Allais, Ellsberg, and Preferences for Hedging*

Mark Dean† and Pietro Ortoleva‡

February 14, 2012

Abstract

We study the relation between ambiguity aversion and the Allais paradox. To this end, we introduce a novel definition of hedging which applies to objective lotteries as well as to uncertain acts, and we use it to define a novel axiom that captures a preference for hedging which generalizes the one of Schmeidler (1989). We argue how this generalized axiom captures both aversion to ambiguity, and attraction towards certainty for objective lotteries. We show that this axiom, together with other standard ones, is equivalent to to two representations both of which generalize the MaxMin Expected Utility model of Gilboa and Schmeidler (1989). In both, the agent reacts to ambiguity using multiple priors, but does not use expected utility to evaluate objective lotteries. In our first representation, the agent treats *objective* lotteries as 'ambiguous objects,' and use a set of priors to evaluate them. In the second, equivalent representation, lotteries are evaluated by distorting probabilities as in the Rank-Dependent Utility model, but using the worst from a set of such distortions. Finally, we show how a preference for hedging is not sufficient to guarantee an Ellsberg-like behavior if the agent violate expected utility for objective lotteries. We then provide an axiom that guarantees that this is the case, and find an associated representation in which the agent first maps acts to an objective lottery using the worst of the priors in a set; then evaluates this lottery using the worst distortion from a set of concave Rank-Dependent Utility functionals.

JEL: *D81*
Keywords: *Ambiguity Aversion, Allais Paradox, Ellsberg Paradox, Hedging, Multiple Priors, Subjective Mixture, Probability Weighting, Rank Dependent Expected Utility*

1 Introduction

In last decades a large amount of empirical and theoretical work has been devoted to the study of two classes of paradox in individual decision making: 1) violations of von-Neumann and Morgestern Expected Utility (EU) for objective risk – most notably the Allais paradox; 2) violations of (Savage) Expected Utility for subjective uncertainty – usually called 'ambiguity aversion' (as demonstrated by the Ellsberg paradox). Together these behaviors constitute two of the most widely studied and robust phenomena in experimental economics and psychology of individual decision making.

*We thank Simone Cerreia-Vioglio, Eddie Dekel, David Dillengerger, Federico Echenique, Paolo Ghirardato, Anna Gumen, Edi Karni, Peter Klibanoff, Fabio Maccheroni, Massimo Marinacci, Stefania Minardi, Efe Ok, Leonardo Pejsachowicz, Philipp Sadowski, Todd Sarver, Andrei Savochkin, Kyoungwon Seo, Marciano Siniscalchi, and the participants of seminars at Bocconi, European University Institute, Northwestern, U. Michigan (Ann Arbor), and Pompeu Fabra for very useful comments and suggestions.

†Department of Economics, Brown University, Robinson Hall, 64 Waterman Street, Providence, RI 02912 USA. Email: mark_dean@brown.edu.

‡Division of the Humanities and Social Sciences, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125 USA. Email: ortoleva@caltech.edu.

These empirical findings have then led to the emergence of two vast theoretical literatures aimed at generalizing standard models to account for each of the two patterns above. However, while there are many theoretical models that address each of the two phenomena separately, much less attention has been devoted to the study of the *relation* between them, and even less to the development of models that allow for a decision maker to exhibit *both* behaviors at the same time. On the one hand the vast majority of models designed to explain Allais-like behavior only look at objective probabilities, thus have nothing to say about behavior under subjective uncertainty. On the other, models that study ambiguity aversion either do not consider objective probabilities at all (as in the setup of Savage (1954)) or, if they do, explicitly assume that the agent follows Expected Utility to assess them.[1]

Despite these largely separate analyses, the idea of a connection between the two classes of behavior has been informally present for decades: loosely speaking, one might expect a decision maker who is 'pessimistic' about the outcome of risky and uncertain events to display *both* violations of EU – both the Allais paradox and ambiguity aversion. This conceptual connection is coupled by a technical one: both phenomena can been seen as violations of some form of linearity/independence of preferences of the agent.[2]

In this paper, we develop a link between ambiguity aversion and Allais-type behavior based on the concept of *preference for hedging*. Since Schmeidler (1989), such preferences have been used as the principle way of capturing ambiguity aversion in situations of subjective uncertainty. The first contribution of this paper is to introduce a generalization of the notion of preference for hedging that can be applied not only to subjective uncertainty, but also to objective risk. We will argue how this generalized notion, which is defined using the concept of subjective (or outcome) mixtures developed by Ghirardato et al. (2003), captures not only aversion to ambiguity, but also Allais-type preferences.

