Differentiating ambiguity and ambiguity attitude

Paolo Ghirardato,[a,*] Fabio Maccheroni,[b] and
Massimo Marinacci[a]

[a] *Dipartimento di Statistica e Matematica Applicata and ICER, Università di Torino, Italy*
[b] *IMQ and IGIER, Università Bocconi, Italy*

We dedicate this paper—an extended version of which was previously circulated with the title "Ambiguity from the Differential Viewpoint"—to Erio Castagnoli on the occasion of his 60th birthday.

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 general decision model. We introduce a relation derived from the DM's preferences, called "unambiguous preference", and show that it can be represented by a set of probabilities. We provide such set with a simple differential characterization, and argue that it is a behavioral representation of the "ambiguity" that the DM may perceive. Given such revealed ambiguity, we provide a representation of ambiguity attitudes. We also characterize axiomatically a special case of our decision model, the "α-maxmin" expected utility model.
© 2003 Elsevier Inc. All rights reserved.

Introduction

When requested to state their maximum willingness to pay for two pairs of complementary bets involving future temperature in San Francisco and Istanbul

*Corresponding author.
E-mail addresses: ghirardato@econ.unito.it (P. Ghirardato), fabio.maccheroni@uni-bocconi.it (F. Maccheroni), massimo@econ.unito.it (M. Marinacci).
URL: http://web.econ.unito.it/gma.

(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 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 [15, Study 4]).

This striking pattern of preferences is by no means peculiar to the inhabitants of the Bay Area. Ever since the seminal thought experiment of Ellsberg [11], 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 [15] being just one of the most recent ones, have found significant evidence of ambiguity affecting decision making (see [25] 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 [23]'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 exhibits no 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 [22], which allows the agent's beliefs to be represented by a set of probabilities, and the "Choquet expected utility" (CEU) model of Schmeidler [34], 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 [13], labor search [32] or voting [16].

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. In fact, as we discuss below, the existing literature has either focused on narrower models, or has not—within the limits of a traditional decision-theoretic setting—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 $(\frac{1}{2})r + (\frac{1}{2})y$ and the

"mixed" act $(\frac{1}{2})b + (\frac{1}{2})y$. Then, we might observe

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

a violation of the independence axiom [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 $(\frac{1}{2})b + (\frac{1}{2})y$. Needless to say, we could observe a DM who displays exactly opposite preferences: she prefers b to r and $(\frac{1}{2})r + (\frac{1}{2})y$ to $(\frac{1}{2})b + (\frac{1}{2})y$ 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 that he may perceive in the problem at hand. Such ambiguity, if at all perceived, 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 cornerstone 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 $\succcurlyeq$. On the other hand, it should be stressed from the outset that such representation is, as every representation of preferences in decision theory, *attributed* to the DM. That is, it is possible that what is going on in the DM's mind may be quite unlike what our mathematical model (and the interpretation that we give to it) suggests—a point to which we shall come back after briefly reviewing our findings.

The revelation 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 [22]'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

---

[1] Therefore, an "act" is a map $f : S \rightarrow X$ 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 $\lambda f(s) + (1 - \lambda)h(s) \in X$.

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 [3]: there is a utility u on X and a set of probabilities $\mathscr{C}$ (nonempty, compact and convex) on S such that

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

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 $\mathscr{C}$. The set $\mathscr{C}$ of probabilistic scenarios represents, as we shall argue presently, the DM's revealed "(perception of) ambiguity." (While we do not carry it around for brevity's sake, the term "perception" serves as a reminder to the reader that no objective meaning is attached to $\mathscr{C}$. That is, nothing precludes two DMs from perceiving different ambiguity in the same decision problem.)

A key motivation for our interpretation of $\mathscr{C}$ as revealed ambiguity is the following analogy. 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 differential of the functional I that represents his preferences.[2] 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 preference functional I is not necessarily Gateaux differentiable. However, it does have a generalized differential—a collection of probabilites—in every point. Such differential is the "Clarke differential," developed by Clarke [9] as an extension of the concept of superdifferential (e.g., [33]) to non-concave functionals. We show that the set $\mathscr{C}$ obtained as the representation of $\succcurlyeq^*$ *is* the Clarke differential of I, analogously to what happens for SEU preferences. Thanks to this differential characterization, we also find that, in a finite state space, $\mathscr{C}$ 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 could rationalize the DM's evaluation of acts, we find $\mathscr{C}$.

Besides its conceptual interest, the differential characterization of $\mathscr{C}$ 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 revealed ambiguity and some very useful results on its mathematical properties.

Armed with the representation of ambiguity, we turn to the issue of formally describing the DM's reaction to the presence of such ambiguity. In our main representation theorem, we show that the DM's preference functional I can be written so as to associate to each act f an ambiguity aversion coefficient $a(f)$ between 0 and 1. The ambiguity aversion function $a(\cdot)$ thus obtained displays

---

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

significantly less variation than we might expect it to. In fact, it turns out that the DM must have identical ambiguity attitude for acts that agree on their ranking of the possible scenarios in $\mathscr{C}$. However, this restriction does not constrain overall ambiguity attitude, which can continuously range from strong attraction to strong aversion.

When the DM's preference $\succcurlyeq$ satisfies MEU, the set $\mathscr{C}$ is shown to be equal to the set of priors that Gilboa and Schmeidler derive in their representation [22], and the corresponding $a(\cdot)$ is uniformly equal to 1. The opposite—i.e., $a(\cdot)$ uniformly equal to 0—happens in the case of a "maxmax EU" preference. We also present the axiomatic characterization of the natural generalization of these two decision rules—a decision rule akin to Hurwicz's α-pessimism rule, known in the literature as the "α-MEU" decision rule (e.g., [27]).

A companion paper [18] analyzes some extensions and applications of the ideas and results in this paper. In particular, we look at a simple dynamic choice setting and show that the unambiguous preference relation allows us to characterize the updating rule that revises every prior in the set $\mathscr{C}$ by Bayes's rule—the so-called "generalized Bayesian updating" rule.

Discussion

It is important to comment on 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 methodological assumption entails some limitations in the accuracy of the terminology we use. First, we attribute no perception of ambiguity to a DM who disregards ambiguity. Indeed, it follows from our definition of unambiguous preference that if the DM never violates the independence axiom, *by definition* he reveals no ambiguity in our sense. Such DM behaves as if he considers only one scenario P to be possible (i.e., his $\mathscr{C} = \{P\}$), maximizing his subjective expected utility with respect to P. Of course, he may just *not* be reacting to the ambiguity he perceives, but we cannot discriminate between these conditions given our observability assumptions. As we are ultimately interested in modelling the ambiguity *as it affects behavior*, we do not believe this to be a serious problem from an economic viewpoint.

Second, we attribute *every* departure from the independence axiom to the presence of ambiguity. That is, following Ghirardato and Marinacci [21] 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—e.g., in experiments with "objective" probabilities—is often at spite with the independence axiom (again, see [25] for a survey). As a result, the relation $\succcurlyeq^*$ we associate with a DM displaying such systematic violations overestimates the DM's possible

perception of ambiguity. His set $\mathscr{C}$ describes behavioral traits that may not be related to ambiguity per se.

As extensively discussed in [21], 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 nonambiguity is *primitively assumed*, in violation of our observability premise. For conceptual reasons outlined in [21], 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 "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 and Aumann 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 et al. [19] that, if the set of outcomes is sufficiently rich, any mixture of state-contingent utility profiles can be obtained subjectively. 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 [22], 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 appears to perceive in a problem. He spelled out this proposal in an unpublished conference presentation of 1996, in which he also presented the characterization of the perceived ambiguity set $\mathscr{C}$ for MEU and CEU preferences when the state space is finite and utility is linear.[3]

In the recent [31], 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 unambiguous preference 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. His analysis mainly differs from ours in two main respects. The first is that his preferences induce an underlying set $\mathscr{C}$ satisfying a range convexity property. The second is that he also investigates preferences that do not satisfy an assumption

---

[3] "Preference and Belief without the Independence Axiom", presented at the LOFT2 conference in Torino (Italy), December 1996.

that he calls "utility sophistication", which is satisfied automatically by the preferences discussed here. A consequence of the range convexity of $\mathscr{C}$ is that CEU preferences can be utility sophisticated only if they maximize SEU, a remarkable result that does not generalize to the preferences we study (whose $\mathscr{C}$ 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 revealed ambiguity. 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 [24].[4] 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 the underlying capacities [28], or for obtaining a more direct computation of Choquet integrals in optimization problems [4].

In a recent paper, Siniscalchi [36] characterizes axiomatically a special case of our preference model—to be later called "piecewise linear" preferences—whose representation also involves a set of probabilities. The relation between his set $\mathscr{P}$ and our $\mathscr{C}$ are clarified in Section 5.2. 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., [14,30]), but do not address the distinction between ambiguity and the DM's reaction to it. We refer the reader to [18] 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 find that f is unambiguously preferred to g, while the "event-based" papers do not. In general, 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 [21]. 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 [21] for discussion of the relation of what we do with other works that address the characterization of ambiguity attitude.

---

[4] The works of Epstein [12] and Machina [26] are more distant from ours, as they take derivatives "with respect to events", rather than "with respect to utility profiles", as we do.

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 theorem and the characterization of ambiguity attitude. The differential interpretation of the set of possible scenarios and related results are presented in Section 5. Section 6 presents the axiomatization of the α-MEU model. Section 7 concludes and briefly reviews the extensions that are presented in detail in [18].

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

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 $\mathscr{F}$ 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 $x \in \mathscr{F}$ the constant act such that $x(s) = x$ for all $s \in S$, thus identifying X with the subset of the constant acts in $\mathscr{F}$. Finally, for $f, g \in \mathscr{F}$ and $A \in \Sigma$, $f A g$ denotes the act which yields $f(s)$ for $s \in A$ and $g(s)$ for $s \in A^c \equiv S \backslash A$.

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 in the classical setting of Anscombe and Aumann [1]. In view of the vector structure of X, for every $f, g \in \mathscr{F}$ and $\lambda \in [0, 1]$, we can thus define the mixed act $\lambda f + (1 - \lambda)g \in \mathscr{F}$ as in footnote 1. We model the DM's preferences on $\mathscr{F}$ by a binary relation $\succcurlyeq$. As usual, $\succ$ and $\sim$ respectively denote 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 $f \in \mathscr{F}$ 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. Since $ba(\Sigma)$ is (isometrically isomorphic to) the norm dual of $B_0(\Sigma)$,[5] all of its subsets inherit a weak* topology, for example, a net P_α of probabilities weak* converges to a probability P if and only if $P_\alpha(A) \to P(A)$ for all $A \in \Sigma$.

---

[5] Provided $ba(\Sigma)$ is endowed with the total variation norm, and $B_0(\Sigma)$ with the sup-norm.

