INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
ICER

WORKING PAPER SERIES

Paolo Ghirardato, Fabio Maccheroni,
and Massimo Marinacci

AMBIGUITY FROM THE DIFFERENTIAL VIEWPOINT

Working Paper no. 17/2002
April 2002

APPLIED MATHEMATICS
WORKING PAPER SERIES

Ambiguity from the Differential Viewpoint*

Paolo Ghirardato
California Institute of Technology,
Università di Torino
and ICER
paolo@hss.caltech.edu

Fabio Maccheroni
Università Bocconi
and ICER
fabio.maccheroni@uni-bocconi.it

Massimo Marinacci
Università di Torino
and ICER
massimo@econ.unito.it

April 2002

Abstract

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

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

In the second part of the paper we provide some extensions and "applications" of these ideas. We present an axiomatic characterization of the α-MEU decision rule. We also consider a simple dynamic choice setting and show the characterization of the updating rule that revises every prior in the afore-mentioned set by Bayes's rule; i.e., the generalized Bayesian updating rule.

*We are grateful to seminar audiences at the Université Paris I, Institute for Advanced Studies (Princeton, NJ), Università di Torino, Columbia University, as well as to Kim Border, Erio Castagnoli, Luigi Montrucchio, Klaus Nehring and Marciano Siniscalchi for helpful comments and discussion. Maccheroni and Marinacci are grateful to the MIUR for financial support.

Introduction

When requested to state their maximum willingness to pay for two pairs of complementary bets involving future temperature in San Francisco and Istanbul (and identical prize of \$ 100 in case of a win) 90 pedestrians on the University of California at Berkeley campus were on average willing to pay about \$ 41 for the two bets on San Francisco temperature, and \$ 25 for the two 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 [10, 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 [6], 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 [10] being just one of the most recent ones, have found significant evidence of ambiguity affecting decision making (see Luce [19] 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 [17]'s "competence hypothesis" experiments). However, the available evidence does show unequivocally that ambiguity matters for choice.

The benchmark decision model of subjective expected utility (SEU) maximization is not equipped to deal with this phenomenon: An agent who maximizes SEU cannot care about ambiguity. Therefore, theory has followed experiment. Several decision models have been proposed which extend SEU in order to allow a role for ambiguity in decision making. Most notable are the "maxmin expected utility with multiple priors" (MEU) model of Gilboa and Schmeidler [16], which allows the agent's beliefs to be represented by a set of probabilities, and the "Choquet expected utility" (CEU) model of Schmeidler [26], which allows the agent's beliefs to be represented by a unique but nonadditive probability. These models have been employed with success in understanding and predicting behavior in activities as diverse as investment (e.g., Epstein and Wang [8]), labor search (Nishimura and Ozaki [24]) or voting (Ghirardato and Katz [12]).

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, as discussed below, in our view has not been fully achieved by the previous literature. The intuition behind our approach can be explained in the context of the "3-color" experiment of Ellsberg. Suppose that a DM is faced with an urn containing 90 balls which are either red, blue or yellow. The DM is told that exactly 30 of the balls are red. If we offer him the choice between a bet r that pays \$ 10 if a red ball is extracted, and the bet b that pays \$ 10 if a blue ball is extracted, he may display the preference

$$r \succ b.$$

On the other hand, let y denote the bet that pays \$ 10 if a yellow ball is extracted, and suppose that we offer him the choice between the "mixed" act $(1/2)r + (1/2)y$ (that yields half the utility of \$ 10 if a blue ball is not extracted) and the "mixed" act $(1/2)b + (1/2)y$ (that yields half the utility of \$ 10 if a red ball is not extracted). Then, we might observe

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

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

In both cases, the presence of ambiguity in the decision problem a DM is facing is revealed to an external observer (who may ignore the information that was given to him 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 an act f with an act g. That is, $f \succcurlyeq g$ and for every act h and weight λ,

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

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

In the first part of the paper (Sections 1–4), we consider a preference $\succcurlyeq$ that satisfies a weak set of axioms, and show that its unambiguous preference relation $\succcurlyeq^*$ (which is clearly incomplete) can be represented by a set $\mathcal{C}$ of probabilities on the state space S. We argue that such set is a representation of the DM's perception of the ambiguity connected to the problem at hand. We do so by invoking concepts from differential calculus, explaining the title of this paper. Next, we show that once ambiguity has thus been isolated, ambiguity attitude can be easily represented formally. This formalization clarifies that the decision model we use is compatible with any type of ambiguity attitude, and is therefore consistent with the rich spectrum of reactions to ambiguity observed in experimental work. The next subsection provides a more detailed overview of these results.

In the second part of the paper (Sections 5–8), we provide some extensions and "applications" of the concepts developed in the first part. In particular, we present an axiomatic characterization of a decision rule akin to Hurwicz's α-pessimism rule, known in the literature as the "α-MEU" decision rule. We also look at a simple dynamic choice setting and show the characterization of the updating rule that revises every prior in the set $\mathcal{C}$ by Bayes's rule, the so-called "generalized Bayesian updating" rule. An overview of this part is presented in the subsection on "Extensions" below.

The Perception of Ambiguity and Differentials

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 a subset of the axioms that characterize Gilboa and Schmeidler [16]'s MEU model. In particular, we do not impose a key axiom that entails a preference for ambiguity hedging (the one they call "uncertainty aversion"), thus obtaining a less restrictive model than MEU. For instance, every preference that satisfies the CEU model satisfies our axioms, while those that satisfy MEU are a strict subclass.

Given such $\succcurlyeq$, we derive from it the unambiguous preference relation $\succcurlyeq^*$ in the manner described by Eq. (1), and show that $\succcurlyeq^*$ has a simple "unanimity" representation in the style of Bewley [3]: there is a utility u on X and a set of probabilities $\mathcal{C}$ (nonempty, closed and convex) on S such that

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

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 that the DM considers possible. The set $\mathcal{C}$ thus obtained represents, as we shall argue presently, the DM's revealed "perception of ambiguity". We use the term "perception" as a reminder to the reader that no objective meaning is attached to $\mathcal{C}$.[2] That is, nothing precludes two DMs from perceiving different ambiguity in the same decision problem.

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

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

It follows from our definition of unambiguous preference that if the DM does not ever violate the independence axiom, *by definition* we attribute to him no perception of ambiguity. Indeed, then $\mathcal{C} = \{P\}$ for some probability P: the DM behaves as if he considers only one scenario P to be possible, maximizing his subjective expected utility with respect to such P. Of course, the DM may just *not* be reacting to the ambiguity he perceives. However, the (standard) assumption that the *only* general observable is the DM's preference over mixed acts does not allow to distinguish between these situations. Ancillary information (e.g., what the DM knows about the decision problem) is needed.

We offer a few reasons why the set $\mathcal{C}$ may be interpreted as a representation of the DM's ambiguity perception. Foremost among these reasons is the following. It is simple to see that if a DM's preference $\succcurlyeq$ has a SEU representation, the DM's probabilistic beliefs P correspond to the derivative (in the sense of Gâteaux) of the functional I that represents his preferences.[3] Intuitively, the probability $P(s)$ is the shadow price for (*ceteris paribus*) changes in the DM's utility in state s. In other words, in the case of a DM who satisfies SEU, we can learn the DM's understanding of the stochastic nature of his decision problem — the collection of all the possible probabilistic scenarios — by calculating the derivative of his preference functional.

If $\succcurlyeq$ does not have a SEU representation, but satisfies our axioms, the functional I that represents $\succcurlyeq$ is not necessarily Gâteaux differentiable. However, it does have a generalized set-valued derivative. It is the notion of "Clarke differential", developed by Clarke [5] as an extension of the concept of superdifferential (e.g., Rockafellar [25]) to functionals that do not satisfy concavity. It is natural to maintain that the generalization of the previous remark should hold in this case: the Clarke differential describes the DM's understanding of the collection of all possible probabilistic scenarios (more than one, in this case). Our most important result shows that, as required, the set $\mathcal{C}$ obtained as the representation of $\succcurlyeq$ *is* the Clarke differential of I.

Interestingly, when the DM's preference functional I is concave — that is, when his $\succcurlyeq$ satisfies MEU — the Clarke differential corresponds to the superdifferential, and the set $\mathcal{C}$ is then equal to the set of priors that Gilboa and Schmeidler derive in their representation [16]. Therefore, the set of priors in MEU represents perceived ambiguity as defined in this paper.

Armed with the notion of perceived ambiguity, we turn to the issue of formally describing the DM's reaction to the presence of ambiguity. In our main representation theorem, we show that it is possible to express the DM's preference functional I so as to associate to each act f an ambiguity aversion coefficient $a(f)$ between 0 and 1. A surprising feature of the ambiguity

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

aversion function $a(\cdot)$ is that it displays significantly less variation than we might expect it to. For instance, the DM must have identical ambiguity attitude for acts that agree on their ranking of the possible scenarios in $\mathcal{C}$. This restriction does not constrain overall ambiguity attitude that, as mentioned earlier, can still range from strong attraction to strong aversion.

In this representation, a MEU preference with set of priors $\mathcal{C}$ is more ambiguity averse (in the relative sense of Ghirardato and Marinacci [15]) than any other preference with identical set of possible scenarios. That is, contrary to what is sometimes believed, MEU preferences *do* represent extreme aversion to ambiguity, a conclusion that could not be drawn without the separate derivation of perceived ambiguity obtained here.

Besides the mentioned conceptual interest, the differential characterization of the set $\mathcal{C}$ is also useful from a purely operational standpoint. By giving access to the large literature on the Clarke differential, it provides a different route to calculate and verify properties of the perceived ambiguity of the DM. For instance, when the state space S is finite, the set $\mathcal{C}$ can be calculated using an explicit formula for the Clarke differential found in Clarke [5].

Extensions

We begin the second part of the paper by considering the consequences of our theory of ambiguity for the classification of events and acts into "ambiguous" and "unambiguous". As should be expected, unambiguous events are the ones whose probability is identical in all possible scenarios, and the collection of unambiguous events is a λ-system. Moreover, for a significant subclass of preferences, such collection coincides with that, easily identifiable, suggested by Ghirardato and Marinacci [15]. As to acts, we call unambiguous any act which is measurable with respect to unambiguous events, and show how unambiguous acts can be characterized in terms of the set $\mathcal{C}$ of possible scenarios.

Next, we study the interesting special case of our decision model in which the ambiguity aversion function $a(\cdot)$ is constantly equal to $\alpha \in [0,1]$. That is, we discuss preferences which evaluate every act according to the rule

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

commonly known as the "α-maxmin expected utility" (α-MEU) rule. We show that α-MEU preferences can be characterized by an axiom that requires that the DM be indifferent between acts which induce the same interval of expected utilities (as we range over the possible scenarios in $\mathcal{C}$).

Finally, we consider a simple dynamic extension of our decision model, and look at the behavior of ambiguity perception as we vary the DM's information as to which event in S

obtains. We require preferences conditional on event A to depend only on acts' behavior on A (an assumption usually dubbed "consequentialism"), and observe that while dynamic consistency of the primitive conditional preference relations may be unduly restrictive in a setting with ambiguity, the same property appears to be more compelling when imposed just on the derived unambiguous preference relations. Interestingly, we find that this weaker dynamic consistency requirement is equivalent to the generalized Bayesian updating rule. We also show that the notion of "rectangularity" introduced by Epstein and Schneider [7] can be given an analogous characterization in the more general decision model we employ.

Discussion

It is perhaps useful to mention from the outset some limitations and peculiarities of our analysis and terminology. We follow the traditional decision-theoretic practice of 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 assume that any other ancillary information will in general be available to the external observer. Hence, we do not use such information in our analysis.

This premise entails a number of limitations in the accuracy of the terminology we use. First, as observed earlier, we may attributing no perception of ambiguity to a DM who is aware of ambiguity but disregards it. As we are ultimately interested in modelling the ambiguity that is reflected in behavior, we do not believe this to be a serious problem.

Second, and more important, we attribute every departure from the independence axiom to the presence of ambiguity. That is, following Ghirardato and Marinacci [15] we implicitly assume that behavior in the absence of ambiguity will be consistent with the SEU model. However, it is well-known that observed behavior in the absence of ambiguity — that is, in experiments with "objective" probabilities — is often at spite with the independence axiom (again, see Luce [19] for a survey). As a result, the relation $\succcurlyeq^*$ we attribute to a DM displaying such systematic violations overestimates the DM's perception of ambiguity. His set $\mathcal{C}$ describes behavioral traits that are not related to ambiguity *per se*.

As extensively discussed in [15], 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 unambiguity is *primitively assumed*, in violation of our observability premise. For conceptual reasons outlined in [15], in the absence of such device we prefer to attribute *all* departures from independence to the presence of ambiguity. However, the reader may prefer to use a different name for what we call "perception of ambiguity". We hope that it will be deemed to be an object of interest regardless of its name.

An aspect of our analysis which may appear to be a limitation is our heavy reliance on the concept of mixed acts. Indeed, the existence of a mixture operation is key to identifying the unambiguous preference relation. As the traditional interpretation of mixtures in the Anscombe-Aumann [1] framework is in terms of "lotteries over acts", it may be believed that our model also relies on an external notion of ambiguity. However, this is not the case, for it has been shown by Ghirardato, Maccheroni, Marinacci and Siniscalchi [13] that, if the set of outcomes is sufficiently rich, for any mixture of acts it is possible to construct an act whose state-contingent utility profile replicates perfectly that of the mixture. Our analysis can be fully reformulated in terms of such "subjective mixtures", and hence requires no external device.

The Related Literature

In addition to the mentioned paper of Gilboa and Schmeidler [16], there are a number of papers that share features, objectives, or methods with this paper.

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

In the recent [23], Nehring develops some of the ideas of the 1996 talk. The first part of that paper moves in a different direction than this paper, as it employs an incomplete relation that reflects probabilistic information *exogenously* available to the DM. The second part is closer to our work. In a setting with infinite states and consequences, Nehring defines a DM's perception of ambiguity by the maximal independent restriction of the primitive preferences over bets. He characterizes such definition and shows that under certain conditions it is equivalent to the one discussed here and in his 1996 talk (see footnote 8 below). His analysis mainly differs from ours in two respects. The first is that his preferences induce an underlying set $\mathcal{C}$ satisfying a range convexity property. The second is that he also investigates preferences that do not satisfy an assumption that he calls "tradeoff consistency", that the preferences discussed here satisfy automatically. A consequence of the range convexity of $\mathcal{C}$ is that CEU

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

preferences can only satisfy tradeoff consistency if they maximize SEU, a remarkable result that does not generalize to the preferences we study (whose $\mathcal{C}$ is not necessarily convex-ranged).

A final major difference between Nehring's mentioned contributions and the present paper is that he does not envision any differential interpretation for the set of probabilities that represents the DM's ambiguity perception. To the best of our knowledge, the only papers to use differential techniques to study ambiguity averse preferences are the recent Carlier and Dana [4] and Marinacci and Montrucchio [21]. Both papers focus on Choquet preference functionals, and they look at the Gâteaux derivatives of Choquet integrals as a device for characterizing the core of such capacities [21], or for obtaining a more direct computation of Choquet integrals in optimization problems [4].

In terms of the representation of preferences, this paper is quite close to a recent work of Siniscalchi [29]. He considers a special case of the preference model we employ, showing that the representation of these preferences can be decomposed in a fashion which also involves a set of probabilities. Both the decomposition and the set of probabilities are in general different from the ones obtained in this paper. On the other hand, he does not explicitly focus on the distinction between ambiguity and ambiguity attitude.

We next review papers that propose behavioral notions of unambiguous events or acts, but do not address the distinction between ambiguity and the DM's reaction to it (Section 5 contains some additional discussion). It is important to underscore an important difference between our "relation-based" approach at modelling ambiguity and the "event-based" approach of these papers. Suppose that f and g are ambiguous acts such that f dominates g statewise. Then we do obtain the conclusion that f is unambiguously preferred to g, but the "event-based" papers do not. That is, there are aspects of ambiguity that a "relation-based" theory can describe, but the "event-based" theories cannot. We are not aware of any instance in which the converse is true.

Epstein and Zhang [9] propose a behavioral notion of unambiguous event, and characterize a family of preferences whose behavior is "probabilistically sophisticated" on the acts measurable with respect to the set of unambiguous events. Their notion of unambiguous event is different to the one we present. For instance, all events are unambiguous in the Epstein-Zhang sense when the DM has a CEU preference whose capacity is a transformation of a probability measure. This is not necessarily the case with our definition. (The discussion in Ghirardato and Marinacci [15] explains why this is consistent with our choice of attributing violations of independence to ambiguity.)

In another offshoot of the mentioned 1996 talk, Nehring [22] looks at CEU preferences in a finite states setting and proposes a notion of unambiguous events for CEU preferences equivalent to the notion we proposed (that he also obtains for a different class of preferences in Nehring [23]). He then shows some interesting results on its characterization. Such notion is different to the one, also for CEU preferences, proposed and discussed by Zhang [30].

Kopylov [18] proposes a notion of unambiguity for acts (that he calls “transparency”) equivalent to our notion of crispness. Given the preference ordering over those acts, he considers the set $\mathcal{M}$ of all the probabilities that represent that ordering, and he characterizes the MEU preferences whose set of priors coincides with $\mathcal{M}$. Outside the MEU model, the set $\mathcal{M}$ he suggests will in general be different from our $\mathcal{C}$.

As to the papers that discuss ambiguity *aversion*, the closest to our work is Ghirardato and Marinacci [15]. That paper also suggests a notion of ambiguity for acts and events that only applies to a subset of preferences, and is more permissive than the one proposed here. 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 [15] for discussion of the relation of what we do with other works that address the characterization of ambiguity attitude.

Outline of the Paper

As explained earlier, the paper is roughly divided into two parts. The first part presents the basic ideas and fundamental results of our approach. After introducing some basic notation and terminology in Section 1, we present the basic axiomatic model in Section 2. Section 3 is the decision-theoretic core of the paper. It discusses the unambiguous preference relation, its characterization by a set of possible scenarios, and closes with the general represenation theorem with the characterization of ambiguity attitude. The differential interpretation of the set of possible scenarios and related results are presented in Section 4.

The second part of the paper starts with the discussion, in Section 5, of the notions of unambiguous event and act that follow from our perspective on ambiguity. In Section 6, we derive from unambiguous preference two additional concepts that are (natural and) useful for the ensuing discussion. The characterization of α-MEU is presented in Section 7. The discussion of the dynamic extension of the model, with the characterizations of generalized Bayesian updating and rectangularity, is found in the closing Section 8

The paper has three appendices. Appendix A presents some results which are employed in almost every other argument, along with further technical detail on Clarke differentials and their properties. Appendix B contains all the proofs of the results in the main body of the paper, in order of appearance. Appendix C shows how to extend the analysis of the first part of the paper to infinitely-valued acts.

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 $\mathfrak{F}$ 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 $x \in \mathfrak{F}$ the constant act such that $x(s) = x$ for all $s \in S$, thus identifying X with the subset of the constant acts in $\mathfrak{F}$. Given $f, g \in \mathfrak{F}$ and $A \in \Sigma$, we denote by $f\,A\,g$ the act in $\mathfrak{F}$ which yields $f(s)$ for $s \in A$ and $g(s)$ for $s \in A^c \equiv S \setminus A$.

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

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

We let $B_0(\Sigma)$ denote the set of all real-valued Σ-measurable simple functions, or equivalently the vector space generated by the indicator functions 1_A of the events $A \in \Sigma$. If $f \in \mathfrak{F}$ and $u : X \to \mathbb{R}$, we denote by $u(f)$ 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 $m(\varphi)$. Elements of $ba(\Sigma)$ which are probabilities are typically denoted with P and Q.

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

2 Invariant Biseparable Preferences

In this section, we introduce the basic preference model that is needed for our later results, which is a generalization of the MEU model of Gilboa and Schmeidler [16].