Our second contribution is to use this axiom, along with other standard ones, to characterize two equivalent models in the classic setup of Anscombe and Aumann (1963) that allow for *both* Allais-like violations of objective EU and ambiguity aversion to be present *at the same time* – to our knowledge the first axiomatized model to do so. Both models are defined in the classical setup of Anscombe and Aumann (1963) and generalize the MaxMin Expected Utility model of Gilboa and Schmeidler (1989) (MMEU). In particular, in both of them the decision maker reacts to subjective uncertainty by minimizing over multiple priors over states, as in MMEU. In the first representation, she also treats objective lotteries as 'ambiguous objects,' and evaluates any given lottey using the worst of a set of priors. In the second equivalent representation, lotteries are evaluated by distorting probabilities as in the Rank-Dependent Utility model, but using the worst from a set of such distortions. As this second representation makes clear, our treatment of objective probability is strict generalization of the Rank Dependent Expected Utility model of Quiggin (1982) with concave (pessimistic) distortions.

While in both representations above the decision maker always has a preference for hedging,

[1] Three exceptions are Wakker (2001), Klibanoff et al. (2005) and Drapeau and Kupper (2010). The first focusses on the link between Allais-type and Ellsberg-type behavior and convex capacities in the Choquet Expected Utility Framework. In the second, a corollary of the main theorem generalizes the representation to the case of non-EU preferences on objective lotteries; this representation, however, is not fully axiomatized, and does not model jointly the attitude towards risk and uncertainty. Drapeau and Kupper (2010) allows for non-Expected Utility behavior on both dimensions in the standard setup of Anscombe and Aumann (1963). However, as we shall discuss, they model violation of Expected Utility which need not conform to the Allais paradox, but rather could exhibit the opposite behavior. We refer to Section 4 for more discussion.

[2] There is indeed a literature that discusses connections between violations of objective and subjective EU, most notably by noticing the formal link between the Choquet Expected Utility model for uncertainty and Rank Dependent Expected Utility model for risk. We refer to Section 4 for an analysis of the literature.

and evaluate acts using the worst from a set of possible priors, this does not necessarily imply that she complies with the Ellsberg paradox: she might display the *opposite* behavior if she distorts objective probabilities more than subjective ones, or, equivalently, if her preference for hedging is stronger for the former than for the latter.[3] In fact, the equivalence between a preference for hedging and the Ellsberg paradox can disappear if the decision maker violates Expected Utility on objective risk. The third contribution of this paper is then to introduce a new axiom that, along with preference for hedging, guarantees that the agent's behavior is in line with the Ellsberg paradox. We then use it to characterize a new representation in which we identify a new set of priors to represent the agent's approach to ambiguity net of any additional distortion of objective probabilities. We argue how this set could be seen, in some sense, as the 'true' set of priors used by the agent.

The paper is organized as follows. The remainder of the introduction presents an overview of our main results. Section 2 presents the formal setup, the axioms, and the first representation theorem. Section 3 discuss the special case of the model in which the agent always distorts subjective probabilities more than objective ones, in line with the Ellsberg paradox, and the second representation theorem. Section 4 discusses the relevant literature. Section 5 concludes. The proofs appear in the appendix.