Given a functional $I : B_0(\Sigma) \to \mathbb{R}$, we say that I is: *monotonic* if $I(\varphi) \geqslant I(\psi)$ for all $\varphi, \psi \in B_0(\Sigma)$ such that $\varphi(s) \geqslant \psi(s)$ 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 $a \geqslant 0$; *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 $f, g, h \in \mathscr{F}$: (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 $f, g \in \mathscr{F}$, $x \in X$, and $\lambda \in (0, 1]$, then*

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

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

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

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

Axiom 5 (Nondegeneracy). *There are $f, g \in \mathscr{F}$ 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 [22] in their characterization of MEU preferences. It requires that independence hold whenever acts are mixed with a constant act x.

The following representation result is easily proved by mimicking the arguments of Gilboa and Schmeidler [22, Lemmas 3.1–3.3].

Lemma 1. *A binary relation $\succcurlyeq$ on $\mathscr{F}$ 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*

$$f \succcurlyeq g \Leftrightarrow I(u(f)) \geqslant I(u(g)). \tag{2}$$

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 [17], 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 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 [20,21]) 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 $A \in \Sigma$,

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

The adjective *invariant* refers to the mentioned invariance of I to utility normalization, which is not necessarily true of the more general preferences in [20] (see [17] 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. These assumptions are summarized in the following axiom:

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

(a) (**Ambiguity neutrality**) $(\frac{1}{2})f + (\frac{1}{2})g \sim g$.
(b) (**Comonotonic ambiguity neutrality**) $(\frac{1}{2})f + (\frac{1}{2})g \sim g$ *if f and g are comonotonic.*[6]
(c) (**Ambiguity hedging**) $(\frac{1}{2})f + (\frac{1}{2})g \succcurlyeq g$.

Axiom 6(c) is due to Schmeidler [34], 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.

It is a matter of modifying known results in the literature to show the consequences of these three properties on 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 $\varphi \in B_0(\Sigma)$, where the integral is taken in the sense of Choquet.*

---

[6] Recall that 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 [21] for an explanation of our departure from that terminology.
[8] We refer the reader to [17,20] for additional examples and properties of invariant biseparable preferences.

- $\succcurlyeq$ *satisfies axiom* 6(c) *if and only if there is a nonempty, weak* compact and convex set $\mathscr{D}$ of probabilities on (S, Σ) such that $I(\varphi) = \min_{P \in \mathscr{D}} \int \varphi \, dP$ for all $\varphi \in B_0(\Sigma)$. Moreover, $\mathscr{D}$ 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 [34], with beliefs given by the capacity ρ.

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 [22] 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 size of $\mathscr{D}$ as representing the ambiguity that the DM may perceive in the decision problem, but a problem with such interpretation is the fact that the set $\mathscr{D}$ appears in Gilboa and Schmeidler's analysis only as a result of the assumption of ambiguity hedging.

It therefore *seems* that the DM's revealed ambiguity cannot be disentangled from his behavioral response to such ambiguity.

In the next section, we show that it is possible to separate a representation of ambiguity from the DM's behavioral reaction to it. For the sake of better assessing such separation, 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 [22] is only motivated by the intention of putting our contribution in sharper focus. The ''subjective mixtures'' of Ghirardato et al. [19] can be employed to extend the analysis in this paper to the case in which X does not have an ''objective'' vector structure (i.e., it is not convex), as long as it is sufficiently rich.

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 revealed 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 $f, g \in \mathscr{F}$. 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 $h \in \mathscr{F}$.

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 $\lambda \in (0, 1]$,*

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

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

$$fAh \succcurlyeq^* gAh \Leftrightarrow f\,A\,h' \succcurlyeq^* gAh'.$$

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 3).[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. Revealed ambiguity

We now show that the unambiguous preference relation $\succcurlyeq^*$ can be represented by a set of probabilities, in the spirit of a well-known result of Bewley [3]. (An analogous result is found in [31].)

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

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

---

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

[10] Nehring [31] independently introduces $\succcurlyeq^*$ and observes, in a setting with infinite states, its equivalence to the approach taken in his 1996 talk. He also provides further motivation for his approach.

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

Remark 1. A natural question that arises in applications is under which conditions the probabilities in the set $\mathscr{C}$ are all countably additive, provided Σ is a σ-algebra. It turns out that the following extension of the monotone continuity property of Villegas [38] and Arrow [2] is necessary and sufficient (see also [6]).[11]

Monotone continuity: *For all $x, y \in X$, if $A_n \downarrow \emptyset$ and $y \succ z$, then $y \succcurlyeq^* x A_n z$ for some n.*

The interpretation is analogous to that given by Villegas and Arrow. For any vanishing sequence of events, there is an event which is so small that it is close to being unambiguously impossible.

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

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

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

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

Proposition 6. *The following statements are equivalent*:

(i) $\succcurlyeq_1$ *reveals more ambiguity than* $\succcurlyeq_2$.
(ii) u_1 *is a positive affine transformation of* u_2 *and* $\mathscr{C}_1 \supseteq \mathscr{C}_2$.

In words, the size of the set $\mathscr{C}$ measures the perception of ambiguity we attribute to a DM. The larger $\mathscr{C}$ is, the more ambiguity the DM appears to perceive in the decision problem. In particular, no DM reveals less ambiguity than one who reveals a singleton set $\mathscr{C} = \{P\}$. In such case, $\succcurlyeq^*$ is complete. It follows that $\succcurlyeq^* = \succcurlyeq$; that is, the DM is a SEU maximizer with subjective probability P.

---

[11] The proof of this claim appears in Section B.3 of the appendix.

Summarizing the results obtained so far, we have shown that $\mathscr{C}$ represents what we call the DM's *revealed ambiguity*, and we have concluded that the DM *reveals some ambiguity* in a decision problem if $\mathscr{C}$ is not a singleton. Such characterization of revealed ambiguity does not rely on any assumption on the DM's *reaction* to it. 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 7. *Let I and u be respectively the functional and utility obtained in Lemma 1, and $\mathscr{C}$ the set obtained in Proposition 5. Then*

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

That is, the functionals on $\mathscr{F}$ defined by $\min_{P \in \mathscr{C}} P(u(\cdot))$ and $\max_{P \in \mathscr{C}} P(u(\cdot))$—that respectively correspond to the "worst-" and "best-case" scenario evaluations within the set $\mathscr{C}$—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 the functional I.

4.1. Crisp acts

It is first of all important to illustrate that revealed ambiguity already partitions $\mathscr{F}$ into sets of acts with "similar ambiguity." The following relation on the set $\mathscr{F}$ is key: For any $f, g \in \mathscr{F}$, 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 8. *For every $f, g \in \mathscr{F}$, the following statements are equivalent*:

(i) $f \asymp g$.

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

$$P(u(f)) = aP(u(g)) + b \quad \text{for all } P \in \mathscr{C}. \tag{7}$$

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

$$P(u(f)) \geqslant Q(u(f)) \Leftrightarrow P(u(g)) \geqslant Q(u(g)).$$

Statement (ii) of the lemma implies that $\asymp$ is an equivalence. Statement (iii) is helpful in interpreting $\asymp$. Two functions are *isotonic* on a set 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. That is, f and g have identical dependence on the ambiguity the DM displays.

As it will be seen presently, the equivalence classes of $\asymp$ play an important role in our representation. Given $f \in \mathscr{F}$, denote by $[f]$ the equivalence class of $\asymp$ that contains f and by $\mathscr{F}_{/\asymp}$ the quotient of $\mathscr{F}$ 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 [24] calls "transparency" (as his terminology suggests, he interprets it differently from us), is key in understanding the structure of $[x]$.

Definition 9. The act $k \in \mathscr{F}$ is called *crisp* if for all $f, g \in \mathscr{F}$ and $\lambda \in (0, 1)$,

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

The set of crisp acts is denoted by $\mathscr{K}$.

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 displays in the decision problem. The following characterization validates this intuition:

Proposition 10. *For every $k \in \mathscr{F}$, the following statements are equivalent*:

  (i) *k is crisp.*
 (ii) *$k \asymp x$ for some $x \in X$.*
(iii) *For every $P, Q \in \mathscr{C}$, $\int u(k)\, dP = \int u(k)\, dQ$.*
(iv) *For every $f \in \mathscr{F}$ 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 $\mathscr{K} = [x]$. Moreover, notice that it follows from statement (iv) of this proposition and the observation after Proposition 6 that if every act is crisp, the DM displays 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 revealed 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 11. *Let $\succcurlyeq$ be a binary relation on $\mathscr{F}$ satisfying axioms 1–5. Then there exist a nonempty, weak* compact and convex set $\mathscr{C}$ of probabilities on Σ, a nonconstant affine function $u : X \to \mathbb{R}$, and a function $a : \mathscr{F}_{/\asymp} \to [0,1]$ such that $\succcurlyeq$ is represented by the functional $I : B_0(\Sigma) \to \mathbb{R}$ defined by*

$$I(u(f)) = a([f]) \min_{P \in \mathscr{C}} \int u(f)\, dP + (1 - a([f])) \max_{P \in \mathscr{C}} \int u(f)\, dP, \qquad (8)$$

and u and $\mathscr{C}$ represent $\succcurlyeq^*$ in the sense of Eq. (4). Moreover, $\mathscr{C}$ is unique, u is unique up to a positive affine transformation, and the restriction of the function a to $\mathscr{F}_{/\asymp} \backslash \mathscr{K}$ is unique.

Clearly, the 1-MEU preference model and more generally the α-MEU preference model in which a is a constant $\alpha \in [0,1]$ (that is characterized axiomatically in Section 6), are special cases of the representation above. Also, observe that when $\mathscr{C} = \{P\}$ 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 $f \in \mathscr{F} \backslash \mathscr{K}$, the coefficient $a([f])$ only depends on the expected utility mapping $\{P(u(f)) : P \in \mathscr{C}\}$ of f on $\mathscr{C}$. As a result, the same is true of DM's evaluation $I(u(f))$ of any act $f \in \mathscr{F}$: The profile of expected utilities of f (as a function over $\mathscr{C}$) completely determines the DM's preference. This is a key feature of our representation, which is also enjoyed by the model studied by Siniscalchi [36].

Remark 2. It is routine to obtain the following converse to Theorem 11. Take a nonempty, weak* compact and convex set $\mathscr{C}$ of probabilities, an affine function u and define, via Eq. (7), an equivalence $\asymp$ on $\mathscr{F}$. Then, given $a : \mathscr{F}_{/\asymp} \to \mathbb{R}$, if the functional I defined by Eq. (8) is monotonic it induces a relation $\succcurlyeq$ which satisfies axioms 1–5.