Axiom 1 (Weak Order) *(a) For all* $f, g \in \mathfrak{F}$, $f \succcurlyeq g$ *or* $g \succcurlyeq f$. *(b) If* $f, g, h \in \mathfrak{F}$, $f \succcurlyeq g$ *and* $g \succcurlyeq h$, *then* $f \succcurlyeq h$.

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

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

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

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

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

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

The following representation result is easily proved by mimicking the arguments of Gilboa and Schmeidler [16, Lemmas 3.1–3.3] (cf. Ghirardato *et al.* [13, Theorem 5]).

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

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

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

This representation is less structured than the MEU representation eventually obtained by Gilboa and Schmeidler. They assume that the preference relation satisfy an additional axiom that we call **ambiguity hedging**:[5] $(1/2)f + (1/2)g \succcurlyeq g$ for all $f, g \in \mathfrak{F}$ such that $f \sim g$. This allows them to prove that the functional I in the lemma is concave. They then show that in such case the functional I can be represented as the minimum expected utility with respect to a set of probabilities $\mathcal{D}$ on Σ, so that the DM chooses according to a "maxmin expected utility" rule. As we shall explain in more detail below, we refer to such preferences as **1-MEU** (rather than just MEU, as we have done so far) preferences. It is natural to interpret the probabilities in $\mathcal{D}$ as a reflection of the ambiguity that the DM perceives in the decision problem, but a problem with such interpretation is the fact that the set $\mathcal{D}$ appears in Gilboa and Schmeidler's analysis only as a result of the assumption of ambiguity hedging.

[5] Gilboa and Schmeidler call this property "uncertainty aversion". See Ghirardato and Marinacci [15] for an explanation of our departure from that terminology.

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

In the next section, we show that it is possible to separate the revealed perception of ambiguity from the DM's reaction to its presence. For that reason, it is important to notice here that axioms 1–5 do not build in any specific reaction to ambiguity (as ambiguity hedging does), a point that will be expounded in greater detail below.

We call a preference $\succcurlyeq$ satisfying axioms 1–5 an **invariant biseparable preference**. The adjective *biseparable* (originating from Ghirardato and Marinacci [15, 14]) is due to the fact that the representation on binary acts of such preferences satisfies the following separability condition (for a proof see [13, Proposition 14]): Let $\rho : \Sigma \to \mathbb{R}$ be defined by $\rho(A) \equiv I(1_A)$. Then, ρ is a normalized and monotone set-function (a **capacity**) and for all $x, y \in X$ such that $x \succcurlyeq y$ and all $A \in \Sigma$,

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

We call ρ a representation the DM's **willingness to bet** on events. The adjective *invariant* refers to the fact that the preferences satisfying axioms 1–5, differently from those discussed in general in [14], are represented by a unique functional I.

As indicated above, 1-MEU preferences (and hence SEU preferences) constitute a subclass of invariant biseparable preferences. Another significant subclass is that of **CEU preferences**, due to Schmeidler [26], which correspond to the special case in which the functional I is the Choquet integral with respect to the willingness to bet ρ. A well-known result of Schmeidler [26, Proposition] shows that the CEU preferences whose willingness to bet ρ is supermodular characterize the intersection of the CEU and 1-MEU classes.[6] We refer the reader to [14] for additional examples of invariant biseparable preferences.

We reiterate that the choice to retain the classical Anscombe-Aumann setting used by Gilboa and Schmeidler is motivated only by the intention of putting our contribution in sharper focus. Ghirardato *et al.* [13] show that if the set X does not have an "objective" vector structure (i.e., it is not convex) but is sufficiently rich, it is still possible to define mixtures in a subjective yet operationally well-defined sense. They use these "subjective mixtures" to provide an axiomatization of invariant biseparable preferences in a fully subjective setting, and they could be similarly used to extend the analysis in this paper.

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

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

3 Priors, Perceived Ambiguity, and Ambiguity Attitude

3.1 Unambiguous Preference

As we explained in the introduction, our point of departure is a relation on $\mathfrak{F}$ derived from $\succcurlyeq$ that says that hedging/speculation considerations do not affect the ranking of acts f and g. (Notice that this relation is defined for any binary relation $\succcurlyeq$ on $\mathfrak{F}$, not only those satisfying axioms 1–5.)

Definition 2 *Let $f, g \in \mathfrak{F}$. 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 $h \in \mathfrak{F}$.

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

Proposition 3 *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 (i.e., reflexive and transitive).*
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 $f, g, h \in \mathfrak{F}$ and $\lambda \in (0,1]$,*

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

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

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

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

[7] That is, if $\succcurlyeq' \subseteq \succcurlyeq$ and $\succcurlyeq'$ satisfies independence then $\succcurlyeq' \subseteq \succcurlyeq^*$.

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

The last point of the proposition shows that if we turned our perspective around and *defined* unambiguous preference as the maximal restriction of $\succcurlyeq$ that satisfies the independence axiom, we would find exactly our $\succcurlyeq^*$. As mentioned earlier, this second approach was suggested by Nehring in a 1996 talk (see footnote 4).[8] While eventually the approaches reach the same conclusions, we prefer the approach taken in this paper as it is directly linked to more basic behavioral considerations about hedging and speculation.

3.2 The Perception of Ambiguity

We now show that the unambiguous preference relation $\succcurlyeq^*$ can be represented by a set of probabilities, thus extending to an infinite state space a result of Bewley [3]. (An alternative generalization is found in Nehring [23].)

Proposition 4 *There exists a unique nonempty, weak* compact and convex set $\mathcal{C}$ of probabilities on Σ such that for all $f, g \in \mathfrak{F}$,*

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

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

Since $\mathcal{C}$ is a set of probabilities, it is natural to interpret it as a representation of the ambiguity the DM sees in the decision problem. In Section 4 we provide further argument in favor of this interpretation by showing the differential nature of $\mathcal{C}$. Here we offer a couple of additional remarks in support of this interpretation.

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 one DM has a *richer* unambiguous preference, it is because he feels better informed about the

[8] Nehring [23] independently introduces $\succcurlyeq^*$ and observes, in a setting with infinite states, its equivalence to the approach taken in 1996 talk. He also provides further motivation for his approach.

decision problem. Formally, $\succcurlyeq_1$ **perceives more ambiguity** than $\succcurlyeq_2$ if for all $f, g \in \mathfrak{F}$:

$$f \succcurlyeq^*_1 g \Longrightarrow f \succcurlyeq^*_2 g.$$

That is, $\succcurlyeq^*_1 \subseteq \succcurlyeq^*_2$. It turns out that this comparative definition of perceived ambiguity is equivalent to the inclusion of the sets of priors $\mathcal{C}$:

Proposition 5 *Given invariant biseparable preferences $\succcurlyeq_1$ and $\succcurlyeq_2$, the following statements are equivalent:*

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

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

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

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

Proposition 6 *The following statements are equivalent:*

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

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

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

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

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

3.3 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 $\mathcal{C}$ the set obtained in Proposition 4. Then*

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

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

It is first of all important to illustrate how the perception of ambiguity already partitions $\mathfrak{F}$ into sets of acts with "similar ambiguity". The following relation on the set $\mathfrak{F}$ is key: For any $f, g \in \mathfrak{F}$, 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. The next result characterizes the relation $\asymp$ in terms of the expected utility mappings of the acts.

Lemma 8 *For every $f, g \in \mathfrak{F}$, the following statements are equivalent:*[9]

(i) $f \asymp g$.

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

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

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

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

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

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

Since $\asymp$ is an equivalence, it is natural to inquire about the structure of its equivalence classes. Given $f \in \mathfrak{F}$, denote by $[f]$ the equivalence class of $\asymp$ that contains f and by $\mathfrak{F}_{/\asymp}$ the quotient of $\mathfrak{F}$ with respect to $\asymp$; i.e., the collection of all the 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. Analogously, 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, which is inspired by a property that Kopylov [18] calls "transparency" (as his terminology suggests, he interprets it differently from us), is key in understanding the structure of $[x]$.

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

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

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

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

(i) *k is crisp.*

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

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

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

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

Statement (ii) shows that $[x]$, the equivalence class of the constants, *is* the collection of all the crisp acts. Moreover, notice that it follows from statement (iv) of this proposition and

(ii) of Proposition 6 that if every act is crisp, the DM perceives no ambiguity (i.e., he satisfies SEU).

We are now ready to formulate our main representation theorem, wherein we achieve the formal separation of perceived ambiguity and the DM's reaction to it. Interestingly, it turns out to be a generalized Hurwicz α-pessimism representation (see footnote 13 below) in which the set of priors is generated endogenously.

Theorem 11 *Let $\succcurlyeq$ be a binary relation on $\mathfrak{F}$. The following statements are equivalent:*

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

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

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

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

In other words, the theorem proves that the functional I derived in Lemma 1, when restricted to noncrisp acts, has the form:

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

Clearly, the 1-MEU preference model and more generally the α-**MEU** preference model (that we characterize axiomatically in Section 7) in which a is a constant $\alpha \in [0,1]$ are special cases of the representation above. Also, observe that when $\mathcal{C} = \{P\}$ every act is crisp. Hence, the function a disappears from the representation, which reduces to SEU.

To understand this representation, it is important to observe that for any $f \in \mathfrak{F} \setminus [x]$, the coefficient $a([f])$ only depends on the expected utility mapping $\{P(u(f)) : P \in \mathcal{C}\}$ of f on $\mathcal{C}$. As a result, the same is true of DM's evaluation $I(u(f))$ of any act $f \in \mathfrak{F}$: The profile of expected utilities of f (as a function over $\mathcal{C}$) determines the DM's preference. This is a key feature of our representation, which is also enjoyed by the model studied by Siniscalchi in [29]. Moreover, if f and g are noncrisp acts and $f \asymp g$, then $a([f]) = a([g])$: The DM's reaction to the ambiguity of f is identical to his reaction to the ambiguity of g.

It is intuitive to interpret the function a as an index of the ambiguity aversion of the DM: The larger $a([f])$, the bigger the weight the DM gives to the "pessimistic" evaluation of f given by $\min_{P\in\mathcal{C}} P(u(f))$. The following simple result verifies this intuition in terms of the relative ambiguity aversion ranking of Ghirardato and Marinacci [15]. In our setting, the latter is formulated as follows: $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$ if for all $f \in \mathfrak{F}$ and all $x \in X$, $x \succcurlyeq_2 f \Rightarrow x \succcurlyeq_1 f$; equivalently, for all f and x, $f \succcurlyeq_1 x \Rightarrow f \succcurlyeq_2 x$.

Proposition 12 *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be invariant biseparable preferences, and suppose that $\succcurlyeq_1$ and $\succcurlyeq_2$ perceive identical ambiguity. Then, $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$* if and only if *$a_1([f]) \geq a_2([f])$ for every $f \in \mathfrak{F} \setminus [x]$.*

(Recall from Proposition 5 that $\succcurlyeq_1$ and $\succcurlyeq_2$ perceive identical ambiguity if and only if $\mathcal{C}_1 = \mathcal{C}_2$ and u_1 and u_2 are equivalent.) We conclude that the function a is a complete description of the DM's **ambiguity attitude** in relation to the perception of ambiguity described by $\mathcal{C}$.

In closing this section, we observe that it follows from Proposition 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 conditions, since they correspond to $a(\cdot)$ constantly equal to 0 and 1 respectively. In a sense, they describe the DM's "ambiguity averse side" and his "ambiguity loving side". However, notice that these DMs do not necessarily satisfy the SEU model, so they may not make the preference ambiguity averse in the sense of Ghirardato and Marinacci [15].

4 Perceived Ambiguity is a Differential

In this section we go back to the functional I on $B_0(\Sigma)$ derived in Lemma 1 and we show that the set $\mathcal{C}$ corresponds to the Clarke differential of the functional I at 0. As remarked earlier, this offers further backing to our interpretation of $\mathcal{C}$, while at the same time yielding a separate, operational, route for constructing a preference's set of possible scenarios.

Recalling our steps in the introduction, we start by assuming that the DM's preferences satisfy axioms 1–5 plus **ambiguity neutrality**: for all $f, g \in \mathfrak{F}$ such that $f \sim g$, $(1/2)f + (1/2)g \sim g$. Then, the functional I is monotonic and *linear*, so that there is a probability P on Σ such that $I(u(f)) = P(u(f))$; i.e., the DM satisfies the SEU model. Such functional has Gâteaux derivative constant and equal to P. The DM's beliefs can therefore be found by calculating the Gâteaux derivative of his preference functional in any point $\varphi \in B_0(\Sigma)$, for instance $\varphi \equiv 0$.

Suppose instead that the DM's preferences satisfy axioms 1–5 plus ambiguity hedging, as described after Lemma 1. Then, the functional I is monotonic, constant linear and *concave*, and as proved by Gilboa and Schmeidler [16] it can be represented by maxmin expected utility over a set of priors $\mathcal{D}$. In this case I is not necessarily Gâteaux differentiable. However, it does everywhere have directional derivatives and a nonempty superdifferential, as defined below (see, e.g., Rockafellar [25]).

Definition 13 *Given a concave functional $I : B_0(\Sigma) \to \mathbb{R}$, its* ***directional derivative*** *in φ in the direction ξ is defined by*

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

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

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

For a concave and constant linear I we have the following generalization of the characterization of I's derivative: the set $\mathcal{D}$ of priors is equal to the superdifferential $\partial I(0)$. Thus, the set of priors corresponds with the set of the possible supergradients of I at 0. The SEU case corresponds to the special case in which $\partial I(0) = \{P\}$, for the superdifferential of I coincides with its Gâteaux derivative when the latter exists.

Turn to preferences that only satisfy axioms 1–5. Now the functional I is only monotonic and constant linear, so that the existence of right-hand derivatives and superdifferentials is not guaranteed. However, I is shown to be Lipschitz (hence supnorm continuous). For such functionals it is customary in the literature on nonsmooth optimization to use the following generalized notions due to Clarke [5]. (See Appendix A for further details.)

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

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

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

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

It can be shown (Prop. 45 in App. A) that Clarke derivatives and differentials exist for every Lipschitz functional, and that if I is also concave, the Clarke differential coincides with the superdifferential. This justifies our usage of the same symbol to denote both sets. Moreover, when I is monotone and constant linear, its Clarke differential is a set of probability charges; that is, all the $L \in \partial I(\varphi)$ are normalized and positive (Prop. 47 in App. A).

The following result shows that the set $\mathcal{C}$ of the possible scenarios that we derived in the previous section corresponds to the Clarke differential of I in 0.

Theorem 15 *Let $\succcurlyeq$ be a binary relation satisfying axioms 1–5, and I and $\mathcal{C}$ respectively the functional and set of probabilities presented in Theorem 11. Then*

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

Thus, the differential characterization of "beliefs" for this class of preferences applies to the set of possible scenarios, providing further validation to our interpretation of $\mathcal{C}$ as a set of possible beliefs. Clearly, this calculus characterization is useful in providing an operational method for assessing a DM's perception of ambiguity $\mathcal{C}$, based on the computation of the Clarke differential at 0. However, it proves enlightening also for purely theoretical reasons.

For instance, from the mentioned equivalence of the Clarke differential and the superdifferential for concave I it immediately follows that $\mathcal{C} = \mathcal{D}$ whenever $\succcurlyeq$ satisfies ambiguity hedging. In other words, for a 1-MEU preference the set of priors corresponds to the set of possible scenarios. We thus generalize a result that was proved for finite S by Nehring, as reported in his 1996 talk (see footnote 4, and cf. his alternative generalization in [23]).

More interestingly, we can use the differential characterization to draw some conclusions on the relation between the comparatively based notion of ambiguity aversion of Ghirardato and Marinacci [15] and the ideas in this paper. Begin by considering the following two subsets of SEU preferences.

Definition 16 *Given a functional $I : B_0(\Sigma) \to \mathbb{R}$, the **core of** I is the set*

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

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

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

As our choice of terminology suggests,[10] when I is a Choquet integral with respect to a capacity ρ (so that it is constant linear), we have [15, Corollary 13] that

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

[10]In Ghirardato and Marinacci [15] these sets are denoted $\mathcal{D}(\succcurlyeq)$ and $\mathcal{E}(\succcurlyeq)$ respectively.

However, these notions apply also to preferences which are not CEU. Indeed, if $\succcurlyeq$ is a 1-MEU preference, then [15, Corollary 14] $Core(I) = \mathcal{D}$. Clearly, both $Core(I)$ and $Anticore(I)$ could be empty, and they are simultaneously nonempty if and only if I is linear.

The elements of $Core(I)$ (resp. $Anticore(I)$) correspond to SEU preferences $\geqslant$ which are less (resp. more) ambiguity averse than $\succcurlyeq$ in the sense of Ghirardato and Marinacci [15]: for all $f \in \mathfrak{F}$ and $x \in X$, $x \geqslant f \Rightarrow x \succcurlyeq f$ (resp. $x \succcurlyeq f \Rightarrow x \geqslant f$). We now show that they correspond to some possible scenario.

Proposition 17 *Let I be a monotonic, constant linear functional. Then*

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

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

The second statement shows that $Core(I)$ contains *all* the possible scenarios if and only if I is concave; that is, $\succcurlyeq$ is a 1-MEU preference with set of priors $\mathcal{D} = Core(I)$.

It follows from the proposition that while Ghirardato and Marinacci's "benchmark measures" of $\succcurlyeq$ (the elements of Core(I)) are all possible scenarios, they exhaust the set $\mathcal{C}$ only when $\succcurlyeq$ is a 1-MEU preference. In particular, the DM may not have any benchmark and yet be *quite* ambiguity averse, in the sense of having a uniformly high (but not constantly 1) ambiguity aversion coefficient. On the other hand, if he *does* have a benchmark measure, then he cannot be too ambiguity loving (say, have $a([f]) \leq 1/2$ for every $f \in \mathfrak{F} \setminus [x]$, with strict inequality for one f) except in the trivial case in which he satisfies SEU.

As an operational consequence of the Clarke differential characterization, consider the special case in which the state space S is finite, $S = \{s_1, s_2, ..., s_n\}$. Then, the Clarke differential can be given a more elementary representation, as it is related to the collection of the standard gradients of I (see Thm. 49 in App. A).

This representation becomes even simpler in the case in which I is a Choquet integral with respect to its willingness to bet ρ. If σ is a permutation of $\{1, 2, ..., n\}$, let

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

and notice that P^{σ} is a probability on S. Now, recall the following property of Choquet integrals. If $\varphi \in B_0(\Sigma)$ is such that $\varphi(s_{\sigma(1)}) \geq \varphi(s_{\sigma(2)}) \geq \cdots \geq \varphi(s_{\sigma(n)})$, then

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

That is, the Choquet integral of a function that orders the states consistently with the permutation σ is equal to a standard integral with respect to P^σ.

The finite representation of $\mathcal{C}$ mentioned above implies that the set $\mathcal{C}$ is the convex hull of the set of all the P^σ; i.e., 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, for the case of linear utility, by Nehring in a 1996 talk (see footnote 4).

Corollary 18 *Let I be a Choquet integral and $S = \{s_1, s_2, ..., s_n\}$. Then*

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

where $Per(n)$ is the set of all the permutations of $\{1, 2, \ldots, n\}$.

Specialize further to the case in which I is concave; i.e., it is a Choquet integral with respect to a *supermodular* ρ. Then it follows from Proposition 17 that $\mathcal{C} = \text{Core}(\rho)$. We can thus use the corollary to obtain as a consequence of our main result the well-known characterization of the core of a supermodular capacity due to Shapley [27].

