neutral purely by looking at the shape of the ϕ's in the respective specifications, irrespective of the u's and μ's involved. However, Theorem 2 allows us to rank the two DMs in terms of their comparative ambiguity aversion on the basis of the *relative* concavity of ϕ only when the DMs share the same μ and, furthermore, shows that we can only compare ambiguity aversion across the two when they share the same risk attitude, u (although there is no restriction on the shared μ and u). As noted earlier, our approach is not unique in having this last feature.

Separation of this kind, with the two limitations mentioned, is nevertheless a significant advance in terms of what it allows in economic modeling. In particular, it allows a comparison of two DMs who share the same beliefs and risk attitude toward acts, but one of whom is sensitive to expected utility (thus is ambiguity neutral) and the other is sensitive to ambiguity, say ambiguity averse. This comparative static is important because it is precisely what is needed to answer a question such as, "What is the *pure* effect of introducing ambiguity sensitivity into a given economic situation?" This is a most basic question and yet this cannot even be posed, let alone answered, within the framework of the other models of decision making that incorporate ambiguity such as MEU, CEU, and α-MEU. (See Section 5.1 for further clarification of this point.)

3.3. *A Regularity Assumption on Ambiguity Attitude*

In this section, we state a restriction on the DM's preferences that requires ambiguity attitude to be "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 over which we require that the DM display either strict ambiguity aversion or strict ambiguity love, but not both. What is ruled out is the possibility that the DM's ambiguity attitude everywhere flits between ambiguity aversion and ambiguity love, continuously from one point to the next. Note that 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 assumption is immediately followed by a proposition that gives an equivalent characterization in terms of ϕ.

ASSUMPTION 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 holds everywhere*.

(ii) *There exists an open interval* $J \subseteq \mathcal{U}$ *such that smooth ambiguity aversion holds strictly at all* (f, Π) *for which* $\text{supp}(\mu_{\Pi,f})$ *is a nonsingleton 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* $\text{supp}(\mu_{\Pi,f})$ *is a nonsingleton subset of* K.

PROPOSITION 4: *Under Assumptions* 1–4, *we have the following*:

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

The following lemma and remark show that if ϕ were twice continuously differentiable, as it is likely to be in any application, then Assumption 5 would actually be implied by the other assumptions and is *not* an additional assumption.

LEMMA 4: *Suppose ϕ is twice continuously differentiable. If ϕ is not linear, then ϕ is either strictly concave or convex over some open interval.*

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

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 two-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 that 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 $yEx \prec yBx$], [$xEy \prec xBy$ 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 assumptions. Although 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 $\delta_x \succ \delta_y$,*

$$xEy \sim xBy \iff yEx \sim yBx, \tag{9}$$

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 Assumption 1, then all events in $\{\Omega\} \times \mathcal{B}_1$ are indeed unambiguous.[10]

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*-nonnull sets* $\Pi', \Pi'' \in \sigma(\Delta)$ *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 Assumptions* 4 *and* 5 *and is not smoothly ambiguity neutral*, *there exists a* $\gamma \in [0, 1]$ *such that* $\pi(E) = \gamma$, μ*-almost-everywhere*.

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 smooth ambiguity averse or strictly smooth 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 4 and Remark 3 showed that under conditions likely to be assumed in any application (twice continuous differentiability of the function ϕ and Assumption 4), ambiguity neutrality is the *only* case where there will fail to be a range of strict ambiguity aversion (or love) and so is the only case where disagreement about an event's probability will not imply that the event is ambiguous.

REMARK 4: 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 that Ghirardato and

[10]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 convex-ranged 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.

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 nonconstant ambiguity attitude as a source of difference from Epstein and Zhang (2001), and further characterizations and discussion of our definition can be found in Klibanoff, Marinacci, and Mukerji (2003). A result relevant to this discussion also proved in that paper is that, given Assumptions 1–4, 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. However, 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 25th percentile rather than the minimum. The MEU sets of priors that represent 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):

$$\hat{V}(f) = \alpha \max_{\pi \in \Pi} \mathbb{E}_\pi (u \circ f) + (1-\alpha) \min_{\pi \in \Pi} \mathbb{E}_\pi (u \circ f). \tag{10}$$

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 that 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. For a recent axiomatization and extension of the α-MEU model and a discussion of the extent to which it may offer a separation between ambiguity and ambiguity attitude, see Ghirardato, Maccheroni, and Marinacci (2004).

We have remarked that a unique contribution of the smooth ambiguity model is that it provides a formal way to compare the choice of two DMs, both of whom share the same information and the same risk attitude toward lotteries, but one of whom is ambiguity sensitive (say, ambiguity averse) while the other is ambiguity neutral (i.e., SEU). This comparative static is of primary importance, since it identifies the pure effect of introducing ambiguity attitude into a model. In contrast, the extent of separation of ambiguity and ambiguity attitude achieved in the α-MEU model is not strong enough to address this comparative static question. To see this, consider the class of preferences represented by $\hat{V}(f)$ which share a *given* set of priors Π but with α ranging over the interval $[0,1]$. This class of preferences may not, in general, include an SEU preference, since for a nonsingleton set Π there may not exist an α that corresponds to an SEU preference. (The same is also true of the generalization of α-MEU in Ghirardato, Maccheroni, and Marinacci (2004).)

