Differentiating Ambiguity and Ambiguity Attitude*

Paolo Ghirardato[a,c] Fabio Maccheroni[b,c] Massimo Marinacci[a,c]

[a]Dipartimento di Statistica e Matematica Applicata - Università di Torino
[b]Istituto di Metodi Quantitativi - Università Bocconi
[c]ICER (Torino)

April 2002
Revised: March 2003

Abstract

The objective of this paper is to show how ambiguity, and a decision maker (DM)'s response to it, can be modelled formally in the context of a very general decision model.

We introduce relation derived from the DM's preferences, called "unambiguous preference", and show that it can be represented by a set of probability measures. We provide such set with a simple differential characterization, and argue that it represents the DM's perception of the "ambiguity" present in the decision problem. Given the notion of ambiguity, we show that preferences can be represented so as to provide an intuitive representation of ambiguity attitudes.

We then argue that these ideas can be applied, e.g., to obtain a behavioral foundation for the "generalized Bayesian" updating rule and for the "α-maxmin" expected utility model.

Introduction

When requested to state their maximum willingness to pay for two pairs of complementary bets involving future temperature in San Francisco and Istanbul (and identical prize of \$ 100 in case of a win) 90 pedestrians on the University of California at Berkeley campus were on average willing to pay about \$ 41 for the two bets on San Francisco temperature, and \$ 25 for the two

*An extended version of this paper is available as a Caltech Social Science Working Paper with the title "Ambiguity from the Differential Viewpoint" [15]. We are grateful to audiences at Arizona State University, Caltech, Columbia University, Cornell University, Institute for Advanced Studies (Princeton), Università della Svizzera Italiana, Università di Messina, Università di Torino, Università di Venezia, Université Paris I, University of Iowa, the 2002 RUD Conference (Gif-sur-Yvette, France), the 2002 School on "Stochastic Methods in Decision and Game Theory" (Erice, Italy), as well as to Kim Border, Erio Castagnoli, Luigi Montrucchio, Klaus Nehring and Marciano Siniscalchi for helpful comments and discussion. We are also grateful to the MIUR for financial support. E-mails: `ghirardato@econ.unito.it`, `fabio.maccheroni@uni-bocconi.it`, and `massimo@econ.unito.it`; URL: `http://web.econ.unito.it/gma/indexi.htm`

bets on Istanbul temperature. That is, on average they would have paid almost \$ 16 more to bet on the (familiar) San Francisco temperature than on the (unfamiliar) Istanbul temperature (Fox and Tversky [11, Study 4]).

This striking pattern of preferences is by no means peculiar to the inhabitants of San Francisco. Ever since the seminal thought experiment of Ellsberg [7], it has been acknowledged that the awareness of missing information, "ambiguity" in Ellsberg's terminology, affects subjects' willingness to bet. And several experimental papers, the cited [11] being just one of the most recent ones, have found significant evidence of ambiguity affecting decision making (see Luce [22] for a survey). Though Ellsberg emphasized the relevance of *aversion* to ambiguity, later work has shown that the reaction to ambiguity is not systematically negative. Examples have been produced in which subjects tend to be ambiguity *loving*, rather than averse (e.g., Heath and Tversky [20]'s "competence hypothesis" experiments). However, the available evidence does show unequivocally that ambiguity matters for choice.

The benchmark decision model of subjective expected utility (SEU) maximization is not equipped to deal with this phenomenon: An agent who maximizes SEU cannot care about ambiguity. Therefore, theory has followed experiment. Several decision models have been proposed which extend SEU in order to allow a role for ambiguity in decision making. Most notable are the "maxmin expected utility with multiple priors" (MEU) model of Gilboa and Schmeidler [19], which allows the agent's beliefs to be represented by a set of probabilities, and the "Choquet expected utility" (CEU) model of Schmeidler [30], which allows the agent's beliefs to be represented by a unique but nonadditive probability. These models have been employed with success in understanding and predicting behavior in activities as diverse as investment (e.g., Epstein and Wang [9]), labor search (Nishimura and Ozaki [28]) or voting (Ghirardato and Katz [13]).

The objective of this paper is to show how to model formally ambiguity, and a decision maker (DM)'s response to it, in the context of a general decision model (that, for instance, encompasses MEU and CEU). It is an objective that in our view has not yet been fully achieved. For, as we discuss below, the existing literature has either focused on narrower models, or has not produced a description of ambiguity as complete as the one offered here.

The intuition behind our approach can be explained in the context of the "3-color" experiment of Ellsberg. Suppose that a DM is faced with an urn containing 90 balls which are either red, blue or yellow. The DM is told that exactly 30 of the balls are red. If we offer him the choice between a bet r that pays \$ 10 if a red ball is extracted, and the bet b that pays \$ 10 if a blue ball is extracted, he may display the preference

$$r \succ b.$$

On the other hand, let y denote the bet that pays \$ 10 if a yellow ball is extracted, and suppose that we offer him the choice between the "mixed" act $(1/2)r + (1/2)y$ and the "mixed" act $(1/2)b + (1/2)y$. Then, we might observe

$$\frac{1}{2}r + \frac{1}{2}y \prec \frac{1}{2}b + \frac{1}{2}y,$$

a violation of the independence axiom (Anscombe and Aumann [1]). The well known rationale is the following: the bet y allows the DM to "hedge" the ambiguity connected with the bet b, but not that connected with r. The DM responds to the ambiguity he perceives in this decision problem by opting for the "ambiguity hedged" positions represented by the acts r and $(1/2)b + (1/2)y$. Needless to say, we could observe a DM who displays exactly opposite preferences: she prefers b to r and $(1/2)r + (1/2)y$ to $(1/2)b + (1/2)y$ because she likes to "speculate" on the ambiguity she perceives, rather than to hedge against it.

In both cases, the presence of ambiguity in the decision problem a DM is facing is revealed to an external observer (who may ignore the information that was given to the DM about the urn composition) in the form of violations of the independence axiom. By comparison, consider a DM who does not violate independence when comparing a given pair of acts f and g. That is, $f \succcurlyeq g$ and for every act h and weight λ,

$$\lambda f + (1-\lambda)h \succcurlyeq \lambda g + (1-\lambda)h. \tag{1}$$

This DM does not appear to find any possibility of hedging against or speculating on the ambiguity of the problem at hand. We therefore conclude that such ambiguity does not affect the comparison of f and g: the DM "unambiguously prefers" f to g, which we denote by $f \succcurlyeq^* g$.

The derived relation $\succcurlyeq^*$ is the stepping stone of this paper. As we now argue, it enables us to obtain an intuitive representation of ambiguity, which in turn yields a simple description of ambiguity attitude. And this without imposing strong restrictions on the DM's primitive preference $\succcurlyeq$.

The Perception of Ambiguity and Ambiguity Attitude

Using the traditional setting of Anscombe and Aumann [1], we consider an arbitrary state space S and a *convex* set of outcomes X.[1] We assume that the DM's preference $\succcurlyeq$ satisfies all the axioms that characterize Gilboa and Schmeidler [19]'s MEU model, with the exception of the key axiom that entails a preference for ambiguity hedging, that they call "uncertainty aversion". By avoiding constraints on the DM's attitude with respect to hedging, we thus obtain a much less restrictive model than MEU. (For instance, every CEU preference satisfies our axioms, while those compatible with the MEU model are a strict subclass.) Indeed, one of the novel contributions of this paper is precisely showing that the preferences satisfying the mentioned axioms have a meaningful representation.

Given such $\succcurlyeq$, we derive from it the unambiguous preference relation $\succcurlyeq^*$ as described in Eq. (1), and show that $\succcurlyeq^*$ has a "unanimity" representation in the style of Bewley [2]: there is a utility u on X and a set of probabilities $\mathcal{C}$ (nonempty, closed and convex) on S such that

$$f \succcurlyeq^* g \quad \text{if and only if} \quad \int_S u(f(s))\, dP(s) \geq \int_S u(g(s))\, dP(s) \quad \text{for all } P \in \mathcal{C}.$$

[1] Therefore, an "act" is a map $f : S \to X$ assigning an outcome $f(s) \in X$ to every state $s \in S$. A "mixed" act $\lambda f + (1-\lambda)h$ assigns to s the outcome $\lambda f(s) + (1-\lambda)h(s) \in X$.

That is, the DM deems f to be unambiguously better than g whenever the expected utility of f is higher than the expected utility of g in *every* probabilistic scenario P in $\mathcal{C}$. The set $\mathcal{C}$ of probabilistic scenarios represents, as we shall argue presently, the DM's revealed "perception of ambiguity". We use the term "perception" as a reminder to the reader that no objective meaning is attached to $\mathcal{C}$.[2] That is, nothing precludes two DMs from perceiving different ambiguity in the same decision problem.

The key motivation for our interpretation of $\mathcal{C}$ as perceived ambiguity is the following. It is simple to see that if a DM's preference $\succcurlyeq$ has a SEU representation, the DM's probabilistic beliefs P correspond to the Gateaux derivative of the functional I that represents his preferences.[3] Intuitively, the probability $P(s)$ is the shadow price for (*ceteris paribus*) changes in the DM's utility in state s. Therefore, in the SEU case we can learn the DM's understanding of the stochastic nature of his decision problem — his subjective probabilistic scenario — by calculating the derivative of his preference functional.

If $\succcurlyeq$ does not have a SEU representation but satisfies our axioms, the functional I that represents $\succcurlyeq$ is not necessarily Gateaux differentiable. However, it does have a generalized derivative, a collection of probability measures, in every point. Such derivative is the "Clarke differential", developed by Clarke [6] as an extension of the concept of superdifferential (e.g., Rockafellar [29]) to non-concave functionals. We show that the set $\mathcal{C}$ obtained as the representation of $\succcurlyeq^*$ *is* the Clarke differential of I at 0. Thus, $\mathcal{C}$ is the (appropriately defined) derivative of I, analogously to what happens for SEU preferences. Thanks to this differential characterization, we also find that in a finite state space $\mathcal{C}$ is (the closed convex hull of) the family of the Gateaux derivatives of I where they exist. That is, if we collect all the probabilistic scenarios that rationalize the DM's evaluation of acts, we find $\mathcal{C}$.

Besides its conceptual import, the differential characterization of $\mathcal{C}$ is useful from a purely operational standpoint. By giving access to the large literature on the Clarke differential, it provides a different route for assessing the DM's perceived ambiguity and some very useful results on its mathematical properties.

Armed with the representation of perceived ambiguity, we turn to the issue of formally describing the DM's reaction to the presence of ambiguity. In our main representation theorem, we show that it is possible to express the DM's preference functional I so as to associate to each act f an ambiguity aversion coefficient $a(f)$ between 0 and 1. A surprising feature of the ambiguity aversion function $a(\cdot)$ is that it displays significantly less variation than we might expect it to. For instance, the DM must have identical ambiguity attitude for acts that agree on their ranking of the possible scenarios in $\mathcal{C}$. This restriction does not constrain overall ambiguity attitude; it can continuously range from strong attraction to strong aversion.

When the DM's preference $\succcurlyeq$ satisfies MEU the set $\mathcal{C}$ is shown to be equal to the set of priors that Gilboa and Schmeidler derive in their representation [19]. This means that a MEU

[2] We do *not* carry around the adjective "revealed". It should be obvious that, since we only use behavioral data, all the aspects of our mathematical representation are revealed (or better, attributed).

[3] That is, I such that $f \succcurlyeq g$ if and only if $I(u(f)) \geq I(u(g))$.

preference with set of priors $\mathcal{C}$ is more averse to ambiguity than any other preference which perceives the same ambiguity; i.e., with the same $\mathcal{C}$. That is, contrary to what is sometimes believed, MEU preferences *do* represent extreme aversion to ambiguity, a conclusion that could not be drawn without the separate derivation of perceived ambiguity obtained here.

In the last section of the paper, we sketch some extensions and "applications" of the ideas and results developed earlier. We look at a simple dynamic choice setting and show that the unambiguous preference relation allows us to obtain a simple characterization of the updating rule that revises every prior in the set $\mathcal{C}$ by Bayes's rule, the so-called "generalized Bayesian updating" rule. Next, we discuss the axiomatic characterization of a decision rule akin to Hurwicz's α-pessimism rule, known in the literature as the "α-MEU" decision rule. We also consider the consequences of our theory of ambiguity for the classification of events and acts into ambiguous and unambiguous. These extensions are fully analyzed in the working paper version [14], to which we refer the interested reader.

Discussion

It is perhaps useful to mention from the outset some limitations and peculiarities of our analysis and terminology. We follow decision-theoretic practice in assuming that only the decision problem (states, outcomes and acts) and the DM's preference over acts are observable to an external observer (e.g., the modeller). We do not know whether other ancillary information may be available to the external observer. Hence, we do not use such information in our analysis.

This assumption entails some limitations in the accuracy of the terminology we use. First, we may end up attributing no perception of ambiguity to a DM who is aware of ambiguity but disregards it. For, it follows from our definition of unambiguous preference that if the DM never violates the independence axiom, *by definition* we attribute to him no perception of ambiguity. Such DM behaves as if he considers only one scenario P to be possible (i.e., his $\mathcal{C} = \{P\}$), maximizing his subjective expected utility with respect to P. He may just *not* be reacting to the ambiguity he perceives, but we cannot distinguish between these situations given our observability assumptions. As we are ultimately interested in modelling the ambiguity that is *reflected in behavior*, we do not believe this to be a serious problem.

Second, and more important, we attribute every departure from the independence axiom to the presence of ambiguity. That is, following Ghirardato and Marinacci [18] we implicitly assume that behavior in the absence of ambiguity will be consistent with the SEU model. However, it is well-known that observed behavior in the absence of ambiguity — that is, in experiments with "objective" probabilities — is often at spite with the independence axiom (again, see Luce [22] for a survey). As a result, the relation $\succcurlyeq^*$ we associate with a DM displaying such systematic violations overestimates the DM's perception of ambiguity. His set $\mathcal{C}$ describes behavioral traits that may not be related to ambiguity *per se*.

As extensively discussed in [18], this overestimation of the role of ambiguity could be avoided by careful filtering of the effects of the behavioral traits unrelated to ambiguity. But such filtering requires an external device (e.g., a rich set of events) whose non-ambiguity is *primitively assumed*, in violation of our observability premise. For conceptual reasons outlined in [18], in the absence of such device we prefer to attribute *all* departures from independence to the presence of ambiguity. However, the reader may prefer to use a different name for what we call "perception of ambiguity". We hope that it will be deemed to be an object of interest regardless of its name.

An aspect of our analysis which may appear to be a limitation is our heavy reliance on the concept of mixed acts. Indeed, the existence of a mixture operation is key to identifying the unambiguous preference relation. As the traditional interpretation of mixtures in the Anscombe-Aumann [1] framework is in terms of "lotteries over acts", it may be believed that our model also relies on an external notion of ambiguity. However, this is not the case, for it has been shown by Ghirardato, Maccheroni, Marinacci and Siniscalchi [16] that, if the set of outcomes is sufficiently rich, for any mixture of acts it is possible to construct an act whose state-contingent utility profile replicates perfectly that of the mixture. Our analysis can be fully reformulated in terms of such "subjective mixtures", and hence requires no external device.

The Related Literature

In addition to the mentioned paper of Gilboa and Schmeidler [19], there are several papers that share features, objectives, or methods with this paper.