1.1 Summary of Results

The first innovation of our paper is to provide a generalized notion of 'preference for hedging' that can be used to model both violations of EU as in the Allais paradox for objective risk, as well as ambiguity aversion for (subjective) uncertainty. Schmeidler (1989) defined a preference for hedging by positing that, for any two acts that are indifferent to each other, the decision maker prefers to each the $\frac{1}{2}$-mixture between them, that is an act that returns in every state a *lottery* that gives with probability $\frac{1}{2}$ what each of the two original acts would give. While intuitive, however, this notion of mixture is based on the use of objective lotteries, and implicitly assumes that the decision maker follows Expected Utility when evaluating them. An alternative approach was proposed by Ghirardato et al. (2003), and is based on the notion of 'outcome mixtures' of prizes instead of probability mixtures: instead of 'mixing' two objects by creating a lottery that returns each of them as a prize, we look for a third object, in our prize space, the utility of which is 'in the middle' of that of the original two.[4] One of the first steps of our paper is then to extend the idea of outcome mixtures to cover both mixtures of lotteries and of acts. To illustrate our approach, consider an agent who has linear utility for money, so the outcome mixture between \$10 and \$0 is \$5, and a lottery p which returns \$10 and \$0 with probability $\frac{1}{2}$. How could we define the 'outcome mixture' of p with itself, i.e. with an identical lottery? Notice that the idea here is to mix the outcomes that this lottery returns. One way to do it is to mix \$0 with \$0 and \$10 with \$10: we obtain exactly the lottery p. But we can also mix \$0 with \$10, and \$10 with \$0, and obtain a lottery that returns, with *probability 1*, the outcome mixture between \$10 and \$0, i.e. \$5. Both of these lotteries could be seen as mixtures between p and itself. In fact, many others mixtures are possible: for example, we can see p as $\frac{1}{4}$ \$0, $\frac{1}{4}$ \$0, $\frac{1}{4}$ \$10, $\frac{1}{4}$ \$10 and derive many other combinations.[5] Following this intuition, we define the set of all possible mixtures between two lotteries p and q. Once these are

[3]Similar observations appear in Epstein (1999), or in Wakker (2001) in the context of Choquet expected utility. Indeed one might wonder whether such decision maker should be defined as ambiguity averse or not: we refer to Section 3 and, in particular, to footnote 34 for more discussion.

[4]For example, to mix \$0 and \$10 we look for an object the utility of which is exactly in the middle. If the utility was linear this would be \$5, while it would be less less in case of diminishing marginal utility.

[5]In this case, we could mix \$0 with \$0, \$0 with \$10, \$10 with \$0, and \$10 with \$10, and obtain $\frac{1}{4}$ \$0, $\frac{1}{4}$ \$10, and with probability $\frac{1}{2}$ the outcome mixture of \$0 and \$10.

defined, we then also define the mixture between two acts point-wise: for any two acts f and g, the set of mixtures is the set of all acts that return in each state a lottery which could be obtained as a mixture between the lotteries returned by f and g in that state.

Endowed with the notion of outcome mixture of lotteries and acts, we introduce our generalized notion of *Hedging*: for any three acts f, g, and h, if f and g are indifferent to each other and if h could be derived as a mixture between f and g, then h must be *weakly preferred to both.* We argue that this axiom captures 'pessimism' about both subjective uncertainty and objective risk. To illustrate, let us go back to our lottery p which returns \$10 and \$0 with probability $\frac{1}{2}$, and think about its mixtures with itself: we have seen that the set of mixtures includes p itself, but also the degenerate lottery that returns, with probability 1, the outcome mixture between \$10 and \$0, as well as many intermediate mixtures. The key observation is that, loosely speaking, any lottery obtained as a mixture of p with itself is more 'concentrated towards the mean' than p itself (in utility terms) – to the point that one of them is a degenerate lottery. While an EU maximizing agent would be indifferent between these lotteries, a 'pessimistic' agent should like this 'pulling towards the mean,' as it reduces her exposure, and should therefore exhibit a preference for hedging, at least with respect to objective lotteries. When we extend this idea to acts, hedging acquires also the advantage of mixing the outcomes that acts return in each state, just like in the Uncertainty Aversion axiom of Schmeidler (1989) – a reduction of subjective uncertainty which should be valued by subjects who are pessimistic in the sense of ambiguity aversion.