5 Ambiguity of Events and Acts

5.1 Unambiguous Events

We have earlier introduced crisp acts, those whose evaluation is not affected by the ambiguity the DM perceives in a problem. Consider in particular crisp *bets*; i.e., acts of the form $x\,A\,y$ for $x \succ y$. We suggest that the event corresponding to a crisp bet be defined unambiguous (Nehring [23] proposes an equivalent definition, and shows that the latter is in turn equivalent to one he earlier studied in [22]):

Definition 19 *An event $A \in \Sigma$ is* ***unambiguous*** *if for some $x \succ y$, the act $x\,A\,y$ is crisp. The collection of all the unambiguous events is denoted by Λ.*

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

Proposition 20 *For any $A \in \Sigma$, the following statements are equivalent:*

(i) A is unambiguous.

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

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

For all invariant biseparable preferences, the collection Λ has a simple and intuitive structure (cf. Zhang [30] and Nehring [22]).

Proposition 21 Λ *is a (finite)* λ*-system. That is: 1)* $S \in \Lambda$*; 2) if* $A \in \Lambda$ *then* $A^c \in \Lambda$*; 3) if* $A, B \in \Lambda$ *and* $A \cap B = \emptyset$ *then* $A \cup B \in \Lambda$.

Nehring [22] shows that, if S is finite and I is a Choquet integral (so that the characterization of $\mathcal{C}$ given in Corollary 18 holds), the set Λ can be further characterized as follows:

$$\Lambda = \{A \in \Sigma : \rho(B) = \rho(B \cap A) + \rho(B \cap A^c) \text{ for all } B \in \Sigma\}.$$

It follows that in this special case Λ is an algebra, a result that shows that Choquet preferences cannot be used to model some potentially interesting ambiguity situations (see for instance the 4-color example in Zhang [30]).

Ghirardato and Marinacci [15] propose a behavioral notion of unambiguous event for a subclass of invariant biseparable preferences, showing that it has a simple characterization terms of the willingness to bet set-function ρ defined in Section 2: an event B is unambiguous in their sense if and only if $\rho(B) + \rho(B^c) = 1$. The definition given above enjoys two main advantages over this earlier proposal: it is more general, applying to any invariant biseparable preference, and, more importantly, it is more accurate, as it allows to distinguish between events which are truly (perceived) unambiguous and those that appear to be because of the behavior of the DM's ambiguity attitude.

To understand the second point, recall the ambiguity aversion index a of Theorem 11, and notice that if $x \succ y$ and $x' \succ y'$, then $x\,B\,y \asymp x'\,B\,y'$. Hence, we can define, with a slight abuse of notation,

$$a([B]) \equiv a([x\,B\,y]),$$

if $x \succ y$. The coefficient $a([B])$ is interpreted as the ambiguity aversion that the DM reveals when betting on the ambiguous event B. Using this coefficient, it is easy to see that the willingness to bet set-function ρ can be written as follows: for all $B \in \Sigma \setminus \Lambda$,

$$\rho(B) = a([B]) \min_{P \in \mathcal{C}} P(B) + (1 - a([B])) \max_{P \in \mathcal{C}} P(B).$$

(Clearly, for $B \in \Lambda$, $\rho(B) = \min_{P \in \mathcal{C}} P(B) = \max_{P \in \mathcal{C}} P(B)$.) The following general result holds:

Proposition 22 *For all $B \in \Sigma$, $\rho(B) + \rho(B^c) = 1$* if and only if *either $B \in \Lambda$ or $B \in \Sigma \setminus \Lambda$ and $a([B]) + a([B^c]) = 1$. In particular, if $\succcurlyeq$ is an α-MEU preference with $\alpha \neq 1/2$, then* $B \in \Lambda$ if and only if $\rho(B) + \rho(B^c) = 1$.

Thus, the fact that $\rho(B) + \rho(B^c) = 1$ allows us to conclude that either the event B is unambiguous, or the DM's ambiguity attitude when betting on or against B exactly compensate (notice that when $a(\cdot) \equiv 1/2$, $\rho(B) + \rho(B^c) = 1$ for every $B \in \Sigma$). Nehring [23] presents an example in which the latter phenomenon obtains, so that complement additivity of ρ holds also for events which are not unambiguous.

While the proposition illustrates how to easily construct such examples, it suggests that they will be quite uncommon. In fact, in addition to the case of α-MEU preferences with $\alpha \neq 1/2$ mentioned in the proposition, it is easy to think of general preference models for which $a([B]) + a([B^c]) \neq 1$ for every ambiguous B.[11] For instance, consider DMs who are consistently ambiguity averse (resp. loving) in the sense of having $a([f]) > 1/2$ (resp. $a([f]) < 1/2$)) for all the $f \in \mathfrak{F} \setminus [x]$ (of which 1-MEU and its symmetric opposite 0-MEU are clearly special cases). For all such preferences, finding the events over which ρ is complement additive provides a convenient way of identifying the set Λ.

5.2 Unambiguous Acts

In view of the fact that Λ is a λ-system, it is natural to call "unambiguous" the acts whose upper level sets are unambiguous events (cf., e.g., Epstein and Zhang [9]).

Definition 23 *Act $f \in \mathfrak{F}$ is* ***unambiguous*** *if its upper sets $\{s \in S : f(s) \succcurlyeq x\}$ belong to Λ for all $x \in X$. The set of all the unambiguous acts is denoted $\mathfrak{U}$.*

Some simple properties of unambiguous acts follow.

Proposition 24 *For any $f \in \mathfrak{F}$, the following statements are equivalent:*

(i) *f is unambiguous.*

(ii) *$\{s \in S : u(f(s)) \geq a\} \in \Lambda$ for all $a \in \mathbb{R}$.*

[11] It is not hard to verify that the condition $a([B]) + a([B^c]) \neq 1$ (that, recall, is defined only for $B \in \Sigma \setminus \Lambda$) is equivalent to the following behavioral condition: for some $x \succ y$,

$$\frac{1}{2}c(x\,B\,y) + \frac{1}{2}c(y\,B\,x) \nsim \frac{1}{2}x + \frac{1}{2}y,$$

where for any $f \in \mathfrak{F}$ we denote by $c(f)$ an arbitrary certainty equivalent of f.

(*iii*) $\{s \in S : u(f(s)) = a\} \in \Lambda$ *for all* $a \in \mathbb{R}$.

(*iv*) $\{s \in S : f(s) \sim x\} \in \Lambda$ *for all* $x \in X$.

(*v*) *There exist a partition* $\{A_1, A_2, ..., A_n\}$ *of* S *in* Λ *and* $x_1 \succ x_2 \succ ... \succ x_n$ *in* X *such that* f *is pointwise indifferent to* $\{x_1, A_1; x_2, A_2; ...; x_n, A_n\}$.

An obvious question to ask at this point is whether crisp acts are unambiguous. It follows from Proposition 20 (*iii*) that every *binary* crisp act is unambiguous. However, the following example shows that this fact does not generalize to acts which pay more than two nonindifferent prizes.

Example 25 *Let* $S = \{s_1, s_2, s_3\}$, *and consider the set* $\mathcal{C}$ *generated by the priors* $P = [1/3, 1/4, 5/12]$ *and* $Q = [1/4, 5/12, 1/3]$. *Clearly the set* $\Lambda = \{\emptyset, S\}$, *so that only the constants belong to* $\mathfrak{U}$. *Consider now an act* $f = \{x, \{s_1\}; y, \{s_2\}; z, \{s_3\}\}$ *such that* $u(x) = 1$, $u(y) = 4$, $u(z) = 7$. *Immediate calculation yields*

$$\int u(f)\, dP = 51/12 = \int u(f)\, dQ. \tag{8}$$

Thus f *is crisp.*

The problem with act f in the example is that Eq. (8) holds because of the specific identity of P and Q, rather than the fact that f is "unambiguous" in any intuitive sense. A simple way to raise a doubt about f's lack of ambiguity is to observe that the act $g = \{y, \{s_1\}; z, \{s_2\}; x, \{s_3\}\}$ that "permutes" payoffs, while being measurable with respect to the same partition as f, does not satisfy the analogue of Eq. (8). This conflicts with the mentioned intuition that ambiguity is a property of the event partition with respect to which an act is measurable.

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

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

Notice that a binary act $x\,B\,y$ satisfies the permutation crispness condition if and only if B is unambiguous. The non-indifference condition in the converse is the reason of the qualifier "basically" above. To see why permutation crispness may fail for unambiguous acts with indifferent outcomes, consider the following:

Example 27 *Let $S = \{s_1, s_2, s_3\}$, and consider the set $\mathcal{C}$ generated by the two probabilities $P = [1/3, 2/3, 0]$ and $Q = [1/3, 0, 2/3]$. Clearly the set $\Lambda = \{\emptyset, \{s_1\}, \{s_2, s_3\}, S\}$. Consider now an act $f = \{x, \{s_1\}; y, \{s_2\}; z, \{s_3\}\}$ such that $u(x) = 1$, $u(y) = u(z) = 0$. By definition f is unambiguous, but when we consider the (permuted) act $f^\sigma = \{y, \{s_1\}; x, \{s_2\}; z, \{s_3\}\}$, we have*

$$\int u(f^\sigma)\, dP = 2/3 > 0 = \int u(f^\sigma)\, dQ.$$

However, it is still possible to use the proposition to obtain a full characterization of the relation between crisp and unambiguous acts. In fact, it can be shown that for any unambiguous f with some indifferent payoffs, there is an act f' with non-indifferent payoffs which is state-wise indifferent to f, whose permutations all have constant expected utility over the set $\mathcal{C}$.

6 Some Additional Derived Concepts

We now introduce some additional concepts derived from the unambiguous preference relation, which, besides being intrinsically interesting, are useful in the following developments of the ideas in the first part of the paper.

6.1 Mixture Certainty Equivalents

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

$$C(f) = \{x \in X : x \sim f\}.$$

In our setting $C(f) \neq \emptyset$ for every $f \in \mathfrak{F}$.

However, there is another notion of certainty equivalent that arises naturally in this context, linked to the relation $\mathrel{\succcurlyeq\!\!\!\!\cup}$. Denote by $N(f)$ the set of the consequences that are "indifferent" to f in the following sense:

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

Intuitively, these are the constants that correspond to possible certainty equivalents of f. (Recall that $x \mathrel{\succcurlyeq\!\!\!\!\cup} y$ if and only if $x \succcurlyeq y$.)

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

Proposition 28 *For every $f \in \mathfrak{F}$,*

$$x \in N(f) \Longleftrightarrow \min_{P \in \mathcal{C}} P(u(f)) \le u(x) \le \max_{P \in \mathcal{C}} P(u(f)).$$

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

It follows immediately from the proposition that $x \in N(f)$ if and only if there is a $P \in \mathcal{C}$ such that $u(x) = P(u(f))$. That is, $u(N(f))$ is the range of the expected utility mapping of f: the collection of the possible expected utilities of f corresponding to the scenarios in $\mathcal{C}$.

It is simple to convince oneself that $N(f) \supseteq C(f)$. This is not surprising, since $N(f)$ corresponds to one possible way of defining the "certainty equivalents" of f according to the incomplete relation $\succcurlyeq^*$, while $C(f)$ is the set of the certainty equivalents of f according to its completion $\succcurlyeq$.

There is another, behaviorally more interesting, way to see that the elements of $N(f)$ are generalized certainty equivalents of f. Consider a consequence $x \in X$ that can be substituted to f as a "payoff" in a given mixture. That is, such that for some $\lambda \in (0, 1]$ and $h \in \mathfrak{F}$,

$$x \,\lambda\, h \sim f \,\lambda\, h.$$

The following result shows that, while not all the elements of the set $N(f)$ can in general be expressed in this fashion, each of them is infinitesimally close (in terms of preference) to a consequence with this property.[12]

Proposition 29 *For every $f \in \mathfrak{F}$, $N(f)$ is the preference closure of the set*

$$\{x \in X : \exists \lambda \in (0, 1],\, \exists h \in \mathfrak{F} \text{ such that } x \,\lambda\, h \sim f \,\lambda\, h\}.$$

In light of this result, we abuse terminology somewhat and call $x \in N(f)$ a **mixture certainty equivalent** of f, and $N(f)$ the **mixture certainty equivalents set** of f.

6.2 Lower and Upper Envelope Preferences

Given the unambiguous preference $\succcurlyeq^*$ induced by $\succcurlyeq$, we can also define the following two relations:

Definition 30 *The **lower envelope preference** is the binary relation $\succcurlyeq^{\downarrow}$ on $\mathfrak{F}$ defined as follows: for all $f, g \in \mathfrak{F}$,*

$$f \succcurlyeq^{\downarrow} g \Longleftrightarrow \{x \in X : f \succcurlyeq^* x\} \supseteq \{x \in X : g \succcurlyeq^* x\}.$$

[12] In the statement by "preference closure" of a subset $Y \subseteq X$, we mean $u^{-1}\left(\overline{u(Y)}\right)$.

The ***upper envelope preference*** *is the binary relation* $\succsim^{\uparrow}$ *on* $\mathfrak{F}$ *defined as follows: for all* $f, g \in \mathfrak{F}$,

$$f \succsim^{\uparrow} g \Longleftrightarrow \{x \in X : x \,\underline{\mathrel{\cup\!\!\!\!\succ}}\, f\} \subseteq \{x \in X : x \,\underline{\mathrel{\cup\!\!\!\!\succ}}\, g\}.$$

The relation $\succsim^{\downarrow}$ describes a "pessimistic" evaluation rule, while $\succsim^{\uparrow}$ an "optimistic" evaluation rule. To see this, notice that $\succsim^{\downarrow}$ ranks acts by the size of the set of consequences that are unambiguously worse than f. In fact, it ranks f exactly as the most valuable consequence that is unambiguously worse than f. The twin relation $\succsim^{\uparrow}$ does the opposite. We denote by $\succ^{\downarrow}$ and $\sim^{\downarrow}$ (resp. $\succ^{\uparrow}$ and $\sim^{\uparrow}$) the asymmetric and symmetric component of $\succsim^{\downarrow}$ (resp. $\succsim^{\uparrow}$) respectively.

This is further clarified by the following result, which shows that the envelope relations can be represented in terms of the set $\mathcal{C}$ derived in the previous section. Given a Lipschitz functional $I : B_0(\Sigma)$, call its **Clarke (upper) derivative** in φ in the direction ξ the functional

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

Proposition 31 *For every* $f, g \in \mathfrak{F}$, *the following statements are equivalent:*

(i) $f \succsim^{\downarrow} g$ *(resp.* $f \succsim^{\uparrow} g$*).*

(ii) $\min_{P \in \mathcal{C}} P(u(f)) \geq \min_{P \in \mathcal{C}} P(u(g))$ *(resp.* $\max_{P \in \mathcal{C}} P(u(f)) \geq \max_{P \in \mathcal{C}} P(u(g))$*).*

(iii) $I_{\circ}(0; u(f)) \geq I_{\circ}(0; u(g))$ *(resp.* $I^{\circ}(0; u(f)) \geq I^{\circ}(0; u(g))$*).*

It follows from this result that $\succsim^{\downarrow}$ is a 1-MEU preference (i.e., it satisfies ambiguity hedging alongside axioms 1–5). Moreover, while for every $x, y \in X$, $x \succsim y$ if and only if $x \succsim^{\downarrow} y$, $\succsim = \succsim^{\downarrow}$ holds if and only if $\succsim$ is 1-MEU, so that $\succsim$ and $\succsim^{\downarrow}$ will be in general distinct. Symmetric observations hold for $\succsim^{\uparrow}$.

The relations between $\succsim^{\downarrow}$, $\succsim^{\uparrow}$ and $\succsim$ can be better understood by recalling the relative ambiguity aversion ranking of Ghirardato and Marinacci [15].

Proposition 32 *The preference relation* $\succsim^{\downarrow}$ *is more ambiguity averse than* $\succsim$, *which is in turn more ambiguity averse than* $\succsim^{\uparrow}$.

Therefore, the envelope relations (and the respective Clarke derivatives) can be interpreted as the "ambiguity averse side" and the "ambiguity loving side" of the DM. Indeed, $\succsim^{\downarrow}$ is ambiguity averse in the absolute sense of Ghirardato and Marinacci [15], while $\succsim^{\uparrow}$ is ambiguity loving.

7 An Axiomatization of α-MEU Preferences

The general characterization of Theorem 11 suggests an interesting type of invariant biseparable preferences: those whose ambiguity aversion index a is constant. The next axiom characterizes this class of preferences.

Axiom 6 *For every $f, g \in \mathfrak{F}$, $N(f) = N(g)$ implies $f \sim g$.*

The interpretation of the axiom is straightforward. For a DM who satisfies axiom 6, the mixture certainty equivalents set of an act contains *all* the information the DM uses in evaluating it, the specific mapping from states to payoffs does not matter.

Notice that the condition $N(f) = N(g)$ 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 $x \succcurlyeq^* g$ (using the unambiguous preference relation), or $f \sim^{\downarrow} g$ and $f \sim^{\uparrow} g$ (using the envelope preference relations).

In terms of the representation in Eq. (7), axiom 6 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.

Theorem 33 *Let $\succcurlyeq$ be a binary relation on $\mathfrak{F}$. The following statements are equivalent:*

(i) $\succcurlyeq$ satisfies axioms 1–6.

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

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

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.

To understand why this result is true, notice that if the evaluation only depends on the range $[\min_{P \in \mathcal{C}} P(u(f)), \max_{P \in \mathcal{C}} P(u(f))]$, the same is true of the ambiguity aversion coefficient $a([f])$. But we observed earlier that $a([f])$ is unaffected by positive affine transformations of (the utility profile of) f. Given any two acts $f, g \in \mathfrak{F}$, there is always a positive affine transformation of g which has the same pair of minima and maxima of f, so that $a([f]) = a([g])$.

This theorem provides a simple axiomatic characterization of the well-known "α-pessimism" model of Hurwicz, where the set of priors is derived endogenously, and is not exogenously fixed to be the set of all the priors on S. We do not know of any previous axiomatization of this preference rule where the DM's perception of the ambiguity in the problem is endogenously derived.[13]

We call any preference with a representation as in statement (*ii*) of the theorem (or equivalently, satisfying axioms 1–6) an α-MEU preference.

7.1 $(\beta, \mathcal{D})$-MEU Preferences and Uniqueness

We have interpreted the coefficient α in the representation in Theorem 33 as a coefficient of the DM's aversion to perceived ambiguity. For this interpretation to make sense, the uniqueness of α, which follows from that of $\mathcal{C}$ as a representation of the relation $\succcurlyeq^*$ (except in the SEU case), is crucial. That is, the coefficient α is indeed well-defined as an index of aversion to the ambiguity described by $\mathcal{C}$.

However, this does not rule out the possibility that the preference $\succcurlyeq$ may have a similar representation with a *different* coefficient and a *different* set of priors. To better understand this point, consider the following definition:

Definition 34 *A binary relation $\succcurlyeq$ which is represented by the functional*

$$f \mapsto \beta \min_{Q \in \mathcal{D}} \int_S u(f)\, dQ + (1-\beta) \max_{Q \in \mathcal{D}} \int_S u(f)\, dQ$$

for some nonconstant affine function u on X, nonempty, weak compact and convex set $\mathcal{D}$ of probabilities on Σ and $\beta \in [0,1]$ is called a $(\beta, \mathcal{D})$-**MEU preference**.*

Given any nonconstant affine u, it can be seen (Lemma 54 in Appendix B) that a preference represented by a functional I is a $(\beta, \mathcal{D})$-MEU preference if and only if there exists a set $\mathcal{D}$ of probabilities such that

$$\min_{P \in \mathcal{D}} P(\psi) \leq I(\psi) \leq \max_{P \in \mathcal{D}} P(\psi)$$

for all $\psi \in B_0(\Sigma)$ and $I(\psi)$ only depends on $[\min_{P \in \mathcal{D}} P(\psi), \max_{P \in \mathcal{D}} P(\psi)]$. For instance, similar preferences could be observed when the set $\mathcal{D}$ corresponds to some objective probabilistic information available to the DM (cf. also Nehring [23]).