Our approach to modelling ambiguity is closely related to that of Klaus Nehring. In particular, Nehring was the first to suggest using the maximal independent restriction of the primitive preference relation, which turns out to be equivalent to our $\succcurlyeq^*$, to model the ambiguity that a DM perceives in a problem. He spelled out this proposal in an unpublished conference presentation of 1996, in which he also presented the characterization of the set $\mathcal{C}$ representing ambiguity perception for MEU and CEU preferences when the state space is finite and utility is linear.[4]

In the recent [27], Nehring develops some of the ideas of the 1996 talk. The first part of that paper moves in a different direction than this paper, as it employs an incomplete relation that reflects probabilistic information *exogenously* available to the DM. The second part is closer to our work. In a setting with infinite states and consequences, Nehring defines a DM's perception of ambiguity by the maximal independent restriction of the primitive preferences over bets. He characterizes such definition and shows that under certain conditions it is equivalent to the one discussed here and in his 1996 talk (see footnote 10 below). His analysis mainly differs from ours in two respects. The first is that his preferences induce an underlying set $\mathcal{C}$ satisfying a range convexity property. The second is that he also investigates preferences that do not satisfy an assumption that he calls "trade-off consistency", that is satisfied automatically by the preferences discussed here. A consequence of the range convexity of $\mathcal{C}$ is that CEU preferences can satisfy trade-off consistency only if they maximize SEU, a remarkable result that does not

[4] "Preference and Belief without the Independence Axiom", presented at the LOFT2 conference in Torino (Italy), December 1996. (The slides are available from the author upon request.)

generalize to the preferences we study (whose $\mathcal{C}$ may not be convex-ranged).

A final major difference between Nehring's mentioned contributions and the present paper is that he does not envision any differential interpretation for the set of probabilities that represents the DM's ambiguity perception. To the best of our knowledge, the only papers that employ differentials of preference functionals in studying ambiguity averse preferences are the recent Carlier and Dana [4] and Marinacci and Montrucchio [25].[5] Both papers focus on Choquet preference functionals, and they look at the Gateaux derivatives of Choquet integrals as a device for characterizing the core of such capacities [25], or for obtaining a more direct computation of Choquet integrals in optimization problems [4].

In a recent paper, Siniscalchi [33] characterizes axiomatically a special case of our preference model (that we later call "piecewise linear" preferences), whose representation also involves a set of probabilities. The relation between his set $\mathcal{P}$ and our $\mathcal{C}$ are clarified in subsections 5.1 and 6.4. He does not explicitly focus on the distinction between ambiguity and ambiguity attitude. On the other hand, unlike us he emphasizes the requirement that each prior in the set yield the unique SEU representation of the DM's preferences over a convex subset of acts.

There exist several papers that propose behavioral notions of unambiguous events or acts (e.g., Nehring [26] and Epstein and Zhang [10]), but do not address the distinction between ambiguity and the DM's reaction to it. We refer the reader to [15] for a more detailed comparison of our notion of unambiguous events and acts with the ones proposed in these papers. Here, we limit ourselves to underscoring an important difference between our "relation-based" approach to modelling ambiguity and the "event-based" approach of these papers. Suppose that f and g are ambiguous acts such that f dominates g statewise. Then we do obtain the conclusion that f is unambiguously preferred to g, but the "event-based" papers do not. That is, there are aspects of ambiguity that a "relation-based" theory can describe, but the "event-based" theories cannot. We are not aware of any instance in which the converse is true.

As to the papers that discuss ambiguity *aversion*, the closest to our work is Ghirardato and Marinacci [18]. They do not obtain a separation of ambiguity and ambiguity attitude, but we show that once that separation is achieved by the technique we propose, their notion of ambiguity attitude is consistent with ours. In light of this, we refer the reader to the introduction of [18] for discussion of the relation of what we do with other works that address the characterization of ambiguity attitude.

Outline of the Paper

The paper is organized as follows. After introducing some basic notation and terminology in Section 1, we present the basic axiomatic model in Section 2. Sections 3 and 4 form the decision-theoretic core of the paper. First, we discuss the unambiguous preference relation and its characterization by a set of possible scenarios. Then, we present a general representation

[5] The works of Epstein [8] and Machina [23] are more distant from ours, as they take derivatives "with respect to sets", rather than "with respect to utility profiles", as we do.

theorem and the characterization of ambiguity attitude. The differential interpretation of the set of possible scenarios and related results are presented in Section 5. Finally, Section 6 sketches the extensions that are presented in detail in the working paper version of this work [14].

The paper has two appendices. Appendix A presents some functional analytic results that are employed in most arguments, along with further technical detail on Clarke differentials and their properties. Appendix B contains proofs for the results in the main body of the paper, in order of appearance.

1 Preliminaries and Notation

Consider a set S of **states of the world**, an algebra Σ of subsets of S called **events**, and a set X of **consequences**. We denote by $\mathcal{F}$ the set of all the **simple acts**: finite-valued Σ-measurable functions $f : S \to X$. Given any $x \in X$, we abuse notation by denoting $x \in \mathcal{F}$ the constant act such that $x(s) = x$ for all $s \in S$, thus identifying X with the subset of the constant acts in $\mathcal{F}$. Finally, for $f, g \in \mathcal{F}$ and $A \in \Sigma$, $f\,A\,g$ denotes the act which yields $f(s)$ for $s \in A$ and $g(s)$ for $s \in A^c \equiv S \setminus A$.

For convenience (see the discussion in the next section), we also assume that X is a convex subset of a vector space. For instance, this is the case if X is the set of all the lotteries on a set of prizes, as it happens in the classical setting of Anscombe and Aumann [1]. In view of the vector structure of X, for every $f, g \in \mathcal{F}$ and $\lambda \in [0, 1]$ as usual we denote by $\lambda f + (1 - \lambda)g$ the act in $\mathcal{F}$ which yields $\lambda f(s) + (1 - \lambda)g(s) \in X$ for every $s \in S$.

We model the DM's preferences on $\mathcal{F}$ by a binary relation $\succcurlyeq$. As usual, $\succ$ and $\sim$ denote respectively the asymmetric and symmetric parts of $\succcurlyeq$.

We let $B_0(\Sigma)$ denote the set of all real-valued Σ-measurable simple functions, or equivalently the vector space generated by the indicator functions 1_A of the events $A \in \Sigma$. If $f \in \mathcal{F}$ and $u : X \to \mathbb{R}$, $u(f)$ is the element of $B_0(\Sigma)$ defined by $u(f)\,(s) = u(f(s))$ for all $s \in S$. We denote by $ba(\Sigma)$ the set of all finitely additive and bounded set-functions on Σ. If $\varphi \in B_0(\Sigma)$ and $m \in ba(\Sigma)$, we write indifferently $\int \varphi\, dm$ or $m(\varphi)$. A nonnegative element of $ba(\Sigma)$ that assigns value 1 to S is called a **probability**, and it is typically denoted by P or Q.

Given a functional $I : B_0(\Sigma) \to \mathbb{R}$, we say that I is: **monotonic** if $I(\varphi) \geq I(\psi)$ for all $\varphi, \psi \in B_0(\Sigma)$ such that $\varphi(s) \geq \psi(s)$ for all $s \in S$; **constant additive** if $I(\varphi + a) = I(\varphi) + a$ for all $\varphi \in B_0(\Sigma)$ and $a \in \mathbb{R}$; **positively homogeneous** if $I(a\varphi) = aI(\varphi)$ for all $\varphi \in B_0(\Sigma)$ and $a \geq 0$; **constant linear** if it is constant additive and positively homogeneous.

2 Invariant Biseparable Preferences

In this section, we introduce the basic preference model that is used throughout the paper, and show that it generalizes all the popular models of ambiguity-sensitive preferences.

The model is characterized by the following five axioms:

Axiom 1 (Weak Order) *For all $f, g, h \in \mathcal{F}$: (1) either $f \succcurlyeq g$ or $g \succcurlyeq f$, (2) if $f \succcurlyeq g$ and $g \succcurlyeq h$, then $f \succcurlyeq h$.*

Axiom 2 (Certainty Independence) *If $f, g \in \mathcal{F}$, $x \in X$, and $\lambda \in (0, 1]$, then*

$$f \succcurlyeq g \Longleftrightarrow \lambda f + (1 - \lambda)x \succcurlyeq \lambda g + (1 - \lambda)x.$$

Axiom 3 (Archimedean Axiom) *If $f, g, h \in \mathcal{F}$, $f \succ g$, and $g \succ h$, then there exist $\lambda, \mu \in (0, 1)$ such that*

$$\lambda f + (1 - \lambda)h \succ g \text{ and } g \succ \mu f + (1 - \mu)h.$$

Axiom 4 (Monotonicity) *If $f, g \in \mathcal{F}$ and $f(s) \succcurlyeq g(s)$ for all $s \in S$, then $f \succcurlyeq g$.*

Axiom 5 (Non-degeneracy) *There are $f, g \in \mathcal{F}$ such that $f \succ g$.*

With the exception of axiom 2, all the axioms are standard and well understood. Axiom 2 was introduced by Gilboa and Schmeidler [19] in their characterization of MEU preferences. It requires independence to hold when the acts being compared are mixed with a constant act x.

The following representation result is easily proved by mimicking the arguments of Gilboa and Schmeidler [19, Lemmas 3.1–3.3].

Lemma 1 *A binary relation $\succcurlyeq$ on $\mathcal{F}$ satisfies axioms 1–5* if and only if *there exists a monotonic, constant linear functional $I : B_0(\Sigma) \to \mathbb{R}$ and a nonconstant affine function $u : X \to \mathbb{R}$ such that*

$$f \succcurlyeq g \Longleftrightarrow I(u(f)) \geq I(u(g)). \tag{2}$$

Moreover, I is unique and u unique up to a positive affine transformation.

Axiom 2 is responsible for the constant linearity of the functional I. As we show in [15], it is also necessary for the independence of the preference functional I from the chosen normalization of u. While the axiom may restrict ambiguity attitude in some fashion, such separation of utility and (generalized) beliefs is key to the analysis in this paper.

We call a preference $\succcurlyeq$ satisfying axioms 1–5 an **invariant biseparable preference**. The adjective *biseparable* (originating from Ghirardato and Marinacci [18, 17]) is due to the fact that the representation on binary acts of such preferences satisfies the following separability condition: Let $\rho : \Sigma \to \mathbb{R}$ be defined by $\rho(A) \equiv I(1_A)$. Then, ρ is a normalized and monotone set-function (a **capacity**) and for all $x, y \in X$ such that $x \succcurlyeq y$ and all $A \in \Sigma$,

$$I(u(x\, A\, y)) = u(x)\, \rho(A) + u(y)\, (1 - \rho(A)). \tag{3}$$

The adjective *invariant* refers to the mentioned invariance of I w.r.t. utility normalization, which is not necessarily true of the more general preferences in [17] (see [15] for details).

Some of the best known models of decision making in the presence of ambiguity employ invariant biseparable preferences. However, these models incorporate additional assumptions on how the DM reacts to ambiguity, i.e., whether he exploits hedging opportunities or not.

Axiom 6 *For all $f, g \in \mathcal{F}$ such that $f \sim g$:*

(a) ***(Ambiguity Neutrality)*** $(1/2)f + (1/2)g \sim g$.

(b) ***(Comonotonic Ambiguity Neutrality)*** $(1/2)f + (1/2)g \sim g$ *if f and g are comonotonic.*[6]

(c) ***(Ambiguity Hedging)*** $(1/2)f + (1/2)g \succcurlyeq g$.

Axiom 6(c) is due to Schmeidler [30], and it says that the DM will in general prefer the mixture, possibly a hedge, to its components. [7] The other two are simple variations on that property.

The following result, which follows immediately from known results in the literature, shows the consequence of these three properties for the structure of the functional I in Lemma 1 (and its restriction ρ).[8]

Proposition 2 *Let $\succcurlyeq$ be a preference satisfying axioms 1–5. Then*

- $\succcurlyeq$ *satisfies axiom 6(a)* if and only if *ρ is a probability on (S, Σ) and $I(\varphi) = \int \varphi \, d\rho$ for all $\varphi \in B_0(\Sigma)$.*
- $\succcurlyeq$ *satisfies axiom 6(b)* if and only if *$I(\varphi) = \int \varphi \, d\rho$ for all $\varphi \in B_0(\Sigma)$.*
- $\succcurlyeq$ *satisfies axiom 6(c)* if and only if *there is a nonempty, closed and convex set $\mathcal{D}$ of probabilities on (S, Σ) such that $I(\varphi) = \min_{P \in \mathcal{D}} \int \varphi \, dP$ for all $\varphi \in B_0(\Sigma)$. Moreover, $\mathcal{D}$ is unique.*

Thus, a DM who satisfies axioms 1–5 and is indifferent to hedging opportunities satisfies the **SEU** model. A DM who is indifferent to hedging opportunities when they involve comonotonic acts (but may care otherwise) satisfies the **CEU** model of Schmeidler [30].

On the other hand, a DM who uniformly likes ambiguity hedging opportunities chooses according to a "maxmin EU" decision rule. Indeed, axioms 1–5 and 6(c) are the axioms proposed by Gilboa and Schmeidler [19] to characterize MEU preferences — that for reasons to be made clear below are henceforth referred to as **1-MEU**. It is natural to interpret the probabilities in $\mathcal{D}$ as a reflection of the ambiguity that the DM perceives in the decision problem, but a problem with such interpretation is the fact that the set $\mathcal{D}$ appears in Gilboa and Schmeidler's analysis only as a result of the assumption of ambiguity hedging. It therefore seems that the DM's perception of ambiguity cannot be disentangled from his behavioral response to such ambiguity.

In the next section, we show that it is possible to separate the revealed perception of ambiguity from the DM's reaction to its presence. For the sake of better assessing such separation,

[6] f and g are comonotonic if there are no states s and s' such that $f(s) \succ f(s')$ and $g(s) \prec g(s')$.

[7] He calls this property "uncertainty aversion". See Ghirardato and Marinacci [18] for an explanation of our departure from that terminology.

[8] We refer the reader to [17] and [15] for additional examples and properties of invariant biseparable preferences.

it is important to notice here that axioms 1–5 do not impose *ex ante* constraints on the DM's reaction to ambiguity (as, say, ambiguity hedging does).

We reiterate that the choice to retain the classical Anscombe-Aumann setting used by Gilboa and Schmeidler is motivated only by the intention of putting our contribution in sharper focus. Ghirardato *et al.* [16] show that if the set X does not have an "objective" vector structure (i.e., it is not convex) but is sufficiently rich, it is still possible to define mixtures in a subjective yet operationally well-defined sense. They use these "subjective mixtures" to provide an axiomatization of invariant biseparable preferences in a fully subjective setting, and they could be similarly used to extend the analysis in this paper.

Unless otherwise indicated, for the remainder of this paper $\succcurlyeq$ is tacitly assumed to be an invariant biseparable preference (i.e., to satisfy axioms 1–5), and I and u are the monotonic, constant linear functional and utility index that represent $\succcurlyeq$ in the sense of Lemma 1.

3 Priors and Perceived Ambiguity

3.1 Unambiguous Preference

As explained in the introduction, our point of departure is a relation derived from $\succcurlyeq$ that formalizes the idea that hedging/speculation considerations do not affect the ranking of acts f and g.