Finally, we remark that it may be helpful to think of another 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 sensitive behavior only when the corresponding indifference curves in the utility space are kinked (behavior that may be called first order ambiguity sensitivity). The model in this paper focuses on incorporating sensitivity to ambiguity even when the indifference curves are not kinked ("second order ambiguity sensitivity"), thus the moniker "a smooth theory."

5.2. *Models that 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 that assumes a unique second order probability over a set of given first order probabilities, but relaxes reduction and weights the probabilities nonlinearly. 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 setup 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 that, 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 nonexpected 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 that map $\Omega_a \times \Omega_b$ 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* nonnull 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. Grant, Kajii, and Polak (2001) gives an interesting application of such a recursive expected utility framework. In this application, reduction is relaxed, replaced by a recursive formulation, to model the idea that agents may not want to "conflate" probabilistic information from two different sources of uncertainty. Halevy and Feltkamp (2005) 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 results in violations of reduction and 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 that lead 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 ambiguity attitude and a comparison with comparative statics in risk 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

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

respectively. The function u is given by

$$u(x)=\begin{cases}1+\dfrac{x^{1-\rho}-1}{2^{1-\rho}-1}, & \text{if } \rho\geq 0, \rho\neq 1,\\ 1+\dfrac{\ln(x)}{\ln(2)}, & \text{if } \rho=1.\end{cases}$$

This utility function displays constant relative risk aversion with ρ as the coefficient of relative risk aversion, and is normalized so that $u(1)=1$ and $u(2)=2$.[11] The function ϕ is given by

$$\phi(x)=\begin{cases}\dfrac{1-e^{-\alpha x}}{1-e^{-\alpha}}, & \text{if } \alpha>0,\\ x, & \text{if } \alpha=0.\end{cases}$$

This function may be said to display *constant ambiguity aversion* with this terminology justified by Proposition 2 in Section 3. Thus α is the *coefficient of ambiguity aversion*.

Table II 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.

Observe that f is an example of an ambiguous act, because 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); and $\delta_{1.15}$ 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%.

TABLE II

GROSS DOLLAR PAYOFF PER DOLLAR INVESTED FOR EACH OF THREE ASSETS

	$\omega_1\times(0,\frac{1}{2}]$	$\omega_1\times(\frac{1}{2},1]$	$\omega_2\times(0,\frac{1}{2}]$	$\omega_2\times(\frac{1}{2},1]$
f	2	2	1	1
l	3	1	3	1
$\delta_{1.15}$	1.15	1.15	1.15	1.15

[11]Notice that we normalized u at 1 and 2 instead of 0 and 1 to avoid the singularity at 0. Formally this would require a corresponding change to Assumption 1 to specify this normalization instead. It is clear, however, that there is no substantive issue involved in this change and that one may specify whatever normalization is convenient as long as one chooses a single normalization for an entire problem.

EXAMPLE 1—Allocating \$1 Between a Safe Asset and an Ambiguous Asset: Consider allocating a dollar across the assets that underlie f and $\delta_{1.15}$. The classic simple example of a static portfolio choice problem is the decision of how to allocate wealth between a safe asset and a purely risky asset. As is well known, for an expected utility DM, an increase in risk aversion leads less wealth to be invested in the risky asset. Here, however, the asset underlying f is ambiguous; $\delta_{1.15}$ is a safe asset. Just as in the case of a purely risky asset, for an expected utility DM ($\alpha = 0$), an increase in risk aversion (ρ) leads less to be invested in the ambiguous asset. Furthermore, holding risk aversion (ρ) fixed, an increase in ambiguity aversion (α) leads less to be invested in the ambiguous asset. Table III gives a numerical illustration of this effect when risk aversion is fixed at $\rho = 2$.[12]

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 and thus implicitly assuming that $\alpha = 0$—then levels of risk aversion may be overestimated. Ambiguity aversion acts like extra risk aversion. 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), and 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 separation between tastes (ρ, α) and beliefs (μ) it provides, which allows one to be confident in doing comparative statics where only tastes (or only beliefs) are being varied.

Our second example will show that ambiguity aversion does not always push behavior in the same direction as increased risk aversion would for an expected utility DM.

TABLE III

OPTIMAL AMOUNT OUT OF \$1 ALLOCATED TO THE AMBIGUOUS ASSET HOLDING RISK AVERSION AT $\rho = 2$

Ambiguity Aversion (α)	Amount Allocated to f
0	1.31521
1	1.07809
2	0.916966
5	0.660381
20	0.418877

[12]The numbers may be larger than 1 due to short sales of the safe asset.

EXAMPLE 2—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 that underlies l is available in addition to the ambiguous and safe assets of the previous example. Notice that l has a higher expected return than f. For an expected utility DM, as risk aversion increases, 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. Hence the ratio of holdings of f to l decreases. Risk aversion and ambiguity aversion are working in opposite directions in terms of the composition of the risky part of the agent's portfolio. Tables IV and V give numerical illustrations of these effects.[13,14]

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 appear to reveal 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. Note that such a story relies on the

TABLE IV

OPTIMAL AMOUNT OUT OF \$1 ALLOCATED TO THE RISKY (l) AND AMBIGUOUS (f) ASSETS AS RISK AVERSION INCREASES ASSUMING AMBIGUITY NEUTRALITY ($\alpha = 0$)

Risk Aversion (ρ)	Amount Allocated to l	Amount Allocated to f	Ratio f/l
0.75	4.47151	2.07989	0.46514
1.25	2.36052	1.86083	0.78831
2	1.19612	1.2136	1.0146
5	0.357606	0.444743	1.2437
20	0.0761992	0.102747	1.3484

[13]It is worth noting that the direction of these numerical comparative statics on the ratio of holdings of f to l as ambiguity aversion increases is unchanged 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). However the risk aversion result for EU is sensitive to the direction of diversification. If f and l have the same expected return, then the ratio is constant in risk aversion under EU. If f has the higher expected return, risk aversion in EU pushes the ratio toward l just as ambiguity aversion does, so that this case would be similar to the first portfolio example.