[13] The original axiomatization of Hurwicz (which apparently delivered the constant α representation) has never been published. What has been is a weaker choice-theoretic axiomatization, developed with Arrow [2], in which the pessimism index is a function of the worst and best outcome of the act.

In general, it is possible that the same preference be represented by different pairs $(\beta_1, \mathcal{D}_1)$, $(\beta_2, \mathcal{D}_2)$. The preferences satisfying axioms 1–6 are no exception, as the following example illustrates:

Example 35 *Given an arbitrary set of states* (S, Σ) *and* $X = [0, 100]$, *let* $\mathcal{D} = \{P : P = \lambda Q + (1-\lambda)Q', \exists \lambda \in [0,1]\}$, *with* Q *and* Q' *two distinct probabilities. Consider the* $(1/2, \mathcal{D})$-*MEU preference* $\succcurlyeq$ *defined by*

$$f \succcurlyeq g \Longleftrightarrow \frac{1}{2}\min_{P\in\mathcal{D}} P(f) + \frac{1}{2}\max_{P\in\mathcal{D}} P(f) \geq \frac{1}{2}\min_{P\in\mathcal{D}} P(g) + \frac{1}{2}\max_{P\in\mathcal{D}} P(g).$$

It is easy to convince oneself that $\succcurlyeq$ *also has the SEU representation*

$$f \succcurlyeq g \Longleftrightarrow \int_S f\, d(\frac{1}{2}Q + \frac{1}{2}Q') \geq \int_S g\, d(\frac{1}{2}Q + \frac{1}{2}Q'),$$

so that it is invariant biseparable and satisfies axiom 6. Hence $\succcurlyeq$ *has a representation with set of priors* $\mathcal{D}$ *and ambiguity aversion parameter 1/2, and another one with the singleton set* $\{(1/2)Q + (1/2)Q'\}$, *which is also the set that represents* $\succcurlyeq^*$ *(the corresponding coefficient* α *is arbitrarily defined, of course).*

The preference in the example can be represented as $(\beta, \mathcal{D})$-MEU in two (indeed, infinitely many) different ways.[14] However, consistently with the espoused view that we expect the DM to satisfy independence in the absence of ambiguity, we find the SEU representation more compelling. Why say that the DM perceives ambiguity when he behaves as if he does not?

In the example $\mathcal{C} \subseteq \mathcal{D}$. This is not a coincidence. As the next result shows, the set $\mathcal{C}$ is contained in every $\mathcal{D}$ that represents a $(\beta, \mathcal{D})$-MEU preference.

Proposition 36 *Let* $\succcurlyeq$ *be a* $(\beta, \mathcal{D})$-*MEU preference. Then* $\succcurlyeq$ *satisfies axioms 1–5 and* $\mathcal{D} \supseteq \mathcal{C}$. *If, moreover,* $\succcurlyeq$ *satisfies axiom 6 and perceives some ambiguity, then* $\alpha \geq \beta$ *if* $\beta > 1/2$ *and* $\alpha \leq \beta$ *if* $\beta < 1/2$.

To understand the second statement, notice that when we use $\mathcal{D} \supset \mathcal{C}$ we are (as in the example above) attributing to the DM an inflated perception of ambiguity. We are thus underestimating the magnitude of his reaction to the perceived ambiguity.

Notice that it is always possible to verify whether we are using the correct representation of perceived ambiguity. In fact, it is easy to verify that if $\mathcal{D} \supset \mathcal{C}$, then there is some act $f \in \mathfrak{F}$

[14] The specific example is chosen for expositional convenience. Nothing in it is knife-edge. In particular, it is not crucial to have that $\beta = 1/2$ and that the preference is SEU.

such that the evaluations given by $x \sim^{\downarrow} f$ and $x \sim^{\uparrow} f$ do not coincide with the ones obtained using the representation $\mathcal{D}$.

On the other hand, we observed in Section 4 that Theorem 15 implies that $\mathcal{D} = \mathcal{C}$ if $\beta = 1$ (or $\beta = 0$). Thus, the uniqueness of the representation of perceived ambiguity is not at issue whenever we observe a DM with an ambiguity aversion coefficient of 0 or 1. For this reason, our reference to such preferences simply as 0- and 1-MEU is warranted.

8 Perceived Ambiguity and Updating

We now consider a simple dynamic extension of our static decision problem. Suppose that our DM has an information structure given by some subclass Π of Σ (say, a partition or a sub-algebra), and assume that we can observe our DM's *ex ante* preference on $\mathfrak{F}$, denoted interchangeably $\succcurlyeq$ or $\succcurlyeq_S$, and his preference on $\mathfrak{F}$ after having been informed that an event $A \in \Pi$ obtained, denoted $\succcurlyeq_A$. For each $A \in \Pi' \equiv \Pi \cup S$, the preference $\succcurlyeq_A$ is assumed to be invariant biseparable, and the utility representing $\succcurlyeq_A$ is denoted by u_A. Clearly, a conditional preference $\succcurlyeq_A$ also induces an unambiguous preference relation $\boxed{\succcurlyeq_A}$, as well as mixture certainty equivalents sets $N_A(\cdot)$ and a lower envelope preference relation $\succcurlyeq^{\downarrow}_A$. Because $\succcurlyeq_A$ is invariant biseparable, it is possible to represent $\boxed{\succcurlyeq_A}$ in the sense of Proposition 4 by a nonempty, weak* compact and convex set of probability measures $\mathcal{C}_A$.

We are interested in preferences conditional on events which are (*ex ante*) unambiguously non-null in the following sense:

Definition 37 *We say that $A \in \Sigma$ is **unambiguously non-null** if $x\,A\,y \succ^{\downarrow} y$ for some (all) $x \succ y$.*

That is, an event is unambiguously non-null if betting on A is unambiguosly better than getting the loss payoff y for sure (notice that this is stronger than the definition of non-null event in [14], which just requires that $x\,A\,y \succ y$). This property is equivalently restated in terms of the possible scenarios $\mathcal{C}$ as follows: $P(A) > 0$ for all $P \in \mathcal{C}$.

We next assume that conditional on being informed of A, the DM only cares about an act's results on A, a natural assumption.

Axiom 7 (Consequentialism) *For every $A \in \Pi$, $f \sim_A f\,A\,g$ for every $f, g \in \mathfrak{F}$.*

Consequentialism extends immediately to the unambiguous and lower envelope preference relations, as the following result shows:

Lemma 38 *For every $A \in \Pi$, the following statements are equivalent:*[15]

(i) $f \sim_A f\,A\,g$ *for every* $f, g \in \mathfrak{F}$.

(ii) $f \underset{\cup}{\sim}_A f\,A\,g$ *for every* $f, g \in \mathfrak{F}$.

(iii) $f \sim_A^{\downarrow} f\,A\,g$ *for every* $f, g \in \mathfrak{F}$.

For the remainder of this section we tacitly assume that all the preferences $\succcurlyeq_A$, for $A \in \Pi'$, satisfy axiom 7 (alongside axioms 1–5).

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

$$f\,A\,g \succcurlyeq g \iff f \succcurlyeq_A g. \tag{9}$$

This property imposes two requirements. The first says that the DM should consistently carry out plans made *ex ante*. The second says that information is valuable to the DM, in the sense that postponing her choice to after knowing whether an event obtained does not make her worse off (see Ghirardato [11] for a more detailed discussion).

It is possible to find several plausible instances in which the presence of ambiguity explains behavior that violates dynamic consistency (see Siniscalchi [28] for elaboration). However, we think that *in the absence of ambiguity* dynamic consistency retains much of its intuitive appeal. It thus seems to be a natural exercise to inquire the effect of requiring dynamic consistency of the unambiguous preference relations $\underset{\cup}{\succcurlyeq}_A$, for $A \in \Pi$, with respect to the *ex ante* $\underset{\cup}{\succcurlyeq}$ (that is, requiring Eq. (9) with $\succcurlyeq$ and $\succcurlyeq_A$ replaced by $\underset{\cup}{\succcurlyeq}$ and $\underset{\cup}{\succcurlyeq}_A$ respectively).

We now show that (for a preference satisfying axiom 7) this is tantamount to assuming that the DM updates all the priors in $\mathcal{C}$, a procedure that we call **generalized Bayesian updating**: For every $A \in \Pi$, the "updated" perception of ambiguity is equal to

$$\mathcal{C}|A \equiv \overline{co}^{w^*}\{P_A : P \in \mathcal{C}\},$$

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

Proposition 39 *Suppose that $A \in \Pi$ is unambiguously non-null. Then the following statements are equivalent:*

[15] In this and the remaining results of this section, we omit the equivalent statements involving the upper envelope preference relation.

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

$$f \boxed{\succcurlyeq_A} g \Longleftrightarrow P_A(u(f)) \geq P_A(u(g)) \text{ for all } P \in \mathcal{C}. \tag{10}$$

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

(*ii*) *The relation* $\boxed{\succcurlyeq}$ *is dynamically consistent with respect to* A. *That is, for every* $f, g \in \mathfrak{F}$:

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

(*iii*) *For every* $x, x' \in X$, $x \succcurlyeq x' \Rightarrow x \succcurlyeq_A x'$. *For every* $f \in \mathfrak{F}$ *and* $x \in X$:

$$x \in N_A(f) \Longleftrightarrow x \in N(f\,A\,x). \tag{12}$$

(*iv*) *For every* $f \in \mathfrak{F}$ *and* $x \in X$:

$$f \succcurlyeq^{\downarrow}_A x \Longleftrightarrow f\,A\,x \succcurlyeq^{\downarrow} x. \tag{13}$$

Alongside the promised equivalence with dynamic consistency of unambiguous preference, this results presents two other characterizations of generalized Bayesian updating. They are inspired by a result of Siniscalchi [28], who shows that when the primitive preference relations $\{\succcurlyeq_A\}_{A \in \Pi'}$ are 1-MEU, generalized Bayesian updating is characterized by (a condition equivalent to)

$$f \succcurlyeq_A (\sim_A)\, x \Longleftrightarrow f\,A\,x \succcurlyeq (\sim)\, x \tag{14}$$

for all $f \in \mathfrak{F}$ and $x \in X$. Statement (*iii*) in the proposition departs from the indifference part of Eq. (14) and applies its logic to the "indifference" notion that is generated by the incomplete preference $\boxed{\succcurlyeq}$. Statement (*iv*) is a direct generalization of Siniscalchi's result to preferences that do not satisfy ambiguity hedging. Notice that Eq. (13) is equivalent to requiring that $f \boxed{\succcurlyeq_A} x$ if and only if $fAx \boxed{\succcurlyeq} x$, a weakening of Eq. (11) that under the assumptions of the proposition is equivalent to it.

It is straightforward to show that dynamic consistency of the primitives $\{\succcurlyeq_A\}_{A \in \Pi'}$ implies condition (*ii*). Thus, dynamic consistency of the primitives is a sufficient condition for generalized Bayesian updating. However, it is easy to verify that it is not necessary (Siniscalchi [28] presents an example with a 1-MEU preference).

A different way of reinforcing the conditions of Proposition 39 is to consider imposing the full strength of dynamic consistency on the lower envelope preference relations, rather than the weaker form seen in Eq. (13). We next show that this leads to the characterization of the notion of rectangularity introduced by Epstein and Schneider [7].

Suppose that the class Π forms a finite partition of S; i.e., $\Pi = \{A_1, \ldots, A_n\}$. Given a set of probabilities $\mathcal{C}$ such that each A_i is unambiguously nonnull, we define

$$[\mathcal{C}] = \left\{ P : \exists Q, P_1, \ldots, P_n \text{ such that } \forall B \in \Sigma,\ P(B) = \sum_{i=1}^{n} P_i(B|A_i)\, Q(A_i) \right\}.$$

We say that $\mathcal{C}$ is Π-**rectangular** if $\mathcal{C} = [\mathcal{C}]$.[16] (We refer the reader to Epstein and Schneider [7] for more discussion of this concept.)

Proposition 40 *Suppose that* Π *is a partition and that every* $A \in \Pi$ *is unambiguously nonnull. Then the following statements are equivalent:*

(*i*) $\mathcal{C}$ *is* Π*-rectangular, and* $u_A = u$ *and* $\mathcal{C}_A = \mathcal{C}|A$ *for every* $A \in \Pi$.

(*ii*) *For every* $f, g \in \mathfrak{F}$ *and* $A \in \Pi$:

$$f \succcurlyeq^{\downarrow}_A g \Longleftrightarrow f\, A\, g \succcurlyeq^{\downarrow} g.$$

The rationale for this result is straightforward: Since the preference $\succcurlyeq^{\downarrow}$ is 1-MEU with set of priors $\mathcal{C}$, it follows from the anlysis of Epstein and Schneider [7] that $\mathcal{C}$ is rectangular and that for every $A \in \Pi$, $\mathcal{C}_A$ is obtained by generalized Bayesian updating. By definition of $\succcurlyeq^{\downarrow}_A$, the sets $\mathcal{C}_A$ are also those that represent the ambiguity perception of the primitive relations $\succcurlyeq_A$.

We have therefore shown that the characterization of rectangularity and generalized Bayesian updating of Epstein and Schneider can be extended to preferences which do not satisfy ambiguity hedging, having taken care to require dynamic consistency of the lower envelope (or equivalently of the upper envelope), rather than of the primitive, preference relations. The relations between dynamic consistency of the primitives $\{\succcurlyeq_A\}_{A\in\Pi'}$ and of the lower envelopes $\{\succcurlyeq^{\downarrow}_A\}_{A\in\Pi'}$ are not obvious and are the object of ongoing research.

[16]We owe this presentation of rectangularity to Marciano Siniscalchi.

A Functional Analysis Toolkit

In this appendix we provide/review some functional analytic results and notions that are used to prove the results in the main text (and in some cases directly mentioned in Section 4).

A.1 Conic Preorders

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

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

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

- a **preorder** if it is reflexive and transitive;
- **continuous** if $\varphi^n \succsim \psi^n$ for all $n \in \mathbb{N}$, $\varphi^n \to \varphi$ and $\psi^n \to \psi$ imply $\varphi \succsim \psi$;
- **conic** if $\varphi \succsim \psi$ implies $\alpha\,\varphi + (1-\alpha)\,\theta \succsim \alpha\,\psi + (1-\alpha)\,\theta$ for all $\theta \in B_0(\Sigma, K)$ and all $\alpha \in [0,1]$; [17]
- **monotonic** if $\varphi \geq \psi$ implies $\varphi \succsim \psi$.
- **nontrivial** if there exists $\varphi, \psi \in B_0(\Sigma, K)$ such that $\varphi \succsim \psi$ but not $\psi \succsim \varphi$.

Next, we have some useful representation results.

Proposition 41 *For $i = 1, 2$, let $\mathcal{C}_i$ be nonempty sets of probability charges on Σ and $\succsim_i$ be the relations defined on $B_0(\Sigma, K)$ by*

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

Then

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

and the following statements are equivalent:

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

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

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

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

Proof. If $\varphi \succsim_i \psi$, then $\int \varphi\, dP \geq \int \psi\, dP$ for all $P \in \mathcal{C}_i$. Hence, for all $n \in \mathbb{N}$, all $P_1, P_2, ..., P_n \in \mathcal{C}_i$, and all $\alpha_1, ..., \alpha_n \geq 0$ such that $\sum_1^n \alpha_i = 1$:

$$\sum_1^n \alpha_i \int_S \varphi\, dP_i \geq \sum_1^n \alpha_i \int_S \psi\, dP_i \quad \text{for all } P \in co(\mathcal{C}_i).$$

That is,

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

Let $P \in \overline{co}^{w^*}(\mathcal{C}_i)$. There exists a net $P_a \in co(\mathcal{C}_i)$ such that $P_a(\theta) \to P(\theta)$ for all $\theta \in B_0(\Sigma)$, in particular $P_a(\varphi) \to P(\varphi)$ and $P_a(\psi) \to P(\psi)$, then $P(\varphi) \geq P(\psi)$. Conversely, it is obvious that $\int \varphi\, dP \geq \int \psi\, dP$ for all $P \in \overline{co}^{w^*}(\mathcal{C}_i)$ implies $\int \varphi\, dP \geq \int \psi\, dP$ for all $P \in \mathcal{C}_i$. That is, $\varphi \succsim_i \psi$.

We can now prove (*i*) $\Rightarrow$ (*ii*). In view of what we have just shown, it is w.l.o.g. to assume that $\mathcal{C}_1$ and $\mathcal{C}_2$ are weak* closed and convex. Let $\varphi \succsim_1 \psi \Longrightarrow \varphi \succsim_2 \psi$ and, by contradiction, assume that $P' \in \mathcal{C}_2 \setminus \mathcal{C}_1$. The cone $\mathcal{K}_1 = \{\alpha P : \alpha \geq 0,\ P \in \mathcal{C}_1\}$ generated by $\mathcal{C}_1$ is weak* closed and convex. In fact, let $\alpha P, \beta Q \in \mathcal{K}_1$ (i.e., $\alpha, \beta \geq 0$ and $P, Q \in \mathcal{C}_1$). If $\alpha = \beta = 0$, then $\alpha P + \beta Q = 0 \in \mathcal{K}_1$; else $\alpha P + \beta Q = (\alpha + \beta)(\frac{\alpha}{\alpha+\beta} P + \frac{\beta}{\alpha+\beta} Q) \in \mathcal{C}_1$. Therefore, $\mathcal{K}_1$ is convex. Let $\{\beta_a P_a\}$ be a net in $\mathcal{K}_1$ (i.e., $\beta_a \geq 0$ and $P_a \in \mathcal{C}_1$ for all a) such that $\beta_a P_a \xrightarrow{*} L$; if $m = 0$, then $L \in \mathcal{K}_1$; else $m = \beta P$ for suitable $\beta > 0$ and $P \in pc(\Sigma)$. In particular, $\beta_a P_a(S) \to \beta P(S)$ and $\beta_a \to \beta$, therefore $\frac{1}{\beta_a}(\beta_a P_a) \xrightarrow{*} \frac{1}{\beta}(\beta P)$ so $m = \beta \lim_a P_a \in \mathcal{K}_1$.

Clearly $P' \notin \mathcal{K}_1$ (else $P' = \alpha P$ for some $\alpha \geq 0$ and $P \in \mathcal{C}_1$). Thus, by the Separating Hyperplane Theorem there exists $\varphi \in B_0(\Sigma) \setminus \{0\}$ such that

$$m(\varphi) \geq 0 > P'(\varphi) \quad \text{for all } m \in \mathcal{K}_1. \tag{15}$$

Consider $\kappa \in K^\circ$ (it must exist since K is not a singleton) and $\alpha > 0$ such that $\alpha\varphi + \kappa \in B_0(\Sigma, K)$. Eq. (15) yields $\alpha\varphi + \kappa \succsim_1 \kappa$ and $\alpha\varphi + \kappa \not\succsim_2 \kappa$, which is absurd. The implication (*ii*) $\Rightarrow$ (*i*) is immediate.

Next, we show that, for all $\varphi \in B_0(\Sigma, K)$,

$$\inf_{P\in\mathcal{C}_i} P(\varphi) = \min_{Q\in\overline{co}^{w^*}(\mathcal{C}_i)} Q(\varphi).$$

Clearly, $\inf_{\mathcal{C}_i} P(\varphi) \geq \min_{\overline{co}^{w^*}(\mathcal{C}_i)} Q(\varphi)$. If it were $\inf_{\mathcal{C}_i} P(\varphi) > \min_{\overline{co}^{w^*}(\mathcal{C}_i)} Q(\varphi)$, there would exists $\varepsilon > 0$ such that $\inf_{\mathcal{C}_i} P(\varphi) > \varepsilon > \min_{\overline{co}^{w^*}(\mathcal{C}_i)} Q(\varphi)$. Thus, for all $n \in \mathbb{N}$, all $P_1, P_2, ..., P_n \in \mathcal{C}_i$ and all $\alpha_1, ..., \alpha_n \geq 0$ such that $\sum_1^n \alpha_i = 1$:

$$\sum_1^n \alpha_i P_i(\varphi) \geq \sum_1^n \alpha_i \varepsilon = \varepsilon.$$

That is,

$$\inf_{co(\mathcal{C}_i)} P(\varphi) \geq \varepsilon.$$

But for all $Q \in \overline{co}^{w^*}(\mathcal{C}_i)$ there exists a net $P_a \in co(\mathcal{C}_i)$ such that $P_a(\varphi) \to Q(\varphi)$. Hence, $Q(\varphi) \geq \varepsilon$, which is a contradiction. Analogously we show that

$$\sup_{P \in \mathcal{C}_i} P(\varphi) = \max_{Q \in \overline{co}^{w^*}(\mathcal{C}_i)} Q(\varphi),$$

for all $\varphi \in B_0(\Sigma, K)$. The implication $(ii) \Rightarrow (iii)$ is now obvious.