Definition 3 *Let $f, g \in \mathcal{F}$. Then, f is* ***unambiguously preferred*** *to g, denoted $f \succcurlyeq^* g$, if*

$$\lambda f + (1-\lambda)h \succcurlyeq \lambda g + (1-\lambda)h$$

for all $\lambda \in (0,1]$ and all $h \in \mathcal{F}$.

The unambiguous preference relation is clearly incomplete in most cases. We collect some of its other properties in the following result.

Proposition 4 *The following statements hold:*

1. *If $f \succcurlyeq^* g$ then $f \succcurlyeq g$.*
2. *For every $x, y \in X$, $x \succcurlyeq^* y$ iff $x \succcurlyeq y$. In particular, $\succcurlyeq^*$ is nontrivial.*
3. *$\succcurlyeq^*$ is a preorder.*
4. *$\succcurlyeq^*$ is monotonic: if $f(s) \succcurlyeq g(s)$ for all $s \in S$, then $f \succcurlyeq^* g$.*
5. *$\succcurlyeq^*$ satisfies* ***independence****: for all $f, g, h \in \mathcal{F}$ and $\lambda \in (0,1]$,*

$$f \succcurlyeq^* g \Longleftrightarrow \lambda f + (1-\lambda)h \succcurlyeq^* \lambda g + (1-\lambda)h.$$

6. $\succcurlyeq^*$ *satisfies the* ***sure-thing principle****: for all* $f, g, h, h' \in \mathcal{F}$ *and* $A \in \Sigma$,

$$f\,A\,h \succcurlyeq^* g\,A\,h \Longleftrightarrow f\,A\,h' \succcurlyeq^* g\,A\,h'.$$

7. $\succcurlyeq^*$ *is the maximal restriction of* $\succcurlyeq$ *satisfying independence.*[9]

Thus, unambiguous preference satisfies both the classical independence conditions. It is a refinement of the state-wise dominance relation, and the maximal restriction of the primitive preference relation satisfying independence.

The last point of the proposition shows that if we turned our perspective around and *defined* unambiguous preference as the maximal restriction of $\succcurlyeq$ that satisfies the independence axiom, we would find exactly our $\succcurlyeq^*$. As mentioned earlier, this second approach was suggested by Nehring in a 1996 talk (see footnote 4).[10] While eventually the approaches reach the same conclusions, we prefer the approach taken in this paper as it is directly linked to more basic behavioral considerations about hedging and speculation.

3.2 The Perception of Ambiguity

We now show that the unambiguous preference relation $\succcurlyeq^*$ can be represented by a set of probabilities, thus extending to an arbitrary state space a result of Bewley [2]. (An alternative generalization is found in Nehring [27].)

Proposition 5 *There exists a unique nonempty, weak* compact and convex set* $\mathcal{C}$ *of probabilities on* Σ *such that for all* $f, g \in \mathcal{F}$,

$$f \succcurlyeq^* g \Longleftrightarrow \int_S u(f)\,dP \geq \int_S u(g)\,dP \quad \text{for all } P \in \mathcal{C}. \tag{4}$$

In words, f is unambiguously preferred to g if and only if every probability $P \in \mathcal{C}$ assigns a higher expected utility to f in terms of the function u obtained in Lemma 1. It is natural to refer to each prior $P \in \mathcal{C}$ as a "possible scenario" that the DM envisions, so that unambiguous preference corresponds to preference *in every scenario*. Given an act $f \in \mathcal{F}$, we will refer to the mapping $\{P(u(f)) : P \in \mathcal{C}\}$ that associates to every probability $P \in \mathcal{C}$ the expected utility of f as the **expected utility mapping of f (on $\mathcal{C}$)**.

In our view, the set $\mathcal{C}$ of probabilities represents formally the ambiguity that the DM sees in the decision problem. Hereafter we offer a couple of remarks in support of this interpretation. In Section 5 we provide further argument in favor of this interpretation by showing the differential nature of $\mathcal{C}$.

Consider two DMs with respective preference relations $\succcurlyeq_1$ and $\succcurlyeq_2$ (whose derived relations are subscripted accordingly). Given our interpretation of $\succcurlyeq^*$, it is natural to posit that if a DM

[9] That is, if $\succcurlyeq^{**} \subseteq \succcurlyeq$ and $\succcurlyeq^{**}$ satisfies independence then $\succcurlyeq^{**} \subseteq \succcurlyeq^*$.

[10] Nehring [27] independently introduces $\succcurlyeq^*$ and observes, in a setting with infinite states, its equivalence to the approach taken in 1996 talk. He also provides further motivation for his approach.

has a *richer* unambiguous preference, it is because he feels better informed about the decision problem. Formally, $\succcurlyeq_1$ **perceives more ambiguity** than $\succcurlyeq_2$ if for all $f, g \in \mathcal{F}$:

$$f \succcurlyeq_1^* g \Longrightarrow f \succcurlyeq_2^* g.$$

It turns out that this comparative definition of perceived ambiguity is equivalent to the inclusion of the sets of priors $\mathcal{C}_i$'s.

Proposition 6 *The following statements are equivalent:*

(i) $\succcurlyeq_1$ *perceives more ambiguity than* $\succcurlyeq_2$.

(ii) u_1 *is a positive affine transformation of* u_2 *and* $\mathcal{C}_1 \supseteq \mathcal{C}_2$.

In words, the size of the set $\mathcal{C}$ measures the DM's perception of ambiguity. The larger $\mathcal{C}$ is, the more ambiguity the DM appears to perceive in the decision problem.

When does a DM behave as if he does not perceive any ambiguity in the decision problem he is facing? Intuitively, it is when his unambiguous preference relation $\succcurlyeq^*$ coincides with his preference relation $\succcurlyeq$. This is clearly tantamount to saying that $\succcurlyeq$ is itself independent. More importantly for our interpretation of the set $\mathcal{C}$, it is also equivalent to saying that there is only one possible scenario:

Proposition 7 *The following statements are equivalent:*

(i) $\succcurlyeq = \succcurlyeq^*$.

(ii) $\succcurlyeq$ *is independent.*

(iii) $\mathcal{C} = \{P\}$.

(iv) $\succcurlyeq$ *has a SEU representation with probability* P.

Summarizing the results obtained so far, we have shown that $\mathcal{C}$ represents what we call the (subjective) **perception of ambiguity** of the DM, and we have concluded that the DM **perceives some ambiguity** in a decision problem if $\mathcal{C}$ is not a singleton. Such characterization of perceived ambiguity does not rely on any assumption on the DM's *reaction* to his perception of ambiguity. We now turn our attention to the latter, which is the force that drives the relation between the expected utility mapping and the DM's evaluation of an act.

4 Enter Ambiguity Attitude: The Representation

We begin our discussion of ambiguity attitude with the following observation.

Proposition 8 *Let I and u be respectively the functional and utility obtained in Lemma 1, and $\mathcal{C}$ the set obtained in Proposition 5. Then*

$$\min_{P\in\mathcal{C}} P(u(f)) \leq I(u(f)) \leq \max_{P\in\mathcal{C}} P(u(f)). \tag{5}$$

That is, the functionals on $\mathcal{F}$ defined by $\min_{P\in\mathcal{C}} P(u(\cdot))$ and $\max_{P\in\mathcal{C}} P(u(\cdot))$ — that respectively correspond to the "worst-" and "best-case" scenario evaluations within the set $\mathcal{C}$ — provide bounds to the DM's evaluation of every act. We now use this sandwiching property to obtain a nontrivial formal description of the ambiguity attitude of the DM, via a decomposition of th functional I.

4.1 Crisp Acts

It is first of all important to illustrate that the perception of ambiguity already partitions $\mathcal{F}$ into sets of acts with "similar ambiguity". The following relation on the set $\mathcal{F}$ is key: For any $f, g \in \mathcal{F}$, write $f \asymp g$ if there exist a pair of consequences $x, x' \in X$ and weights $\lambda, \lambda' \in (0,1]$ such that

$$\lambda\, f + (1-\lambda)\, x \sim^* \lambda'\, g + (1-\lambda')\, x', \tag{6}$$

where $\sim^*$ denotes the symmetric component of the unambiguous preference relation. Such relation $\asymp$ can be simply characterized in terms of the expected utility mappings of the acts:

Lemma 9 *For every $f, g \in \mathcal{F}$, the following statements are equivalent:*[11]

(*i*) $f \asymp g$.

(*ii*) *The expected utility mappings $\{P(u(f)) : P \in \mathcal{C}\}$ and $\{P(u(g)) : P \in \mathcal{C}\}$ are a positive affine transformation of each other: there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that*

$$P(u(f)) = \alpha P(u(g)) + \beta \quad \textit{for all } P \in \mathcal{C}.$$

(*iii*) *The expected utility mappings $\{P(u(f)) : P \in \mathcal{C}\}$ and $\{P(u(g)) : P \in \mathcal{C}\}$ are isotonic: for all $P, Q \in \mathcal{C}$,*

$$P(u(f)) \geq Q(u(f)) \Longleftrightarrow P(u(g)) \geq Q(u(g)).$$

Statement (*ii*) of the lemma implies that $\asymp$ is an equivalence. Statement (*iii*) is helpful in interpreting $\asymp$. Two functions on a set are isotonic if they order its elements identically. Therefore, $f \asymp g$ is tantamount to saying that f and g order *possible* scenarios identically: the best scenario for f is best for g, the worst for g is worst for f, etc. From the vantage point of the DM's perception of ambiguity, f and g have identical dependence on the existing ambiguity.

As it will be seen presently, the equivalence classes of $\asymp$ play a key role in our representation. Given $f \in \mathcal{F}$, denote by $[f]$ the equivalence class of $\asymp$ that contains f and by $\mathcal{F}_{/\asymp}$ the quotient

[11] As inspection of the proof quickly reveals, the result is true under the assumption that there exist a function u and a set $\mathcal{C}$ that represent $\succcurlyeq^*$ in the sense of Eq. (4), without any additional conditions on the primitive $\succcurlyeq$.

of $\mathcal{F}$ with respect to $\asymp$; i.e., the collection of all equivalence classes. Clearly, $[f]$ contains all acts that are unambiguously indifferent to f (take $\lambda = 1$ in Eq. (6)), but it may contain many more acts.

It follows immediately from the lemma above that all constants are $\asymp$-equivalent; that is, for all $x, y \in X$, we have $y \in [x]$. However, the class $[x]$ contains also acts which are not constants. The following behavioral property of acts, inspired by a property that Kopylov [21] calls "transparency" (as his terminology suggests, he interprets it differently from us), is key in understanding the structure of $[x]$.

Definition 10 *The act $k \in \mathcal{F}$ is called **crisp** if for all $f, g \in \mathcal{F}$ and $\lambda \in (0,1)$,*

$$f \sim g \Longrightarrow \lambda f + (1-\lambda)\, k \sim \lambda g + (1-\lambda)\, k.$$

That is, an act is crisp if it cannot be used for hedging other acts. Intuitively, this suggests that a crisp act's evaluation is not affected by the ambiguity the DM perceives in the decision problem. The following characterization validates this intuition:

Proposition 11 *For every $k \in \mathcal{F}$, the following statements are equivalent:*

(i) *k is crisp.*

(ii) *$k \asymp x$ for some $x \in X$.*

(iii) *For every $P, Q \in \mathcal{C}$, $\int u(k)\, dP = \int u(k)\, dQ$.*

(iv) *For every $f \in \mathcal{F}$ and $\lambda \in [0,1]$,*

$$I[u(\lambda\, k + (1-\lambda)\, f)] = \lambda\, I(u(k)) + (1-\lambda)\, I(u(f)).$$

Statement *(ii)* shows that $[x]$, the equivalence class of the constants, *is* the collection of all the crisp acts. Moreover, notice that it follows from statement *(iv)* of this proposition and *(ii)* of Proposition 7 that if every act is crisp, the DM perceives no ambiguity (i.e., he satisfies SEU).

4.2 The Representation Theorem

We now have all the necessary elements to formulate our main representation theorem, wherein we achieve the formal separation of perceived ambiguity and the DM's reaction to it. Interestingly, it turns out to be a generalized Hurwicz α-pessimism representation in which the set of priors is generated endogenously.

Theorem 12 *Let $\succcurlyeq$ be a binary relation on $\mathcal{F}$. The following statements are equivalent:*

(i) *$\succcurlyeq$ satisfies axioms 1–5.*

(*ii*) *There exist a nonempty, weak* compact and convex set* $\mathcal{C}$ *of probabilities on* Σ *and a nonconstant affine function* $u : X \to \mathbb{R}$ *that represent the induced* $\succcurlyeq^*$ *in the sense of Eq. (4). There exists a function* $a : (\mathcal{F}_{/\asymp} \setminus \{[x]\}) \to [0,1]$ *such that* $\succcurlyeq$ *is represented by the monotonic functional* $I : B_0(\Sigma) \to \mathbb{R}$ *defined by*

$$I(u(f)) = \begin{cases} a([f]) \min_{P\in\mathcal{C}} \int u(f)\, dP + (1 - a([f])) \max_{P\in\mathcal{C}} \int u(f)\, dP & \text{if } f \notin [x] \\ \int u(f)\, dP \text{ for some } P \in \mathcal{C} & \text{if } f \in [x]. \end{cases}$$

Moreover, $\mathcal{C}$ *is unique,* u *is unique up to a positive affine transformation, and* a *is unique if* $\mathcal{C}$ *is not a singleton.*

In other words, the theorem proves that the functional I derived in Lemma 1 has the form:

$$I(u(f)) = a([f]) \min_{P\in\mathcal{C}} \int_S u(f)\, dP + (1 - a([f])) \max_{P\in\mathcal{C}} \int_S u(f)\, dP \tag{7}$$

when restricted to noncrisp acts. Clearly, the 1-MEU preference model and more generally the α-**MEU** preference model (that we characterize axiomatically in Section 6), in which a is a constant $\alpha \in [0,1]$, are special cases of the representation above. Also, observe that when $\mathcal{C} = \{P\}$ every act is crisp. Hence, the function a disappears from the representation, which reduces to SEU.

Two analytical observations on this representation are in order. First, notice that if f and g are noncrisp acts and $f \asymp g$, then $a([f]) = a([g])$: If f and g have identical dependence on ambiguity, the DM's reaction to the ambiguity of f is identical to his reaction to the ambiguity of g. Second, observe that for any $f \in \mathcal{F} \setminus [x]$, the coefficient $a([f])$ only depends on the expected utility mapping $\{P(u(f)) : P \in \mathcal{C}\}$ of f on $\mathcal{C}$. As a result, the same is true of DM's evaluation $I(u(f))$ of any act $f \in \mathcal{F}$: The profile of expected utilities of f (as a function over $\mathcal{C}$) completely determines the DM's preference. This is a key feature of our representation, which is also enjoyed by the model studied by Siniscalchi in [33].

4.3 An Index of Ambiguity Aversion

It is intuitive to interpret the function a as an index of the ambiguity aversion of the DM: The larger $a([f])$, the bigger the weight the DM gives to the "pessimistic" evaluation of f given by $\min_{P\in\mathcal{C}} P(u(f))$. The following simple result verifies this intuition in terms of the relative ambiguity aversion ranking of Ghirardato and Marinacci [18]. In our setting, the latter is formulated as follows: $\succcurlyeq_1$ is **more ambiguity averse than** $\succcurlyeq_2$ if for all $f \in \mathcal{F}$ and all $x \in X$, $f \succcurlyeq_1 x$ implies $f \succcurlyeq_2 x$.