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

TABLE V

OPTIMAL AMOUNT OUT OF $1 ALLOCATED TO THE RISKY (l) AND AMBIGUOUS (f) ASSETS AS AMBIGUITY AVERSION INCREASES, HOLDING RISK AVERSION AT $\rho = 2$

Ambiguity Aversion (α)	Amount Allocated to l	Amount Allocated to f	Ratio f/l
0	1.19612	1.2136	1.0146
1	1.20169	1.04758	0.87176
2	1.2052	0.922631	0.76554
5	1.21016	0.694279	0.57371
20	1.2139	0.407139	0.3354

assets that risk aversion would push one to diversify into being perceived as more ambiguous than other assets. 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 this paper. First, 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 while allowing a separation of tastes and beliefs, and a full range of ambiguity attitudes (including ambiguity neutrality) for any given beliefs. Second, the paper also shows how familiar techniques from the literature on risk and risk attitude may be used to analyze ambiguity and ambiguity attitude. Third, 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. Finally, it offers a model that is smooth. Rather than the minimum operator, which generates kinks, here the model allows for smooth operators that are much easier to use in economic applications.

Dept. of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, U.S.A.; peterk@kellogg.northwestern.edu,

Dip. di Statistica e Matematica Applicata and ICER, Università di Torino, Piazza Arbarello 8, 10122, Torino, Italy; massimo.marinacci@unito.it,

and
Dept. of Economics, University of Oxford, Manor Road Building, Manor Road, Oxford OX1 3UQ, U.K.; and ICER, Torino, Italy; sujoy.mukerji@economics.ox.ac.uk.

Manuscript received April, 2003; final revision received May, 2005.

APPENDIX: PROOFS AND RELATED MATERIAL

A.1. *Preliminaries*

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

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

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

LEMMA 5: *Given a continuous and nondecreasing function $\varphi : [0, 1] \to \mathbb{R}$ and a $\psi \in B(\Sigma)$, the map*

$$\pi \mapsto \int \psi \, d\varphi(\pi),$$

from Δ to $\mathbb{R}$, is $\sigma(\Delta)$-measurable.

PROOF: The case $\varphi(x) = x$ for each $x \in [0, 1]$ is a standard result (see, e.g., Aliprantis and Border (1999, p. 483)). Let $\varphi : [0, 1] \to \mathbb{R}$ be any continuous and nondecreasing function. Given $\psi \in B(\Sigma)$, let $I^{\psi} : \Delta \to \mathbb{R}$ and $L^{\psi} : \Delta \to \mathbb{R}$ be given by $I^{\psi}(\pi) = \int \psi \, d\varphi(\pi)$ and $L^{\psi}(\pi) = \int \psi \, d\pi$, respectively. Suppose first that $\psi = 1_E$ for $E \in \Sigma$. In this case $I^{1_E}(\pi) = \varphi(\pi(E))$ and $L^{1_E}(\pi) = \pi(E)$, and so $I^{1_E} = \varphi \circ L^{1_E}$. Hence, I^{1_E} is $\sigma(\Delta)$-measurable since L^{1_E} is.

Now, suppose ψ is a simple Σ-measurable function. Then ψ can be written as $\psi = \alpha_0 + \sum_{i=1}^n \alpha_i 1_{E_i}$ with $\{\alpha_i\}_{i=1}^n \subseteq \mathbb{R}_+$ and $E_1 \subseteq \cdots \subseteq E_n$. Hence,

$$\begin{aligned} I^{\psi}(\pi) = \int \psi \, d\varphi(\pi) &= \alpha_0 + \sum_{i=1}^n \alpha_i \varphi(\pi(E_i)) \\ &= \alpha_0 + \sum_{i=1}^n \alpha_i I^{1_{E_i}}(\pi), \quad \forall \pi \in \Delta, \end{aligned}$$

and so I^{ψ} is $\sigma(\Delta)$-measurable since each $I^{1_{E_i}}$ is.