As to the implication $(iii) \Rightarrow (ii)$, it is again w.l.o.g. to assume that $\mathcal{C}_1$ and $\mathcal{C}_2$ are weak* closed and convex, so that the result follows from an application of the Separating Hyperplane Theorem. □

Proposition 42 $\succsim$ *is a nontrivial, continuous, conic, and monotonic preorder on* $B_0(\Sigma, K)$ *if and only if there exists a nonempty subset* $\mathcal{C}$ *of* $pc(\Sigma)$ *such that*

$$\varphi \succsim \psi \Longleftrightarrow \int_S \varphi \, dP \geq \int_S \psi \, dP \quad \text{for all } P \in \mathcal{C}. \tag{16}$$

Moreover, $\overline{co}^{w^*}(\mathcal{C})$ *is the unique weak* closed and convex subset of* $pc(\Sigma)$ *representing* $\succsim$ *in the sense of Eq. (16).*

Proof. Let $\kappa_0 \in K^\circ$, for all $\varphi, \psi \in B_0(\Sigma, K - \kappa_0)$ set $\varphi \succsim' \psi \Leftrightarrow \varphi + \kappa_0 \succsim \psi + \kappa_0$. It is easy to verify that $\succsim'$ is a nontrivial, continuous, conic, and monotonic preorder on $B_0(\Sigma, K - \kappa_0)$.
Claim. If $\varphi, \psi \in B_0(\Sigma, K - \kappa_0)$, the following facts are equivalent:

(i) $\varphi \succsim' \psi$,

(ii) there exists $\alpha > 0$ such that $\alpha\varphi, \alpha\psi \in B_0(\Sigma, K - \kappa_0)$ and $\alpha\varphi \succsim' \alpha\psi$,

(iii) for all $\alpha > 0$ such that $\alpha\varphi, \alpha\psi \in B_0(\Sigma, K - \kappa_0) : \alpha\varphi \succsim' \alpha\psi$.

Proof of the Claim. $(i) \Rightarrow (ii)$ and $(iii) \Rightarrow (i)$ are obvious. We show $(ii) \Rightarrow (iii)$. By contradiction, assume there exists $\alpha > 0$ such that $\alpha\varphi, \alpha\psi \in B_0(\Sigma, K - \kappa_0)$ and $\alpha\varphi \gtrsim' \alpha\psi$. If $0 < \beta \le \alpha$, then $\beta\varphi = \frac{\beta}{\alpha}\alpha\varphi + (1 - \frac{\beta}{\alpha})0 \gtrsim' \frac{\beta}{\alpha}\alpha\psi + (1 - \frac{\beta}{\alpha})0 = \beta\psi$. Therefore, if there exists $\beta > 0$ such that $\beta\varphi, \beta\psi \in B_0(\Sigma, K - \kappa_0)$ and *not* $\beta\varphi \gtrsim' \beta\psi$, then it have to be $\beta > \alpha > 0$. Let $\bar{\varphi} = \beta\varphi$ and $\bar{\psi} = \beta\psi$, $\bar{\varphi}, \bar{\psi} \in B_0(\Sigma, K - \kappa_0)$ and *not* $\bar{\varphi} \gtrsim' \bar{\psi}$, but $\frac{\alpha}{\beta}\bar{\varphi} \gtrsim' \frac{\alpha}{\beta}\bar{\psi}$. Let $\bar{\gamma} = \sup\{\gamma \in [0,1] : \frac{\alpha}{\beta}\bar{\varphi} \gtrsim' \frac{\alpha}{\beta}\bar{\psi}\}$,by continuity, $\bar{\gamma}\bar{\varphi} \gtrsim' \bar{\gamma}\bar{\psi}$, hence $\delta\bar{\varphi} \gtrsim' \delta\bar{\psi}$ for all $\delta \in [0, \bar{\gamma}]$. Then,

$$\tfrac{1}{1+\bar{\gamma}}\bar{\gamma}\bar{\varphi} + \tfrac{\bar{\gamma}}{1+\bar{\gamma}}\bar{\varphi} \gtrsim' \tfrac{1}{1+\bar{\gamma}}\bar{\gamma}\bar{\psi} + \tfrac{\bar{\gamma}}{1+\bar{\gamma}}\bar{\varphi} = \tfrac{1}{1+\bar{\gamma}}\bar{\gamma}\bar{\varphi} + \tfrac{\bar{\gamma}}{1+\bar{\gamma}}\bar{\psi} \gtrsim' \tfrac{1}{1+\bar{\gamma}}\bar{\gamma}\bar{\psi} + \tfrac{\bar{\gamma}}{1+\bar{\gamma}}\bar{\psi}$$

so that $\frac{2\bar{\gamma}}{1+\bar{\gamma}}\bar{\varphi} \gtrsim' \frac{2\bar{\gamma}}{1+\bar{\gamma}}\bar{\psi}$. But by definition of $\bar{\gamma}$, $\frac{2\bar{\gamma}}{1+\bar{\gamma}} \le \bar{\gamma}$, that is, $\bar{\gamma}^2 - \bar{\gamma} \ge 0$. Since $\bar{\gamma} > 0$, therefore, we have $\bar{\gamma} = 1$, and hence $\bar{\gamma}\bar{\varphi} \gtrsim' \bar{\gamma}\bar{\psi}$, which is absurd. ■

For all $\varphi, \psi \in B_0(\Sigma)$, set $\varphi \gtrsim'' \psi \iff \alpha\varphi \gtrsim' \alpha\psi$ for some (all) $\alpha > 0$ such that $\alpha\varphi, \alpha\psi \in B_0(\Sigma, K - \kappa_0)$. Clearly, if $\varphi, \psi \in B_0(\Sigma, K - \kappa_0)$, then $\varphi \gtrsim'' \psi \iff \varphi \gtrsim' \psi$. It is also easy to verify that $\varphi \gtrsim'' \psi$ is a nontrivial, continuous, conic, and monotonic preorder on $B_0(\Sigma)$.

Let $\mathcal{K}$ be the set of charges on Σ defined by

$$\mathcal{K} = \{m \in ba(\Sigma) : m(\varphi) \ge 0 \text{ for all } \varphi \gtrsim'' 0\}.$$

Clearly $0 \in \mathcal{K}$ and it is trivial to verify that $\mathcal{K}$ is a convex cone. Let $\{m_a\}$ be a net in $\mathcal{K}$ which weak* converges to m, for all $\psi \in B_0(\Sigma)$, $m_a(\psi) \to m(\psi)$, in particular, if $\varphi \gtrsim'' 0$, $m(\varphi) = \lim_a m_a(\varphi) \ge 0$. Hence $\mathcal{K}$ is $\sigma(ba(\Sigma), B_0(\Sigma))$-closed.

If $\varphi \gtrsim'' \psi$, then $\varphi - \psi \gtrsim'' 0$. Hence, $m(\varphi - \psi) \ge 0$ $\forall m \in \mathcal{K}$; that is, $m(\varphi) \ge m(\psi)$ $\forall m \in \mathcal{K}$. Conversely, assume $m(\varphi') \ge m(\psi')$, $\forall m \in \mathcal{K}$ and not $\varphi' \gtrsim'' \psi'$. Consider the closed convex cone defined by

$$C(\gtrsim'') \equiv \{\varphi \in B_0(\Sigma) : \varphi \gtrsim'' 0\}.$$

Then $\xi' = \varphi' - \psi' \notin C(\gtrsim'')$ and, by the Separating Hyperplane Theorem there exists $m' \in ba(\Sigma)$ such that

$$m'(\varphi) \ge 0 > L'(\xi') \quad \forall \varphi \in C(\gtrsim'').$$

That is, $m' \in \mathcal{K}$ and $m'(\xi') < 0$, so $m'(\varphi') < m'(\psi')$, which is absurd. We have thus shown that

$$\varphi \gtrsim'' \psi \iff m(\varphi) \ge m(\psi) \quad \text{for all } m \in \mathcal{K}.$$

Since for all $A \in \Sigma$, $1_A \ge 0$, monotonicity implies that $m(A) = m(1_A) \ge 0$. Thus, $\mathcal{K}$ consists of nonnegative elements of $ba(\Sigma)$. Since $\mathcal{K} = \{0\}$ contradicts nontriviality, we have $\mathcal{K} \ne \{0\}$ and

$$\varphi \gtrsim'' \psi \iff m(\varphi) \ge m(\psi) \quad \text{for all } m \in \mathcal{K} \setminus \{0\}.$$

That is,

$$\varphi \succsim'' \psi \Longleftrightarrow \frac{m(\varphi)}{m(S)} \geq \frac{m(\psi)}{m(S)} \quad \text{for all } m \in \mathcal{K} \setminus \{0\}.$$

Letting $\mathcal{C}$ denote the set of all the probability charges in $\mathcal{K}$, we have

$$\varphi \succsim'' \psi \Longleftrightarrow P(\varphi) \geq P(\psi) \quad \text{for all } P \in \mathcal{C}.$$

$\mathcal{C}$ is weak* closed and convex since $\mathcal{C} = \mathcal{K} \cap pc(\Sigma)$, the intersection of two weak* closed and convex sets. Moreover, for all $\varphi, \psi \in B_0(\Sigma, K)$,

$$\begin{aligned} \varphi \succsim \psi \quad &\Longleftrightarrow \quad \varphi - \kappa_0 \succsim' \psi - \kappa_0 \\ &\Longleftrightarrow \quad P(\varphi - \kappa_0) \geq P(\psi - \kappa_0) \quad \text{for all } P \in \mathcal{C} \\ &\Longleftrightarrow \quad P(\varphi) \geq P(\psi) \quad \text{for all } P \in \mathcal{C}, \end{aligned}$$

as wanted. Uniqueness follows from Proposition 41. □

A.2 Clarke Derivatives and Differentials

Recall that a functional $I : B_0(\Sigma) \to \mathbb{R}$ is:

- **monotonic** if $I(\varphi) \geq I(\psi)$ for all $\varphi \geq \psi$;
- **constant additive** if $I(\varphi + \alpha) = I(\varphi) + \alpha$ for all $\varphi \in B_0(\Sigma)$ and $\alpha \in \mathbb{R}$;
- **positively homogeneous** if $I(\alpha\varphi) = \alpha I(\varphi)$ for all $\varphi \in B_0(\Sigma)$ and $\alpha \in \mathbb{R}_+$;
- **constant linear** if it is constant additive and positively homogeneous.

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

Thus, given a monotonic constant linear functional $I : B_0(\Sigma) \to \mathbb{R}$, we can study its Clarke derivatives and Clarke differentials as defined in Clarke [5]:

Definition 43 *The **Clarke (upper) derivative** of I in φ in the direction v is*

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

Notice that the expression on the r.h.s. is equal to

$$\lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\psi-\varphi\|\le\delta\varepsilon\\ 0<t<\gamma\varepsilon}}\frac{I(\psi+t\upsilon)-I(\psi)}{t}\right)$$

with δ and γ fixed positive numbers (which is easily shown to be independent of δ and γ).

Definition 44 *The **Clarke differential** of I in φ is the set*

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

The next result sums up some basic results about the Clarke derivative and differential drawn from Clarke [5].

Proposition 45 *Let $I: B_0(\Sigma)\to\mathbb{R}$ be Lipschitz of rank r. Then:*

1. *For all $\varphi\in B_0(\Sigma)$, the function $I^\circ(\varphi;\cdot): B_0(\Sigma)\to\mathbb{R}$ is finite, sublinear and Lipschitz of rank r. Moreover, if $\varphi_i\to\varphi$ and $\upsilon_i\to\upsilon$, then*

$$\limsup_{i\to\infty} I^\circ(\varphi_i;\upsilon_i)\le I^\circ(\varphi;\upsilon).$$

2. *For all $\varphi,\upsilon\in B_0(\Sigma)$, $I^\circ(\varphi;-\upsilon)=(-I)^\circ(\varphi;\upsilon)$.*

3. *$\partial I(\varphi)$ is a nonempty, convex, weak* compact subset of $ba(\Sigma)$ and $\|m\|_{BV}\le r$ for every $m\in\partial I(\varphi)$.*

4. *For every $\upsilon\in B_0(\Sigma)$,*

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

5. *For all $\varphi,\psi\in B_0(\Sigma)$, there exists $\gamma\in(0,1)$ such that*

$$I(\varphi)-I(\psi)\in\left\langle\partial I(\gamma\varphi+(1-\gamma)\psi),\varphi-\psi\right\rangle.$$

6. *If I is also convex, then Clarke derivatives are the usual directional derivatives and Clarke differentials are subdifferentials. That is, $I^\circ(\varphi;\upsilon)=\lim_{t\downarrow 0}t^{-1}(I(\varphi+t\upsilon)-I(\varphi))$ for all $\varphi,\upsilon\in B_0(\Sigma)$, and $\partial I(\varphi)=\{m\in ba(\Sigma): m(\upsilon)-m(\varphi)\le I(\upsilon)-I(\varphi)\ \forall\upsilon\in B_0(\Sigma)\}$ for all $\varphi\in B_0(\Sigma)$.*

7. *If $J: B_0(\Sigma)\to\mathbb{R}$ is another Lipschitz functional and $\lambda,\mu>0$, then*

$$\partial(\lambda I+\mu J)(\varphi)\subseteq\lambda\partial I(\varphi)+\mu\partial J(\varphi).$$

It will be recalled that in Section 4 we also defined a Clarke lower derivative, and defined the Clarke differential in terms of that. The next result shows that this (presentational) choice does not make a difference.

Lemma 46 *For every $\varphi \in B_0(\Sigma)$, $\upsilon \in B_0(\Sigma)$,*

$$I_\circ(\varphi; -\upsilon) = -I^\circ(\varphi; \upsilon).$$

In particular,

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

Proof. For every $\varphi, \upsilon \in B_0(\Sigma)$, we have

$$\begin{aligned}
I^\circ(\varphi; -\upsilon) &= (-I)^\circ(\varphi; \upsilon) = \\
&= \lim_{\varepsilon \to 0} \left(\sup_{\substack{\|\psi - \varphi\| \le \delta\varepsilon \\ 0 < t < \gamma\varepsilon}} \frac{-I(\psi + t\upsilon) + I(\psi)}{t} \right) \\
&= \lim_{\varepsilon \to 0} \left(- \inf_{\substack{\|\psi - \varphi\| \le \delta\varepsilon \\ 0 < t < \gamma\varepsilon}} \frac{I(\psi + t\upsilon) - I(\psi)}{t} \right) \\
&= - \lim_{\varepsilon \to 0} \left(\inf_{\substack{\|\psi - \varphi\| \le \delta\varepsilon \\ 0 < t < \gamma\varepsilon}} \frac{I(\psi + t\upsilon) - I(\psi)}{t} \right) \\
&= -I_\circ(\varphi; \upsilon).
\end{aligned}$$

□

It follows that the set defined in Section 4 coincides with the Clarke differential as defined above. For easier reference to the existing literature in the rest of this subsection we use the traditional I°, but the lemma can be used to adapt all the results to $I_\circ$.

We next prove some additional properties of I° and $\partial I(\cdot)$ that we use below, in particular for our case of interest of a monotonic and constant linear I.

Proposition 47 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a Lipschitz functional. Then:*

1. *$I^\circ(\varphi; \cdot) = I^\circ(\alpha\varphi; \cdot)$ for all $\alpha > 0$, and $\partial I(\varphi) \subseteq \partial I(0)$ for all $\varphi \in B_0(\Sigma)$.*

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

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

Proof. 1.) For all $\alpha > 0$,

$$\begin{aligned}
I^\circ(\alpha\varphi; \upsilon) &= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\psi-\alpha\varphi\|\le\delta\varepsilon\\ 0<t<\gamma\varepsilon}} I\left(\frac{\psi}{t}+\upsilon\right) - I\left(\frac{\psi}{t}\right)\right)\\
&= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\theta-\varphi\|\le\frac{\delta}{\alpha}\varepsilon\\ 0<t<\gamma\varepsilon}} I\left(\frac{\alpha\theta}{t}+\upsilon\right) - I\left(\frac{\alpha\theta}{t}\right)\right)\\
&= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\theta-\varphi\|\le\frac{\delta}{\alpha}\varepsilon\\ 0<t<\gamma\varepsilon}} I\left(\frac{\theta}{\alpha^{-1}t}+\upsilon\right) - I\left(\frac{\theta}{\alpha^{-1}t}\right)\right)\\
&= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\theta-\varphi\|\le\frac{\delta}{\alpha}\varepsilon\\ 0<t'<\frac{\gamma}{\alpha}\varepsilon}} I\left(\frac{\theta}{t'}+\upsilon\right) - I\left(\frac{\theta}{t'}\right)\right) = I^\circ(\varphi;\upsilon).
\end{aligned}$$

That is, $I^\circ(\alpha\varphi;\upsilon) = I^\circ(\varphi;\upsilon)$ for all $\varphi, \upsilon \in B_0(\Sigma)$ and all $\alpha > 0$. Therefore,

$$\begin{aligned}
\max_{m\in\partial I(0)} m(\upsilon) = I^\circ(0;\upsilon) &= I^\circ\left(\lim_{n\to\infty}\frac{1}{n}\varphi;\upsilon\right)\\
&\ge \limsup_{n\to\infty} I^\circ\left(\frac{1}{n}\varphi;\upsilon\right) = I^\circ(\varphi;\upsilon) = \max_{m\in\partial I(\varphi)} m(\upsilon),
\end{aligned}$$

whence $\partial I(\varphi) \subseteq \partial I(0)$.

2.) If $\upsilon \le 0$, then $\psi + t\upsilon \le \psi$ for all $\psi \in B_0(\Sigma)$ and all $t > 0$. Hence, $(I(\psi+t\upsilon) - I(\psi))/t \le 0$ for all $\psi \in B_0(\Sigma)$ and all $t > 0$, so that

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

If $\upsilon \ge 0$ and $m \in \partial I(\varphi)$, $-\upsilon \le 0$ and $m(-\upsilon) \le I^\circ(\varphi;-\upsilon) \le 0$, hence $m(\upsilon) \ge 0$. If $\upsilon \ge \upsilon'$, then

$$I^\circ(\varphi;\upsilon) = \max_{m\in\partial I(\varphi)} m(\upsilon) \ge \max_{m\in\partial I(\varphi)} m(\upsilon') = I^\circ(\varphi;\upsilon').$$

3.) Since $I^\circ(\varphi;\cdot)$ is sublinear, it suffices to prove that it is constant additive. For all $\varphi, \upsilon \in B_0(\Sigma)$ and all $\beta \in \mathbb{R}$,

$$\begin{aligned} I^\circ(\varphi;\upsilon+\beta) &= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\psi-\varphi\|\le\delta\varepsilon \\ 0<t<\gamma\varepsilon}} \frac{I(\psi+t(\upsilon+\beta))-I(\psi)}{t}\right) \\ &= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\psi-\varphi\|\le\delta\varepsilon \\ 0<t<\gamma\varepsilon}} \frac{I(\psi+t\upsilon)+t\beta-I(\psi)}{t}\right) \\ &= \lim_{\varepsilon\to 0}\left(\sup_{\substack{\|\psi-\varphi\|\le\delta\varepsilon \\ 0<t<\gamma\varepsilon}} \frac{I(\psi+t\upsilon)-I(\psi)}{t}\right)+\beta \\ &= I^\circ(\varphi;\upsilon)+\beta \end{aligned}$$

and $I^\circ(\varphi;\cdot)$ is constant additive. In particular, $I^\circ(\varphi;1_S)+I^\circ(\varphi;-1_S)=I^\circ(\varphi;1_S-1_S)=0$. For all $m\in\partial I(\varphi)$, $m(S)\le I^\circ(\varphi;1_S)=-I^\circ(\varphi;-1_S)\le -m(-1_S)=m(S)$. That is, $m(S)=I^\circ(\varphi;1_S)=1$. □

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

The next lemma characterizes the Clarke derivatives in terms of increments of the functional I.