Proposition 13 *Let* $\succcurlyeq_1$ *and* $\succcurlyeq_2$ *be invariant biseparable preferences, and suppose that* $\succcurlyeq_1$ *and* $\succcurlyeq_2$ *perceive identical ambiguity. Then,* $\succcurlyeq_1$ *is more ambiguity averse than* $\succcurlyeq_2$ if and only if $a_1([f]) \geq a_2([f])$ *for every* $f \in \mathcal{F} \setminus [x]$.

(Recall from Proposition 6 that $\succcurlyeq_1$ and $\succcurlyeq_2$ perceive identical ambiguity if and only if $\mathcal{C}_1 = \mathcal{C}_2$ and u_1 and u_2 are equivalent.) We conclude that the function a is a complete description of the DM's **ambiguity attitude** in relation to the perception of ambiguity described by $\mathcal{C}$.

In closing this section, we observe that it follows from Proposition 13 that there are always DMs which are more and less ambiguity averse than the DM whose preference is $\succcurlyeq$. In fact, the best- and worst-case scenario evaluations define invariant biseparable preferences that satisfy these conditions, since they correspond to $a(\cdot)$ constantly equal to 0 and 1 respectively. In a sense, they describe the DM's "ambiguity averse side" and his "ambiguity loving side". However, as these DMs do not necessarily satisfy the SEU model, they may not make the preference ambiguity averse in the sense of Ghirardato and Marinacci [18].

5 Perceived Ambiguity is a Differential

In this section we turn back to the set $\mathcal{C}$ derived in Proposition 5, showing that it is equal to the Clarke differential at 0 of the functional I obtained in Lemma 1. This provides further backing to our interpretation of $\mathcal{C}$, and at the same time yields a separate, operational, route for constructing a preference's set of possible scenarios.

Suppose first that the DM's preferences satisfy axioms 1–5 and 6(a). In such a case, the DM satisfies the SEU model; i.e., there is a probability P on Σ such that $I(u(f)) = P(u(f))$. Being linear, such I is Gateaux differentiable everywhere, with derivative constantly equal to P. The DM's beliefs can thus be found by calculating the Gateaux derivative of I in any $\varphi \in B_0(\Sigma)$, for instance in $\varphi \equiv 0$.

In contrast, if the DM's preferences only satisfy axioms 1–5, the functional I may not be Gateaux differentiable everywhere, and where it is the Gateaux derivatives may be different. That is, because of the presence of ambiguity, the shadow price for state s could depend on the structure of the act being evaluated.

One way out of this difficulty is to generalize the notion of derivative used. For instance, suppose that the DM's preferences satisfy axioms 1–5 and 6(c), so that as shown in Proposition 2 it can be represented by maxmin expected utility with a set of priors $\mathcal{D}$. Then, the functional I is monotonic, constant linear and *concave*. It is not necessarily Gateaux differentiable, but it does possess directional derivatives and a nonempty superdifferential, as defined below (see, e.g., Rockafellar [29]).

Definition 14 *Given a concave functional $I : B_0(\Sigma) \to \mathbb{R}$, its **directional derivative** in φ in the direction ξ is defined by*

$$dI(\varphi; \xi) \equiv \lim_{t \downarrow 0} \frac{I(\varphi + t\xi) - I(\varphi)}{t}.$$

*The **superdifferential** of I at φ is the set of linear functionals that dominate the directional derivative $dI(\varphi; \cdot)$. That is,*

$$\partial I(\varphi) \equiv \{m \in ba(\Sigma) : m(\xi) \geq dI(\varphi; \xi),\ \forall \xi \in B_0(\Sigma)\}.$$

Interestingly, given a monotonic, constant linear and concave I, if we calculate its superdifferential $\partial I(0)$ at 0, we find that $\mathcal{D} = \partial I(0)$. That is, the superdifferential of I at 0 (which contains $\partial I(\varphi)$ for every $\varphi \in B_0(\Sigma)$) corresponds with the set of priors obtained in the representation of I of Proposition 2. In this perspective, as the superdifferential of such I coincides with its Gateaux derivative when the latter exists, SEU corresponds to the special case in which $\partial I(0) = \{P\}$.

The question, though, is what to do with a preference $\succcurlyeq$ that only satisfies axioms 1–5. The functional I that represents $\succcurlyeq$ is only monotonic and constant linear, so that the existence of right-hand derivatives and superdifferentials is not guaranteed. As such a functional must be Lipschitz, we can use the following generalized notions due to Clarke [6], which are well known in the literature on nonsmooth optimization. (See Appendix A for further details.)

Definition 15 *Given a Lipschitz functional $I : B_0(\Sigma) \to \mathbb{R}$, its **Clarke (lower) directional derivative** in φ in the direction ξ is defined by*

$$I_\circ(\varphi; \xi) = \liminf_{\substack{\psi \to \varphi \\ t \downarrow 0}} \frac{I(\psi + t\xi) - I(\psi)}{t}.$$

*The **Clarke differential** of I in φ is the set of linear functionals that dominate the Clarke derivative $I_\circ(\varphi; \cdot)$. That is,*

$$\partial I(\varphi) = \{m \in ba(\Sigma) : m(\xi) \geq I_\circ(\varphi; \xi),\ \forall \xi \in B_0(\Sigma)\}.$$

Clarke differentials are nonempty for every Lipschitz functional, and for concave functionals they coincide with superdifferentials (see Clarke [6]). (This justifies our usage of the same symbol to denote both sets.) Moreover, when I is monotone and constant linear, its Clarke differential is a set of probability charges; that is, all the $m \in \partial I(\varphi)$ are normalized and positive (Prop. 31 in App. A).

We now show that the set $\mathcal{C}$ is equal to the Clarke differential of I in 0.

Theorem 16 *Let $\succcurlyeq$ be a binary relation satisfying axioms 1–5, and I and $\mathcal{C}$ respectively the functional and set of probabilities presented in Theorem 12. Then*

$$\mathcal{C} = \partial I(0).$$

Thus we see that, if we employ the appropriate notion of derivative, the set $\mathcal{C}$ of possible scenarios coincides with the derivative of I at 0 (which again contains $\partial I(\varphi)$ for every $\varphi \in B_0(\Sigma)$). That is, our generalized notion of "beliefs" can be obtained from the functional I by differentiation, much in the same way as we did in the SEU case.

Clearly, this calculus characterization is useful in providing an operational method for assessing a DM's perception of ambiguity $\mathcal{C}$, based on the computation of the Clarke differential at 0. However, it proves enlightening also for purely theoretical reasons. We next discuss these features in more detail.

5.1 An Operational Consequence

To present a useful operational consequence of the characterization of $\mathcal{C}$ as a Clarke differential, we consider the important special case in which the state space S is finite; i.e., $S = \{s_1, s_2, ..., s_n\}$. (However, we remark that an analogous result can be proved in the infinite case, see subsection 6.4 for details.)

When S is finite, the Clarke differential at 0 can be given the following sharp representation in terms of the standard gradients of I (see Theorem 32 in App. A):

$$\partial I(0) = \overline{co}\{\nabla I(\varphi) : \varphi \in \Omega\}, \tag{8}$$

where Ω is any subset of $\mathbb{R}^n$ such that I is differentiable on Ω and Ω^c has Lebesgue measure zero — by Rademacher's Theorem it can simply be the domain of differentiability of I.

Eq. (8) provides us with further motivation for our interpretation of the set $\mathcal{C}$ as perceived ambiguity. For, given a functional I that is not Gateaux differentiable and has (different) Gateaux derivatives almost everywhere, it is natural to interpret each such derivative as a "possible probabilistic scenario" that is implicitly used in evaluating a subset of acts. Thus, the collection of the Gateaux derivatives of I is a set-valued "belief" associated with the preference functional. Alongside Theorem 16, Eq. (8) shows that the set $\mathcal{C}$ is also a "belief" in this sense.

To illustrate the usefulness of Eq. (8), consider a **piecewise linear** functional I. That is, a functional for which there exists a family $\{C_l\}_{l \in L}$ of convex cones such that:

- $\mathbb{R}^n = \bigcup_l C_l$,
- $\operatorname{int} C_l \neq \emptyset \quad$ for each l,
- $\operatorname{int} C_l \cap \operatorname{int} C_h = \emptyset \quad$ if $l \neq h$,
- I is linear on each C_l.

For instance, Choquet integrals on finite spaces are piecewise linear functionals. The same is said of the preference functionals studied by Castagnoli *et al.* [5] and Siniscalchi [33], when defined on finite state spaces.

Eq. (8) enables us to simply describe the set $\partial I(0)$ for this class of functionals. As I is linear on each cone C_l, there is a probability vector P_l corresponding to the unique linear extension of $I_{|C_l}$ to $\mathbb{R}^n$. By Eq. (8), we then have (see after Theorem 32 in App. A) that

$$\partial I(0) = \overline{co}\{P_l : l \in L\}. \tag{9}$$

This shows that there exists a simple direct connection between our $\mathcal{C}$ and the collection of probabilities P_l derived in the cited [5] and [33].

Indeed, for Choquet preferences we can use Eq. (9) to retrieve $\mathcal{C}$ from the capacity ρ, as explained in the next example.

Example 17 *Let I be a Choquet integral with respect to a capacity ρ. Set*

$$C_\sigma = \left\{\varphi \in \mathbb{R}^n : \varphi(s_{\sigma(1)}) \geq \varphi(s_{\sigma(2)}) \geq \cdots \geq \varphi(s_{\sigma(n)})\right\}$$

for each permutation σ of $\{1, ..., n\}$ and observe that I is linear on each convex cone C_σ. In fact,

$$I(\varphi) = \int_S \varphi \, d\rho = \int_S \varphi \, dP^\sigma,$$

where P^σ is the probability defined by

$$P^\sigma(s_{\sigma(i)}) = \rho(\{s_{\sigma(1)}, s_{\sigma(2)}, ..., s_{\sigma(i)}\}) - \rho(\{s_{\sigma(1)}, s_{\sigma(2)}, ..., s_{\sigma(i-1)}\})$$

for each $i = 1, ..., n$. Hence, I is piecewise linear with respect to the collection $\{C_\sigma\}_{\sigma \in Per(n)}$, where $Per(n)$ is the set of all the permutations of $\{1, \ldots, n\}$. By Eq. (9), we then have

$$\mathcal{C} = co\{P^\sigma : \sigma \in Per(n)\}. \tag{10}$$

In other words, in the Choquet case (with finite states) the set $\mathcal{C}$ is simply the convex hull of the set of all the P^σ; that is, the convex hull generated by the probabilities used in calculating the Choquet integral as we vary the monotonicity of the act being evaluated. We thus generalize a result obtained, in the case of linear utility, by Nehring in a 1996 talk (see footnote 4).

Indeed, when the functional I is also concave — i.e., when $I(\varphi) = \int \varphi \, d\rho$, with ρ supermodular[12] — Eq. (10) easily reduces to $\mathcal{C} = Core(\rho)$. Thus, the well-known characterization of the core of a supermodular capacity due to Shapley [31] is also a consequence of Theorem 16. Proposition 19 in the next subsection says more on this.

5.2 Other Consequences

We next discuss some consequences of Theorem 16 of a more theoretical nature. First of all, from the mentioned equivalence of the Clarke differential and the superdifferential for concave I it follows immediately that $\mathcal{C} = \mathcal{D}$ whenever $\succcurlyeq$ satisfies ambiguity hedging. In other words, for a 1-MEU preference the set of priors corresponds to the set of possible scenarios. We thus generalize a result that was proved for finite S by Nehring, as reported in his 1996 talk (see footnote 4, and cf. his alternative generalization in [27]).

We can also use the differential characterization to draw some conclusions on the relation between the comparatively based notion of ambiguity aversion of Ghirardato and Marinacci [18] and the ideas in this paper. Begin by considering the following two subsets of SEU preferences.

Definition 18 *Given a functional $I : B_0(\Sigma) \to \mathbb{R}$, the* ***core of*** *$I$ is the set*

$$Core(I) \equiv \{m \in ba(\Sigma) : m(\xi) \geq I(\xi), \ \forall \xi \in B_0(\Sigma)\}.$$

The ***anti-core of*** *I is the set*

$$Eroc(I) \equiv \{m \in ba(\Sigma) : m(\xi) \leq I(\xi), \ \forall \xi \in B_0(\Sigma)\}.$$

[12] A capacity ρ is supermodular if $\rho(A \cup B) + \rho(A \cap B) \geq \rho(A) + \rho(B)$ for every $A, B \in \Sigma$.

As our choice of terminology suggests,[13] when I is a Choquet integral with respect to a capacity ρ, we have [18, Corollary 13] that

$$Core(I) = Core(\rho) \quad \text{and} \quad Eroc(I) = Eroc(\rho).$$

However, these notions apply also to preferences which are not CEU. Indeed, if $\succcurlyeq$ is a 1-MEU preference, then [18, Corollary 14] $Core(I) = \mathcal{D}$. Clearly, both $Core(I)$ and $Eroc(I)$ could be empty, and they are simultaneously nonempty if and only if I is linear.

The elements of $Core(I)$ (resp. $Eroc(I)$) correspond to SEU preferences $\geqslant$ which are less (resp. more) ambiguity averse than $\succcurlyeq$ in the sense of Ghirardato and Marinacci [18]: for all $f \in \mathcal{F}$ and $x \in X$, $x \geqslant f \Rightarrow x \succcurlyeq f$ (resp. $x \succcurlyeq f \Rightarrow x \geqslant f$). We now show that they correspond to some possible scenario.

Proposition 19 *Let I be a monotonic, constant linear functional. Then*

$$Core(I) \cup Eroc(I) \subseteq \partial I(0).$$

Moreover, $Core(I) = \partial I(0)$ if and only if *I is concave, while $Eroc(I) = \partial I(0)$* if and only if *I is convex.*

The second statement shows that $Core(I)$ contains *all* the possible scenarios if and only if I is concave; that is, $\succcurlyeq$ is a 1-MEU preference with set of priors $\mathcal{D} = Core(I)$. Hence, while Ghirardato and Marinacci's "benchmark measures" of $\succcurlyeq$ (the elements of $Core(I)$) are all possible scenarios, they exhaust the set $\mathcal{C}$ only when $\succcurlyeq$ is a 1-MEU preference.

In particular, the DM may not have any benchmark and yet be *quite* ambiguity averse, in the sense of having a uniformly high (but not constantly 1) ambiguity aversion coefficient. On the other hand, if he *does* have a benchmark measure, then he cannot be too ambiguity loving (say, have $a([f]) \leq 1/2$ for every $f \in \mathcal{F} \setminus [x]$, with strict inequality for one f), except in the trivial case in which he satisfies SEU.

6 Extensions

In this final section we sketch some extensions of our work. We refer the interested reader to [14] for a thorough discussion of these extensions, as well as for the proofs of the results reviewed here.