Finally, suppose ψ is any function in $B(\Sigma)$. Then there exists a sequence $\{\psi_n\}_n$ of simple Σ-measurable functions that uniformly converge to ψ. Hence,

$\lim_n I^{\psi_n}(\pi) = I^{\psi}(\pi)$ for each $\pi \in \Delta$ and so I^{ψ} is $\sigma(\Delta)$-measurable since each I^{ψ_n} is. *Q.E.D.*

The lottery act $l_f(\pi)$ of Lemma 1 is constructed as follows: define a Borel probability measure $\rho : \mathcal{B} \to [0, 1]$ by $\rho(B) = \pi_f(B \cap \mathcal{C})$ for all Borel subsets of $\mathbb{R}$ and set

$$F(x) = \rho((-\infty, x]) \quad \text{for all } x \in \mathbb{R}.$$

Define its generalized inverse $F^{-1} : (0, 1] \to \mathbb{R}$ by

$$F^{-1}(r) = \inf\{x \in \mathbb{R} : F(x) \geq r\} \quad \text{for each } r \in (0, 1].$$

Since f is bounded, $\text{supp}(\pi_f) = \text{supp}(\rho)$ is a compact subset of $\mathbb{R}$ and so $F^{-1}(r)$ is well defined for all $r \in (0, 1]$. The desired lottery $l_f(\pi)$ is then 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, $\text{supp}(\pi_f)$ is finite, say $\text{supp}(\pi_f) = \{x_1, \ldots, x_n\}$, with $x_1 < \cdots < x_n$. It is easily seen that here $l_f(\pi)$ is given by

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

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

PROOF OF LEMMA 1: For each $x \in \mathbb{R}$ we have $F^{-1}(r) \leq x$ if and only if $F(x) \geq r$, so that

$$\lambda\big(\{r \in (0, 1] : F^{-1}(r) \leq x\}\big) = \lambda\big((0, F(x)]\big) = F(x) = \rho((-\infty, x]).$$

In turn, this implies $\lambda(\{r \in (0, 1] : F^{-1}(r) \in B\}) = \rho(B)$ for all $B \in \mathcal{B}$, and so

$$\lambda\big(\{r \in (0, 1] : F^{-1}(r) \in B\}\big) = \pi_f(B) \quad \text{for all } B \in \mathcal{B}_c$$

(see, e.g., Aliprantis and Border (1999, p. 611)). 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 nondecreasing and so it is Riemann integrable). *Q.E.D.*

A.2. *Representation Theorem*

We state the more comprehensive representation result mentioned right after Theorem 1 in which the assumptions 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 are equivalent*:

(i) *Assumptions* 1, 2, *and* 3 *hold*.

(ii) *There exists a continuous*, *strictly increasing* $\phi : \mathcal{U} \to \mathbb{R}$, *a unique countably additive probability* $\mu : \sigma(\Delta) \to [0, 1]$, *and 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*

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

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

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

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*

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. *Q.E.D.*

We next state a lemma (see Theorems 88 and 91 in Hardy, Littlewood, and Polya (1952)) useful in several proofs to follow.

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* $x \neq y$,

$$\phi(\lambda x + (1 - \lambda)y) \geq (>)\ \lambda\phi(x) + (1 - \lambda)\phi(y). \qquad (11)$$

PROOF OF PROPOSITION 1: Part (i) implies (iii): By the Jensen inequality, $\phi(\int x\, d\mu_{\Pi,f}) \geq \int \phi(x)\, d\mu_{\Pi,f}$. Thus, $\phi(e(\mu_{\Pi,f})) \geq \int \phi(x)\, d\mu_{\Pi,f}$, which in turn implies $\delta_{u^{-1}(e(\mu_{\Pi,f}))} \succeq_{\Pi} f$ by Theorem 1.

Part (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,

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

so there exists $\lambda \in (0, 1)$ such that, given any $x, y \in \mathcal{U}$, (12) holds. By Lemma 6, ϕ is concave. Finally, by Assumption 4, ϕ is independent of the choice of Π above.

That part (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. *Q.E.D.*

PROOF OF THEOREM 2: The "if" part follows easily from the Jensen inequality.

As to the "only if" part, we first show that A and B share the same vN–M utility function u. Let $\mathcal{L}^*$ be the set of all lottery acts that are step functions of the form $\sum_{i=1}^n x_i 1_{(r_{i-1}, r_i]}$, with $r_0 = 0$, $r_n = 1$, $x_i \in \mathcal{C}$ for each $i = 1, \dots, n$, and $x_1 < \cdots < x_n$. Such lottery acts are in one-to-one correspondence with simple probability measures p on $\mathcal{C}$, i.e., measures such that $p(A) = 1$ for some finite set $A \subseteq \mathcal{C}$. In fact, each such p induces a unique $l_p \in \mathcal{L}^*$ given by

$$l_p(\omega, r) = \begin{cases} x_1, & \text{if } r \in (0, p(x_1)], \\ x_2, & \text{if } r \in \big(p(x_1), p(x_2) + p(x_1)\big], \\ \vdots & \vdots \qquad\qquad \vdots \\ x_n, & \text{if } r \in \left(\sum_{i=1}^{n-1} p(x_i), \sum_{i=1}^{n} p(x_i) = 1\right], \end{cases}$$

where $\{x_1, \dots, x_n\}$ is the support of p, with $x_1 < \cdots < x_n$. On the other hand, each $l \in \mathcal{L}^*$ induces a unique simple probability measure p_l on $\mathcal{C}$ with support $\{x_1, \dots, x_n\}$ given by $p(x_i) = \lambda(l^{-1}(x_i))$ for each $i = 1, \dots, n$.

Let $\mathcal{P}$ be the set of all simple probability measures on $\mathcal{C}$. Define $\succsim_\Pi^A$ on $\mathcal{P}$ by

$$p \succsim_\Pi^A q \iff l_p \succeq_\Pi^A l_q. \tag{13}$$

In a similar way, define $\succsim_\Pi^B$ on $\mathcal{P}$. Both $\succsim_\Pi^A$ and $\succsim_\Pi^B$ are well defined by what was just observed. Moreover, define $U_A : \mathcal{P} \to \mathbb{R}$ by $U_A(p) = \sum_{x \in \mathrm{supp}(p)} u_A(x) \times$

$p(x)$, where u_A is the vN–M utility index given by Assumption 1. In a similar way, define $U_B : \mathcal{P} \to \mathbb{R}$.