Lemma 48 *Let $I : B_0(\Sigma)\to\mathbb{R}$ be a monotonic, constant linear functional. Then for all $\varphi\in B_0(\Sigma)$,*

$$I^\circ(0;\varphi)=\sup_{\psi\in B_0(\Sigma)} I(\psi+\varphi)-I(\psi) \quad \textit{and} \quad I_\circ(0;\varphi)=\inf_{\psi\in B_0(\Sigma)} I(\psi+\varphi)-I(\psi) \tag{17}$$

Proof. By homogeneity,

$$\begin{aligned} I^\circ(0;\varphi) &= \limsup_{\substack{\psi\to 0\\ t\downarrow 0}} \frac{I(\psi+t\varphi)-I(\psi)}{t} \\ &= \limsup_{\substack{\psi\to 0\\ t\downarrow 0}} I\left(\frac{\psi}{t}+\varphi\right)-I\left(\frac{\psi}{t}\right)\le \sup_{\psi\in B_0(\Sigma)}(I(\psi+\varphi)-I(\psi)). \end{aligned}$$

On the other hand, by the Mean Value Theorem (point 4 of Proposition 45), for each $\psi\in B_0(\Sigma)$ there exists $\gamma\in(0,1)$ such that, letting $\xi=\gamma(\psi+\varphi)+(1-\gamma)\psi$,

$$I(\psi+\varphi)-I(\psi)\in\langle\partial I(\xi),\varphi\rangle\,.$$

By point 1 of Proposition 47, $\partial I(\xi) \subseteq \partial I(0)$. Hence, there is $L \in \partial I(0)$ such that $I(\psi + \varphi) - I(\psi) = L(\varphi)$. Hence, $I(\psi + \varphi) - I(\psi) \leq I^\circ(0; \varphi)$ for all $\psi \in B_0(\Sigma)$, which implies $\sup_{\psi \in B_0(\Sigma)}(I(\psi + \varphi) - I(\psi)) \leq I^\circ(0; \varphi)$. We conclude that the equation on the l.h.s. of Eq. (17) holds. The proof for the other equation is analogous. $\square$

We conclude this appendix by recalling Clarke's characterization of Clarke differentials in finite dimensional spaces. Here we assume that $B_0(\Sigma) = \mathbb{R}^n$ and denote the standard gradients by ∇ as customary.

Theorem 49 (Clarke, 2.5.1.) *Let $I : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz functional, and let Ω be any set of Lebesgue measure 0 in $\mathbb{R}^n$ such that I is differentiable on Ω^c. Then*

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

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

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

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

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

B.1 Proof of Proposition 3

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

Next, we prove point 5. Given $f, g, h \in \mathfrak{F}$ and $\lambda \in (0,1)$, suppose that $f \succcurlyeq^{\cup} g$. Then for every $\mu \in (0,1]$ and every $k \in \mathfrak{F}$, 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 $\succcurlyeq^{\cup}$. 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 $\lambda f + (1-\lambda)h \succcurlyeq^{\cup} \lambda g + (1-\lambda)h$, since the choice of μ and k was arbitrary. The case $\lambda = 1$ is trivial. Point 6 follows immediately from the following Proposition 4. (It is not used in the proof of that proposition.)

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

B.2 Proof of Proposition 4

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

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

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

We have thus shown that $\gtrsim$ is a conic, continuous, monotonic, nontrivial preorder on $B_0(\Sigma, u(X))$. By Lemma 42 it follows that there exists a nonempty, weak* closed and convex set $\mathcal{C}$ of probability charges on Σ such that

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

which immediately yields the statement.

B.3 Proof of Proposition 5

Lemma 50 *Let Y be a vector space and u, v be two nonzero linear functionals on Y. One and only one of the following statements is true:*

- $u = av$ *for some* $a > 0$.

- $\exists y \in Y : u(y)\, v(y) < 0$.

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

Then $Y = [uv > 0] \cup [u = 0] \cup [v = 0] = [uv > 0] \cup \ker u \cup \ker v$. $\ker u$ and $\ker v$ are maximal subspaces of Y, hence $Y = \langle z \rangle \oplus \ker u$ for some $z \in Y$ such that $u(z) > 0$. If $\ker u = \ker v$: for all $y \in Y$, exist $b \in \mathbb{R}$, $x \in \ker u$ such that $y = bz + x$, whence $u(y) = b\, u(z) = \dfrac{u(z)}{v(z)} b\, v(z) = \dfrac{u(z)}{v(z)} v(y)$, 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. □

Corollary 51 *Let X be a nonempty convex subset of a vector space and u, v be two non-constant affine functionals on X. There exist $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$ such that $u = av + b$ iff $u(x_1) \geq u(x_2) \Longrightarrow v(x_1) \geq v(x_2)$ for every $x_1, x_2 \in X$.*

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

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

is a vector space and the functionals

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

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

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

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

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

so

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

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

Proof of Proposition 5.

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

$$u_1(x) \geq u_1(y) \Longleftrightarrow x \succcurlyeq_1 y \Longrightarrow x \succcurlyeq^*_1 y \Longrightarrow x \succcurlyeq^*_2 y \Longrightarrow x \succcurlyeq_2 y \Longleftrightarrow u_2(x) \geq u_2(y).$$

By Corollary 51, this implies that we can assume $u_1 = u_2 = u$. Moreover, for all $f, g \in \mathfrak{F}$, $f \succcurlyeq^*_1 g \Longrightarrow f \succcurlyeq^*_2 g$. That is,

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

which by Lemma 41 (applied to $B_0(\Sigma, u(X))$) implies $\mathcal{C}_2 \subseteq \mathcal{C}_1$.

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

B.4 Proof of Proposition 6

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

B.5 Proof of Proposition 7

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

Lemma 52 *For all $f \in \mathfrak{F}$,*

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

and

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

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

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

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

or

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

Therefore,

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

from which we obtain

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

Finally,

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

Analogously,

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

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

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

Then,

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

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

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

Analogously,

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

which concludes the proof.

B.6 Proof of Lemma 8

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

$$\lambda\, f + (1-\lambda)\, x \Cup \lambda'\, g + (1-\lambda')\, x',$$

which, applying Eq. (6) of Proposition 4, is equivalent to

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

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

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

so that we get the conclusion by letting

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

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

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

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

$$f \Cup \lambda\, g + (1-\lambda)\, x.$$

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

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

and proceeding as above to get

$$\lambda\, f + (1-\lambda)\, x \Cup g.$$

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

$$\frac{1}{2} f + \frac{1}{2} x \Cup \frac{1}{2} g + \frac{1}{2} x'.$$

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

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

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

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

B.7 Proof of Proposition 10

$(i) \Rightarrow (iii)$: By Lemma 52,

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

and

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

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

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

That is,

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

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

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

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

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

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

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

so that again

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

We conclude that if k is crisp

$$\underline{\mathcal{C}}(u(k)) = \inf_{\varphi\in B_0(\Sigma)} \{I(u(k)+\psi)-I(\psi)\} = \sup_{\varphi\in B_0(\Sigma)} \{I(u(k)+\psi)-I(\psi)\} = \overline{\mathcal{C}}(u(k)). \qquad (18)$$

$(iii) \Rightarrow (iv)$: From Eq. (18) (which is (iii)) we obtain

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

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

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

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 4, is equivalent to

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

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

$(iii) \Rightarrow (ii)$: Since $P(u(k)) = \gamma$ for every $P \in \mathcal{C}$, we just need to choose $x \in X$ such that $u(x) = \gamma$, and then apply Proposition 4 to see that $k \sim^* x$, yielding (ii).

B.8 Proof of Theorem 11

$(i) \Rightarrow (ii)$: Suppose that $\succcurlyeq$ satisfies axioms 1-5. Let I and u respectively be the preference functional and utility that represent $\succcurlyeq$ obtained in Lemma 1, and $\mathcal{C}$ the weak* compact and convex set of probabilities on Σ that represents $\succcurlyeq^*$ obtained in Proposition 4.

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

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

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

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

If we now recall the consequence of Lemma 8 and Proposition 10 that $[x]$ is the set of all crisp acts, we see that the function $a(\cdot)$ provides the sought representation. We are therefore done if we prove that a can be defined on $\mathfrak{F}_{/\asymp} \setminus \{\mathfrak{C}\} = \mathfrak{F}_{/\asymp} \setminus \{[x]\}$.

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

$$\lambda\, f + (1 - \lambda)\, x \,\bigcup\!\!\!\!\smile\, \lambda'\, g + (1 - \lambda')\, x'. \tag{19}$$

It follows from point 1 of Proposition 3 that Eq. (19) 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. (19) also implies that for every $P \in \mathcal{C}$,

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

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

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

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

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

B.9 Proof of Proposition 12

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

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

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

That is,

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

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

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

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

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

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

B.10 Proof of Theorem 15

For all $f \in \mathfrak{F}$, Lemma 52 yields

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

while Lemma 48 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 $f \in \mathfrak{F}$ such that $\lambda\varphi = u(f)$. Hence,

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

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

B.11 Proof of Proposition 17

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

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

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

for all $\xi \in B_0(\Sigma)$, so $I^\circ(0;\cdot) = I(\cdot)$ and I is convex. Conversely, if I is convex, point 6 of Proposition 45 guarantees that $\partial I(0) = \mathit{Anticore}(I)$. (The concave case is analogous.)

B.12 Proof of Corollary 18

Notice that the set $\Omega = \{\varphi \in B_0(\Sigma) : \varphi(s_i) = \varphi(s_j) \text{ for some } i \neq j\}$ is the union of $n!$ hyperplanes, and hence its Lebesgue measure is 0, moreover, $\Omega^c = \{\varphi \in B_0(\Sigma) : \varphi \text{ is injective}\}$ is an open and dense subset of $B_0(\Sigma)$. For all $\varphi \in \Omega^c$, there exist a unique $\sigma \in$Per(n) s.t. $\varphi(s_{\sigma(1)}) > \varphi(s_{\sigma(2)}) > \cdots > \varphi(s_{\sigma(n)})$ and a number $\varepsilon > 0$ s.t. $\psi(s_{\sigma(1)}) > \psi(s_{\sigma(2)}) > \cdots > \psi(s_{\sigma(n)})$ for all $\psi \in U(\varphi, \varepsilon)$ (the ε neighborhood of φ in the uniform metric). Therefore,

$$\lim_{\psi \to \varphi} \frac{I(\psi) - I(\varphi) - P^\sigma(\psi - \varphi)}{\|\psi - \varphi\|} = 0$$

and $\nabla I(\varphi) = P^\sigma$. For any sequence $\varphi_i \in \Omega^c$ s.t. $\varphi_i \to 0$, $\nabla I(\varphi_i)$ is a sequence taking (a finite number of) values in $\{P^\sigma : \sigma \in \text{Per}(n)\}$, hence

$$\left\{\lim_{i\to\infty} \nabla I(\varphi_i) : \varphi_i \in \Omega^c,\, \varphi_i \to \varphi, \text{ and } \nabla I(\varphi_i) \text{ converges}\right\} \subseteq \{P^\sigma : \sigma \in \text{Per}(n)\}.$$

Conversely, for each $\sigma \in$Per(n), choose $\varphi \in B_0(\Sigma)$ s.t. $\varphi(s_{\sigma(1)}) > \varphi(s_{\sigma(2)}) > \cdots > \varphi(s_{\sigma(n)})$. Then, the sequence $\frac{1}{i}\varphi$ yields

$$\left\{\lim_{i\to\infty} \nabla I(\varphi_i) : \varphi_i \in \Omega^c,\, \varphi_i \to \varphi, \text{ and } \nabla I(\varphi_i) \text{ converges}\right\} \supseteq \{P^\sigma : \sigma \in \text{Per}(n)\}.$$

The result now follows from Theorem 49.

B.13 Proof of Proposition 20

$(i) \Rightarrow (ii)$: Let $x \succ y$ be s.t. $x\,A\,y$ is crisp. Then,

$$(u(x) - u(y))\,P(A) + u(y) = P(u(x\,A\,y)) = Q(u(x\,A\,y)) = (u(x) - u(y))\,Q(A) + u(y)$$

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

$(ii) \Rightarrow (iii)$: Let $x \nsim y$. Then,

$$P(u(x\,A\,y)) = (u(x) - u(y))\,P(A) + u(y) = (u(x) - u(y))\,Q(A) + u(y) = Q(u(x\,A\,y))$$

for all $P, Q \in \mathcal{C}$. That is, $x\,A\,y$ is crisp.

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

B.14 Proof of Proposition 21

1) For all $P, Q \in \mathcal{C}$, $P(S) = 1 = Q(S)$, hence $S \in \Lambda$. 2) If $A \in \Lambda$, for all $P, Q \in \mathcal{C}$, $P(A^c) = 1 - P(A) = 1 - Q(A) = Q(A^c)$, hence $A^c \in \Lambda$. 3) If $A, B \in \Lambda$ and $A \cap B = \emptyset$, for all $P, Q \in \mathcal{C}$, $P(A \cup B) = P(A) + P(B) = Q(A) + Q(B) = Q(A \cup B)$, hence $A \cup B \in \Lambda$.

B.15 Proof of Proposition 22

Notice that, for all $B \in \Sigma$,

$$\min_{P \in \mathcal{C}} P(B^c) = 1 - \max_{P \in \mathcal{C}} P(B) \quad \text{and} \quad \max_{P \in \mathcal{C}} P(B^c) = 1 - \min_{P \in \mathcal{C}} P(B).$$

Hence,

$$\begin{aligned} \rho(B) &= a([B])\left(\min_{P \in \mathcal{C}} P(B) - \max_{P \in \mathcal{C}} P(B)\right) + \max_{P \in \mathcal{C}} P(B) \\ \rho(B^c) &= a([B^c])\left(\min_{P \in \mathcal{C}} P(B) - \max_{P \in \mathcal{C}} P(B)\right) + 1 - \min_{P \in \mathcal{C}} P(B). \end{aligned}$$

It follows that $\rho(B) + \rho(B^c) = 1$ iff $(a([B]) + a([B^c]) - 1)\,(\min_{P \in \mathcal{C}} P(B) - \max_{P \in \mathcal{C}} P(B)) = 0$, as wanted.

B.16 Proof of Proposition 24

$(i) \Rightarrow (ii)$: Notice that $u(X)$ is an interval. Let $a \in \mathbb{R}$. If $a \in u(X)$, say $a = u(x')$, then $\{s \in S : u(f(s)) \geq a\} = \{s \in S : f(s) \succcurlyeq x'\} \in \Lambda$. Else, either $a < t$ for all $t \in u(X)$, and then $\{s \in S : u(f(s)) \geq a\} = S \in \Lambda$, or $a > t$ for all $t \in u(X)$, and then $\{s \in S : u(f(s)) \geq a\} = \emptyset \in \Lambda$.

$(ii) \Rightarrow (iii)$: Let $u(f) = \Sigma_{i=1}^{n} a_i 1_{A_i}$, with $\{A_1, A_2, ..., A_n\}$ a partition of S in Σ and $a_1 > a_2 > ... > a_n$. If $a \notin \{a_1, a_2, ..., a_n\}$, then $\{s \in S : u(f(s)) = a\} = \emptyset \in \Lambda$. The set $A_1 = \{s \in S : u(f(s)) = a_1\} = \{s \in S : u(f(s)) \geq a_1\} \in \Lambda$. For all $i \geq 2$, then $\Lambda \ni \{s \in S : u(f(s)) \geq a_i\} = \{s \in S : u(f(s)) \in \{a_1, a_2, ..., a_i\}\} = \bigcup_{j=1}^{i} \{s \in S : u(f(s)) = a_j\} = A_1 \cup A_2 \cup ... \cup A_i$. Therefore, for all $i \geq 2$, $\{s \in S : u(f(s)) = a_i\} = A_i = (A_1 \cup A_2 \cup ... \cup$

$A_i) \setminus (A_1 \cup A_2 \cup ... \cup A_{i-1}) \in \Lambda$ (remember that if Λ is a λ-system, $B, C \in \Lambda$, and $C \subseteq B$, then $B \setminus C \in \Lambda$).

(iii) $\Rightarrow$ *(iv)*: For all $x \in X$, $\{s \in S : f(s) \sim x\} = \{s \in S : u(f(s)) = u(x)\} \in \Lambda$.

(iv) $\Rightarrow$ *(v)*: Let $u(f) = \Sigma_{i=1}^n a_i 1_{A_i}$ with $\{A_1, A_2, ..., A_n\}$ a partition of S in Σ and $a_1 > a_2 > ... > a_n$. For all $i = 1, 2, ..., n$ choose $s_i \in A_i$ and set $x_i = f(s_i)$, clearly $u(x_i) = u(f(s_i)) = a_i$. Therefore:

- for all $i = 1, 2, ..., n$, $A_i = \{s \in S : u(f(s)) = a_i\} = \{s \in S : f(s) \sim x_i\} \in \Lambda$;
- $x_1 \succ x_2 \succ ... \succ x_n$ (since $a_1 > a_2 > ... > a_n$);
- if $g = \{x_1, A_1; x_2, A_2; ...; x_n, A_n\}$, for all $s \in S$, say $s \in A_j$, $u(f(s)) = a_j = u(x_j) = u(g(s))$. That is, $f(s) \sim g(s)$.

(v) $\Rightarrow$ *(i)*: Let $g = \{x_1, A_1; x_2, A_2; ...; x_n, A_n\}$. For all $x \in X$, $\{s \in S : f(s) \succcurlyeq x\} = \{s \in S : g(s) \succcurlyeq x\} = \bigcup_{i=1}^n \{s \in A_i : g(s) \succcurlyeq x\} = \bigcup_{i=1}^n \{s \in A_i : x_i \succcurlyeq x\}$. But for all $i = 1, 2, ..., n$, $\{s \in A_i : x_i \succcurlyeq x\}$ either coincides with A_i or is empty. Hence, $\{s \in S : f(s) \succcurlyeq x\}$ is a disjoint union of elements of Λ, which is a λ-system.

B.17 Proof of Proposition 26

A lemma first:

Lemma 53 *Let $a_1, a_2, ..., a_n, b_1, b_2, ..., b_n, c \in \mathbb{R}$ be such that $\sum_{h=1}^n a_h b_{\sigma(h)} = c$ for all permutations $\sigma \in Per(n)$. Then either $a_1 = a_2 = ... = a_n$ or $b_1 = b_2 = ... = b_n$.*

Proof. By contradiction, assume that there exist $i, j \in \{1, \ldots, n\}$ such that $a_i \neq a_j$ and $k, l \in \{1, \ldots, n\}$ such that $b_k \neq b_l$. Consider a permutation σ such that $\sigma(i) = k$ and $\sigma(j) = l$, and the permutation $\sigma' = \sigma(kl)$ obtained applying σ and then switching around k and l. It follows that

$$a_i b_k + a_j b_l + \sum_{h \neq i,j} a_h b_{\sigma(h)} = \sum_{h=1}^n a_h b_{\sigma(h)} = c = \sum_{h=1}^n a_h b_{\sigma'(h)} = a_i b_l + a_j b_k + \sum_{h \neq i,j} a_h b_{\sigma(h)},$$

whence $a_i b_k + a_j b_l = a_i b_l + a_j b_k$. That is, $a_i(b_k - b_l) = a_j(b_k - b_l)$, which implies $a_i = a_j$, a contradiction. $\square$

Proof of Proposition 26.