6.1 Perceived Ambiguity and Updating

First, we have a simple dynamic extension of our static decision problem. Suppose that we can observe our DM's *ex ante* preference on $\mathcal{F}$, denoted by $\succcurlyeq$, and his preference on $\mathcal{F}$ after having been informed that an event $A \in \Sigma$ obtained, denoted by $\succcurlyeq_A$. Further assume that

[13] In Ghirardato and Marinacci [18] these sets are denoted $\mathcal{D}(\succcurlyeq)$ and $\mathcal{E}(\succcurlyeq)$ respectively.

the preference $\succcurlyeq$ and $\succcurlyeq_A$ are invariant biseparable, and that A is not unambiguously null.[14] The utility representing $\succcurlyeq_A$ on X is denoted by u_A. Clearly, the conditional preference $\succcurlyeq_A$ also induces an unambiguous preference relation $\succcurlyeq_A^*$. Because $\succcurlyeq_A$ is invariant biseparable, it is possible to represent $\succcurlyeq_A^*$ in the sense of Proposition 5 by a nonempty, weak* compact and convex set of probability measures $\mathcal{C}_A$.

An important property linking *ex ante* and *ex post* preferences is **dynamic consistency**: for all $f, g \in \mathcal{F}$,

$$f\,A\,g \succcurlyeq g \Longleftrightarrow f \succcurlyeq_A g. \tag{11}$$

This property imposes two requirements. The first says that the DM should consistently carry out plans made *ex ante*. The second says that information is valuable to the DM, in the sense that postponing her choice to after knowing whether an event obtained does not make her worse off (see Ghirardato [12] for a more detailed discussion).

It is possible to find several plausible instances in which the presence of ambiguity explains behavior that violates dynamic consistency (see Siniscalchi [32] for elaboration). However, we think that *in the absence of ambiguity* dynamic consistency retains much of its intuitive appeal. It thus seems to be a natural exercise to inquire the effect of requiring dynamic consistency of the unambiguous preference relations $\succcurlyeq_A^*$, with respect to the *ex ante* $\succcurlyeq^*$ (that is, requiring Eq. (11) with $\succcurlyeq$ and $\succcurlyeq_A$ replaced by $\succcurlyeq^*$ and $\succcurlyeq_A^*$ respectively).

We show that dynamic consistency of the unambiguous preference relations is tantamount to assuming that the DM updates *all* the priors in $\mathcal{C}$, a well known procedure that we call **generalized Bayesian updating**: the "updated" perception of ambiguity is equal to

$$\mathcal{C}|A \equiv \overline{co}^{w^*}\{P_A : P \in \mathcal{C} \text{ such that } P\,(A) \neq 0\},$$

where P_A denotes the posterior of P conditional on A, and $\overline{co}^{w^*}$ stands for the weak* closure of the convex hull.

Proposition 20 *Let $\succcurlyeq$ and $\succcurlyeq_A$ be invariant biseparable preferences and suppose that A is not unambiguously null. Then the following statements are equivalent:*

(i) *For every $f, g \in \mathcal{F}$,*

$$f \succcurlyeq_A^* g \Longleftrightarrow P_A(u(f)) \geq P_A(u(g)) \textit{ for all } P \in \mathcal{C} \textit{ such that } P\,(A) \neq 0.$$

Equivalently, $\mathcal{C}_A = \mathcal{C}|A$ and $u_A = u$.

(ii) *The relations $\succcurlyeq^*$ and $\succcurlyeq_A^*$ are dynamically consistent.*

We thus have a further reason for the interest of the unambiguous preference relation: it yields a simple and intuitive foundation for the generalized Bayesian updating rule.

[14] It is not the case that $x\,A\,y \sim^* y$ for some (all) $x \not\sim y \in X$.

6.2 α-MEU preferences

As we observed just after Theorem 12, an interesting class of invariant biseparable preferences are those whose ambiguity aversion index a is constant, the α-MEU preferences. Here we show their behavioral characterization.

For any act $f \in \mathcal{F}$, we can denote by $C(f)$ the set of the certainty equivalents of f for $\succcurlyeq$. That is,

$$C(f) \equiv \{x \in X : \text{ for all } y \in X,\ y \succcurlyeq f \text{ implies } y \succcurlyeq x,\ f \succcurlyeq y \text{ implies } x \succcurlyeq y\}.$$

As $\succcurlyeq$ is complete, we have $C(f) = \{x \in X : x \sim f\}$. We analogously define the set $C^*(f)$ of the certainty equivalents of f for the relation $\succcurlyeq^*$:

$$C^*(f) \equiv \{x \in X : \text{ for all } y \in X,\ y \succcurlyeq^* f \text{ implies } y \succcurlyeq^* x,\ f \succcurlyeq^* y \text{ implies } x \succcurlyeq^* y\}.$$

Intuitively, these are the constants that correspond to "possible" certainty equivalents of f. (Recall that $x \succcurlyeq^* y$ if and only if $x \succcurlyeq y$.)

The following result provides the characterization of $C^*(f)$ in terms of the expected utilities mapping on $\mathcal{C}$:

Proposition 21 *For every $f \in \mathcal{F}$,*

$$x \in C^*(f) \Longleftrightarrow \min_{P\in\mathcal{C}} P(u(f)) \leq u(x) \leq \max_{P\in\mathcal{C}} P(u(f)).$$

Moreover, $u(C^*(f)) = [\min_{P\in\mathcal{C}} P(u(f)), \max_{P\in\mathcal{C}} P(u(f))]$.

Thus, $x \in C^*(f)$ if and only if there is a $P \in \mathcal{C}$ such that $u(x) = P(u(f))$. That is, $u(C^*(f))$ is the image of the expected utility mapping of f: the set of possible expected utilities of f as we range over the scenarios in $\mathcal{C}$.

We can now present the axiom that characterizes α-MEU preferences.

Axiom 7 *For every $f, g \in \mathcal{F}$, $C^*(f) = C^*(g)$ implies $f \sim g$.*

The interpretation of the axiom is straightforward. For a DM who satisfies axiom 7, the set of certainty equivalents of f with respect to $\succcurlyeq^*$ contains *all* the information the DM uses in evaluating f. Notice that the condition $C^*(f) = C^*(g)$ in the axiom could also be rewritten as follows: for every $x \in X$, $f \succcurlyeq^* x$ if and only if $g \succcurlyeq^* x$, and $x \succcurlyeq^* f$ if and only if $x \succcurlyeq^* g$.

In terms of the representation in Eq. (7), axiom 7 clearly guarantees that the DM's evaluation $I(u(f))$ of act f depends only on the range $[\min_{P\in\mathcal{C}} P(u(f)), \max_{P\in\mathcal{C}} P(u(f))]$ of the expected utility mapping $\{P(u(f)) : P \in \mathcal{C}\}$, rather than on the expected utility mapping itself. More surprisingly, such dependence must be linear.

Proposition 22 *Let $\succcurlyeq$ be a binary relation on $\mathcal{F}$. The following statements are equivalent:*

(i) $\succcurlyeq$ *satisfies axioms 1–5 and 7.*

(*ii*) *There exist a nonempty, weak* compact and convex set $\mathcal{C}$ of probabilities on Σ, a nonconstant affine function $u : X \to \mathbb{R}$ and $\alpha \in [0,1]$ such that $\succcurlyeq$ is represented by the monotonic preference functional $I : B_0(\Sigma) \to \mathbb{R}$ defined by*

$$I(u(f)) = \alpha \min_{P\in\mathcal{C}} \int_S u(f)\, dP + (1-\alpha) \max_{P\in\mathcal{C}} \int_S u(f)\, dP,$$

and u and $\mathcal{C}$ represent $\succcurlyeq^*$ in the sense of Eq. (4).

Moreover, $\mathcal{C}$ is unique, u is unique up to a positive affine transformation, and α is unique if $\mathcal{C}$ is not a singleton.

6.3 Ambiguity of Events and Acts

We have earlier introduced crisp acts, those whose evaluation is unaffected by the ambiguity the DM perceives in a problem. Consider in particular crisp *bets*; i.e., acts of the form $x\,A\,y$ for $x \succ y$. We propose that the event corresponding to a crisp bet be defined unambiguous (Nehring [27] gives an equivalent definition and shows that the latter is in turn equivalent to one he earlier studied in [26]):

Definition 23 *An event $A \in \Sigma$ is **unambiguous** if for some $x \succ y$, the act $x\,A\,y$ is crisp. The collection of all the unambiguous events is denoted by Λ.*

The next result shows that unambiguous events have a simple and intuitive characterization in terms of the probabilities in $\mathcal{C}$, and that if $x\,A\,y$ is crisp for some $x \succ y$ then $x'\,A\,y'$ is crisp for every $x' \not\sim y'$. This conforms with our intuition that ambiguity is property of *events* (more accurately, event partitions), not acts.

Proposition 24 *For any $A \in \Sigma$, the following statements are equivalent:*

(*i*) *A is unambiguous.*

(*ii*) *$P(A) = Q(A)$ for all $P, Q \in \mathcal{C}$.*

(*iii*) *For* every *$x \not\sim y$, the act $x\,A\,y$ is crisp.*

For all invariant biseparable preferences, the collection Λ is a **(finite) λ-system** (cf. Zhang [35] and Nehring [26]): 1) $S \in \Lambda$; 2) if $A \in \Lambda$ then $A^c \in \Lambda$; 3) if $A, B \in \Lambda$ and $A \cap B = \emptyset$ then $A \cup B \in \Lambda$. Nehring [26] offers further results on the structure of Λ in the CEU case.

Ghirardato and Marinacci [18] propose a behavioral notion of unambiguous event for a subclass of invariant biseparable preferences, showing that it has a simple characterization in terms of the capacity ρ defined just before Eq. (3): an event B is unambiguous in their sense if and only if $\rho(B) + \rho(B^c) = 1$. It can be shown that in general the set Λ is a subset of the set of events satisfying this condition, but for a large class of invariant biseparable preferences (e.g., those having $a(f) > 1/2$ for all $f \in \mathcal{F}$) the two sets coincide.

In view of the fact that Λ is a λ-system, it is natural to call "unambiguous" the acts whose upper level sets are unambiguous events (cf., e.g., Epstein and Zhang [10]).

Definition 25 *Act $f \in \mathcal{F}$ is **unambiguous** if its upper sets $\{s \in S : f(s) \succcurlyeq x\}$ belong to Λ for all $x \in X$.*

An obvious question to ask at this point is whether crisp acts are unambiguous. It follows from Proposition 24 (*iii*) that every *binary* crisp act is unambiguous. However, it can be seen that this fact does not generalize to acts which yield more than two nonindifferent prizes. The issue is that permuting the prizes of a crisp (nonbinary) act may produce a noncrisp act. This conflicts with our intuition that ambiguity is a property of the event partition with respect to which an act is measurable.

Indeed, it turns out that the unambiguous acts are basically those acts whose crispness is not affected by permuting payoffs.

Proposition 26 *Let $f = \{x_i; A_i\}_{i=1}^n$, with $x_1, x_2, ..., x_n \in X$ and $\{A_1, A_2, ..., A_n\}$ a partition of S in Σ. If for each permutation σ of $\{1, 2, \dots, n\}$ the act $f^\sigma = \{x_{\sigma(i)}; A_i\}_{i=1}^n$ is crisp, then f is unambiguous. The converse is true whenever $x_i \nsim x_j$ for every $i \neq j$ in $\{1, \dots, n\}$.*

Notice that a binary act $x\,B\,y$ satisfies the permutation crispness condition if and only if B is unambiguous. The non-indifference condition in the converse is the reason of the qualifier "basically" above. However, it is still possible to use the proposition to obtain a full characterization of the relation between crisp and unambiguous acts. In fact, it can be shown that for any unambiguous f with some indifferent payoffs, there exists an act f' with non-indifferent payoffs which is state-wise indifferent to f.

6.4 More on the Operational Consequences

In this subsection we show how to extend the results of Subsection 5.1 when S is infinite.

Let S be a compact metric space, Σ its Borel σ-algebra, $B(\Sigma)$ the set of all bounded Σ-measurable functions, and $C(S)$ the set of all continuous functions on S. Denote by $I_{|C}$ the restriction of I to $C(S)$. Let $\mathcal{F}_{\succcurlyeq}$ be the set all Σ-measurable and $\succcurlyeq$-bounded functions from S into X; that is, $f \in \mathcal{F}_{\succcurlyeq}$ if both $\{s \in S : f(s) \succcurlyeq x\}$ and $\{s \in S : f(s) \succ x\}$ belong to Σ for every $x \in X$, and if there exist $x, y \in X$ such that $x \succcurlyeq f(s) \succcurlyeq y$ for all $s \in S$. Notice that the finite state space case considered in Subsection 5.1 is a special case of this setting. (The finite S is compact when endowed with its discrete topology; moreover, $\mathcal{F} = \mathcal{F}_{\succcurlyeq}$ and $C(S) = B_0(\Sigma) = B(\Sigma) = \mathbb{R}^n$.)

Following Gilboa and Schmeidler [19], it is possible to show that if $\succcurlyeq$ satisfies axioms 1–5 on $\mathcal{F}$, it has a *unique* extension which satisfies axioms 1–5 on $\mathcal{F}_{\succcurlyeq}$. Such relation, also denoted by $\succcurlyeq$, is represented by the unique extension $\hat{I}$ of I to $B(\Sigma)$, and the associated subset $\mathcal{C}_{\succcurlyeq}$ satisfies $\mathcal{C}_{\succcurlyeq} = \mathcal{C} = \partial\hat{I}(0)$. For this reason, we abuse notation and write I and $\mathcal{C}$ instead of $\hat{I}$ and $\mathcal{C}_{\succcurlyeq}$.

Next, we impose a monotone continuity assumption (see Marinacci, Maccheroni, Chateauneuf and Tallon [24]), where $\succ^*$ denotes the asymmetric component of $\succcurlyeq^*$.

Axiom 8 (Monotone Continuity) *For all $x, y, z \in X$, if $A_n \downarrow \emptyset$ and $y \succ^* z$, then eventually $y \succ^* x\,A_n\,z$.*

We can now state our result, based on a "D-representation" of the Clarke differential due to Thibault [34] (see also [3]).

Proposition 27 *Assume that $\succcurlyeq$ satisfies axioms 1–5 and 8. Then,*

$$\mathcal{C} = \partial I(0) = \overline{co}^{w^*}\{\nabla I_{|C}(\varphi) : \varphi \in D\}, \tag{12}$$

where $\nabla I_{|C}$ is the Gateaux derivative of $I_{|C}$, and $D \subseteq C(S)$ is the domain of Gateaux differentiability of $I_{|C}$.

According to this result, the set $\mathcal{C}$ can be described by just taking the closed convex hull of the collection of the Gateaux derivatives of the restriction $I_{|C}$. In this regard, it is important to observe that $\nabla I_{|C}$ is more likely to exist (and simpler to compute) than ∇I, the Gateaux derivative of I on the *entire* space $B(\Sigma)$. Therefore, the fact that only the derivatives $\nabla I_{|C}$ appear in Eq. (12) is a positive feature of Proposition 27.

Similarly to what we did for the finite setting, we can illustrate our result by considering piecewise linear functionals. Here, we say that I is **properly piecewise linear** if there exists a family $\{C_l\}_{l\in L}$ of convex cones such that:

- $B(\Sigma) = \bigcup_l C_l$,
- $\mathrm{int}\, C_l \cap C(S) \neq \emptyset \quad$ for each l,
- $\mathrm{int}\, C_l \cap \mathrm{int}\, C_h \cap C(S) = \emptyset \quad$ if $l \neq h$,
- I is linear on each C_l.

Interiors are taken in the supnorm topology. We remark that in the finite case this definition reduces to that introduced in Subsection 5.1, and similarly refer the reader to Castagnoli *et al.* [5] and Siniscalchi [33] for decision models with properly piecewise linear representations.