Clearly, $U_A(p) = \int_{(0,1]} u(l_p(r))\,dr$ for all $p \in \mathcal{P}$, and so Assumption 1 and (13) imply that $p \succsim_{\Pi}^{A} q$ if and only if $U_A(p) \geq U_A(q)$ for all $p, q \in \mathcal{P}$.

Let $l, l' \in \mathcal{L}^*$. By (7),

$$l \succeq_{\Pi}^{A} l' \quad \Longrightarrow \quad l \succeq_{\Pi}^{B} l'.$$

Hence, for all $p, q \in \mathcal{P}$,

$$p \succsim_{\Pi}^{A} q \quad \Longrightarrow \quad p \succsim_{\Pi}^{B} q.$$

Since U_A and U_B are nonconstant affine functionals on $\mathcal{P}$, by Corollary B.3 of Ghirardato, Maccheroni, and Marinacci (2004) there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $U_A = \alpha U_B + \beta$. Hence, $u_A = \alpha u_B + \beta$. Along with the normalization $u_A(0) = u_B(0) = 0$ and $u_A(1) = u_B(1) = 1$, this implies $u_A = u_B$, as desired.

Set $u = u_A = u_B$ and $h(x) = (\phi_A \circ \phi_B^{-1})(x)$ for all $x \in \mathcal{U}$. The function h is clearly strictly increasing. Moreover, since $(\phi_A^{-1} \circ \phi_A)(x) = x = (\phi_B^{-1} \circ \phi_B)(x)$ 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 (7), $\int \phi_B\,d\mu_f \geq \phi_B(u(x_f))$. In turn this implies that, for all $f \in \mathcal{F}$,

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

and so

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

Let $\phi_B(x), \phi_B(y) \in \phi_B(\mathcal{U})$. By proceeding as in the proof of Proposition 1, there is a set 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, (14) reduces to

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

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

PROOF OF PROPOSITION 2: Without loss of generality, 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 $f \succeq_{\Pi}^{k} g$ 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 $x \in \mathcal{U}_k$. For any $f \in \mathcal{F}^k$, $l \in \mathcal{L} \cap \mathcal{F}^k$ and corresponding f' and l' as in Definition 6, we have

$$\begin{aligned} f \succeq_{\Pi}^{k} l \quad &\Longleftrightarrow \quad \int \phi_k\left(\int u(f(s))\,d\pi\right)d\mu_\Pi \geq \phi_k\left(\int_{(0,1]} u(l(r))\,dr\right) \\ &\Longleftrightarrow \quad \int \phi\left(\int \big(u(f(s)) + k\big)\,d\pi\right)d\mu_\Pi \\ &\qquad\qquad \geq \phi\left(\int_{(0,1]} \big(u(l(r)) + k\big)\,dr\right) \\ &\Longleftrightarrow \quad \int \phi\left(\int u(f'(s))\,d\pi\right)d\mu_\Pi \geq \phi\left(\int_{(0,1]} u(l'(r))\,dr\right) \\ &\Longleftrightarrow \quad f' \succeq_\Pi l' \quad \Longleftrightarrow \quad f \succeq_\Pi l, \end{aligned}$$

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 $x \in [0, 1-k]$,

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

Since k was arbitrary, we conclude that the functional equation (15) holds for all $k \in (0,1)$ and all $x \in (0,1)$ such that $x + k \leq 1$. 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 $\phi(x) = -\frac{1}{\alpha}e^{-\alpha x}$, $\alpha \neq 0$. *Q.E.D.*

PROOF OF PROPOSITION 3: We begin with a couple of lemmas. The first one is proved in Maccheroni, Marinacci, and Rustichini (2004), and is a variation on a result of Donsker and Varadhan (see, e.g., Proposition 1.4.1 of Dupuis and Ellis (1997)).

LEMMA 7: *Given any bounded $\sigma(\Delta)$-measurable function $\psi : \Delta \to \mathbb{R}$ and any finitely additive probability measure η on $\sigma(\Delta)$, we have*

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

Observe that, if we set $\phi_n(x) = -e^{-nx}$ for all $x \in \psi(\Delta)$, we can write

$$\phi_n^{-1}\left(\int_\Delta \phi_n(\psi)\,d\eta\right) = -\frac{1}{n}\log\int_\Delta e^{-n\psi}\,d\eta,$$

where $\phi_n^{-1}(x) = -(1/n)\log(-x)$ for all $x \in \phi_n(\psi(\Delta)) \subseteq (-\infty, 0)$. Next we generalize Lemma 7 to general sequences of functions ϕ_n.