Using this generalized notion of hedging, together with other standard axioms, we derive two, equivalent representations in the standard setup of Anscombe and Aumann (1963). Both representation generalize the MMEU model of Gilboa and Schmeidler (1989),[6] and in both the decision maker evaluates acts in a way similar way: she has a set of priors Π over the states of the world, and she evaluates each act by taking the expectation using the *worst* of the priors in Π. Where both representations differ from MMEU is that the utility of objective lotteries need not follow Expected Utility. In our first representation, for any lottery p in $\Delta(X)$ our agent acts as follows. First, she maps each lottery into an act defined on a hypothetical urn which contains a measure 1 of 'balls,' with the the fraction of balls giving a particular prize equal to the probability of that prize under p.[7] An expected utility agent would use the uniform prior on $[0, 1]$ to evaluate such acts – the Lebesgue measure ℓ. By contrast, our agent has a *set* of priors on $[0, 1]$, which contains ℓ, and evaluates lotteries using the *worst* one of them. This leads her to distort the probabilities of objective lotteries in a 'pessimistic' fashion, while leaving the ranking of degenerate lotteries unchanged, thus exhibiting Allais-style violations of EU.

While in the representation above the decision maker distorts objective probabilities when she evaluates lotteries, the procedure she uses is rather different from others forms of distortion suggested in the literature. Our second representation, which is equivalent to the first, will instead distort probabilities using a procedure in line with the Rank Dependent Expected Utility Model (henceforth RDEU) of Quiggin (1982). In this model, the agent uses a rule similar to Expected Utility, but applies a weighting function to the *cumulative* probability distribution of each lottery. Depending on the shape of this function, the behavior of the agent can be either exhibit Allais-type violations of EU (concave weighting), or the opposite (convex weighting). Then, in our second representation the agent considers a *set* of *concave* probability weightings, the worst of which will be used to evaluate any given lottery. It is easy to see how this is a strict generalization of RDEU

[6] More precisely, they generalize the special case of MMEU in which the utility function over consequences is continuous.

[7] For example, when she faces the lottery $p = \frac{1}{2}x + \frac{1}{2}y$, the agent could think that at some point this lottery will be executed by taking some urn with many balls and saying, for example, that if one the first half of the balls is extracted, then the outcome is x, while if one of the second half is extracted, the outcome is y.

with concave distortion (as the set of distortions can be a made of only one element); in Section 4 we argue how this generalization allows our model to capture some interesting features of pessimism which the RDEU model cannot capture for lotteries with more than two elements in their support.

In both representations there is a sense in which the decision maker could be thought of as ambiguity averse, as she evaluates acts by using the most pessimistic of a set of priors. However, this does *not* necessarily mean that the decision maker will exhibit Ellsberg paradox. If the distortions of objective lotteries are 'stronger' than those of the subjective ones then the opposite of Ellsberg behavior may occur – a possibility which has been noted by Epstein (1999) and Wakker (2001). We introduce a novel additional axiom that rules out this possibility, and show that the addition of this axiom allows us to characterize a third representation. In this representation, the agent again evaluates objective lotteries using the worst RDEU distortion from a set. What now differs is how she evaluates acts. Again, she has a set of priors over the states of the world, $\hat{\Pi}$, but she evaluates each act in two steps: first, she transforms each act into a lottery using the worst prior in $\hat{\Pi}$; second, she evaluate this lottery as she does with other lotteries, using the worst distortion in a set of concave RDEU distortions. Finally, we argue that this set of priors $\hat{\Pi}$ is the *correct* one if we are looking for the set of 'models of the world' used by the decision maker to reduce subjective uncertainty to an objective one – a set which is important to identify if we wish to understand how the decision maker approaches uncertainty, for example if we wish to study how she reacts to new information.[8] We believe this emphasizes the importance of studying the preferences of the agent not only for (Savage) acts, often the object of interest, but for objective lotteries as well. In particular, the latter are no longer used just for mathematical convenience, as was the case when the setup of Anscombe and Aumann (1963) was introduced, but rather are now essential to identify the correct set of models of the world used by the agent to evaluate purely subjective acts.