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 $\min_{P \in \mathscr{C}} P(u(f))$. The following simple result verifies this intuition in terms of the relative ambiguity aversion ranking of Ghirardato and Marinacci [21]. In our setting, the latter is formulated as follows: $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$ if for all $f \in \mathscr{F}$ and all $x \in X$, $f \succcurlyeq_1 x$ implies $f \succcurlyeq_2 x$.

Proposition 12. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be invariant biseparable preferences, and suppose that $\succcurlyeq_1$ and $\succcurlyeq_2$ reveal identical ambiguity.[12] Then, $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$ if and only if $a_1([f]) \geqslant a_2([f])$ for every $f \in \mathscr{F} \backslash \mathscr{K}$.*

We conclude that the function a is a complete description of the DM's *ambiguity attitude* in relation to the revealed ambiguity described by $\mathscr{C}$.

In closing this section, we observe that it follows from Proposition 12 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 descriptions, since they correspond to a 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 [21].

5. Revealed ambiguity is a differential

In this section we turn back to the set $\mathscr{C}$ 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 support to our interpretation of $\mathscr{C}$, 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); i.e., there is a probability P on Σ such that $I(u(f)) = P(u(f))$. Being linear, I is Gateaux differentiable with derivative everywhere equal to P.[13] The DM's beliefs can thus be obtained by calculating the Gateaux derivative of I at any $\varphi \in B_0(\Sigma)$, for instance at $\varphi \equiv 0$. Using economic terminology (and assuming that S is finite) this is restated as follows: the probability $P(s)$ of state s gives the shadow price for increases of the DM's utility in state s.[14]

In contrast, if the DM's preferences do not satisfy axiom 6(a), the functional I may not be Gateaux differentiable everywhere, and even where it is, the Gateaux derivatives may differ from one point to another. Intuitively, because of the presence of ambiguity the shadow price for state s could depend on the structure of the act being evaluated. There are many different shadow prices.

A natural theoretical solution to this nondifferentiability problem is to allow for a more general notion of differentiability. For instance, suppose that the DM's preferences satisfy axioms 1–5 and 6(c), so that as shown in Proposition 2 they can be represented by maxmin expected utility with a set of priors $\mathscr{D}$. Then, the functional I is monotonic, constant linear and *concave*, so that, while not necessarily Gateaux

---

[12] Recall from Proposition 6 that $\succcurlyeq_1$ and $\succcurlyeq_2$ reveal identical ambiguity if and only if $\mathscr{C}_1 = \mathscr{C}_2$ and u_1 and u_2 are equivalent.

[13] In this discussion, we abuse terminology and identify the linear functional $P(\cdot)$—which is the real Gateaux derivative of I—with the probability P that induces it.

[14] In the case of monetary payoffs, looking at derivatives gurantees that we can ignore the shape of the utility function: The range of payoffs is infinitesimal.

differentiable, it does have a nonempty superdifferential (see, e.g., [33]). One could therefore think of using the superdifferential $\partial I(\varphi)$ of I at $\varphi \equiv 0$ (which contains $\partial I(\varphi)$ for every $\varphi \in B_0(\Sigma)$) as a possible description of the collection of shadow prices compatible with the DM's preferences.

Interestingly, calculating the superdifferential $\partial I(0)$ of such I yields exactly the set of priors $\mathscr{D}$. That is, the set of probabilities $\mathscr{D}$ of Gilboa and Schmeidler [22] can be obtained as derivative of the preference representation I. In this perspective, as the superdifferential of such an I coincides with its Gateaux derivative when the latter exists, SEU corresponds to the special case in which $\partial I(0) = \{P\}$.

For a preference $\succcurlyeq$ that only satisfies axioms 1–5—and therefore does not necessarily induce a concave I—we can use a generalization of the superdifferential due to Clarke [9], which is widely used in the literature on nonsmooth optimization.

Definition 13. Given a locally Lipschitz functional $I : B_0(\Sigma) \to \mathbb{R}$, its *Clarke (lower) directional derivative* at 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 at φ is the set of linear functionals that dominate the Clarke derivative $I_\circ(\varphi; \cdot)$. That is,

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

A monotonic and constant linear functional I, such as that obtained in Lemma 1, is Lipschitz and hence has a nonempty Clarke differential. Indeed, for an I with such properties Clarke differentials are sets of probabilities; that is, all the $m \in \partial I(\varphi)$ are normalized and positive. If I is also concave, then its Clarke differentials and its superdifferentials coincide (see [9]). This justifies our usage of the same symbol ∂I to denote both sets.[15]

We now show that the set $\mathscr{C}$ is equal to the Clarke differential of I at 0 (which contains $\partial I(\varphi)$ for every $\varphi \in B_0(\Sigma)$). Thus, the set of possible scenarios coincides with the appropriately generalized notion of derivative of the preference functional. That is, analogously to what happens in the SEU and 1-MEU case, our generalized "beliefs" can be obtained from the functional I by differentiation.

Theorem 14. *Let $\succcurlyeq$ be a binary relation satisfying axioms 1–5, and I and $\mathscr{C}$ the functional and set of probabilities obtained in Lemma 1 and Proposition 5, respectively. Then*

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

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

---

[15] Appendix A further discusses Clarke differentials and their properties.

5.1. Some theoretical consequences

First of all, from the mentioned equivalence of the Clarke differential and the superdifferential for concave I it follows immediately that $\mathscr{C} = \mathscr{D}$ whenever $\succcurlyeq$ satisfies axiom 6(c). In other words, for a 1-MEU preference the set of priors corresponds to the set of possible scenarios. (A result that was proved for finite S by Nehring, as reported in his 1996 talk; see footnote 3, and cf. his different generalization in [31]).

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 [21] and the ideas in this paper. Begin by considering the following two subsets of priors.

Definition 15. Given a functional $I : B_0(\Sigma) \to \mathbb{R}$, the core of I is the set

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

The anti-core of I is the set

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

As our choice of terminology suggests,[16] when I is a Choquet integral with respect to a capacity ρ, we have [21, 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 [21, Corollary 14] $Core(I) = \mathscr{D}$. 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. $Eroc(I)$) are the possible beliefs of SEU preferences $\geqslant$ which are less (resp. more) ambiguity averse than $\succcurlyeq$ in the sense of Ghirardato and Marinacci [21]: for all $f \in \mathscr{F}$ and $x \in X$, $f \succcurlyeq x \Rightarrow f \geqslant x$ (resp. $f \geqslant x \Rightarrow f \succcurlyeq x$). The next result shows that they also describe possible scenarios in the sense of this paper.

Proposition 16. *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 of the proposition 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 $\mathscr{D} = Core(I)$. Differently put, while Ghirardato and Marinacci's "benchmark measures" of $\succcurlyeq$ (the elements of $Core(I)$) are possible scenarios, they exhaust the set $\mathscr{C}$ only when $\succcurlyeq$ has extreme aversion to revealed ambiguity.

---

[16] In [21] these sets are denoted $\mathscr{D}(\succcurlyeq)$ and $\mathscr{E}(\succcurlyeq)$, respectively.

5.2. An operational consequence

A useful operational consequence of the characterization of $\mathscr{C}$ as a Clarke differential can be obtained the special case in which the state space S is finite; i.e., $S = \{s_1, s_2, \ldots, s_n\}$. (This is mainly for expositional purposes. The result can be extended to an S which is a compact metric space; see Appendix A for the details.)

In such a case, the Clarke differential at 0 can be given the following sharp representation in terms of the standard gradients of I (see Corollary A.5 in Appendix A):

$$\partial I(0) = \overline{co}\{\nabla I(\varphi) : \varphi \in \Omega\}, \tag{9}$$

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.)

We mention in passing that Eq. (9) provides further motivation for our interpretation of the set $\mathscr{C}$ as revealed ambiguity. For, given a functional I that has Gateaux derivatives almost everywhere (possibly different across points), each derivative can be interpreted as a "possible probabilistic scenario" that is implicitly used when evaluating a certain subset of acts. Thus, we can look at the collection of the Gateaux derivatives of the preference functional I as a set-valued "belief" associated with I. Alongside Theorem 14, Eq. (9) shows that the set $\mathscr{C}$ also fits this definition of "belief."

To see the operational import of Eq. (9), assume that the preference functional I is also *piecewise linear*. That is, there exists a countable family $\{C_l\}_{l \in L}$ of convex cones such that:

- $\mathbb{R}^n = \bigcup_l C_l$,
- $\mathrm{int}\, C_l \neq \emptyset$ for each l,
- I is linear on each C_l.

On finite state spaces, Choquet integrals are piecewise linear functionals; the same is true of the preference functionals studied by Castagnoli et al. [5] and Siniscalchi [36].[17]

Given a piecewise linear I, it is simple to describe its $\partial I(0)$. 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. (9), we then have (see Corollaries A.6 and A.8 in Appendix A) that

$$\partial I(0) = \overline{co}\{P_l : l \in L\}. \tag{10}$$

This equation shows that there exists a simple connection between our $\mathscr{C}$ and the collections of probabilities $\{P_l : l \in L\}$ derived in [5,36]. For CEU preferences, Eq. (10) enables us to retrieve $\mathscr{C}$ from the capacity ρ, as explained in the next example.

---

[17] Notice that in [36] the set L is countable because of the condition $\mathrm{int}\,(C_l \cap C_h) = \emptyset$ for all $l \neq h$.

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

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

for each permutation σ of $\{1, \ldots, n\}$ 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

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

for each $i = 1, \ldots, n$. 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. (10), we then have

$$\mathscr{C} = co\{P^\sigma : \sigma \in Per(n)\}. \tag{11}$$

In other words, in the Choquet case (with finite states) the set $\mathscr{C}$ 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 3).

When the functional I is also concave—i.e., when $I(\varphi) = \int \varphi \, d\rho$, with ρ supermodular[18]—Proposition 16 and Eq. (11) imply that $Core(\rho) = co\{P^\sigma : \sigma \in Per(n)\}$. Thus, Shapley's [35] well-known characterization of the core of a supermodular capacity can also be obtained as a consequence of Theorem 14.

6. A special case: α-MEU preferences

As we observed just after Theorem 11, an interesting class of invariant biseparable preferences are those whose ambiguity aversion index a is constant, the so-called α-MEU preferences. Here we show their behavioral characterization.

For any act $f \in \mathscr{F}$, denote by $C(f)$ the set of the certainty equivalents of f for $\succcurlyeq$; i.e., the elements $x \in X$ such that $x \sim f$. It is easy to see that

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

We analogously set

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

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 $\mathscr{C}$:

---

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

Proposition 18. *For every $f \in \mathcal{F}$,*

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

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

Thus, $u(C^*(f))$ 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. (8), 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 19. *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 preference functional $I : B_0(\Sigma) \to \mathbb{R}$ defined by*

$$I(u(f)) = \alpha \min_{P \in \mathcal{C}} \int_S u(f)\, dP + (1-\alpha) \max_{P \in \mathcal{C}} \int_S u(f)\, dP, \tag{12}$$

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.