LEMMA 8: *Let ψ be a bounded $\sigma(\Delta)$-measurable function $\psi : \Delta \to \mathbb{R}$ and let η be a finitely additive probability measure on $\sigma(\Delta)$. Suppose $\{\phi_n\}_n$ is a sequence of real-valued functions $\phi_n : I \to \mathbb{R}$ defined on an interval I of $\mathbb{R}$ with Arrow–Pratt coefficients $\alpha_n : I \to \mathbb{R}$ such that* $\lim_{n\to\infty}(\inf_{x\in I}\alpha_n(x)) = +\infty$ *and $\alpha_n(x) \le \alpha_{n+1}(x)$ for each $x \in I$ and each n. Then*

$$\lim_{n\to+\infty}\phi_n^{-1}\left(\int_\Delta \phi_n(\psi)\,d\eta\right) = \operatorname{ess\,inf}\psi. \tag{16}$$

PROOF: Given any strictly increasing function $\chi : I \to \mathbb{R}$, we have $\int_\Delta \chi(\psi) \times d\eta \ge \chi(\operatorname{ess\,inf}\psi)$ and so

$$\chi^{-1}\left(\int_\Delta \chi(\psi)\,d\eta\right) \ge \operatorname{ess\,inf}\psi. \tag{17}$$

Now let $\chi_1 : I \to \mathbb{R}$ and $\chi_2 : I \to \mathbb{R}$ be any two concave and strictly increasing functions with Arrow–Pratt coefficients α_1 and α_2 such that $\alpha_1(x) \ge \alpha_2(x)$ for each $x \in I$. Since I is an interval and $\alpha_1 \ge \alpha_2$, by Theorem 1 of Pratt (1964) the function $(\chi_1 \circ \chi_2^{-1})(t)$ is concave in t. Hence, by the Jensen inequality,

$$\begin{aligned}\int_\Delta \chi_1(\psi)\,d\eta &= \int_\Delta (\chi_1 \circ \chi_2^{-1} \circ \chi_2)(\psi)\,d\eta = \int_\Delta (\chi_1 \circ \chi_2^{-1})(\chi_2(\psi))\,d\eta \\ &\le (\chi_1 \circ \chi_2^{-1})\left(\int_\Delta \chi_2(\psi)\,d\eta\right),\end{aligned}$$

so that, being χ_1 strictly increasing,

$$\chi_1^{-1}\left(\int_\Delta \chi_1(\psi)\,d\eta\right) \le \chi_2^{-1}\left(\int_\Delta \chi_2(\psi)\,d\eta\right). \tag{18}$$

In particular, given any m and n with $m > n$, (18) implies

$$\phi_m^{-1}\left(\int_\Delta \phi_m(\psi)\,d\eta\right) \le \phi_n^{-1}\left(\int_\Delta \phi_n(\psi)\,d\eta\right). \tag{19}$$

Hence, $\lim_{n\to+\infty}\phi_n^{-1}(\int_\Delta \phi_n(\psi)\,d\eta)$ exists.

By (17) with $\chi(x) = -e^{-nx}$, we have $-\frac{1}{n}\log\int_\Delta e^{-n\psi}\,d\eta \geq \operatorname{ess\,inf}\psi$ for each $n \geq 1$. Hence, given any $\varepsilon > 0$, by Lemma 7 there is $n_\varepsilon \geq 1$ such that

$$-\frac{1}{n_\varepsilon}\log\int_\Delta \exp(-n_\varepsilon\psi)\,d\eta - \operatorname{ess\,inf}\psi \leq \varepsilon.$$

Since $\lim_{n\to\infty}(\inf_{x\in\mathcal{U}}\alpha_n(x)) = +\infty$, there exists $\overline{n}_\varepsilon \geq 1$ such that $\inf_{x\in\mathcal{U}}\alpha_n(x) \geq n_\varepsilon$ for each $n \geq \overline{n}_\varepsilon$. Hence, if $n \geq \overline{n}_\varepsilon$, the Arrow–Pratt coefficient α_n of each ϕ_n is larger than n_ε, which is the Arrow–Pratt coefficient of $-\exp(-n_\varepsilon x)$. As a result, (18) implies

$$\phi_n^{-1}\left(\int_\Delta \phi_n(\psi)\,d\eta\right) \leq -\frac{1}{n_\varepsilon}\log\int_\Delta \exp(-n_\varepsilon\psi)\,d\eta, \quad \forall n \geq \overline{n}_\varepsilon,$$

and so

$$\phi_n^{-1}\left(\int_\Delta \phi_n(\psi)\,d\eta\right) - \operatorname{ess\,inf}\psi$$
$$\leq -\frac{1}{n_\varepsilon}\log\int_\Delta \exp(-n_\varepsilon\psi)\,d\eta - \operatorname{ess\,inf}\psi \leq \varepsilon, \quad \forall n \geq \overline{n}_\varepsilon.$$

Since ε was arbitrary, this implies

$$\lim_{n\to+\infty}\phi_n^{-1}\left(\int_\Delta \phi_n(\psi)\,d\eta\right) \leq \operatorname{ess\,inf}\psi.$$

On the other hand, since each $\phi_n : I \to \mathbb{R}$ is strictly increasing, by (17) we have $\phi_n^{-1}(\int_\Delta \phi_n(\psi)\,d\eta) \geq \operatorname{ess\,inf}\psi$ for each $n \geq 1$ and we conclude that (16) holds. *Q.E.D.*