By assumption, we have that for every $P, Q \in \mathcal{C}$ and each permutation $\sigma \in \mathrm{Per}(n)$,

$$\sum_{i=1}^n [P(A_i) - Q(A_i)] \, u\left(x_{\sigma(i)}\right) = 0. \tag{20}$$

Therefore, either $P(A_1) - Q(A_1) = P(A_2) - Q(A_2) = \cdots = P(A_n) - Q(A_n) = b$ or $u(x_1) = u(x_2) = \ldots = u(x_n)$. In the former case $1 = \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{n} Q(A_i) + nb = 1 + nb$. Therefore $b = 0$ and $A_i \in \Lambda$ for all $i = 1, 2, \ldots, n$. It follows that for all $x \in X$,

$$\{s \in S : f(s) \sim x\} = \bigcup_{i=1}^{n} \{s \in A_i : f(s) \sim x\} = \bigcup_{i=1}^{n} \{s \in A_i : x_i \sim x\}.$$

But for all $i = 1, 2, \ldots, n$, $\{s \in A_i : x_i \sim x\}$ either coincides with A_i or is empty, so that $\{s \in S : f(s) \sim x\}$ is a disjoint union of elements of Λ.

Conversely, suppose that f is unambiguous and $x_i \nsim x_j$ for every $i \neq j$ in $\{1, \ldots, n\}$. Then, for any permutation $\sigma \in \mathrm{Per}(N)$ and each $i = 1, 2, \ldots, n$, $\{s \in S : f^{\sigma}(s) \sim x_{\sigma(i)}\} = A_i \in \Lambda$. As a consequence, for all $P, Q \in \mathcal{C}$,

$$\int u(f^{\sigma})\, dP = \sum_{i=1}^{n} u\left(x_{\sigma(i)}\right) P(A_i) = \sum_{i=1}^{n} u\left(x_{\sigma(i)}\right) Q(A_i) = \int u(f^{\sigma})\, dQ.$$

B.18 Proof of Proposition 28

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

$$u(x_*) \leq \underline{\mathcal{C}}(u(f)) \leq \overline{\mathcal{C}}(u(f)) \leq u(x^*).$$

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

If $x \notin N(f)$, either there exists $y \in X$ such that $x \succ y \trianglerighteq f$, or there exists $y \in X$ such that $f \trianglerighteq y \succ x$. That is, either $u(x) > u(y) \geq \overline{\mathcal{C}}(u(f))$ or $u(x) < u(y) \leq \underline{\mathcal{C}}(u(f))$. Thus, $u(x) \notin [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]$.

Conversely, if $u(x) \notin [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]$, either $u(x) > \overline{\mathcal{C}}(u(f))$ or $u(x) < \underline{\mathcal{C}}(u(f))$. Take $x_{\max}$ and $x_{\min}$ in X such that $u(x_{\max}) = \overline{\mathcal{C}}(u(f))$ and $u(x_{\min}) = \underline{\mathcal{C}}(u(f))$, to obtain either $x \succ x_{\max} \trianglerighteq f$ or $f \trianglerighteq x_{\min} \succ x$. Thus, $x \notin N(f)$.

We conclude that $x \in N(f)$ iff $u(x) \in [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]$ or

$$N(f) = u^{-1}\left([\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]\right).$$

This implies $u(N(f)) \subseteq [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]$. Conversely, if $t \in [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]$, there exists x_t such that $u(x_t) = t$. Clearly, $x_t \in N(f)$ and $t \in u(N(f))$.

B.19 Proof of Proposition 29

Since the map from $B_0(\Sigma)$ to $\mathbb{R}$ defined by

$$\psi \mapsto I(u(f)+\psi) - I(\psi)$$

is continuous and $B_0(\Sigma)$ is connected, the set

$$\begin{aligned} J &= \{I(u(f)+\psi) - I(\psi) : \psi \in B_0(\Sigma)\} \\ &= \left\{ I\left(u(f) + \frac{1-\lambda}{\lambda}u(g)\right) - I\left(\frac{1-\lambda}{\lambda}u(g)\right) : g \in \mathfrak{F},\ \lambda \in (0,1] \right\} \end{aligned}$$

is connected. That is, it is an interval. From the lemma, it follows that

$$\overline{J} = [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))].$$

Let

$$M(f) = \{x \in X : \exists \lambda \in (0,1], \exists h \in \mathfrak{F} \text{ such that } x\,\lambda\,h \sim f\,\lambda\,h\}.$$

We have $x \in M(f)$ iff

$$u(x) = I\left(u(f) + \frac{1-\lambda}{\lambda}u(h)\right) - I\left(\frac{1-\lambda}{\lambda}u(h)\right)$$

iff $x \in u^{-1}(J)$. Hence, $u(M(f)) \subseteq J$. Conversely, if $t \in J$, $t \in [\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))]$ and there exists $x \in X$ such that $u(x) = t$. Clearly, $x \in M(f)$, whence $u(M(f)) = J$. We conclude observing that

$$N(f) = u^{-1}\left([\min_{P\in\mathcal{C}} P(u(f)), \max_{P\in\mathcal{C}} P(u(f))]\right) = u^{-1}\left(\overline{J}\right) = u^{-1}\left(\overline{u(M(f))}\right).$$

B.20 Proof of Proposition 31

To prove the statement for $\succcurlyeq^{\downarrow}$ (that for $\succcurlyeq^{\uparrow}$ is proved analogously), we only need to show that

$$f \succcurlyeq^{\downarrow} g \iff \underline{\mathcal{C}}(u(f)) \geq \underline{\mathcal{C}}(u(g)).$$

Applying the definition of $\succcurlyeq^{\downarrow}$ and Proposition 4, we have that $f \succcurlyeq^{\downarrow} g$ iff for every $x \in X$,

$$P(u(g)) \geq u(x) \text{ for all } P \in \mathcal{C} \Rightarrow P(u(f)) \geq u(x) \text{ for all } P \in \mathcal{C}.$$

That is, iff for every $x \in X$,

$$\underline{\mathcal{C}}(u(g)) \geq u(x) \Rightarrow \underline{\mathcal{C}}(u(f)) \geq u(x).$$

This is equivalent to $\underline{\mathcal{C}}(u(f)) \geq \underline{\mathcal{C}}(u(g))$, concluding the proof.

B.21 Proof of Proposition 32

$(i) \Leftrightarrow (ii)$: We have proved in Proposition 31 that $\succcurlyeq^{\downarrow}$ is represented by the functional $\underline{\mathcal{C}}(u(\cdot))$, and $\succcurlyeq^{\uparrow}$ by $\overline{\mathcal{C}}(u(\cdot))$. Consider $\succcurlyeq$ and $\succcurlyeq^{\downarrow}$, and recall Prop. 7. It is clear that $\underline{\mathcal{C}}(u(f)) \leq I(u(f))$ is tantamount to saying that for every $x \in X$,

$$x \succcurlyeq f \Rightarrow x \succcurlyeq^{\downarrow} f.$$

The argument for $\succcurlyeq$ and $\succcurlyeq^{\uparrow}$ is analogous.

$(ii) \Leftrightarrow (iii)$: This follows immediately from Lemmas 48 and 52.

B.22 Proof of Theorem 33

The proof of the theorem builds on the following lemma.

Lemma 54 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a monotonic constant linear functional, and $\mathcal{D}$ a set of probabilities such that*

$$\min_{P \in \mathcal{D}} P(\psi) \leq I(\psi) \leq \max_{P \in \mathcal{D}} P(\psi)$$

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

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

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

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

$$\min_{P \in \mathcal{D}} P(\psi) \leq I(\psi) \leq \max_{P \in \mathcal{D}} P(\psi)$$

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

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

a little algebra yields:

$$\begin{aligned} \beta(\varphi) &= \frac{I(\varphi) - \max_{P\in\mathcal{D}} P(\varphi)}{\min_{P\in\mathcal{D}} P(\varphi) - \max_{P\in\mathcal{D}} P(\varphi)} = \\ &= -\frac{I(\varphi) - \max_{P\in\mathcal{D}} P(\varphi)}{\max_{P\in\mathcal{D}} P(\varphi) - \min_{P\in\mathcal{D}} P(\varphi)} = -I\left(\frac{\varphi - \max_{P\in\mathcal{D}} P(\varphi)}{\max_{P\in\mathcal{D}} P(\varphi) - \min_{P\in\mathcal{D}} P(\varphi)}\right). \end{aligned}$$

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

$$\begin{aligned}
\max_{P\in\mathcal{D}}\left(\frac{\varphi - \max_{P\in\mathcal{D}} P(\varphi)}{\max_{P\in\mathcal{D}} P(\varphi) - \min_{P\in\mathcal{D}} P(\varphi)}\right) &= 0 \text{ and} \\
\min_{P\in\mathcal{D}}\left(\frac{\varphi - \max_{P\in\mathcal{D}} P(\varphi)}{\max_{P\in\mathcal{D}} P(\varphi) - \min_{P\in\mathcal{D}} P(\varphi)}\right) &= -1;
\end{aligned}$$

therefore,

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

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

Proof of Theorem 33. It is enough to show that, for all $\varphi \in B_0(\Sigma)$, $I(\varphi)$ depends only on $\min_{P\in\mathcal{C}} P(\varphi)$ and $\max_{P\in\mathcal{C}} P(\varphi)$. For, then we can set $T(\min_{P\in\mathcal{C}} P(\varphi), \max_{P\in\mathcal{C}} P(\varphi)) = I(\varphi)$ and apply Lemma 54.

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

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

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

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

then

$$N(f) = u^{-1}\left(\left[\underline{\mathcal{C}}(u(f)), \overline{\mathcal{C}}(u(f))\right]\right) = u^{-1}\left(\left[\underline{\mathcal{C}}(u(g)), \overline{\mathcal{C}}(u(g))\right]\right) = N(g)$$

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

B.23 Proof of Proposition 36

The uniqueness of I descending from Lemma 1 guarantees that

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

for all $\varphi \in B_0(\Sigma)$. Then, by points 6 and 7 of Proposition 45 it follows that

$$\begin{aligned}
\mathcal{C} &= \partial I(0) \\
&= \partial(\beta \min_{Q\in\mathcal{D}} Q(\cdot) + (1-\beta) \max_{Q\in\mathcal{D}} Q(\cdot))(0) \\
&\subseteq \beta\, \partial(\min_{Q\in\mathcal{D}} Q(\cdot))(0) + (1-\beta)\, \partial(\max_{Q\in\mathcal{D}} Q(\cdot))(0) \\
&= \beta\mathcal{D} + (1-\beta)\mathcal{D}.
\end{aligned}$$

If, moreover, $\succcurlyeq$ satisfies axiom 6, then

$$I(\varphi) = \alpha \min_{P\in\mathcal{C}} P(\varphi) + (1-\alpha) \max_{P\in\mathcal{C}} P(\varphi).$$

That is,

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

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

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

Let $c = \frac{1}{2}\underline{c} + \frac{1}{2}\overline{c}$ to obtain

$$\begin{aligned} I(\varphi) &= \alpha\underline{c} + (1-\alpha)\overline{c} \\ &= c + \alpha\underline{c} + (1-\alpha)\overline{c} - \frac{1}{2}\underline{c} - \frac{1}{2}\overline{c} \\ &= c + \left(\frac{1}{2} - \alpha\right)(\overline{c} - \underline{c}) \end{aligned}$$

and, analogously,

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

Therefore,

$$\left(\frac{1}{2} - \beta\right)(\overline{d} - \underline{d}) = \left(\frac{1}{2} - \alpha\right)(\overline{c} - \underline{c}).$$

If $\beta > 1/2$, then $\alpha > 1/2$ and

$$\left(\frac{1}{2} - \beta\right)(\overline{d} - \underline{d}) = \left(\frac{1}{2} - \alpha\right)(\overline{c} - \underline{c}) > \left(\frac{1}{2} - \alpha\right)(\overline{d} - \underline{d})$$

yields $\alpha > \beta$. Analogously, if $\beta < 1/2$, then $\alpha < 1/2$ and

$$\left(\frac{1}{2} - \beta\right)(\overline{d} - \underline{d}) = \left(\frac{1}{2} - \alpha\right)(\overline{c} - \underline{c}) < \left(\frac{1}{2} - \alpha\right)(\overline{d} - \underline{d})$$

yields $\alpha < \beta$.

B.24 Proof of Lemma 38

$(i) \Leftrightarrow (ii)$: Assume that, for every $A \in \Pi$, $f \sim_A f\,A\,g$ for every $f, g \in \mathfrak{F}$, hence, for every $h \in \mathfrak{F}$ $[\lambda f + (1-\lambda)h] \sim_A [\lambda f + (1-\lambda)h]A[\lambda g + (1-\lambda)h]$, that is $\lambda f + (1-\lambda)h \sim_A \lambda f A g + (1-\lambda)h$, thus $f \overset{\sim}{\cup}_A f\, A\, g$. Conversely, if $f \overset{\sim}{\cup}_A f\, A\, g$ for every $f, g \in \mathfrak{F}$, then in particular $f \sim_A f\, A\, g$ for every $f, g \in \mathfrak{F}$.

$(ii) \Leftrightarrow (iii)$: By Proposition 4, $f \overset{\sim}{\cup}_A fAg$ iff $P(u_A(f)) = P(u_A(fAg))$ for all $P \in \mathcal{C}_A$. It immediately follows that $\underline{\mathcal{C}_A}(u_A(f)) = \underline{\mathcal{C}_A}(u_A(fAg))$. By Proposition 31, this is equivalent to $f \sim_A^{\downarrow} fAg$.

Conversely, suppose that $f \sim_A^{\downarrow} fAg$ for every $f, g \in \mathfrak{F}$. Consider $x \succ_A y$. Since $x \sim_A^{\downarrow} xAy$, it follows from Proposition 31 that

$$\begin{aligned} u_A(x) &= \min_{P \in \mathcal{C}_A} \left[u_A(x)\, P(A) + u_A(y)\, (1 - P(A)) \right] \\ &= u_A(x) \min_{P \in \mathcal{C}_A} P(A) + u_A(y)\, (1 - \min_{P \in \mathcal{C}_A} P(A)). \end{aligned}$$

Since $u_A(x) > u_A(Y)$, this implies that $\min_{P \in \mathcal{C}_A} P(A) = 1$, or equivalently, that $P(A) = 1$ for all $P \in \mathcal{C}_A$. It follows that $P(u_A(f)) = P(u_A(fAg))$ for all $P \in \mathcal{C}_A$, which is equivalent to $f \overset{\sim}{\cup}_A fAg$.

B.25 Proof of Proposition 39

First, we observe that the fact that Eq. (10) implies $\mathcal{C}_A = \mathcal{C}|A$ is a consequence of Proposition 41. That it implies $u_A = u$ is seen by taking $f = x$ and $g = x'$ to show that $x \succcurlyeq_A x' \Leftrightarrow x \succcurlyeq x'$. The converse is trivial.

$(i) \Leftrightarrow (ii)$: $fAg \overset{\succsim}{\cup} g$ for all $P \in \mathcal{C}$ iff $\int_A u(f)\, dP + \int_{A^c} u(g)\, dP \geq \int_A u(g)\, dP + \int_{A^c} u(g)\, dP$ for all $P \in \mathcal{C}$ iff $\int_A u(f)\, dP \geq \int_A u(g)\, dP$ for all $P \in \mathcal{C}$ iff $P_A(u(f)) \geq P_A(u(g))$ for all $P \in \mathcal{C}$.

$(i) \Rightarrow (iii)$: Suppose that $u = u_A$. We first observe that it follows from Proposition 28 that for every $f \in \mathfrak{F}$ and $x \in X$, with obvious notation,

$$x \in N_A(f) \Longleftrightarrow \underline{\mathcal{C}_A}(u(f)) \leq u(x) \leq \overline{\mathcal{C}_A}(u(f)). \tag{21}$$

Next, we prove that for every $f \in \mathfrak{F}$ and $x \in X$, again with obvious notation,

$$x \in N(f\, A\, x) \Longleftrightarrow \underline{\mathcal{C}|A}(u(f)) \leq u(x) \leq \overline{\mathcal{C}|A}(u(f)). \tag{22}$$

To see this, apply again Proposition 28 to find

$$x \in N(f\, A\, x) \Longleftrightarrow \underline{\mathcal{C}}(u(f\, A\, x)) \leq u(x) \leq \overline{\mathcal{C}}(u(f\, A\, x)).$$

That is, $x \in N(fAx)$ iff both

$$\min_{P \in \mathcal{C}} \int_S u(f\, A\, x)\, dP \leq u(x) \tag{23}$$

and

$$u(x) \leq \max_{P \in \mathcal{C}} \int_S u(f\,A\,x)\,dP. \tag{24}$$

Denote resp. $\underline{P}$ and $\overline{P}$ the probabilities in $\mathcal{C}$ that attain the extrema in Eqs. (23) and (24). Then we can rewrite Eq. (23) as follows:

$$u(x) \geq \frac{1}{\underline{P}(A)} \int_A u(f)\,d\underline{P},$$

which is equivalent to saying that

$$u(x) \geq \min_{P \in \mathcal{C}} \int_S u(f)\,dP_A = \underline{\mathcal{C}|A}(u(f)).$$

Analogously, Eq. (24) can be rewritten as

$$u(x) \leq \frac{1}{\overline{P}(A)} \int_A u(f)\,d\overline{P},$$

which is equivalent to

$$u(x) \leq \max_{P \in \mathcal{C}} \int_S u(f)\,dP_A = \overline{\mathcal{C}|A}(u(f)).$$

This ends the proof of Eq. (22).

To prove (i), notice that $x \succcurlyeq x' \Rightarrow x \succcurlyeq_A x'$ obviously follows from the assumption $u = u_A$, and that, given Eqs. (21) and (22), Eq. (12) follows immediately from the assumption $\mathcal{C}|A = \mathcal{C}_A$.

$(iii) \Rightarrow (i)$: First, observe that the assumption that $x \succcurlyeq x' \Rightarrow x \succcurlyeq_A x'$ implies $u = u_A$ by Corollary 51. Hence, it follows from Eqs. (21) and (22) above that Eq. (12) is equivalent to

$$\underline{\mathcal{C}_A}(u(f)) \leq u(x) \leq \overline{\mathcal{C}_A}(u(f)) \Longleftrightarrow \underline{\mathcal{C}|A}(u(f)) \leq u(x) \leq \overline{\mathcal{C}|A}(u(f)).$$

In particular, this implies that for every $\varphi \in B_0(\Sigma, u(X))$,

$$\min_{P \in \mathcal{C}|A} P(\varphi) = \min_{Q \in \mathcal{C}_A} Q(\varphi) \tag{25}$$

The result that $\mathcal{C}|A = \mathcal{C}_A$ now follows from two applications of Proposition 41.

$(i) \Rightarrow (iv)$: By Proposition 31 and the assumption that $u_A = u$, for every $f \in \mathfrak{F}$ and $x \in X$ we have that $f \succcurlyeq^{\downarrow}_A x$ iff $Q(u(f)) \geq u(x)$ for all $Q \in \mathcal{C}_A$. Next, we show that

$$f\,A\,x \succcurlyeq^{\downarrow} x \Longleftrightarrow P(u(f)) \geq u(x) \quad \text{for all } P \in \mathcal{C}|A, \tag{26}$$

so that the result follows from the assumption that $\mathcal{C}|A = \mathcal{C}_A$.