The interpretation of α as the DM's coefficient of aversion to ambiguity hinges crucially on its uniqueness, which follows from the fact that $\mathcal{C}$ represents the relation $\succcurlyeq^*$. Such uniqueness does not rule out the possibility (see, e.g., [27]) that the preference $\succcurlyeq$ may have a similar representation with a *different* coefficient β and a *different* set of priors $\mathcal{D}$. That is,

$$I(u(f)) = \beta \min_{P \in \mathcal{D}} \int_S u(f)\, dP + (1-\beta) \max_{P \in \mathcal{D}} \int_S u(f)\, dP. \tag{13}$$

However, the next result shows that in such a case the set $\mathscr{C}$ must be included in the set $\mathscr{D}$.

Proposition 20. *Let $\succcurlyeq$ be a preference that can be simultaneously represented as in Eqs.* (12) *and* (13). *Then $\mathscr{D} \supseteq \mathscr{C}$, $\alpha \geqslant \beta$ if $\beta > 1/2$ and $\alpha \leqslant \beta$ if $\beta < 1/2$.*

To understand the relation between α and β, notice that when we use $\mathscr{D} \supset \mathscr{C}$ we are attributing to the DM an inflated perception of ambiguity. We are thus underestimating the magnitude of his reaction to the perceived ambiguity.

Summing up, among all the possible representations, the representation obtained in Proposition 19 is made special by two considerations: (1) it is the only one yielding a set $\mathscr{C}$ which represents $\succcurlyeq^*$, (2) it yields the smallest set of possible probabilistic scenarios, i.e., it offers the closest approximation to SEU that can be obtained.

7. Conclusions

We have introduced the notion of unambiguous preference, and proved that such a notion can be helpful in separating the DM's preference representation in "revealed" ambiguity and "revealed" ambiguity aversion. We have also shown that the DM's revealed ambiguity can be seen as the (properly defined) generalized differential of the DM's preference representation, analogously to what happens in the SEU case.

It is our hope that such separation—though artificial as any representation of preferences by mathematical means—will be helpful in analyzing the impact of ambiguity and ambiguity aversion/love in decision making situations of different sorts.

It is worth remarking that some interesting consequences of the results in this paper are already drawn in a companion paper [18]. For instance, we discuss a natural dynamic extension of our static choice setting, and show that dynamic consistency of the unambiguous preference relation—a property arguably more defensible than dynamic consistency of the DM's preferences in the presence of ambiguity—characterizes exactly the so-called *generalized Bayesian* updating rule, whereby *all* the probabilities in the set $\mathscr{C}$ are revised by Bayes's rule.

From a more theoretical perspective, in [18] we also consider the issue of defining unambiguous events and acts, which was briefly touched upon in our discussion of crisp acts in this paper. We argue that, while it is natural to define unambiguous events as those which correspond to crisp bets, the same is not necessarily true of general (nonbinary) acts. In fact, if unambiguous acts are those which are measurable with respect to a partition of unambiguous events, then any act which is obtained by permuting the payoffs of an unambiguous act should also be unambiguous. This is not in general true for crisp acts.

An important issue that is stimulated by our analysis and awaits further inspection is the following "integrability" question: Given a set $\mathscr{C}$ of priors and the associated relation $\asymp$, which functions can be ambiguity aversion indices for an invariant

biseparable preference that has $\mathscr{C}$ as its revealed ambiguity? The characterization of the set of such functions is made important by the (numerous) potential applications in which external considerations dictate the structure of the set $\mathscr{C}$.

Appendix 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 are directly mentioned in Section 5. Some of the proofs are standard, and are thus omitted.

A.1. Conic preorders

Given a non singleton interval K in the real line, we denote by $B_0(\Sigma, K)$ the subset of the functions in $B_0(\Sigma)$ taking values in K.

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 $\alpha \in [0, 1]$;[19]
- *monotonic* if $\varphi \geqslant \psi$ implies $\varphi \succsim \psi$.
- *nontrivial* if there exists $\varphi, \psi \in B_0(\Sigma, K)$ such that $\varphi \succsim \psi$ but not $\psi \succsim \varphi$.

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

$$\varphi \succsim_i \psi \Leftrightarrow \int_S \varphi \, dP \geqslant \int_S \psi \, dP \quad \text{for all } P \in \mathscr{C}_i.$$

Then

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

and the following statements are equivalent:

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

---

[19] 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 $\theta \in B_0(\Sigma, K)$ and all $\alpha \in \mathbb{R}_+$.

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

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

Proposition A.2. $\succsim$ *is a nontrivial, continuous, conic, and monotonic preorder on* $B_0(\Sigma, K)$ *if and only if there exists a nonempty set* $\mathscr{C}$ *of probabilities such that*

$$\varphi \succsim \psi \Leftrightarrow \int_S \varphi \, dP \geqslant \int_S \psi \, dP \quad \text{for all } P \in \mathscr{C}. \tag{A.1}$$

Moreover, $\overline{co}^{w^*}(\mathscr{C})$ *is the unique weak* compact and convex set of probabilities representing* $\succsim$ *in the sense of Eq.* (A.1).

A.2. Clarke derivatives and differentials: preliminary properties

We denote by $B(\Sigma)$ the closure in the supnorm of $B_0(\Sigma)$, whose norm dual is isometrically isomorphic to $ba(\Sigma)$. If S is a compact metric space, we denote by $C(S)$ the set of all continuous functions on S; in this case, we always assume Σ to be the Borel σ-algebra. The norm dual of $C(S)$ is isometrically isomorphic to the subset $ca(S)$ of $ba(\Sigma)$, consisting of all countably additive set functions.

In what follows, Φ denotes either $B_0(\Sigma)$, or $B(\Sigma)$, or $C(S)$; Φ^* denotes Φ's norm dual.

A monotonic constant linear functional $I : \Phi \to \mathbb{R}$ is Lipschitz of rank 1. In fact, given $\varphi, \psi \in \Phi$, $\varphi \leqslant \psi + ||\varphi - \psi||$ implies $I(\varphi) \leqslant I(\psi) + ||\varphi - \psi||$, hence $I(\varphi) - I(\psi) \leqslant ||\varphi - \psi||$; switching φ and ψ yields $|I(\varphi) - I(\psi)| \leqslant ||\varphi - \psi||$. Thus, given a monotonic constant linear functional $I : \Phi \to \mathbb{R}$, we can study its Clarke derivatives and Clarke differentials (as defined in Section 5). For easier reference to the existing literature we remind that—instead of the Clarke lower directional derivative—many authors use the *Clarke upper directional derivative*, defined by

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

for every $\varphi, \xi \in \Phi$ and define the Clarke differential at φ by

$$\partial I(\varphi) = \{m \in \Phi^* : m(\xi) \leqslant I^\circ(\varphi; \xi), \quad \forall \xi \in \Phi\}.$$

The observation that $I^\circ(\varphi; \xi) = -I_\circ(\varphi; -\xi)$ for every $\varphi, \xi \in \Phi$ shows that the two approaches are completely equivalent.

We refer to Clarke [9] for properties of the Clarke derivative and differential. Among them, the following are especially important:

1. For every $\varphi, \xi \in \Phi$ $I_\circ(\varphi; \xi) = \min_{m \in \partial I(\varphi)} m(\xi)$ and $I^\circ(\varphi; \xi) = \max_{m \in \partial I(\varphi)} m(\xi)$.
2. (**Lebourg Mean Value Theorem**) For all $\varphi, \psi \in \Phi$, there exist $\gamma \in (0, 1)$ and $m \in \partial I(\gamma \varphi + (1 - \gamma)\psi)$ such that $I(\varphi) - I(\psi) = m(\varphi - \psi)$.

Some additional properties of $I_\circ$ and $\partial I(\cdot)$ that we use below are stated next.

Proposition A.3. *Let $I : \Phi \to \mathbb{R}$ be a locally 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 $\varphi \in \Phi$. Moreover, $I_\circ(0; \xi) = \inf_{\psi \in \Phi} I(\psi + \xi) - I(\psi)$ and $I^\circ(0; \xi) = \sup_{\psi \in \Phi} I(\psi + \xi) - I(\psi)$ for all $\xi \in \Phi$.*
2. *If I is monotonic, then for all $\varphi \in \Phi$ the function $I_\circ(\varphi; \cdot)$ is monotonic, and m is positive for all $m \in \partial I(\varphi)$.*
3. *If I is constant additive, then for all $\varphi \in \Phi$ the functional $I_\circ(\varphi; \cdot)$ 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 for all $\varphi \in B_0(\Sigma)$ we have $\partial I(\varphi) \subseteq \partial I(0)$ and $\partial I(0)$ consists of probabilities.

A.3. Clarke differentials: representation on compact metric S

Suppose now that S is a compact metric state space. Notice that, for S finite, we have

$$B_0(\Sigma) = C(S) = B(\Sigma) = \mathbb{R}^{|S|},$$

while for S infinite, $C(S)$ is the only separable Banach space of the four.

A Borel subset N of $C(S)$ is *Haar-null* if there exists a (not necessarily unique) probability measure π on the Borel σ-algebra of $C(S)$, such that $\pi(\varphi + N) = 0$ for each $\varphi \in C(S)$. More generally, a subset N of $C(S)$ is a *Haar-null* set if it is contained in a Borel Haar-null set. Haar-null sets are closed under translation and countable unions; see Christensen [7]. In finite dimensions (i.e., for finite S), Haar-nulls sets coincide with the sets of Lebesgue measure 0. Using this terminology, Christensen [8] shows that each real-valued locally Lipschitz function defined on a non-empty open subset Ω of $C(S)$ is Gateaux differentiable except on an Haar-null subset of Ω. In fact, the following even stronger result is known.[20] Here, ∇ denotes a Gateaux derivative.

Theorem A.4 (Thibault [37, Proposition 2.2]). *Let J be a locally Lipschitz functional defined on a non-empty open subset Ω of $C(S)$ and let $D \equiv \{\varphi \in \Omega : \nabla J(\varphi) \text{ exists}\}$. Then for each Haar-null set $N \subseteq C(S)$ and each $\varphi \in \Omega$ we have that*

$$\partial J(\varphi) = \overline{co}^{w^*} \left\{ w^* - \lim_{i \to \infty} \nabla J(\varphi_i) : \varphi_i \in D \backslash N, \varphi_i \to \varphi \right\}.$$

Corollary A.5. *Let J be a locally Lipschitz and positively homogeneous functional defined on $C(S)$ and let $D \equiv \{\varphi \in C(S) : \nabla J(\varphi) \text{ exists}\}$. Then, for each Haar-null set $N \subseteq C(S)$, we have that $\partial J(0) = \overline{co}^{w^*}\{\nabla J(\varphi) : \varphi \in D \backslash N\}$.*

---

[20] The results of Christensen and Thibault are stated for separable Banach spaces rather than for $C(S)$.