We can now prove Proposition 3. Let $\succeq_\Pi^{A_n}$ be the orderings on $\mathcal{F}$ that share the same measures μ_Π. Since, for all n, A_{n+1} is more ambiguity averse than A_n, Theorem 2 implies that all the $\succeq_\Pi^{A_n}$ share the same vN–M utility function u. 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 $u(f)$, $u(g)$ are in $B(\Sigma)$, by Lemma 5 both F and G are bounded $\sigma(\Delta)$-measurable functions. Suppose that, for some n_0, $f \succeq_\Pi^{A_n} g$ for each $n \geq n_0$. Set $\phi_n = \phi_{A_n}$. By Theorem 1,

$$\int_\Delta \phi_n(F(\pi))\,d\mu_\Pi \geq \int_\Delta \phi_n(G(\pi))\,d\mu_\Pi, \quad \forall n \geq n_0$$

and so

$$\phi_n^{-1}\left(\int_\Delta \phi_n(F(\pi))\,d\mu_\Pi\right) \geq \phi_n^{-1}\left(\int_\Delta \phi_n(G(\pi))\,d\mu_\Pi\right), \quad \forall n \geq n_0.$$

Since $\mathcal{U}$ is an interval, we can apply Lemma 8 and so we have $\operatorname{ess\,inf}_{\Pi} F \geq \operatorname{ess\,inf}_{\Pi} G$, as desired.

To complete the proof, suppose $\operatorname{ess\,inf}_{\Pi} F > \operatorname{ess\,inf}_{\Pi} G$. By Lemma 8, there exists n_0 large enough so that, for all $n \geq n_0$,

$$\left|\phi_n^{-1}\left(\int_{\Delta} \phi_n(G(\pi))\, d\mu_{\Pi}\right) - \operatorname{ess\,inf}_{\Pi} G\right| < \frac{\operatorname{ess\,inf}_{\Pi} F - \operatorname{ess\,inf}_{\Pi} G}{2}.$$

Hence, since $\phi_n^{-1}(\int_{\Delta} \phi_n(F(\pi))\, d\mu_{\Pi}) \geq \operatorname{ess\,inf}_{\Pi} F$,

$$\phi_n^{-1}\left(\int_{\Delta} \phi_n(F(\pi))\, d\mu_{\Pi}\right) > \phi_n^{-1}\left(\int_{\Delta} \phi_n(G(\pi))\, d\mu_{\Pi}\right), \quad \text{for all } n \geq n_0,$$

which in turn implies $f \succ_{\Pi}^{A_n} g$ for all $n \geq n_0$. *Q.E.D.*

PROOF OF LEMMA 3: 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., Theorem 14.5 in Aliprantis and Border (1999)). Clearly, $\mu(\mathbb{E}_{\pi} u(f) \geq \operatorname{ess\,inf}_{\Pi} \mathbb{E}_{\pi} u(f)) = 1$. Since 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 Δ 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. *Q.E.D.*

PROOF OF PROPOSITION 4: To prove part 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 part 2, where ϕ strictly concave on an open interval $J \subseteq \mathcal{U}$ implies $\phi(\int x\, d\mu_{\Pi,f}) > \int \phi(x)\, d\mu_{\Pi,f}$ for all $\mu_{\Pi,f}$ with nonsingleton $\operatorname{supp}(\mu_{\Pi,f}) \subseteq J$ by the strict version of Jensen's inequality. Thus $\phi(e(\mu_{\Pi,f})) > \int \phi(x)\, d\mu_{\Pi,f}$, which in turn implies $\delta_{u^{-1}(e(\mu_{\Pi}))} \succ_{\Pi} f$ for all (f, Π) with nonsingleton $\operatorname{supp}(\mu_{\Pi,f}) \subseteq J$ 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 by 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 $x, y \in K \subseteq \mathcal{U}$. *Q.E.D.*

PROOF OF LEMMA 4: Suppose $\phi:\mathcal{U}\to\mathbb{R}$ is twice continuously differentiable and is not linear. There exists $x_0\in\mathcal{U}$ such that $\phi''(x_0)\neq 0$. 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. *Q.E.D.*

A.4. *Results on Ambiguity*

PROOF OF PROPOSITION 5: Suppose (9) 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

$$\phi\big(u(x)\beta^*+u(y)(1-\beta^*)\big)=V(xEy). \tag{20}$$

Since λ is nonatomic, there is $\{\Omega\}\times\mathcal{B}_1\ni B^*\supseteq B$ such that $\pi(B^*)=\beta^*$ for all $\pi\in\Pi$. Hence, by (20) and by Theorem 1, $xEy\sim xB^*y$. By (9), this implies that $yEx\sim yB^*x$. Whereas 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

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

A similar argument proves the converse implication and so

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

Finally, again a similar argument shows that

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

as desired. This completes the proof because the "only if" part is trivial. *Q.E.D.*

PROOF OF THEOREM 3: By assumption, $\succeq$ satisfies the conditions in Theorem 1 and so the representation there applies. Fix an event $E\in\Sigma$. 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 $\pi(B)$ (which is the same for all $\pi\in\Delta$). If $\pi(E)$ were equal to some fixed $\alpha\in[0,1]$ for μ-almost-all π, then, by the representation, for all $w,z\in\mathcal{C}$,

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

However, this makes it impossible for E to be ambiguous. Therefore $\pi(E)$ must vary across Δ. Specifically, if $\gamma = \int_\Delta \pi(E)\,d\mu$, then there exist μ-nonnull sets $\Pi' \in \sigma(\Delta)$ and $\Pi'' \in \sigma(\Delta)$ such that $\pi(E) < \gamma$ for $\pi \in \Pi'$ and $\pi(E) > \gamma$ for $\pi \in \Pi''$, and the first claim in the theorem is proved.