To see why Eq. (26) holds, notice that by Prop. 31, $fAx \succcurlyeq^{\downarrow} x$ iff $P(u(fAx)) \geq u(x)$ for all $P \in \mathcal{C}$. Equivalently, for every $P \in \mathcal{C}$,

$$\int_A u(f)\, dP + (1 - P(A))\, u(x) \geq u(x),$$

which holds iff $P_A(u(f)) \geq u(x)$ (recall that $P(A) > 0$ for all $P \in \mathcal{C}$). In turn, the latter is equivalent to saying that $P(u(f)) \geq u(x)$ for every $P \in \mathcal{C}|A$.

$(iv) \Rightarrow (i)$: We first show (mimicking an argument of Siniscalchi [28]) that Eq. (13) implies that $u_A = u$. To see this, notice that we have $u_A(x) \geq u_A(x')$ iff $x \succcurlyeq_A x'$ iff $x\, A\, x' \succcurlyeq^{\downarrow} x'$ iff

$$\min_{P \in \mathcal{C}} \left[u(x)P(A) + u(x')(1 - P(A)) \right] \geq u(x').$$

Using the assumption that $\min_{P \in \mathcal{C}} P(A) > 0$, the latter is equivalent to $u(x) \geq u(x')$, proving that $u_A = u$.

We now show that $\mathcal{C}|A = \mathcal{C}_A$ by showing that Eq. (25) holds for every $\varphi \in B_0(\Sigma, u(X))$, so that the result follows again from Lemma 41. As argued above, Eq. (13) holds for f and x iff

$$P(u(f)) \geq u(x) \quad \text{for all } P \in \mathcal{C}|A \Longleftrightarrow Q(u(f)) \geq u(x) \quad \text{for all } Q \in \mathcal{C}_A.$$

Fix $\varphi \in B_0(\Sigma, u(X))$ and suppose that, in violation of Eq. (25), $\alpha \equiv \min_{P \in \mathcal{C}|A} P(\varphi) > \min_{Q \in \mathcal{C}_A} Q(\varphi) \equiv \beta$. Then there exists $\gamma \in (\beta, \alpha)$. Let x denote the consequence such that $u(x) = \gamma$. By the assumption we just made, we have that (if $f \in \mathfrak{F}$ is such that $u(f) = \varphi$),

$$\min_{P \in \mathcal{C}|A} P(u(f)) > u(x) \quad \text{and} \quad u(x) > \min_{Q \in \mathcal{C}_A} Q(u(f)),$$

which, as proved above (the proof for strict preference works *mutatis mutandis* as that for weak preference, recalling that $\mathcal{C}|A$ is weak* compact), is equivalent to $fAx \succ^{\downarrow} x$ and $f \prec_A^{\downarrow} x$, a contradiction. Suppose instead that $\alpha < \beta$ and let $\gamma \in (\alpha, \beta)$. In this case we obtain

$$\min_{P \in \mathcal{C}|A} P(u(f)) < u(x) \quad \text{and} \quad u(x) < \min_{Q \in \mathcal{C}_A} Q(u(f)),$$

which is equivalent to $f \succ_A^{\downarrow} x$ and $x \succ^{\downarrow} fAx$ (to see the latter, let $P^* \in \mathcal{C}$ be the probability whose posterior minimizes the left-hand inequality; it follows that $\min_{P \in \mathcal{C}} P(u(fAx)) \leq P^*(u(fAx)) < u(x)$), again a contradiction. This concludes the proof of Eq. (25).

C The Extension to Infinite Acts

We now briefly discuss how the analysis in the first part of the paper (Sections 2–4) extends to the case in which the choice set that the DM faces contains acts which are not finite-valued.

Here, we start by assuming that $\succcurlyeq$ is a weak order on $\mathfrak{F}$, and we let $\mathfrak{F}'$ denote the set of all the **Σ-measurable** and **preference-bounded** functions from S into X. That is, $f \in \mathfrak{F}'$ if: $\{s \in S : f(s) \succcurlyeq x\} \in \Sigma$ and $\{s \in S : f(s) \succ x\} \in \Sigma\}$ for every $x \in X$ and there exist $x, y \in X$ such that $x \succcurlyeq f(s) \succcurlyeq y$ for all $s \in S$. Following the example of Gilboa and Schmeidler [16], we then show that if $\succcurlyeq$ satisfies axioms 1–5 (as stated), it has a *unique* extension $\succcurlyeq'$ that also satisfies axioms 1–5 (with $\mathfrak{F}$ substituted with $\mathfrak{F}'$). That is, the DM's preference on simple acts determine his preferences on preference-bounded acts. Moreover, such extended preference is represented by the utility function u and a functional $I' : \mathfrak{F}' \to \mathbb{R}$ that is monotonic and constant linear.

As recalled earlier, $ba(\Sigma)$, endowed with the total variation norm, is isometrically isomorphic to the norm dual of $B(\Sigma)$. Since $B_0(\Sigma)$ is dense in $B(\Sigma)$, $ba(\Sigma)$ is also isometrically isomorphic to the norm dual of $B_0(\Sigma)$. Moreover, for bounded subsets of $ba(\Sigma)$ the $\sigma(ba(\Sigma), B(\Sigma))$ and the $\sigma(ba(\Sigma), B_0(\Sigma))$ topologies coincide (for a proof of this result, see Maccheroni and Marinacci [20]):

Lemma 55 *Let μ_a be a bounded net in $ba(\Sigma)$, where Σ is a field of subsets of S. The following facts are equivalent:*

(i) The net μ_a converges to μ in the $\sigma(ba(\Sigma), B(\Sigma))$ topology,

(ii) The net μ_a converges to μ in the $\sigma(ba(\Sigma), B_0(\Sigma))$ topology,

(iii) $\lim_a \mu_a(A) = \mu(A)$ for all $A \in \Sigma$.

In view of this result, we refer to either one of $\sigma(ba(\Sigma), B(\Sigma))$ and $\sigma(ba(\Sigma), B_0(\Sigma))$ as "the weak* topology".

We now show that a monotonic constant linear functional defined on $B_0(\Sigma)$ (as the I obtained in Lemma 1) has a unique monotonic constant linear extension to $B(\Sigma)$.

Definition 56 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a monotonic constant linear functional. The functional $I_* : B \to \mathbb{R}$ defined by*

$$I_*(\xi) = \sup\{I(\varphi) : \varphi \in B_0(\Sigma), \varphi \leq \xi\}$$

*is called the **lower extension** of I. The functional $I^* : B \to \mathbb{R}$ defined by*

$$I^*(\xi) = \inf\{I(\varphi) : \varphi \in B_0(\Sigma), \varphi \geq \xi\}$$

*is called the **upper extension** of I. Finally, ξ is said to be **I-exhaustible** if*

$$I_*(\xi) = I^*(\xi).$$

Remark 1 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a monotonic constant linear functional, $\xi \in B(S)$. Then:*

1. $-\infty < I_*(\xi) \le I^*(\xi) < \infty$;
2. *If $\xi \in B_0(\Sigma)$, ξ is I-exhaustible, and $I(\xi) = I_*(\xi) = I^*(\xi)$.*

Hence, if ξ is I-exhaustible, we can set $I(\xi) = I_*(\xi)$. The set of I-exhaustible functions is denoted by $B(I)$.

Proposition 57 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a monotonic constant linear functional, $\xi \in B(S)$. The following facts are equivalent:*

(i) $\xi \in B(I)$.

(ii) *For all $\varepsilon > 0$, there exist $\varphi, \psi \in B_0(\Sigma)$ such that $\varphi \le \xi \le \psi$ and $I(\psi) - I(\varphi) \le \varepsilon$.*

(iii) *There exist $\{\varphi_n\}, \{\psi_n\} \subseteq B_0(\Sigma)$ such that $\varphi_n \le \varphi_{n+1} \le \xi \le \psi_{n+1} \le \psi_n$ and $I(\psi_n) - I(\varphi_n) \downarrow 0$.*

In this case, $I(\xi) = \lim_{n\to\infty} I(\psi_n) = \lim_{n\to\infty} I(\varphi_n)$.

Proof. $(i) \Rightarrow (iii)$: $I(\xi) = \sup\{I(\varphi) : \varphi \in B_0(\Sigma), \varphi \le \xi\}$. Therefore, there exist $\{\widehat{\varphi}_n\} \subseteq B_0(\Sigma)$ such that $\widehat{\varphi}_n \le \xi$ for all $n \in \mathbb{N}$, and $\lim_{n\to\infty} I(\widehat{\varphi}_n) = I(\xi)$.

Let $\varphi_n = \max\{\widehat{\varphi}_1, \widehat{\varphi}_2, ..., \widehat{\varphi}_n\}$ for all $n \in \mathbb{N}$, $\{\varphi_n\} \subseteq B_0(\Sigma)$, and $\widehat{\varphi}_n \le \varphi_n \le \varphi_{n+1} \le \xi$ for all $n \in \mathbb{N}$. $I(\widehat{\varphi}_n) \le I(\varphi_n) \le I(\varphi_{n+1}) \le I(\xi)$ for all $n \in \mathbb{N}$, thus $I(\varphi_n) \uparrow I(\xi)$. Similarly, we can build $\{\psi_n\} \subseteq B_0(\Sigma)$ such that $\xi \le \psi_{n+1} \le \psi_n$ and $I(\psi_n) \downarrow I(\xi)$.

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

$(ii) \Rightarrow (i)$: For all $\varepsilon > 0$, there exist $\varphi, \psi \in B_0(\Sigma)$ such that $\varphi \le \xi \le \psi$ and $I(\psi) \le I(\varphi) + \varepsilon$. Then, $I(\varphi) \le I_*(\xi) \le I^*(\xi) \le I(\psi) \le I(\varphi) + \varepsilon$, so $0 \le I^*(\xi) - I_*(\xi) \le \varepsilon$. □

Proposition 58 *For any monotonic constant linear functional I on $B_0(\Sigma)$, $B(I)$ is a cone, it is closed w.r.t. addition of constant functions, and the functional $I : B(I) \to \mathbb{R}$ is constant linear, monotone and continuous.*

Proof. See Maccheroni and Marinacci [20] . □

Theorem 59 *Let $I : B_0(\Sigma) \to \mathbb{R}$ be a monotonic constant linear functional, $\xi \in B(S)$. Then, $\xi \in B(\Sigma)$* if and only if *ξ is I-exhaustible.*

Proof. Let $I : B_0(\Sigma) \to \mathbb{R}$ be a monotonic constant linear functional and $\xi \in B(\Sigma)$. There exist $\{\varphi_n\} \subseteq B_0(\Sigma)$ and $\{\delta_n\} \subseteq \mathbb{R}^+$ such that $\delta_n \to 0$ and $\varphi_n - \delta_n \leq \xi \leq \varphi_n + \delta_n$ for all $n \in \mathbb{N}$. For all $\varepsilon > 0$, there exist n_ε such that $2\delta_{n_\varepsilon} \leq \varepsilon$. The functions $\varphi_{n_\varepsilon} - \delta_{n_\varepsilon}, \varphi_{n_\varepsilon} + \delta_{n_\varepsilon} \in B_0(\Sigma)$ satisfy $\varphi_{n_\varepsilon} - \delta_{n_\varepsilon} \leq \xi \leq \varphi_{n_\varepsilon} + \delta_{n_\varepsilon}$ and $I(\varphi_{n_\varepsilon} + d_{n_\varepsilon}) - I(\varphi_{n_\varepsilon} - \delta_{n_\varepsilon}) = 2\delta_{n_\varepsilon} \leq \varepsilon$. So, by Proposition 57, ξ is I-exhaustible. For a proof of the converse, see Maccheroni and Marinacci [20]. □

We therefore obtain the sought result:

Corollary 60 *If I is a monotonic constant linear functional I on $B_0(\Sigma)$, then there exists a unique monotonic constant linear extension of I to $B(\Sigma)$.*

To extend the representation obtained in Lemma 1 to the preference-bounded acts in $\mathfrak{F}'$, we first notice that $B(\Sigma)$ is the set of utility images of acts in $\mathfrak{F}'$ and consider a preference relation $\succcurlyeq$ whose restriction to $\mathfrak{F}$ satisfies axioms 1–5. Hence, it can be represented by a monotonic constant linear functional I on $B_0(\Sigma)$. By Corollary 60, I has a unique monotonic constant linear extension I' to $B(\Sigma)$. Let $\succcurlyeq'$ on $\mathfrak{F}'$ be defined as follows:

$$f \succcurlyeq' g \Longleftrightarrow I'(u(f)) \geq I'(u(g)).$$

It follows that $\succcurlyeq'$ is an extension of $\succcurlyeq$ to $\mathfrak{F}'$ that satisfies axioms 1–5 (in which every instance of $\mathfrak{F}$ is substituted with $\mathfrak{F}'$), and it is uniquely defined.

Having thus extended the representation of Lemma 1, we may wonder whether the other results proved in sections 2–4 generalize to the case in which acts belong to $\mathfrak{F}'$. The simple answer is "yes", for it is straightforward to check that the results in Appendix A hold with $B(\Sigma, K)$ (resp. I') substituting $B_0(\Sigma, K)$ (resp. I). Therefore, we can adapt the proofs in Appendix B to show that the results in the main text can be extended to this case.

More interestingly, we also find that this generalization does not affect our representation of the DM's ambiguity perception. This occurs for three main reasons.

First, suppose that given the "extended" preference relation $\succcurlyeq'$, we define the "extended" unambiguous preference relation $\underline{\succcurlyeq}'$. It is immediate to use the continuity of the functional I' (which follows from the fact that I' is Lipschitz) to verify that for all $f, g \in \mathfrak{F}$,

$$f \,\underline{\succcurlyeq}\, g \Longleftrightarrow f \,\underline{\succcurlyeq}'\, g.$$

That is, the extended unambiguous preference is a proper extension of the "narrow" unambiguous preference. Consider now this immediate corollary of Proposition 42:

Corollary 61 *If $\gtrsim$ is a nontrivial, continuous, conic, and monotonic preorder on $B_0(\Sigma, K)$ represented by a nonempty weak* closed and convex subset $\mathcal{C}$ of $pc(\Sigma)$, then there exists a* *unique nontrivial, continuous, conic, and monotonic preorder* $\gtrsim'$ *on* $B(\Sigma)$ *such that for all* $\varphi, \psi \in B_0(\Sigma, K)$

$$\varphi \gtrsim \psi \Longleftrightarrow \varphi \gtrsim' \psi,$$

and $\gtrsim'$ *is represented by* $\mathcal{C}$*.*

It follows that the set $\mathcal{C}$ represents also $\succcurlyeq^{*\prime}$: For all $f, g \in \mathfrak{F}'$,

$$f \succcurlyeq^{*\prime} g \Longleftrightarrow \int_S u(f)\, dP \geq \int_S u(g)\, dP \quad \text{for all } P \in \mathcal{C}. \tag{27}$$

Second, notice that if we define the "extended" mixture certainty equivalents set $N'(f)$, for$f \in \mathfrak{F}'$ we find, again by the continuity of I', that $N'(f)$ is the preference closure of the set

$$\{x \in X : \exists \lambda \in (0, 1], \exists h \in \mathfrak{F} \text{ such that } x\, \lambda\, h \sim f\, \lambda\, h\}.$$

That is, simple acts suffice in finding the mixture certainty equivalents of f. Third, it is analogously sufficient to use simple acts in deciding whether $f \in \mathfrak{F}'$ is crisp. Thus the set of "extended" crisp acts $[x]'$ contains the "narrow" $[x]$.

Thus, we conclude that not only can the previous results be equivalently stated with $\mathfrak{F}'$, $\succcurlyeq'$, I', $\succcurlyeq^{*\prime}$, N' and $[x]'$ in place of their unprimed counterparts. We also have that every conclusion that can be drawn by restricting our attention to simple acts remains valid in this extended problem. In particular, Eq. (27) implies that the set $\mathcal{C}$ obtained from $\succcurlyeq^{*}$ provides us with a complete description of his perception of ambiguity, so that $\partial I(0) = \partial I'(0)$.

References

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

[2] Kenneth J. Arrow and Leonid Hurwicz. An optimality criterion for decision making under ignorance. In C.F. Carter and J.L. Ford, editors, *Uncertainty and Expectations in Economics*. Basil Blackwell, Oxford, 1972.

[3] Truman Bewley. Knightian decision theory: Part I. Technical Report 807, Cowles Foundation, Yale University, 1986.

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

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

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

[7] Larry G. Epstein and Martin Schneider. Recursive multiple-priors. Mimeo, University of Rochester and UCLA, August 2001.

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

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

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

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

[12] Paolo Ghirardato and Jonathan N. Katz. Indecision theory: Explaining selective abstention in multiple elections. Social Science Working Paper 1106, Caltech, November 2000. `http://www.hss.caltech.edu/~paolo/voting.pdf`.

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

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

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

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

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

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

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

[20] Fabio Maccheroni and Massimo Marinacci. Notes on the theory of capacities. Unpublished lecture notes, Università Bocconi and Università di Torino, November 1999.

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

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

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

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

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

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

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

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

[29] Marciano Siniscalchi. A vector-valued representation of preferences under uncertainty. Mimeo, Princeton University, March 2002.

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

INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
APPLIED MATHEMATICS WORKING PAPER SERIES

1. Luigi Montrucchio and Fabio Privileggi, "On Fragility of Bubbles in Equilibrium Asset Pricing Models of Lucas-Type," *Journal of Economic Theory* 101, 158-188, 2001 (ICER WP 2001/5).
2. Massimo Marinacci, "Probabilistic Sophistication and Multiple Priors," *Econometrica*, forthcoming (ICER WP 2001/8).
3. Massimo Marinacci and Luigi Montrucchio, "Subcalculus for Set Functions and Cores of TU Games," April 2001 (ICER WP 2001/9).
4. Juan Dubra, Fabio Maccheroni, and Efe Ok, "Expected Utility Theory without the Completeness Axiom," April 2001 (ICER WP 2001/11).
5. Adriana Castaldo and Massimo Marinacci, "Random Correspondences as Bundles of Random Variables," April 2001 (ICER WP 2001/12).
6. Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi, "A Subjective Spin on Roulette Wheels," July 2001 (ICER WP 2001/17).
7. Domenico Menicucci, "Optimal Two-Object Auctions with Synergies," July 2001 (ICER WP 2001/18).
8. Paolo Ghirardato and Massimo Marinacci, "Risk, Ambiguity, and the Separation of Tastes and Beliefs," *Mathematics of Operations Research* 26, 864-890, 2001 (ICER WP 2001/21).
9. Andrea Roncoroni, "Change of Numeraire for Affine Arbitrage Pricing Models Driven By Multifactor Market Point Processes," September 2001 (ICER WP 2001/22).
10. Maitreesh Ghatak, Massimo Morelli, and Tomas Sjoström, "Credit rationing, wealth inequality, and allocation of talent", September 2001 (ICER WP 2001/23).
11. Fabio Maccheroni and William H. Ruckle, "BV as a Dual Space," *Rendiconti del Seminario Matematico dell'Università di Padova*, forthcoming (ICER WP 2001/29).
12. Fabio Maccheroni, "Yaari Dual Theory without the Completeness Axiom," October 2001 (ICER WP 2001/30).
13. Umberto Cherubini and Elisa Luciano, "Multivariate Option Pricing with Copulas," January 2002 (ICER WP 2002/5).
14. Umberto Cherubini and Elisa Luciano, "Pricing Vulnerable Options with Copulas," January 2002 (ICER WP 2002/6).
15. Steven Haberman and Elena Vigna, "Optimal investment strategies and risk measures in defined contribution pension schemes," *Insurance: Mathematics and Economics*, forthcoming (ICER WP 2002/10).
16. Enrico Diecidue and Fabio Maccheroni, "Coherence without additivity," *Journal of Mathematical Psychology*, forthcoming (ICER WP 2002/11).
17. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, "Ambiguity from the differential viewpoint," April 2002 (ICER WP 2002/17).