Proof. Suppose J is Gateaux differentiable at $\varphi \in C(S)$, then $\nabla\varphi \in \partial J(\varphi)$ and, by positive homogeneity, $\partial J(\varphi) \subseteq \partial J(0)$. This proves that $\partial J(0) \supseteq \overline{co}^{w^*}\{\nabla J(\varphi) : \varphi \in D\backslash N\}$. Conversely, by the above theorem we have

$$\partial J(0) = \overline{co}^{w^*}\left\{ w^* - \lim_{i \to \infty} \nabla J(\varphi_i) : \varphi_i \in D\backslash N, \varphi_i \to 0 \right\}.$$

But, for all $\varphi_i \in D\backslash N$ such that $\varphi_i \to 0$ and $w^* - \lim_{i \to \infty} \nabla J(\varphi_i)$ exists, we have

$$w^* - \lim_{i \to \infty} \nabla J(\varphi_i) \in \overline{\{\nabla J(\varphi) : \varphi \in D\backslash N\}}^{w^*} \subseteq \overline{co}^{w^*}\{\nabla J(\varphi) : \varphi \in D\backslash N\}.$$

We conclude that $\partial J(0) \subseteq \overline{co}^{w^*}\{\nabla J(\varphi) : \varphi \in D\backslash N\}$. $\square$

Notice that the definition of piecewise linear functional of Section 5 can be naturally be extended to functionals defined on $C(S)$ (or $B_0(\Sigma)$, or $B(\Sigma)$). Obviously, a piecewise linear functional is positively homogeneous, and if it is also locally Lipschitz, then for all $l \in L$ there exists a unique $m_l \in ca(\Sigma)$ such that $J_{|C_l} = m_l$. In the sake of brevity, call $\{C_l, m_l\}_{l \in L}$ a *linear decomposition* of J.

Corollary A.6. *Let J be a locally Lipschitz and piecewise linear function defined on $C(S)$, and $\{C_l, m_l\}_{l \in L}$ a linear decomposition of J. Then,*

$$\partial J(0) = \overline{co}^{w^*}\{m_l : l \in L\}.$$

Proof. Clearly, J is Gateaux differentiable in $\text{int}\,C_l$ for each l, and $\nabla J(\varphi) = m_l$ for each $\varphi \in int\,C_l$. In particular, $\overline{co}^{w^*}\{m_l : l \in L\} \subseteq \partial J(0)$.

For each l, $\text{bnd}\,C_l$ is Haar-null (see [29, p. 1794]), hence $N = \bigcup_{l \in L}$ and C_l is Haar-null. Let $D \equiv \{\varphi \in C(S) : \nabla J(\varphi) \text{ exists}\}$ and observe that $D\backslash N \subseteq C(S)\backslash N \subseteq \bigcup_{l \in L} \text{int}\,C_l$. Therefore, $\{\nabla J(\varphi) : \varphi \in D\backslash N\} \subseteq \{m_l : l \in L\}$ and $\partial J(0) \subseteq \overline{co}^{w^*}\{m_l : l \in L\}$. $\square$

Lemma A.7. *Let $H : B(\Sigma) \to \mathbb{R}$ be a monotonic, positively homogeneous and locally Lipschitz functional. Denote by I (resp. J) the restriction of H to $B_0(\Sigma)$ (resp. $C(S)$), and by ∂I (resp. ∂J, resp. ∇J) the Clarke differential of I (resp. Clarke differential of J, resp. Gateaux derivative of J) relative to $B_0(\Sigma)$ (resp. relative to $C(S)$). Then $\partial H(0) = \partial I(0)$.*

Moreover, provided $\partial I(0) \subseteq ca(\Sigma)$,

$$\partial I(0) = \partial J(0) = \overline{co}\{\nabla J(\varphi) : \varphi \in \Omega\}, \tag{A.2}$$

where Ω is any subset of $C(S)$ on which J is Gateaux differentiable and such that $C(S)\backslash\Omega$ is Haar-null, and the closure is with respect to any one of the following weak topologies*: $\sigma(ca(\Sigma), C(S))$, $\sigma(ca(\Sigma), B_0(\Sigma))$, $\sigma(ca(\Sigma), B(\Sigma))$.

Notice that, if I is obtained from an invariant biseparable preference $\succcurlyeq$ such that $\succcurlyeq^*$ is monotone continuous, then $\partial I(0) \subseteq ca(\Sigma)$.

Proof. Just notice that, since $B_0(\Sigma)$ is dense in $B(\Sigma)$ and H is continuous, then $H^\circ(0;\xi) = \sup_{\psi \in B(\Sigma)} H(\psi + \xi) - H(\psi) = \sup_{\varphi \in B_0(\Sigma)} I(\varphi + \xi) - I(\varphi) = I^\circ(0;\xi)$ for all $\xi \in B_0(\Sigma)$. Then, $\partial I(0) = \{m \in ba(\Sigma) : m(\xi) \leqslant H^\circ(0;\xi), \ \forall \xi \in B_0(\Sigma)\}$. Continuity of H° and density of $B_0(\Sigma)$ in $B(\Sigma)$ yield

$$\partial I(0) = \{m \in ba(\Sigma) : m(\xi) \leqslant H^\circ(0;\xi), \ \forall \xi \in B(\Sigma)\} = \partial H(0).$$

Next, assume $\partial I(0) \subseteq ca(\Sigma)$. Notice that monotonicity of H implies that $\partial H(0) = \partial I(0)$ consists of positive countably additive set functions. Therefore, $\varphi_n, \varphi \in B(\Sigma)$ and $\varphi_n \uparrow \varphi$ or $\varphi_n \downarrow \varphi$ imply $H(\varphi_n) \to H(\varphi)$ and $H^\circ(0;\varphi_n) \to H^\circ(0;\varphi)$.[21]

For all $\xi \in C(S)$,

$$H^\circ(0;\xi) = \sup_{\psi \in B(\Sigma)} H(\psi + \xi) - H(\psi) \geqslant \sup_{\varphi \in C(S)} J(\varphi + \xi) - J(\varphi) = J^\circ(0;\xi).$$

On the other hand, for all $\xi \in C(S)$ the set $\{\psi \in B(\Sigma) : H(\psi + \xi) - H(\psi) \leqslant J^\circ(0;\xi)\}$ contains $C(S)$ and it is closed under monotone pointwise limits, so that it coincides with $B(\Sigma)$. It follows that $H^\circ(0;\xi) = J^\circ(0;\xi)$ if $\xi \in C(S)$.

As a consequence, $\partial J(0) = \{m \in ca(\Sigma) : m(\xi) \leqslant H^\circ(0;\xi), \ \forall \xi \in C(S)\}$. If $m \in \partial J(0)$, $\{\xi \in B(\Sigma) : m(\xi) \leqslant H^\circ(0;\xi)\}$ is a set containing $C(S)$ and closed under monotone pointwise limits, so that it also coincides with $B(\Sigma)$. We can conclude that, since $\partial I(0) \subseteq ca(\Sigma)$, $\partial J(0) = \{m \in ca(\Sigma) : m(\xi) \leqslant H^\circ(0;\xi), \ \forall \xi \in B(\Sigma)\} = \partial H(0) = \partial I(0)$.

Finally, $\partial J(0) = \partial I(0) = \partial H(0)$ is compact Hausdorff in the topologies $\sigma(ca(\Sigma), C(S))$, $\sigma(ca(\Sigma), B_0(\Sigma))$, and $\sigma(ca(\Sigma), B(\Sigma))$. Since $\sigma(ca(\Sigma), B(\Sigma))$ is finer than the others, they all coincide on $\partial H(0)$ and Corollary A.5 concludes the proof. $\square$

Let $H : B(\Sigma) \to \mathbb{R}$ be a monotonic, locally Lipschitz functional. Say that H is *properly piecewise linear* if there exists a countable family $\{C_l\}_{l \in L}$ of convex cones such that:

- $B(\Sigma) = \bigcup_l C_l$,
- int $C_l \cap C(S) \neq \emptyset$ for each l,
- I is linear on each C_l.

Corollary A.8. *Let $H : B(\Sigma) \to \mathbb{R}$ be a monotonic, locally Lipschitz, properly piecewise linear functional such that $\partial H(0) \subseteq ca(\Sigma)$, and $\{C_l, m_l\}_{l \in L}$ a linear decomposition*

---

[21] For all $\psi \in B(\Sigma)$, the function $\langle \psi, \cdot \rangle : \partial H(0) \to \mathbb{R}$ defined by $\langle \psi, m \rangle = \int_S \psi \, dm$ for all $m \in \partial H(0)$ is $\sigma(ba(\Sigma), B(\Sigma))$-continuous. If $\varphi_n, \varphi \in B(\Sigma)$ and $\varphi_n \uparrow \varphi$, then $\langle \varphi_n, \cdot \rangle \uparrow \langle \varphi, \cdot \rangle$ (Levi's Monotone Convergence Theorem), therefore $\langle \varphi_n, \cdot \rangle$ uniformly converges to $\langle \varphi, \cdot \rangle$ (Dini's Theorem and the fact that $\partial H(0)$ is $\sigma(ba(\Sigma), B(\Sigma))$-compact). Then, for all $\varepsilon > 0$ there exists $\bar{n} \in \mathbb{N}$ such that $|\langle \varphi_n, m \rangle - \langle \varphi, m \rangle| \leqslant \varepsilon$ for all $m \in \partial H(0)$ and all $n \geqslant \bar{n}$. By the Lebourg Mean Value Theorem, for all $n \geqslant \bar{n}$ there exist $\gamma_n \in (0,1)$ and $m_n \in \partial H(\gamma_n \varphi_n + (1 - \gamma_n)\varphi) \subseteq \partial H(0)$ such that $|H(\varphi_n) - H(\varphi)| = |\langle \varphi_n, m_n \rangle - \langle \varphi, m_n \rangle| \leqslant \varepsilon$, and $H(\varphi_n) \to H(\varphi)$. Moreover, since $\langle \varphi_n, \cdot \rangle$ converges to $\langle \varphi, \cdot \rangle$ uniformly on $\partial H(0)$ implies that

$$H^\circ(0;\varphi_n) = \max_{m \in \partial H(0)} \langle \varphi_n, m \rangle \to \max_{m \in \partial H(0)} \langle \varphi, m \rangle = H^\circ(0;\varphi).$$

The case $\varphi_n \downarrow \varphi$ is analogous.

of H. Then

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

where the closure is taken with respect to the $\sigma(ca(\Sigma), B(\Sigma))$ topology.