Next, suppose that $\succeq$ are not smoothly ambiguity neutral, Assumptions 4 and 5 hold and E is unambiguous. Proposition 4 implies that ϕ is strictly concave (or strictly convex) on a nonempty open interval $(u_1, u_2) \subseteq \mathcal{U}$. Fix $k, l \in \mathcal{U}$ such that $u_1 < k < l < u_2$. Let $\gamma = \int_\Delta \pi(E)\,d\mu$. One can think of γ as the DM's "expected" probability of the event E. According to our representation of preferences, the following equalities are true:

$$V\big(u^{-1}(l)\{\Omega\} \times (0,\gamma]u^{-1}(k)\big) = \phi(\gamma l + (1-\gamma)k),$$

$$V\big(u^{-1}(l)Eu^{-1}(k)\big) = \int_\Delta \phi\big(\pi(E)l + (1-\pi(E))k\big)\,d\mu,$$

$$V\big(u^{-1}(k)\{\Omega\} \times (0,\gamma]u^{-1}(l)\big) = \phi(\gamma k + (1-\gamma)l),$$

$$V\big(u^{-1}(k)Eu^{-1}(l)\big) = \int_\Delta \phi\big(\pi(E)k + (1-\pi(E))l\big)\,d\mu.$$

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

$$V\big(u^{-1}(l)Eu^{-1}(k)\big) \leq V\big(u^{-1}(l)\{\Omega\} \times (0,\gamma]u^{-1}(k)\big)$$

and

$$V\big(u^{-1}(k)Eu^{-1}(l)\big) \leq V\big(u^{-1}(k)\{\Omega\} \times (0,\gamma]u^{-1}(l)\big)$$

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

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

and

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

implying that E is ambiguous, a contradiction. Therefore, it must be that $\pi(E) = \gamma$ μ-almost-everywhere and the second claim in the theorem is proved. *Q.E.D.*

REFERENCES

ACZEL, J. (1966): *Lectures on Functional Equations and Their Applications*. New York: Academic Press.

ALIPRANTIS, C., AND K. BORDER (1999): *Infinite Dimensional Analysis*. Berlin: Springer-Verlag.

ARROW, K. J. (1971): *Essays in the Theory of Risk-Bearing*. Chicago: Markham.

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

CHEW, H. S., AND J. S. SAGI (2003): "Small Worlds: Modeling Attitudes Towards Sources of Uncertainty," Mimeo, available at *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*. New York: John Wiley & Sons.

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

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

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, 283–322.

EPSTEIN, L. G., AND J. ZHANG (2001): "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69, 265–306.

EPSTEIN, L. G., AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

ERGIN, H., AND F. GUL (2002): "A Subjective Theory of Compound Lotteries," Discussion Paper, Princeton University.

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

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

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, 45–57.

GRANT, S., A. KAJII, AND B. POLAK (2001): "Third Down with a Yard to Go: Recursive Expected Utility and the Dixit–Skeath Conundrum," *Economics Letters*, 73, 275–286.

HALEVY, Y. (2004): "Ellsberg Revisited: An Experimental Study," available at *http://www.econ.ubc.ca/halevy*.

HALEVY, Y., AND V. FELTKAMP (2005): "A Bayesian Approach to Uncertainty Aversion," *Review of Economic Studies*, 72, 449–466.

HANSEN, L., AND T. SARGENT (2000): "Wanting Robustness in Macroeconomics," available at *http://home.uchicago.edu/~lhansen/*.

HANSEN, L. P., T. J. SARGENT, AND T. D. TALLARINI, JR. (1999): "Robust Permanent Income and Pricing," *Review of Economic Studies*, 66, 873–907.

HARDY, G., J. E. LITTLEWOOD, AND G. POLYA (1952): *Inequalities*. Cambridge, U.K.: Cambridge University Press.

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

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2003): "A Smooth Model of Decision Making under Ambiguity," Discussion Paper 113, Department of Economics, Oxford University, available at *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, 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, available at *http://web.econ.unito.it/gma/massimo.htm*.

MACHINA, M. (2004): "Almost-Objective Uncertainty," *Economic Theory*, 24, 1–54.
NAU, R. (2001): "Uncertainty Aversion with Second-Order Probabilities and Utilities," in *Second International Symposium on Imprecise Probabilities and Their Application*. Ithaca, NY, 273–283; available at *http://www.sipta.org/~isipta01/proceedings/index.html*.
——— (2003): "Uncertainty Aversion with Second-Order Probabilities and Utilities," *Management Science*, forthcoming.
NEILSON, W. S. (1993): "Ambiguity Aversion: An Axiomatic Approach Using Second Order Probabilities," Mimeo, Texas A&M University.
PRATT, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.
QUIGGIN, J. (1993): *Generalized Expected Utility Theory: The Rank-Dependent Expected Utility Model*. Dordrecht: Kluwer.
SARIN, R., AND P. WAKKER (1992): "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica*, 60, 1255–1272.
——— (1997): "A Single-Stage Approach to Anscombe and Aumann's Expected Utility," *Review of Economic Studies*, 64, 399–409.
SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 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, 349–377.
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*. Dordrecht: Kluwer.