Given a preference $\succcurlyeq$ represented by a properly piecewise linear I, Eq. (12) gives a representation of $\partial I(0)$ which is identical to the one obtained in Eq. (9) for the finite case. In fact, we have

$$\partial I(0) = \overline{co}^{w^*}\{P_l : l \in L\},$$

where P_l is the probability representing I on the cone C_l. Therefore, we again find that the sets of priors obtained in the representation by Castagnoli *et al.* [5] and Siniscalchi [33] are, up to convex closure, equal to our $\mathcal{C}$.

Appendices

A Functional Analysis Mini-Kit

In this appendix we provide/review some functional analytic results and notions that are used to prove the results in the main text (and in some cases directly mentioned in Section 5). Most of the proofs are standard, and are thus omitted.[15]

A.1 Conic Preorders

We recall that $B_0(\Sigma)$ is the vector space generated by the indicator functions of the elements of Σ. We denote by $ba(\Sigma)$ the set of the bounded, finitely additive set functions on Σ, and by $pc(\Sigma)$ the set of the probability charges on Σ. As it is well known, $ba(\Sigma)$, endowed with the total variation norm, is isometrically isomorphic to the norm dual of $B_0(\Sigma)$.

Given a non singleton interval K in the real line (whose interior is denoted K°) we denote by $B_0(\Sigma, K)$ the subset of the functions in $B_0(\Sigma)$ taking values in K. Clearly, $B_0(\Sigma) = B_0(\Sigma, \mathbb{R})$.

We recall that a binary relation $\succsim$ on $B_0(\Sigma, K)$ is:

- a **preorder** if it is reflexive and transitive;
- **continuous** if $\varphi^n \succsim \psi^n$ for all $n \in \mathbb{N}$, $\varphi^n \to \varphi$ and $\psi^n \to \psi$ imply $\varphi \succsim \psi$;
- **conic** if $\varphi \succsim \psi$ implies $\alpha\,\varphi + (1-\alpha)\,\theta \succsim \alpha\,\psi + (1-\alpha)\,\theta$ for all $\theta \in B_0(\Sigma, K)$ and all $\alpha \in [0,1]$;[16]
- **monotonic** if $\varphi \geq \psi$ implies $\varphi \succsim \psi$.
- **nontrivial** if there exists $\varphi, \psi \in B_0(\Sigma, K)$ such that $\varphi \succsim \psi$ but not $\psi \succsim \varphi$.

Next, we have some useful representation results.

Proposition 28 *For $i = 1, 2$, let $\mathcal{C}_i$ be nonempty sets of probability charges on Σ and $\succsim_i$ be the relations defined on $B_0(\Sigma, K)$ by*

$$\varphi \succsim_i \psi \iff \int_S \varphi \, dP \geq \int_S \psi \, dP \quad \text{for all } P \in \mathcal{C}_i.$$

Then

$$\varphi \succsim_i \psi \iff \int_S \varphi \, dP \geq \int_S \psi \, dP \quad \text{for all } P \in \overline{co}^{w^*}(\mathcal{C}_i),$$

and the following statements are equivalent:

(i) $\varphi \succsim_1 \psi \Longrightarrow \varphi \succsim_2 \psi$ *for all φ and ψ in $B_0(\Sigma, K)$.*

[15] They are available from the authors upon request.

[16] Notice that if $K = \mathbb{R}$ or $\mathbb{R}_+$ and $\succsim$ is a preorder, then $\succsim$ is conic iff $\varphi \succsim \psi$ implies $\alpha\varphi + \theta \succsim \alpha\psi + \theta$ for all $\theta \in B_0(\Sigma)$ and all $\alpha \in \mathbb{R}_+$.

(ii) $\overline{co}^{w^*}(\mathcal{C}_2) \subseteq \overline{co}^{w^*}(\mathcal{C}_1)$.

(iii) $[\inf_{P\in\mathcal{C}_2} P(\varphi), \sup_{P\in\mathcal{C}_2} P(\varphi)] \subseteq [\inf_{P\in\mathcal{C}_1} P(\varphi), \sup_{P\in\mathcal{C}_1} P(\varphi)]$ *for all* $\varphi \in B_0(\Sigma, K)$.

Proposition 29 $\succsim$ *is a nontrivial, continuous, conic, and monotonic preorder on* $B_0(\Sigma, K)$ *if and only if there exists a nonempty subset* $\mathcal{C}$ *of* $pc(\Sigma)$ *such that*

$$\varphi \succsim \psi \Longleftrightarrow \int_S \varphi \, dP \geq \int_S \psi \, dP \quad \text{for all } P \in \mathcal{C}. \tag{13}$$

Moreover, $\overline{co}^{w^*}(\mathcal{C})$ *is the unique weak* closed and convex subset of* $pc(\Sigma)$ *representing* $\succsim$ *in the sense of Eq. (13).*

A.2 Clarke Derivatives and Differentials

A monotonic constant linear functional $I : B_0(\Sigma) \to \mathbb{R}$ is Lipschitz of rank 1. For, given $\varphi, \psi \in B_0(\Sigma)$, $\varphi \leq \psi + \|\varphi - \psi\|$ implies $I(\varphi) \leq I(\psi) + \|\varphi - \psi\|$, hence $I(\varphi) - I(\psi) \leq \|\varphi - \psi\|$; switching φ and ψ yields $|I(\varphi) - I(\psi)| \leq \|\varphi - \psi\|$. It follows that I is also uniformly continuous.

Thus, given a monotonic constant linear functional $I : B_0(\Sigma) \to \mathbb{R}$, we can study its Clarke derivatives and Clarke differentials (as defined in Clarke [6]):

Definition 30 *The* ***Clarke (upper) directional derivative*** *of* I *in* φ *in the direction* υ *is*

$$I^\circ(\varphi; \upsilon) = \limsup_{\substack{\psi \to \varphi \\ t \downarrow 0}} \frac{I(\psi + t\upsilon) - I(\psi)}{t}.$$

The ***Clarke differential*** *of* I *in* φ *is the set*

$$\partial I(\varphi) = \{m \in ba(\Sigma) : m(\upsilon) \leq I^\circ(\varphi; \upsilon), \ \forall \upsilon \in B_0(\Sigma)\}.$$

We refer to Clarke [6] for properties of the Clarke derivative and differential. Among them, the following are especially important:

1. For every $\upsilon \in B_0(\Sigma)$:
$$I^\circ(\varphi; \upsilon) = \max_{m \in \partial I(\varphi)} m(\upsilon).$$

2. (Mean Value Theorem) For all $\varphi, \psi \in B_0(\Sigma)$, there exist $\gamma \in (0, 1)$ and $m \in \partial I(\gamma\varphi + (1 - \gamma)\psi)$ such that
$$I(\varphi) - I(\psi) = m(\varphi - \psi).$$

In Section 5 we defined a Clarke lower derivative, and defined the Clarke differential in terms of that. The observation that for every $\varphi, \upsilon \in B_0(\Sigma)$

$$I_\circ(\varphi; \upsilon) = -I^\circ(\varphi; -\upsilon) = \min_{m \in \partial I(\varphi)} m(\upsilon).$$

shows that our presentational choice does not make a difference since the set defined in Section 5 coincides with the Clarke differential as defined above. For easier reference to the existing literature in the rest of this subsection we mostly use the traditional I°.

The following are additional properties of I° and $\partial I(\cdot)$ that we use below, and hold for a monotonic and constant linear I.

Proposition 31 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a Lipschitz functional. Then:*

1. *If I is positively homogenous, $I^{\circ}(\varphi;\cdot) = I^{\circ}(\alpha\varphi;\cdot)$ for all $\alpha > 0$, and $\partial I(\varphi) \subseteq \partial I(0)$ for all $\varphi \in B_0(\Sigma)$. Moreover,*

$$I^{\circ}(0;\varphi) = \sup_{\psi \in B_0(\Sigma)} I(\psi+\varphi) - I(\psi) \quad \text{and} \quad I_{\circ}(0;\varphi) = \inf_{\psi \in B_0(\Sigma)} I(\psi+\varphi) - I(\psi)$$

   for all $\varphi \in B_0(\Sigma)$.

2. *If I is monotone, then for all $\varphi \in B_0(\Sigma)$ the function $I^{\circ}(\varphi;\cdot)$ is monotone, and m is positive for all $m \in \partial I(\varphi)$.*

3. *If I is constant additive, then for all $\varphi \in B_0(\Sigma)$ the function $I^{\circ}(\varphi;\cdot)$ is constant linear, and $m(S) = 1$ for all $m \in \partial I(\varphi)$.*

Notice that it follows from this proposition that if I is monotonic and constant linear, then $\partial I(\varphi) \subseteq \partial I(0) \subseteq pc(\Sigma)$ for all $\varphi \in B_0(\Sigma)$. That is, the Clarke differential only contains probabilities.

We conclude this appendix by recalling Clarke's characterization of Clarke differentials in finite dimensional spaces, and drawing a consequence for positively homogeneous functionals. Here, $B_0(\Sigma) = \mathbb{R}^n$ and, as usual, ∇ denotes a standard gradient.

Theorem 32 *Let $I : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz functional, and let Ω be any subset of $\mathbb{R}^n$ such that I is differentiable on Ω and Ω^c has Lebesgue measure 0. Then*

$$\partial I(\varphi) = co\left\{\lim_{i\to\infty} \nabla I(\varphi_i) : \varphi_i \in \Omega,\ \varphi_i \to \varphi,\ \text{and } \nabla I(\varphi_i) \text{ converges}\right\}. \tag{14}$$

If, in addition, I is positively homogeneous, then

$$\partial I(0) = \overline{co}\left\{\nabla I(\varphi) : \varphi \in \Omega\right\}. \tag{15}$$

Proof. The first statement is a classical result (see, e.g., Theorem 2.5.1 of [6]). As to the second, suppose that I is differentiable at φ. Then $\nabla I(\varphi) \in \partial I(\varphi)$ and, by positive homogeneity, $\partial I(\varphi) \subseteq \partial I(0)$. This proves the inclusion $\supseteq$ in Eq. (15).

As to other inclusion, by Eq. (14) we have

$$\partial I\left(0\right) = co\left\{\lim_{i\to\infty} \nabla I(\varphi_i) : \varphi_i \in \Omega,\ \varphi_i \to 0,\ \text{and } \nabla I(\varphi_i) \text{ converges}\right\}.$$

But, for all $\varphi_i \in \Omega$ such that $\varphi_i \to 0$, we have

$$\lim_{i\to\infty} \nabla I(\varphi_i) \in \overline{\{\nabla I(\varphi) : \varphi \in \Omega\}} \subseteq \overline{co}\,\{\nabla I(\varphi) : \varphi \in \Omega\},$$

as desired. □

This result allows us to prove Eq. (9) in Section 5 for a piecewise linear I. Let m_l be the unique linear extension of $I_{|C_l}$ to $\mathbb{R}^n$; by (15), we have

$$\partial I\,(0) = \overline{co}\,\{m_l : l \in L\}.$$

In fact, if $\varphi \in \operatorname{int} C_l$ for some l, then $\nabla I(\varphi) = m_l$. Set $\Omega = \bigcup_l \operatorname{int} C_l$. Then $\Omega^c \subseteq \bigcup_l \partial C_l$. Notice that L is at most countable – $\{\operatorname{int} C_l\}$ is a disjoint family of nonempty open sets – and so Ω^c has zero Lebesgue measure since each ∂C_l does.

B Proofs of the Results in the Main Text

We begin with two preliminary remarks and a piece of notation, that are used throughout this appendix. First, given the representation in Lemma 1, we observe without proof that

$$\{u(f) : f \in \mathcal{F}\} \equiv \{\varphi \in B_0(\Sigma, \mathbb{R}) : \varphi = u(f),\ \text{for some } f \in \mathcal{F}\} = B_0(\Sigma, u(X)).$$

Second, notice that it is w.l.o.g. to assume that $u(X) \supseteq [-1, 1]$. Finally, given a nonempty, convex and weak* compact set $\mathcal{C}$ of probability charges on (S, Σ), we denote for every $\varphi \in B_0(\Sigma)$,

$$\underline{\mathcal{C}}(\varphi) = \min_{P\in\mathcal{C}} P(\varphi), \qquad \overline{\mathcal{C}}(\varphi) = \max_{P\in\mathcal{C}} P(\varphi).$$

B.1 Proof of Proposition 4

Taking $\lambda = 1$ in the definition proves point 1. Next we prove that $\succcurlyeq^*$ is monotonic (point 4). Suppose that $f(s) \succcurlyeq g(s)$ for all $s \in S$. By axiom 2, for every $h \in \mathcal{F}$ and $\lambda \in (0, 1]$, $\lambda f(s) + (1 - \lambda)h(s) \succcurlyeq \lambda g(s) + (1 - \lambda)h(s)$ for all $s \in S$. Using axiom 4, we thus obtain that $\lambda f + (1 - \lambda)h \succcurlyeq \lambda g + (1 - \lambda)h$. This shows that $f \succcurlyeq^* g$. If $x \succcurlyeq y$, then the monotonicity of $\succcurlyeq^*$ yields $x \succcurlyeq^* y$. Along with point 1, this proves point 2. As to point 3, reflexivity also follows from monotonicity. To show transitivity, suppose that $f \succcurlyeq^* g$ and $g \succcurlyeq^* h$. Then for all $k \in \mathcal{F}$ and all $\lambda \in (0, 1]$, we have

$$\lambda f + (1 - \lambda)k \succcurlyeq \lambda g + (1 - \lambda)k \succcurlyeq \lambda h + (1 - \lambda)k.$$

This shows that $f \succcurlyeq^* h$.

Next, we prove the implication $\Longrightarrow$ of point 5 (The other implication follows immediately from the following Proposition 5, and it is not used in the proof of that proposition). Given $f, g, h \in \mathcal{F}$ and $\lambda \in (0, 1)$, suppose that $f \succcurlyeq^* g$. Then for every $\mu \in (0, 1]$ and every $k \in \mathcal{F}$, we have

$$(\lambda\mu)f + (1 - \lambda\mu)\left[\frac{(1-\lambda)\mu}{1-\lambda\mu}h + \frac{1-\mu}{1-\lambda\mu}k\right] \succcurlyeq (\lambda\mu)g + (1 - \lambda\mu)\left[\frac{(1-\lambda)\mu}{1-\lambda\mu}h + \frac{1-\mu}{1-\lambda\mu}k\right]$$

by definition of $\succcurlyeq^*$. Rearranging terms, we find

$$\mu(\lambda f + (1-\lambda)h) + (1-\mu)k \succcurlyeq \mu(\lambda g + (1-\lambda)h) + (1-\mu)k,$$

which implies $\lambda f + (1-\lambda)h \succcurlyeq^* \lambda g + (1-\lambda)h$, since the choice of μ and k was arbitrary. The case $\lambda = 1$ is trivial. Point 6 follows immediately from the following Proposition 5. (It is not used in the proof of that proposition.)

Finally, assume that $\succcurlyeq^{**}$ is an independent binary relation such that $f \succcurlyeq^{**} g$ implies $f \succcurlyeq g$. Then $f \succcurlyeq^{**} g$ implies $\lambda f + (1-\lambda)h \succcurlyeq^{**} \lambda g + (1-\lambda)h$ for all $h \in \mathcal{F}$ and $\lambda \in (0,1]$, hence $\lambda f + (1-\lambda)h \succcurlyeq \lambda g + (1-\lambda)h$ for all $h \in \mathcal{F}$ and $\lambda \in (0,1]$, finally $f \succcurlyeq^* g$. This proves 7.