Proof. Let J be the restriction of H to $C(S)$ and $K_l = C_l \cap C(S)$. Clearly, (K_l, m_l) is a linear decomposition of J. Lemma A.7 and Corollary A.6 yield $\partial H(0) = \partial J(0) = \overline{co}\{m_l : l \in L\}$. □

Appendix 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

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

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 probabilities on Σ, we denote for every $\varphi \in B_0(\Sigma)$,

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

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 $\Rightarrow$ 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 $h \in \mathscr{F}$ and $\lambda \in (0, 1]$, hence $\lambda f + (1 - \lambda)h \succcurlyeq \lambda g + (1 - \lambda)h$ for all $h \in \mathscr{F}$ and $\lambda \in (0, 1]$, finally $f \succcurlyeq^* g$. This proves 7.

B.2. Proof of Proposition 5

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

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

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 $\varphi_n \gtrsim \psi_n$ for all $n \in \mathbb{N}$, $\varphi_n \to \varphi$, $\psi_n \to \psi$, then $I(\lambda \varphi_n + (1 - \lambda)\theta) \geqslant I(\lambda \psi_n + (1 - \lambda)\theta)$, 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 $\varphi \gtrsim \psi$.

We have thus shown that $\gtrsim$ is a conic, continuous, monotonic, nontrivial preorder on $B_0(\Sigma, u(X))$. By Proposition A.2 it follows that there exists a unique nonempty, weak* compact and convex set $\mathscr{C}$ of probabilities on Σ such that

$$\varphi \gtrsim \psi \Leftrightarrow \int_S \varphi \, dP \geqslant \int_S \psi \, dP \quad \text{for all} \quad P \in \mathscr{C},$$

which immediately yields the statement.

B.3. Proof of Remark 1

The following result is the claim in the remark. Recall that Σ is here assumed to be a σ-algebra.

Proposition B.1. *Let $\succcurlyeq$ be an invariant biseparable preference. Then the following statements are equivalent*:

(i) *For all $x, y, z \in X$ such that $y \succ z$, and all sequences of events $\{A_n\}_{n \geqslant 1} \subseteq \Sigma$ with $A_n \downarrow \emptyset$, there exists $\bar{n} \in \mathbb{N}$ such that $y \succcurlyeq^* x A_{\bar{n}} z$.*
(ii) *$\mathscr{C}$ consists of countably additive probabilities.*

Proof. (i) $\Rightarrow$ (ii): Let $A_n \downarrow \emptyset$ and let $y, z \in X$ be such that $y \succ z$. W.l.o.g. assume $u(y) = 1$ and $u(z) = 0$ and let $z_k = (1/k)y + (1 - (1/k))z$ so that $u(z_k) = 1/k$ (hence $z_k \succ z$). By monotone continuity, for all $k \in \mathbb{N}$ there exists $\bar{n} \in \mathbb{N}$ such that $z_k \succcurlyeq^* y A_{\bar{n}} z$.

Whence $1/k \geqslant P(A_{\bar{n}})$ for all $P \in \mathscr{C}$, but $P(A_n)$ is decreasing, and thus $\lim_{n \to \infty} P(A_n) \leqslant 1/k$. Clearly this implies that all the Ps belonging to $\mathscr{C}$ are countably additive.

(ii) $\Rightarrow$ (i): As $\mathscr{C}$ is a weak* compact set of countably additive probabilities, it is weak compact. By Theorem IV.9.1 of Dunford and Schwartz [10] it follows that if $\varepsilon > 0$ and $A_n \downarrow \emptyset$ there exists $\bar{n}$ such that $P(A_n) < \varepsilon$ for all $n \geqslant \bar{n}$ and all $P \in \mathscr{C}$. Now, let $x, y, z \in X$ be such that $y \succ z$. If $z \geqslant x$, we have $y \succcurlyeq^* z \succcurlyeq^* x A_n z$ for all n (z statewise dominates xA_nz). If $x \succ z$, there exists $\bar{n}$ such that $P(A_n) < [u(y) - u(z)]/[u(x) - u(z)]$ for all $n \geqslant \bar{n}$ and all $P \in \mathscr{C}$. That is, $u(y) > (u(x) - u(z))P(A_n) + u(z) = P(u(xA_nz))$. We conclude that $y \succcurlyeq^* xA_{\bar{n}}z$. $\quad\square$

B.4. Proof of Proposition 6

Lemma B.2. *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 $u(y)v(y) \geqslant 0$ 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 $u(z) > 0$. Were $\ker u = \ker v$, since for all $y \in Y$ there exist $b \in \mathbb{R}$ and $x \in \ker u$ such that $y = bz + x$, it would follow that $u(y) = b\, u(z) = \frac{u(z)}{v(z)} bv(z) = \frac{u(z)}{v(z)} v(y)$, 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. $\quad\square$

Corollary B.3. *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 $u(x_1) \geqslant u(x_2) \Rightarrow v(x_1) \geqslant v(x_2)$ 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

$$\hat{u} : t(x_1 - x_2) \mapsto t(u(x_1) - u(x_2)),$$

$$\hat{v} : t(x_1 - x_2) \mapsto t(v(x_1) - v(x_2))$$

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

$$\hat{u}(t(x_1 - x_2)) \geqslant 0 \Rightarrow u(x_1) \geqslant u(x_2) \Rightarrow v(x_1) \geqslant v(x_2) \Rightarrow \hat{v}(t(x_1 - x_2)) \geqslant 0.$$

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 $x_0 \in X$, for all $x \in X$

$$u(x) - u(x_0) = \hat{u}(1(x - x_0)) = a\hat{v}(1(x - x_0)) = av(x) - av(x_0)$$

so $u(x) = av(x) + [u(x_0) - av(x_0)]$, set $b = [u(x_0) - av(x_0)]$.   □

Proof of Proposition 6. (i) $\Rightarrow$ (ii): For all $x, y \in X$,

$$u_1(x) \geqslant u_1(y) \Rightarrow x \succcurlyeq_1 y \Rightarrow x \succcurlyeq_1^* y \Rightarrow x \succcurlyeq_2^* y \Rightarrow x \succcurlyeq_2 y \Rightarrow u_2(x) \geqslant u_2(y).$$

By Corollary B.3, this implies that we can assume $u_1 = u_2 = u$. Moreover, for all $f, g \in \mathscr{F}$, $f \succcurlyeq_1^* g \Rightarrow f \succcurlyeq_2^* g$. That is,

$$P(u(f)) \geqslant P(u(g)) \quad \forall P \in \mathscr{C}_1 \Rightarrow P(u(f)) \geqslant P(u(g)) \quad \forall P \in \mathscr{C}_2,$$

which by Proposition A.1 (applied to $B_0(\Sigma, u(X))$) implies $\mathscr{C}_2 \subseteq \mathscr{C}_1$.
   (ii) $\Rightarrow$ (i): Obvious.

B.5. Proof of Proposition 7

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

Lemma B.4. *For all $f \in \mathscr{F}$,*

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

$$= \inf_{\psi \in B_0(\Sigma)} \{ I(u(f) + \psi) - I(\psi) \}$$

and

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

$$= \sup_{\psi \in B_0(\Sigma)} \{ I(u(f) + \psi) - I(\psi) \}.$$

Proof. Clearly $\left\{ \frac{(1-\lambda)}{\lambda} u(g) : g \in \mathscr{F}, \lambda \in (0,1] \right\} \subseteq B_0(\Sigma)$. Conversely, for all $\psi \in B_0(\Sigma)$ there exists $\alpha \in (0,1)$ and $g \in \mathscr{F}$ 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.

Take $x_{\min} \in X$ that satisfies $u(x_{\min}) = \underline{\mathscr{C}}(u(f))$.[22] We have $f \succcurlyeq^* x_{\min}$; that is, for all $g \in \mathscr{F}$ and $\lambda \in (0,1]$:

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

or

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

Therefore,

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

from which we obtain

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

Finally,

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

Conversely, let $x_{\inf} \in X$ be such that[23]

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

Then,

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

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

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

which concludes the proof. $\square$

B.6. Proof of Lemma 8

(i) $\Rightarrow$ (ii): Suppose that for some λ, λ' and x, x', $\lambda f + (1 - \lambda) x \sim^* \lambda' g + (1 - \lambda')x'$. Applying Eq. (6) of Proposition 5, this is equivalent to

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

---

[22] Notice that such $x_{\min}$ exists. In fact, there exist $x', x'' \in X$ such that $x' \succcurlyeq f(s) \succcurlyeq x''$ for all $s \in S$, then $x' \succcurlyeq^* f \succcurlyeq^* x''$ and $u(x') \geqslant P(u(f)) \geqslant u(x'')$ for all $P \in \mathscr{C}$. Finding $x_{\min}$ is now trivial.

[23] Again, notice that such $x_{\inf}$ exists. In fact, choosing $x', x'' \in X$ such that $x' \succcurlyeq f(s) \succcurlyeq x''$, it follows

$$u(x') \geqslant I(u(f)) \geqslant \inf_{\substack{g \in \mathscr{F} \\ \lambda \in (0,1]}} \left\{ I\left(u(f) + \frac{1 - \lambda}{\lambda}u(g)\right) - I\left(\frac{1 - \lambda}{\lambda}u(g)\right) \right\} \geqslant u(x_{\min}).$$

Finding $x_{\inf}$ is now trivial.

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

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

so that we get the conclusion by letting $a = \frac{\lambda'}{\lambda}$ and $b = \frac{1}{\lambda}[(1 - \lambda')u(x') - (1 - \lambda)u(x)]$.

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

$$P(u(f)) = aP(u(g)) + b \quad \text{for all} \quad P \in \mathscr{C}.$$

Suppose first that $a < 1$. Then, let $\lambda = a$. By renormalizing the utility function if necessary, we can assume that $b/(1 - \lambda) \in u(X)$, so that there is $x \in X$ for which $u(x) = b/(1 - \lambda)$. It follows that $f \sim^* \lambda g + (1 - \lambda)x$. The case of $a > 1$ is dealt with by rewriting the equation as follows:

$$P(u(g)) = \frac{1}{a}P(u(f)) - \frac{b}{a} \quad \text{for all} \quad P \in \mathscr{C},$$

and proceeding as above to get $\lambda f + (1 - \lambda)x \sim^* g$. Finally, suppose that $a = 1$. Having chosen (renormalizing utility if necessary) $x, x' \in X$ such that $u(x) = 0$ and $u(x') = b$, 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

$$P \mapsto P(u(f)),$$

$$P \mapsto P(u(g))$$

are affine functionals on $\mathscr{C}$. 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 10

(i) $\Rightarrow$ (iii): 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

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

By Lemma B.4, we conclude that if k is crisp

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

(iii) $\Rightarrow$ (iv): Notice that (iii) and Lemma B.4 imply

$$\inf_{\varphi \in B_0(\Sigma)} \{I(u(k) + \psi) - I(\psi)\} = \underline{\mathscr{C}}(u(k)) = \overline{\mathscr{C}}(u(k)) = \sup_{\varphi \in B_0(\Sigma)} \{I(u(k) + \psi) - I(\psi)\},$$

thus $I(u(k) + \psi) - I(\psi) = I(u(k))$ for all $\psi \in B_0(\Sigma)$, whence for all $\lambda \in (0, 1]$ and all $g \in \mathscr{F}$:

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

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

$$\lambda P(u(k)) + (1 - \lambda)u(y) = \lambda' u(x) + (1 - \lambda')u(y'),$$

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

(iii) $\Rightarrow$ (ii): Since $P(u(k)) = \gamma$ for every $P \in \mathscr{C}$, 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 11

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 $\mathscr{C}$ the weak* compact and convex set of probabilities on Σ that represents $\succcurlyeq^*$ obtained in Proposition 5.