B.2 Proof of Proposition 5

Notice that $f \succcurlyeq^* g$ iff $I(\lambda u(f) + (1-\lambda)u(h)) \geq I(\lambda u(g) + (1-\lambda)u(h))$ for all $h \in \mathcal{F}$ and all $\lambda \in (0,1]$. Define $\gtrsim$ on $B_0(\Sigma, u(X))$ by setting

$$\varphi \gtrsim \psi \Longleftrightarrow I(\lambda\varphi + (1-\lambda)\theta) \geq I(\lambda\psi + (1-\lambda)\theta),\ \forall\theta \in B_0(\Sigma, u(X)),\ \forall\lambda \in (0,1].$$

Clearly, $f \succcurlyeq^* g$ iff $u(f) \gtrsim u(g)$. It is routine to show, either using the properties of $\succcurlyeq^*$ or those of I, that $\gtrsim$ is a nontrivial, monotonic and conic preorder on $B_0(\Sigma, u(X))$. Moreover, if $\varphi_1^n \gtrsim \varphi_2^n$ for all $n \in \mathbb{N}$, $\varphi_1^n \to \varphi_1$, $\varphi_2^n \to \varphi_2$, then $I(\lambda\varphi_1^n + (1-\lambda)\theta) \geq I(\lambda\varphi_2^n + (1-\lambda)\theta)$, for all $\lambda \in (0,1]$, all $\theta \in B_0(\Sigma, u(X))$, and all $n \in \mathbb{N}$. Since I is supnorm continuous, it follows that $\varphi_1 \gtrsim \varphi_2$.

We have thus shown that $\gtrsim$ is a conic, continuous, monotonic, nontrivial preorder on $B_0(\Sigma, u(X))$. By Lemma 29 it follows that there exists a nonempty, weak* closed and convex set $\mathcal{C}$ of probability charges on Σ such that

$$\varphi \gtrsim \psi \Longleftrightarrow \int_S \varphi\, dP \geq \int_S \psi\, dP \quad \text{for all } P \in \mathcal{C},$$

which immediately yields the statement.

B.3 Proof of Proposition 6

Lemma 33 *Let Y be a vector space and u, v be two nonzero linear functionals on Y. One and only one of the following statements is true:*

- *$u = av$ for some $a > 0$.*
- *$\exists y \in Y : u(y)\, v(y) < 0$.*

Proof. Clearly the two statements cannot be both true. Assume, by contradiction that both are false. That is: there exist u, v nonzero linear functionals on Y such that $u \neq av$ for all $a > 0$, and $u(y)\, v(y) \geq 0$ for all $y \in Y$.

Then $Y = [uv > 0] \cup [u = 0] \cup [v = 0] = [uv > 0] \cup \ker u \cup \ker v$. $\ker u$ and $\ker v$ are maximal subspaces of Y, hence $Y = \langle z \rangle \oplus \ker u$ for some $z \in Y$ such that $u(z) > 0$. If

$\ker u = \ker v$: for all $y \in Y$, exist $b \in \mathbb{R}$, $x \in \ker u$ such that $y = bz + x$, whence $u(y) = b\,u(z) = \frac{u(z)}{v(z)} b\, v(z) = \frac{u(z)}{v(z)} v(y)$, which is absurd. Else: $\ker u \neq \ker v$, so there exist $y' \in \ker u \backslash \ker v$ and $y'' \in \ker v \backslash \ker u$ ($\ker u$ and $\ker v$ are maximal subspaces), we can choose y' and y'' such that $v(y') > 0$ and $u(y'') < 0$. Finally, $u(y' + y'')v(y' + y'') = u(y'')v(y') < 0$, which is absurd. $\square$

Corollary 34 *Let X be a nonempty convex subset of a vector space and u, v be two nonconstant affine functionals on X. There exist $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$ such that $u = av + b$ iff $u(x_1) \geq u(x_2) \Longrightarrow v(x_1) \geq v(x_2)$ for every $x_1, x_2 \in X$.*

Proof. Necessity being trivial, we only prove sufficiency. Notice that

$$Y = \{t(x_1 - x_2) : t \in \mathbb{R}_{++},\, x_1, x_2 \in X\}$$

is a vector space and the functionals

$$\begin{aligned} \hat{u} &: t(x_1 - x_2) \mapsto t(u(x_1) - u(x_2)), \\ \hat{v} &: t(x_1 - x_2) \mapsto t(v(x_1) - v(x_2)) \end{aligned}$$

are well defined, nonzero, and linear on Y. Moreover,

$$\hat{u}(t(x_1 - x_2)) \geq 0 \Longrightarrow u(x_1) \geq u(x_2) \Longrightarrow v(x_1) \geq v(x_2) \Longrightarrow v(t(x_1 - x_2)) \geq 0.$$

Therefore $\nexists y \in Y$ such that $\hat{u}(y)\,\hat{v}(y) < 0$. By the previous lemma, there exists $a > 0$ such that $\hat{u} = a\hat{v}$. Finally, fix $x^\circ \in X$, for all $x \in X$

$$u(x) - u(x^\circ) = \hat{u}(1(x - x^\circ)) = a\hat{v}(1(x - x^\circ)) = av(x) - av(x^\circ)$$

so

$$u(x) = av(x) + [u(x^\circ) - av(x^\circ)],$$

set $b = [u(x^\circ) - av(x^\circ)]$. $\square$

Proof of Proposition 6.

$(i) \Rightarrow (ii)$: For all $x, y \in X$,

$$u_1(x) \geq u_1(y) \Longleftrightarrow x \succcurlyeq_1 y \Longrightarrow x \succcurlyeq_1^* y \Longrightarrow x \succcurlyeq_2^* y \Longrightarrow x \succcurlyeq_2 y \Longleftrightarrow u_2(x) \geq u_2(y).$$

By Corollary 34, this implies that we can assume $u_1 = u_2 = u$. Moreover, for all $f, g \in \mathcal{F}$, $f \succcurlyeq_1^* g \Longrightarrow f \succcurlyeq_2^* g$. That is,

$$P(u(f)) \geq P(u(g)) \quad \forall P \in \mathcal{C}_1 \Longrightarrow P(u(f)) \geq P(u(g)) \quad \forall P \in \mathcal{C}_2,$$

which by Lemma 28 (applied to $B_0(\Sigma, u(X))$) implies $\mathcal{C}_2 \subseteq \mathcal{C}_1$.

$(ii) \Rightarrow (i)$: Obvious.

B.4 Proof of Proposition 7

The fact that $(iii) \Rightarrow (iv)$ follows from the observation that for all $f \in \mathcal{F}$, if $c(f) \in C(f)$, $I(u(f)) = u(c(f)) \in u(N(f)) = \{P(u(f)) : P \in \mathcal{C}\}$. On the other hand, if $\succcurlyeq$ has a SEU representation with probability P (statement (iv)), then $\succcurlyeq$ satisfies independence (statement (ii)), which implies that $f \succcurlyeq g$ implies $f \succcurlyeq^* g$ for all $f, g \in \mathcal{F}$, so that $\succcurlyeq = \succcurlyeq^*$ (statement (i)). By the uniqueness of the representation in Eq. (4), it follows that $\mathcal{C} = \{P\}$ (statement (iii)), closing the chain.

B.5 Proof of Proposition 8

The result follows immediately (take $\psi \equiv 0$) from the following lemma, that will be of further use.

Lemma 35 *For all $f \in \mathcal{F}$,*

$$\begin{aligned}\underline{\mathcal{C}}(u(f)) &= \inf_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)\right\} \\ &= \inf_{\psi\in B_0(\Sigma)} \{I(u(f)+\psi) - I(\psi)\}\end{aligned}$$

and

$$\begin{aligned}\overline{\mathcal{C}}(u(f)) &= \sup_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u\left(g\right)\right)\right\} \\ &= \sup_{\psi\in B_0(\Sigma)} \{I(u(f)+\psi) - I(\psi)\}.\end{aligned}$$

Proof. Clearly $\{\frac{(1-\lambda)}{\lambda}u(g) : g \in \mathcal{F},\, \lambda \in (0,1]\} \subseteq B_0(\Sigma)$. Conversely, for all $\psi \in B_0(\Sigma)$ there exists $\alpha \in (0,1)$ and $g \in \mathcal{F}$ such that $\alpha\psi = u(g)$ hence $\psi = \frac{1}{\alpha}u(g)$. Since $\frac{(1-\lambda)}{\lambda}$ ranges from 0 to ∞ (recall that $\lambda \in (0,1]$), there exists λ^* such that $\frac{1}{\alpha} = \frac{(1-\lambda^*)}{\lambda^*}$ and $\psi = \frac{(1-\lambda^*)}{\lambda^*}u(g)$. We have thus proved the second equality in both equations.

Given $x_{\min} \in X$ that satisfies $u(x_{\min}) = \underline{\mathcal{C}}(u(f))$, we have $f \succcurlyeq^* x_{\min}$. That is, for all $g \in \mathcal{F}$ and $\lambda \in (0,1]$:

$$I(u(\lambda x_{\min} + (1-\lambda)g)) \leq I(u(\lambda f + (1-\lambda)g))$$

or

$$I(\lambda u(x_{\min}) + (1-\lambda)u(g)) \leq I(\lambda u(f) + (1-\lambda)u(g)).$$

Therefore,

$$\lambda u(x_{\min}) + I((1-\lambda)u(g)) \leq I(\lambda u(f) + (1-\lambda)u(g))$$

from which we obtain

$$u(x_{\min}) \leq I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right).$$

Finally,

$$\underline{\mathcal{C}}(u(f)) \leq \inf_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)\right\}.$$

Analogously,

$$\sup_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)\right\} \leq \overline{\mathcal{C}}(u(f)).$$

Conversely, let $x_{\text{inf}} \in X$ be such that

$$u(x_{\text{inf}}) = \inf_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)\right\}.$$

Then,

$$u(x_{\text{inf}}) \leq I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)$$

for all $g \in \mathcal{F}$ and $\lambda \in (0,1]$, whence $f \succcurlyeq^* x_{\text{inf}}$. That is, $u(x_{\text{inf}}) \leq \underline{\mathcal{C}}(u(f))$, or

$$\inf_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)\right\} \leq \min_{P\in\mathcal{C}} P(u(f)).$$

Analogously,

$$\sup_{\substack{g\in\mathcal{F}\\ \lambda\in(0,1]}} \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right)\right\} \geq \max_{P\in\mathcal{C}} P(u(f)),$$

which concludes the proof.

B.6 Proof of Lemma 9

$(i) \Rightarrow (ii)$: Suppose that for some λ, λ' and $x, x' \in X$,

$$\lambda\, f + (1-\lambda)\, x \sim^* \lambda'\, g + (1-\lambda')\, x',$$

which, applying Eq. (6) of Proposition 5, is equivalent to

$$\lambda\, P(u(f)) + (1-\lambda)\, u(x) = \lambda'\, P(u(g)) + (1-\lambda')\, u(x') \quad \text{for all } P \in \mathcal{C}.$$

It follows that for all $P \in \mathcal{C}$,

$$P(u(f)) = \frac{\lambda'}{\lambda}P(u(g)) + \frac{1}{\lambda}[(1-\lambda')u(x') - (1-\lambda)u(x)],$$

so that we get the conclusion by letting

$$\alpha = \frac{\lambda'}{\lambda} \quad \text{and} \quad \beta = \frac{1}{\lambda}[(1-\lambda')u(x') - (1-\lambda)u(x)].$$

$(ii) \Rightarrow (ii)$: Suppose that

$$P(u(f)) = \alpha P(u(g)) + \beta \quad \text{for all } P \in \mathcal{C}.$$

Suppose first that $\alpha < 1$. Then, let $\lambda = \alpha$. By renormalizing the utility function if necessary, we can assume that $\beta/(1-\lambda) \in u(X)$, so that there is $x \in X$ for which $u(x) = \beta/(1-\lambda)$. It follows that

$$f \sim^* \lambda\, g + (1-\lambda)\, x.$$

The case of $\alpha > 1$ is dealt with by rewriting the equation as follows:

$$P(u(g)) = \frac{1}{\alpha} P(u(f)) - \beta \quad \text{for all } P \in \mathcal{C},$$

and proceeding as above to get

$$\lambda\, f + (1-\lambda)\, x \sim^* g.$$

Finally, suppose that $\alpha = 1$. Having chosen (renormalizing utility if necessary) $x, x' \in X$ such that $u(x) = 0$ and $u(x') = \beta$, it follows that

$$\frac{1}{2} f + \frac{1}{2} x \sim^* \frac{1}{2} g + \frac{1}{2} x'.$$

$(ii) \Rightarrow (iii)$: Obvious.

$(iii) \Rightarrow (ii)$: Notice that the expected utility mappings

$$\begin{aligned} P &\mapsto P(u(f)) \\ P &\mapsto P(u(g)) \end{aligned}$$

are affine functionals on $\mathcal{C}$. Therefore, (by the standard uniqueness properties of affine representations) they are isotonic iff one is a positive affine transformation of the other.

B.7 Proof of Proposition 11

$(i) \Rightarrow (iii)$: By Lemma 35,

$$\begin{aligned} \underline{\mathcal{C}}(u(k)) &= \inf_{\substack{g \in \mathcal{F} \\ \lambda \in (0,1]}} \left\{ I\left(u(k) + \frac{1-\lambda}{\lambda} u(g)\right) - I\left(\frac{1-\lambda}{\lambda} u(g)\right) \right\} \\ &= \inf_{\varphi \in B_0(\Sigma)} \{ I(u(k) + \psi) - I(\psi) \} \end{aligned}$$

and

$$\begin{aligned} \overline{\mathcal{C}}(u(k)) &= \sup_{\substack{g \in \mathcal{F} \\ \lambda \in (0,1]}} \left\{ I\left(u(k) + \frac{1-\lambda}{\lambda} u(g)\right) - I\left(\frac{1-\lambda}{\lambda} u\left(g\right)\right) \right\} \\ &= \sup_{\varphi \in B_0(\Sigma)} \{ I(u(k) + \psi) - I(\psi) \}. \end{aligned}$$

Suppose that k is crisp. Then for all $f \sim g$ and $\lambda \in (0,1]$,

$$\lambda k + (1-\lambda) f \sim \lambda k + (1-\lambda) g.$$

That is,

$$I(\lambda u(k) + (1-\lambda)u(f)) = I(\lambda u(k) + (1-\lambda)u(g)),$$

or, equivalently since $I(u(f)) = I(u(g))$,

$$I\left(u(k) + \frac{1-\lambda}{\lambda}u(f)\right) - I\left(\frac{1-\lambda}{\lambda}u(f)\right) = I\left(u(k) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right).$$

Therefore, for all $\psi, \theta \in B_0(\Sigma)$ such that $I(\psi) = I(\theta)$,

$$I(u(k) + \psi) - I(\psi) = I(u(k) + \theta) - I(\theta).$$

If $I(\psi) \neq I(\theta)$, set $a = I(\psi) - I(\theta)$. Then, $I(\psi) = I(\theta + a)$, whence

$$I(u(k) + \psi) - I(\psi) = I(u(k) + \theta + a) - I(\theta + a),$$

so that again

$$I(u(k) + \psi) - I(\psi) = I(u(k) + \theta) - I(\theta).$$

We conclude that if k is crisp

$$\underline{\mathcal{C}}(u(k)) = \inf_{\varphi \in B_0(\Sigma)} \{I(u(k) + \psi) - I(\psi)\} = \sup_{\varphi \in B_0(\Sigma)} \{I(u(k) + \psi) - I(\psi)\} = \overline{\mathcal{C}}(u(k)). \tag{16}$$

$(iii) \Rightarrow (iv)$: From Eq. (16) (which is (iii)) we obtain

$$I(u(k) + \psi) - I(\psi) = I(u(k))$$

for all $\psi \in B_0(\Sigma)$, whence for all $\lambda \in (0,1]$ and all $g \in \mathcal{F}$:

$$I\left(u(k) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right) = I(u(k))$$

or

$$I(\lambda u(k) + (1-\lambda)u(g)) = \lambda I(u(k)) + (1-\lambda)I(u(g)).$$

Finally, notice that the above equation is trivially true if $\lambda = 0$.

$(iv) \Rightarrow (i)$: If $f \sim g$ and $\lambda \in (0,1)$, it follows from (iv) that

$$\begin{aligned} I(\lambda u(k) + (1-\lambda)u(f)) &= \lambda I(u(k)) + (1-\lambda)I(u(f)) \\ &= \lambda I(u(k)) + (1-\lambda)I(u(g)) \\ &= I(\lambda u(k) + (1-\lambda)u(g)), \end{aligned}$$

whence

$$\lambda k + (1-\lambda)f \sim \lambda k + (1-\lambda)g.$$

$(ii) \Rightarrow (iii)$: Since $k \asymp x$, there exist λ, λ' and y, y' such that

$$\lambda\, k + (1-\lambda)\, y \sim^* \lambda'\, x + (1-\lambda')\, y',$$

which, applying Proposition 5, is equivalent to

$$\lambda\, P(u(k)) + (1-\lambda)\, u(y) \sim^* \lambda'\, u(x) + (1-\lambda')\, u(y'),$$

for every $P \in \mathcal{C}$. This immediately implies (iii).

$(iii) \Rightarrow (ii)$: Since $P(u(k)) = \gamma$ for every $P \in \mathcal{C}$, we just need to choose $x \in X$ such that $u(x) = \gamma$, and then apply Proposition 5 to see that $k \sim^* x$, yielding (ii).

B.8 Proof of Theorem 12

$(i) \Rightarrow (ii)$: Suppose that $\succcurlyeq$ satisfies axioms 1-5. Let I and u respectively be the preference functional and utility that represent $\succcurlyeq$ obtained in Lemma 1, and $\mathcal{C}$ the weak* compact and convex set of probabilities on Σ that represents $\succcurlyeq^*$ obtained in Proposition 5.

We have observed in Proposition 8 that $\underline{C}(u(f)) \leq I(u(f)) \leq \overline{C}(u(f))$ for all $f \in \mathcal{F}$. Hence, if f is crisp then $I(u(f)) = P(u(f))$ for every $P \in \mathcal{C}$. If f is not crisp, then there exists $a(u(f)) \in [0,1]$ such that

$$I(u(f)) = a(u(f))\,\underline{C}(u(f)) + (1 - a(u(f)))\,\overline{C}(u(f)).$$

Such $a(u(f))$ is unique, for

$$a(u(f)) = \frac{I(u(f)) - \overline{C}(u(f))}{\underline{C}(u(f)) - \overline{C}(u(f))}.$$

If we now recall the consequence of Lemma 9 and Proposition 11 that $[x]$ is the set of all crisp acts, we see that the function $a(\cdot)$ provides the sought representation. We are therefore done if we prove that a can be defined on $\mathcal{F}_{/\asymp} \setminus \{[x]\}$.

Suppose that $f \asymp g$. Then, there exist a pair of constants $x, x' \in X$ and weights $\lambda, \lambda' \in (0,1]$ such that

$$\lambda\, f + (1 - \lambda)\, x \sim^* \lambda'\, g + (1 - \lambda')\, x'. \tag{17}$$

It follows from point 1 of Proposition 4 that Eq. (17) implies

$$I(\lambda u(f) + (1 - \lambda)u(x)) = I(\lambda' u(g) + (1 - \lambda')u(x'))$$

so that, by the constant linearity of I:

$$\lambda I(u(f)) + (1 - \lambda)\, u(x) = \lambda' I(u(g)) + (1 - \lambda')\, u(x').$$

As a consequence,

$$I(u(f)) = \frac{\lambda'}{\lambda} I(u(g)) + \frac{1}{\lambda'}[(1 - \lambda')\, u(x') - (1 - \lambda)\, u(x)].$$

If we set $\beta = \frac{1}{\lambda}[(1 - \lambda')u(x')' - (1 - \lambda)u(x)]$ and $\alpha = \lambda'/\lambda$, we then obtain

$$I(u(f)) = \alpha I(u(g)) + \beta.$$

Notice that Eq. (17) also implies that for every $P \in \mathcal{C}$,

$$\lambda P(u(f)) + (1 - \lambda)\, u(x) = \lambda' P(u(g)) + (1 - \lambda')\, u(x').$$

That is, $P(u(f)) = \alpha P(u(g)) + \beta$ for every $P \in \mathcal{C}$. We conclude that

$$\begin{aligned}
a(u(f)) &= \frac{I(u(f)) - \overline{C}(u(f))}{\underline{C}(u(f)) - \overline{C}(u(f))} \\
&= \frac{\alpha I(u(g)) + \beta - \max_{P \in \mathcal{C}}(\alpha P(u(g)) + \beta)}{\min_{P \in \mathcal{C}}(\alpha P(u(g)) + \beta) - \max_{P \in \mathcal{C}}(\alpha P(u(g)) + \beta)} \\
&= a(u(g)).
\end{aligned}$$

Therefore, $a(u(f)) = a(u(g))$ whenever $f \asymp g$. If, with a little abuse of notation, we let $a([f]) = a(u(f))$, we find that $a : (\mathcal{F}_{/\asymp} \setminus \{[x]\}) \to [0,1]$, as claimed.

$(ii) \Rightarrow (i)$: Obvious.

B.9 Proof of Proposition 13

Since $\succcurlyeq_1$ and $\succcurlyeq_2$ perceive identical ambiguity, we have $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}$ and we can assume $u_1 = u_2 = u$. If $\mathcal{C}$ is a singleton, then $\succcurlyeq_1$ and $\succcurlyeq_2$ coincide, hence $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$ and $a_1([f]) \geq a_2([f])$ for every $f \in \mathcal{F} \setminus [x] = \emptyset$. Therefore, we assume $|\mathcal{C}| > 1$.

Suppose that $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$. Fix $f \in \mathcal{F} \setminus [x]$, and let $x \in X$ be indifferent to f for $\succcurlyeq_2$. We have:

$$\begin{aligned} a_2([f])\,\underline{\mathcal{C}}(u(f)) + (1 - a_2([f])\,\overline{\mathcal{C}}(u(f)) &= u(x) \\ &\geq a_1([f])\,\underline{\mathcal{C}}(u(f)) + (1 - a_1([f])\,\overline{\mathcal{C}}(u(f)). \end{aligned}$$

That is,

$$a_2([f])\left(\underline{\mathcal{C}}(u(f)) - \overline{\mathcal{C}}(u(f))\right) + \overline{\mathcal{C}}(u(f)) \geq a_1([f])\left(\underline{\mathcal{C}}(u(f)) - \overline{\mathcal{C}}(u(f))\right) + \overline{\mathcal{C}}(u(f)),$$

whence $a_1([f]) \geq a_2([f])$.

Conversely, suppose that $a_1([f]) \geq a_2([f])$ for every $f \in \mathcal{F} \setminus [x]$. For all $x \in X$,

$$\begin{aligned} x \succcurlyeq_2 f &\Leftrightarrow u(x) \geq a_2(u(f))\ \left(\underline{\mathcal{C}}(u(f)) - \overline{\mathcal{C}}(u(f))\right) + \overline{\mathcal{C}}(u(f)) \\ &\Rightarrow u(x) \geq a_1(u(f))\ \left(\underline{\mathcal{C}}(u(f)) - \overline{\mathcal{C}}(u(f))\right) + \overline{\mathcal{C}}(u(f)) \\ &\Rightarrow x \succcurlyeq_1 f. \end{aligned}$$

On the other hand, for all $f \in [x]$ and all $x \in X$, we can take $P \in \mathcal{C}$ to obtain:

$$\begin{aligned} x \succcurlyeq_2 f &\Leftrightarrow u(x) \geq P(u(f)) \\ &\Leftrightarrow x \succcurlyeq_1 f. \end{aligned}$$

B.10 Proof of Theorem 16

For all $f \in \mathcal{F}$, Lemma 35 yields

$$\max_{P \in \mathcal{C}} P(u(f)) = \sup_{\varphi \in B_0(\Sigma)} \{I(u(f) + \psi) - I(\psi)\},$$

while item 1 of Proposition 31 yields

$$\sup_{\psi \in B_0(\Sigma)} I(u(f) + \psi) - I(\psi) = I^\circ(0; u(f)) = \max_{P \in \partial I(0)} P(u(f)).$$

But, for all $\varphi \in B_0(\Sigma)$, there exist $\lambda \in (0,1)$ and $f \in \mathcal{F}$ such that $\lambda\varphi = u(f)$. Hence,

$$\max_{P \in \mathcal{C}} P(\varphi) = \max_{P \in \mathcal{C}} P\left(\frac{1}{\lambda}u(f)\right) = \max_{P \in \partial I(0)} P\left(\frac{1}{\lambda}u(f)\right) = \max_{P \in \partial I(0)} P(\varphi).$$

Since both $\mathcal{C}$ and $\partial I(0)$ are weak*-compact and convex subsets of $ba(\Sigma)$, we conclude that $\mathcal{C} = \partial I(0)$.

B.11 Proof of Proposition 19

If $m \in \mathit{Core}(I)$, then $m(\xi) \geq I(\xi) \geq \inf_{\psi \in B_0(\Sigma)} I(\psi + \xi) - I(\psi) = I_\circ(0;\xi)$. Analogously, if $m \in \mathit{Eroc}(I)$, then $m(\xi) \leq I(\xi) \leq \sup_{\psi \in B_0(\Sigma)} I(\psi + \xi) - I(\psi) = I^\circ(0;\xi)$.

If $\mathit{Eroc}(I) = \partial I(0)$, then

$$I^\circ(0;\xi) = \max_{m \in \partial I(0)} m(\xi) = \max_{m \in \mathit{Core}(I)} m(\xi) \leq I(\xi) \leq I^\circ(0;\xi)$$

for all $\xi \in B_0(\Sigma)$, so $I^\circ(0;\cdot) = I(\cdot)$ and I is convex. Conversely, if I is convex, a standard result (see Clarke [6]) guarantees that $\partial I(0) = \mathit{Eroc}(I)$. (The concave case is analogous.)

References

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

[2] Truman Bewley. Knightian decision theory: Part I. *Decisions in Economics and Finance*, 25(2):79–110, November 2002. (first version 1986).

[3] Jonathan M. Borwein and Warren B. Moors. Essentially smooth Lipschitz functions. *Journal of Functional Analysis*, 149:305–351, 1997.

[4] G. Carlier and Rose-Anne Dana. Core of convex distortions of a probability on a nonatomic space. Mimeo, CEREMADE, U. Paris Dauphine, July 2001.

[5] Erio Castagnoli, Fabio Maccheroni, and Massimo Marinacci. Expected utility with multiple priors. Mimeo, Università Bocconi and Università di Torino, July 2000.

[6] Frank H. Clarke. *Optimization and Nonsmooth Analysis*. J. Wiley, New York, 1983.

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

[8] Larry G. Epstein. A definition of uncertainty aversion. *Review of Economic Studies*, 66:579–608, 1999.

[9] Larry G. Epstein and Tan Wang. Intertemporal asset pricing under Knightian uncertainty. *Econometrica*, 62:283–322, 1994.

[10] Larry G. Epstein and Jiankang Zhang. Subjective probabilities on subjectively unambiguous events. *Econometrica*, 69:265–306, March 2001.

[11] Craig R. Fox and Amos Tversky. Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics*, 110:585–603, August 1995.

[12] Paolo Ghirardato. Revisiting Savage in a conditional world. *Economic Theory*, 20:83–92, March 2002.

[13] Paolo Ghirardato and Jonathan N. Katz. Indecision theory: Quality of information and voting behavior. Social Science Working Paper 1106R, Caltech, March 2002. http://www.hss.caltech.edu/∼paolo/voting.pdf.

[14] Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci. Ambiguity from the differential viewpoint. Social Science Working Paper 1130, Caltech, 2002. http://www.hss.caltech.edu/∼paolo/differential.pdf.

[15] Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci. Certainty independence and the separation of utility and beliefs. Applied Mathematics Working Paper 40/02, ICER, 2002. http://web.econ.unito.it/gma/icer-wps.htm.

[16] Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi. A subjective spin on roulette wheels. Social Science Working Paper 1127, Caltech, August 2001. http://www.hss.caltech.edu/∼paolo/invariant.pdf.

[17] Paolo Ghirardato and Massimo Marinacci. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research*, 26:864–890, November 2001.

[18] Paolo Ghirardato and Massimo Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102:251–289, February 2002.

[19] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[20] Chip Heath and Amos Tversky. Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4:5–28, 1991.

[21] Igor Kopylov. Procedural rationality in the multiple priors model. Mimeo, University of Rochester, April 2001.

[22] R. Duncan Luce. *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*. Lawrence Erlbaum, London, 2000.

[23] Mark J. Machina. Almost-objective uncertainty. Mimeo, UC San Diego, July 2001.

[24] Massimi Marinacci, Fabio Maccheroni, Alain Chateauneuf, and Jean-Marc Tallon. Monotone continuous multiple-priors. Mimeo, Univ. di Torino, Univ. Bocconi and Univ. Paris I, November 2002.

[25] Massimo Marinacci and Luigi Montrucchio. A characterization of the core of convex games through Gateaux derivatives. Mimeo, Università di Torino, April 2002.

[26] Klaus Nehring. Capacities and probabilistic beliefs: A precarious coexistence. *Mathematical Social Sciences*, 38:197–213, 1999.

[27] Klaus Nehring. Ambiguity in the context of probabilistic beliefs. Mimeo, UC Davis, November 2001.

[28] Kiyohiko G. Nishimura and Hiroyuki Ozaki. Search and Knightian uncertainty. Mimeo, University of Tokyo and Tohoku University, April 2001.

[29] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[30] David Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.

[31] Lloyd S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1:11–26, 1971.

[32] Marciano Siniscalchi. Bayesian updating for general maxmin expected utility preferences. Mimeo, Princeton University, September 2001.

[33] Marciano Siniscalchi. A behavioral characterization of plausible priors. Mimeo, Northwestern University, December 2002.

[34] Lionel Thibault. On generalized differentials and subdifferentials of Lipschitz vector-valued functions. *Nonlinear Analysis*, 6:1037–1053, 1982.

[35] Jiankang Zhang. Subjective ambiguity, probability and capacity. *Economic Theory*, 20:159–181, March 2002.