We have observed in Proposition 7 that $\underline{\mathscr{C}}(u(f)) \leqslant I(u(f)) \leqslant \overline{\mathscr{C}}(u(f))$ for all $f \in \mathscr{F}$. Hence, if f is crisp then $I(u(f)) = P(u(f))$ for every $P \in \mathscr{C}$. If f is not crisp, then there exists $a(u(f)) \in [0, 1]$ such that

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

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

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

We see that the function $a(\cdot)$ provides the sought representation. We are therefore done if we prove that a can be defined on $\mathscr{F}_{/\asymp}\backslash\mathscr{K}$.

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

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

It follows from point 1 of Proposition 4 that Eq. (16) 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. (B.1) also implies that for every $P \in \mathscr{C}$,

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

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

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

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 $a : (\mathscr{F}_{/\asymp}\backslash\mathscr{K}) \to [0, 1]$, as claimed.

B.9. Proof of Proposition 12

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

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

$$a_1([f])\,\underline{\mathscr{C}}(u(f)) + (1 - a_1([f])\overline{\mathscr{C}}(u(f)) = u(x) \leqslant a_2([f])\,\underline{\mathscr{C}}(u(f)) + (1 - a_2([f])\overline{\mathscr{C}}(u(f)).$$

That is,

$$a_2([f])(\underline{\mathscr{C}}(u(f)) - \overline{\mathscr{C}}(u(f))) + \overline{\mathscr{C}}(u(f)) \geqslant a_1([f])(\underline{\mathscr{C}}(u(f)) - \overline{\mathscr{C}}(u(f))) + \overline{\mathscr{C}}(u(f)),$$

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

Conversely, suppose that $a_1([f]) \geqslant a_2([f])$ for every $f \in \mathscr{F} \backslash \mathscr{K}$. For all $x \in X$,

$$f \succcurlyeq_1 x \Leftrightarrow a_1(u(f))(\underline{\mathscr{C}}(u(f)) - \overline{\mathscr{C}}(u(f))) + \overline{\mathscr{C}}(u(f)) \geqslant u(x)$$

$$\Rightarrow a_2(u(f))(\underline{\mathscr{C}}(u(f)) - \overline{\mathscr{C}}(u(f))) + \overline{\mathscr{C}}(u(f)) \geqslant u(x)$$

$$\Rightarrow f \succcurlyeq_2 x.$$

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

$$f \succcurlyeq_1 x \Leftrightarrow P(u(f)) \geqslant u(x) \Leftrightarrow f \succcurlyeq_2 x.$$

B.10. Proof of Theorem 14

For all $f \in \mathscr{F}$, Lemma B.4 yields

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

while item 1 of Proposition A.3 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 $f \in \mathscr{F}$ such that $\lambda\varphi = u(f)$. Hence,

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

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

B.11. Proof of Proposition 16

If $m' \in Core(I)$, then $m'(\xi) \geqslant I(\xi) \geqslant \inf_{\psi \in B_0(\Sigma)} I(\psi + \xi) - I(\psi) = I_\circ(0; \xi) = \min_{m \in \partial I(0)} m(\xi)$ for all $\xi \in B_0(\Sigma)$, which implies $m' \in \partial I(0)$.

If $Core(I) = \partial I(0)$, then $I_\circ(0; \xi) = \min_{m \in \partial I(0)} m(\xi) = \min_{m \in Core(I)} m(\xi) \geqslant I(\xi) \geqslant I_\circ(0; \xi)$ for all $\xi \in B_0(\Sigma)$, so $I_\circ(0; \cdot) = I(\cdot)$ and I is concave. Conversely, if I is concave a standard result (see [9]) guarantees that $\partial I(0) = Core(I)$. (The convex case is analogous.)

B.12. Proof of Proposition 18

As we observed earlier, if $y \in X$, $y \succcurlyeq^* f$ iff $u(y) = P(u(y)) \geqslant P(u(f))$ for all $P \in \mathscr{C}$ iff $u(y) \geqslant \overline{\mathscr{C}}(u(f))$. Similarly, $f \succcurlyeq^* y$ iff $\underline{\mathscr{C}}(u(f)) \geqslant u(y)$. Let $x', x'' \in X$ be such that $x' \succcurlyeq f(s) \succcurlyeq x''$ for all $s \in S$. Since $\succcurlyeq^*$ is monotonic, $x' \succcurlyeq^* f \succcurlyeq^* x''$, so that

$$u(x'') \leqslant \underline{\mathscr{C}}(u(f)) \leqslant \overline{\mathscr{C}}(u(f)) \leqslant u(x').$$

Hence for all $t \in [\underline{\mathscr{C}}(u(f)), \overline{\mathscr{C}}(u(f))]$ there exists x_t such that $u(x_t) = t$ (recall that u is affine and X is convex).

Let $x \in X$ satisfy $\underline{\mathscr{C}}(u(f)) \leqslant u(x) \leqslant \overline{\mathscr{C}}(u(f))$. If $y \succcurlyeq^* f$ then $u(y) \geqslant \overline{\mathscr{C}}(u(f)) \geqslant u(x)$ and $y \succcurlyeq^* x$; analogously, if $f \succcurlyeq^* y$, then $x \succcurlyeq^* y$. We can conclude that $x \in C^*(f)$.

Conversely, let $x \in C^*(f)$, and take $x_{\min}, x_{\max} \in X$ such that $u(x_{\min}) = \underline{\mathscr{C}}(u(f))$ and $u(x_{\max}) = \overline{\mathscr{C}}(u(f))$; $x_{\max} \succcurlyeq^* f \succcurlyeq^* x_{\min}$,   hence $x_{\max} \succcurlyeq^* x \succcurlyeq^* x_{\min}$.   That   is, $\overline{\mathscr{C}}(u(f)) \geqslant u(x) \geqslant \underline{\mathscr{C}}(u(f))$. This concludes the proof, as it amounts to saying that

$$C^*(f) = \{x \in X : \underline{\mathscr{C}}(u(f)) \leqslant u(x) \leqslant \overline{\mathscr{C}}(u(f))\},$$

while the existence of $x_{\min}, x_{\max} \in X$ such that $u(x_{\min}) = \underline{\mathscr{C}}(u(f))$ and $u(x_{\max}) = \overline{\mathscr{C}}(u(f))$ guarantees that $u(C^*(f)) = [\underline{\mathscr{C}}(u(f)), \overline{\mathscr{C}}(u(f))]$.

B.13. Proof of Theorem 19

The proof of the theorem builds on the following lemma.

Lemma B.5. *Let* $I : B_0(\Sigma) \to \mathbb{R}$ *be a monotonic constant linear functional, and* $\mathscr{D}$ *a set of probabilities such that*

$$\min_{P \in \mathscr{D}} P(\psi) \leqslant I(\psi) \leqslant \max_{P \in \mathscr{D}} P(\psi)$$

for all $\psi \in B_0(\Sigma)$. *If* $I(\psi) = T(\min_{P \in \mathscr{D}} P(\psi), \max_{P \in \mathscr{D}} P(\psi))$ *for all* $\psi \in B_0(\Sigma)$, *then there exists* $\beta \in [0, 1]$ *such that*

$$I(\psi) = \beta \min_{P \in \mathscr{D}} P(\psi) + (1 - \beta) \max_{P \in \mathscr{D}} P(\psi)$$

for all $\psi \in B_0(\Sigma)$. *If* $\mathscr{D}$ *is not a singleton,* β *is unique.*

Proof. If $\mathscr{D}$ is a singleton the result is trivial, so assume it is not. Since $\mathscr{D}$ is such that

$$\min_{P \in \mathscr{D}} P(\psi) \leqslant I(\psi) \leqslant \max_{P \in \mathscr{D}} P(\psi)$$

for all $\psi \in B(\Sigma)$, for all φ such that $\min_{P \in \mathscr{D}} P(\varphi) < \max_{P \in \mathscr{D}} P(\varphi)$ there exists a unique $\beta(\varphi) \in [0, 1]$ for which

$$I(\varphi) = \beta(\varphi) \min_{P \in \mathscr{D}} P(\varphi) + (1 - \beta(\varphi)) \max_{P \in \mathscr{D}} P(\varphi),$$

a little algebra yields:

$$\beta(\varphi) = \frac{I(\varphi) - \max_{P \in \mathscr{D}} P(\varphi)}{\min_{P \in \mathscr{D}} P(\varphi) - \max_{P \in \mathscr{D}} P(\varphi)}$$

$$= -\frac{I(\varphi) - \max_{P \in \mathscr{D}} P(\varphi)}{\max_{P \in \mathscr{D}} P(\varphi) - \min_{P \in \mathscr{D}} P(\varphi)} = -I\left(\frac{\varphi - \max_{P \in \mathscr{D}} P(\varphi)}{\max_{P \in \mathscr{D}} P(\varphi) - \min_{P \in \mathscr{D}} P(\varphi)}\right).$$

But, $I(\psi) = T(\min_{P \in \mathscr{D}} P(\psi), \max_{P \in \mathscr{D}} P(\psi))$ for all $\psi \in B(\Sigma)$. Moreover,

$$\max_{P \in \mathscr{D}}\left(\frac{\varphi - \max_{P \in \mathscr{D}} P(\varphi)}{\max_{P \in \mathscr{D}} P(\varphi) - \min_{P \in \mathscr{D}} P(\varphi)}\right) = 0$$

and

$$\min_{P\in\mathscr{D}}\left(\frac{\varphi - \max_{P\in\mathscr{D}} P(\varphi)}{\max_{P\in\mathscr{D}} P(\varphi) - \min_{P\in\mathscr{D}} P(\varphi)}\right) = -1;$$

therefore,

$$\beta(\varphi) = -I\left(\frac{\varphi - \max_{P\in\mathscr{D}} P(\varphi)}{\max_{P\in\mathscr{D}} P(\varphi) - \min_{P\in\mathscr{D}} P(\varphi)}\right) = -T(-1,0).$$

That is $\beta(\varphi) \equiv \beta$ does not depend on φ. $\square$

Proof of Theorem 19. Let I, u and $\mathscr{C}$ be the objects obtained in Lemma 1 and in Proposition 5. It is enough to show that, for all $\varphi \in B_0(\Sigma)$, $I(\varphi)$ depends only on $\min_{P\in\mathscr{C}} P(\varphi)$ and $\max_{P\in\mathscr{C}} P(\varphi)$. For, then we can set $T(\min_{P\in\mathscr{C}} P(\varphi), \max_{P\in\mathscr{C}} P(\varphi)) = I(\varphi)$ and apply Lemma B.5.

Let $\varphi, \psi \in B_0(\Sigma)$ be such that

$$\min_{P\in\mathscr{C}} P(\varphi) = \min_{P\in\mathscr{C}} P(\psi) \quad \text{and} \quad \max_{P\in\mathscr{C}} P(\varphi) = \max_{P\in\mathscr{C}} P(\psi).$$

Take $\alpha > 0$ such that $\alpha\varphi, \alpha\psi \in B(\Sigma, u(X))$ and $f, g \in \mathscr{F}$ such that $u(f) = \alpha\varphi$ and $u(g) = \alpha\psi$. Clearly,

$$\min_{P\in\mathscr{C}} P(u(f)) = \min_{P\in\mathscr{C}} P(u(g)) \quad \text{and} \quad \max_{P\in\mathscr{C}} P(u(f)) = \max_{P\in\mathscr{C}} P(u(g)).$$

By Proposition 18,

$$C^*(f) = u^{-1}([\underline{\mathscr{C}}(u(f)), \overline{\mathscr{C}}(u(f))]) = u^{-1}([\underline{\mathscr{C}}(u(g)), \overline{\mathscr{C}}(u(g))]) = C^*(g)$$

and Axiom 6 yields $f \sim g$, so that $I(\alpha\varphi) = I(u(f)) = I(u(g)) = I(\alpha\psi)$ and $I(\varphi) = I(\psi)$. The converse is trivial.

B.14. Proof of Proposition 20

The uniqueness of I descending from Lemma 1 guarantees that

$$I(\varphi) = \beta \min_{Q\in\mathscr{D}} Q(\varphi) + (1 - \beta) \max_{Q\in\mathscr{D}} Q(\varphi)$$

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

$$\mathscr{C} = \partial I(0)$$

$$= \partial\left(\beta \min_{Q\in\mathscr{D}} Q(\cdot) + (1 - \beta) \max_{Q\in\mathscr{D}} Q(\cdot)\right)(0)$$

$$\subseteq \beta \, \partial\left(\min_{Q\in\mathscr{D}} Q(\cdot)\right)(0) + (1 - \beta) \, \partial\left(\max_{Q\in\mathscr{D}} Q(\cdot)\right)(0)$$

$$= \beta\mathscr{D} + (1 - \beta)\mathscr{D}.$$

On the other hand, $I(\varphi) = \alpha\min_{P\in\mathscr{C}} P(\varphi) + (1 - \alpha)\max_{P\in\mathscr{C}} P(\varphi)$ for all $\varphi \in B_0(\Sigma)$.

That is,

$$\beta \min_{Q \in \mathscr{D}} Q(\cdot) + (1 - \beta) \max_{Q \in \mathscr{D}} Q(\cdot) = \alpha \min_{P \in \mathscr{C}} P(\cdot) + (1 - \alpha) \max_{P \in \mathscr{C}} P(\cdot).$$

If $\mathscr{C} = \mathscr{D}$, then clearly $\alpha = \beta$, and we are done. So suppose that $\mathscr{C} \subset \mathscr{D}$. Let φ be such that $\underline{c} = \min_{P \in \mathscr{C}} P(\varphi) < \max_{P \in \mathscr{C}} P(\varphi) = \bar{c}$. A fortiori, $\underline{d} = \min_{Q \in \mathscr{D}} Q(\varphi) < \max_{Q \in \mathscr{D}} Q(\varphi) = \bar{d}$. Moreover,

$$\tfrac{1}{2}\underline{c} + \tfrac{1}{2}\bar{c} = \tfrac{1}{2}I(\varphi) - \tfrac{1}{2}I(-\varphi) = \tfrac{1}{2}\underline{d} + \tfrac{1}{2}\bar{d}.$$

Let $c = \tfrac{1}{2}\underline{c} + \tfrac{1}{2}\bar{c}$ to obtain

$$I(\varphi) = \alpha\underline{c} + (1 - \alpha)\bar{c} = c + \alpha\underline{c} + (1 - \alpha)\bar{c} - \tfrac{1}{2}\underline{c} - \tfrac{1}{2}\bar{c} = c + \left(\tfrac{1}{2} - \alpha\right)(\bar{c} - \underline{c})$$

and, analogously,

$$I(\varphi) = c + \left(\tfrac{1}{2} - \beta\right)(\bar{d} - \underline{d}).$$

Therefore,

$$\left(\tfrac{1}{2} - \beta\right)(\bar{d} - \underline{d}) = \left(\tfrac{1}{2} - \alpha\right)(\bar{c} - \underline{c}). \tag{B.3}$$

If $\beta > 1/2$ then $\alpha > 1/2$ and $\left(\tfrac{1}{2} - \alpha\right)(\bar{c} - \underline{c}) > \left(\tfrac{1}{2} - \alpha\right)(\bar{d} - \underline{d})$, so that Eq. (B.3) implies $\alpha > \beta$. Analogously, if $\beta < 1/2$ then $\alpha < 1/2$ and $\left(\tfrac{1}{2} - \alpha\right)(\bar{c} - \underline{c}) < \left(\tfrac{1}{2} - \alpha\right)(\bar{d} - \underline{d})$, so that Eq. (B.3) implies $\alpha < \beta$. This concludes the proof.

References

[1] F.J. Anscombe, R.J. Aumann, A definition of subjective probability, Ann. of Math. Stat. 34 (1963) 199–205.

[2] K.J. Arrow, Exposition of a theory of choice under uncertainty, in: Essays in the Theory of Risk-Bearing, North-Holland, Amsterdam, 1974. (Chapter 2). (Part of the Yriö Jahnssonin Säätio lectures in Helsinki, 1965.)

[3] T. Bewley, Knightian decision theory: Part I, Decisions Econom. Finance 25 (2002) 79–110 (First version: 1986.).

[4] G. Carlier, R.A. Dana, Core of convex distortions of a probability, J. Econom. Theory 113 (2003) 119–222.

[5] E. Castagnoli, F. Maccheroni, M. Marinacci, Expected utility with multiple priors, in: J.-M. Bernard, T. Seidenfeld and M. Zaffalon (Eds.), Proceedings of the Third International Symposium on Imprecise Probabilities and their Applications Carleton Scientific, Waterloo, 2003, 121–132.

[6] A. Chateauneuf, F. Maccheroni, M. Marinacci, J.-M. Tallon, Monotone continuous multiple-priors, Econom. Theory, forthcoming.

[7] J.P.R. Christensen, On sets of Haar measure zero in abelian Polish groups, Israel J. Math. 13 (1972) 255–260.

[8] J.P.R. Christensen, Topology and Borel Structure, North-Holland, Amsterdam, 1974.

[9] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[10] N. Dunford, J.T. Schwartz, Linear Operators; Part I: General Theory, Wiley, New York, 1958.

[11] D. Ellsberg, Risk, ambiguity, the Savage axioms, Quart. J. Econom. 75 (1961) 643–669.

[12] L.G. Epstein, A definition of uncertainty aversion, Rev. Econom Stud. 66 (1999) 579–608.

[13] L.G. Epstein, T. Wang, Intertemporal asset pricing under Knightian uncertainty, Econometrica 62 (1994) 283–322.

[14] L.G. Epstein, J. Zhang, Subjective probabilities on subjectively unambiguous events, Econometrica 69 (2001) 265–306.

[15] C.R. Fox, A. Tversky, Ambiguity aversion and comparative ignorance, Quart. J. Econom. 110 (1995) 585–603.

[16] P. Ghirardato, J.N. Katz, Indecision theory: quality of information and voting behavior, Social Science Working Paper 1106R, Caltech, 2002. http://papers.ssrn.com/.

[17] P. Ghirardato, F. Maccheroni, M. Marinacci, Certainty independence and the separation of utility and beliefs, J. Econom. Theory, forthcoming.

[18] P. Ghirardato, F. Maccheroni, M. Marinacci, Revealed ambiguity and its consequences, Mimeo, Università Bocconi, Università di Torino, 2003.

[19] P. Ghirardato, F. Maccheroni, M. Marinacci, M. Siniscalchi, A subjective spin on roulette wheels, Econometrica 17 (2003) 1897–1908.

[20] P. Ghirardato, M. Marinacci, Risk, ambiguity and the separation of utility and beliefs, Math. Oper. Res. 26 (2001) 864–890.

[21] P. Ghirardato, M. Marinacci, Ambiguity made precise: a comparative foundation, J. Econom. Theory 102 (2002) 251–289.

[22] I. Gilboa, D. Schmeidler, Maxmin expected utility with a non-unique prior, J. Math. Econom. 18 (1989) 141–153.

[23] C. Heath, A. Tversky, Preference and belief: ambiguity and competence in choice under uncertainty, J. Risk Uncertainty 4 (1991) 5–28.

[24] I. Kopylov, Procedural rationality in the multiple priors model, Mimeo, University of Rochester, 2001.

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

[26] M.J. Machina, Almost-objective uncertainty, Mimeo, UC San Diego, 2001.

[27] M. Marinacci, Probabilistic sophistication and multiple priors, Econometrica 70 (2002) 755–764.

[28] M. Marinacci, L. Montrucchio, A characterization of the core of convex games through Gateaux derivatives, J. Econom. Theory, forthcoming.

[29] E. Matouskova, Convexity and Haar null sets, Proc. Amer. Math. Soc. 125 (1997) 1793–1799.

[30] K. Nehring, Capacities and probabilistic beliefs: a precarious coexistence, Math. Soc. Sci. 38 (1999) 197–213.

[31] K. Nehring, Ambiguity in the context of probabilistic beliefs, Mimeo, UC Davis, 2001.

[32] K.G. Nishimura, H. Ozaki, Search and Knightian uncertainty, Mimeo, University of Tokyo, Tohoku University, 2001.

[33] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, New Jersey, 1970.

[34] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.

[35] L.S. Shapley, Cores of convex games, Int. J. Game Theory 1 (1971) 11–26.

[36] M. Siniscalchi, A behavioral characterization of plausible priors, Mimeo, Northwestern University, 2002.

[37] L. Thibault, On generalized differentials and subdifferentials of Lipschitz vector-valued functions, Nonlinear Anal. 6 (1982) 1037–1053.

[38] C. Villegas, On qualitative probability σ-algebras, Ann. Math. Stat. 35 (1964) 1787–1796.