2 The Model

2.1 Formal Setup

We consider a standard Anscombe-Aumann setup with the additional restrictions that the set of consequences is both connected and compact. More precisely, consider a finite (non-empty) set Ω of *states* of the world, an algebra Σ of subsets of Ω called *events*, and a (non-empty) set X of *consequences*, which we assume to be a connected and compact subset of a metric space.[9] As usual, by $\Delta(X)$ we define the set of simple probability measures over X, while by $\mathcal{F}$ we denote the set of simple Anscombe-Aumann acts: finite-valued, Σ-measurable functions $f : \Omega \to \Delta(X)$. We metrize $\Delta(X)$ in such a way that metric convergence on it coincides with weak convergence of Borel probability measures. Correspondingly, we metrize $\mathcal{F}$ using point-wise convergence.

We use some additional standard notation. For every consequence $x \in X$ we denote by δ_x the degenerate lottery in $\Delta(X)$ which returns x with probability 1. For any $x, y \in X$ and $\alpha \in (0, 1)$, we denote by $\alpha x + (1 - \alpha)y$ the lottery that returns x with probability α, and y with probability $(1 - \alpha)$. For any $p \in \Delta(X)$, we denote by c_p, certainty equivalent of p, the elements of X such that $p \sim \delta_{c_p}$. For any $p \in \Delta(X)$, with the usual slight abuse of notation we denote the constant act in

[8]In fact, it is reasonable to expect that the arrival of new information about the state of the world affects the agent's set of priors over the states, but not how she reacts to objective lotteries. We should therefore expect her to update her set of models of the world $\hat{\Pi}$, but nothing else, thus making the identification of $\hat{\Pi}$ an important step.

[9]It is standard practice to generalize our analysis to the case in which X is a connected and compact topological space. Similarly, our analysis could be also be easily generalized to the case in which the state space is infinite, although in this case the Continuity axiom would have to be adapted: see Section 2.2.1, and specifically the discussion after Axiom 3.

$\mathcal{F}$ such that $p(\omega) = p$ for all $\omega \in \Omega$. Finally, given some $p, q \in \Delta(X)$ and some $E \in \Sigma$, pEq denotes the acts that yields lottery p if E is realized, and q otherwise.

Our primitive is a complete, non degenerate preference relation $\succeq$ on $\mathcal{F}$, whose symmetric and asymmetric components are denoted $\sim$ and $\succ$.

When $|\Omega| = 1$, this setup coincides with a standard preference over vNM lotteries. This is a special case of particular interest for us, and one which we will discuss at length, because our analysis will introduce a new representations for this special case as well, in which the agent is pessimistic in her evaluation of objective lotteries.

We use the setup of Anscombe and Aumann (1963) for convenience, since it allows for the contemporaneous presence of both objective lotteries and of uncertain acts, and because of its widespread use in the literature on ambiguity aversion, thus simplifying the comparison with other models. However, this setup has two features that go beyond simply allowing for both risk and uncertainty: first, it allows them to appear in one object at the same time; second, it entails a specific order in which the two are resolved. It is important to emphasize that our analysis does not depend on either of these features. In particular, as we believe will be easy to see from our analysis, it is straightforward to translate our results into an alternative setup in which preferences are defined over the union of simple vNM lotteries on X, and Savage acts with consequences X, i.e. preferences over $\Delta(X) \cup X^{\Omega}$.

2.2 Axioms and Subjective and Objective Mixtures

2.2.1 Basic Axioms

We start by imposing some basic axioms on our preference relation. To this end, we use the following standard definition of First Order Stochastic Dominance (FOSD).[10]

Definition 1. For any $p, q \in \Delta(X)$, we say that p *First Order Stochastically Dominates* q, denoted $p \trianglerighteq_{FOSD} q$, if $p(\{x : \delta_x \succeq \delta_z\}) \geq q(\{x : \delta_x \succeq \delta_z\})$ for all $z \in X$. We say $p \triangleright_{FOSD} q$, if $p \trianglerighteq_{FOSD} q$ and $p(\{x : \delta_x \succeq \delta_z\}) > q(\{x : \delta_x \succeq \delta_z\})$ for some $z \in X$.

We are now ready to posit some basic standard postulates.

A.1 (FOSD). *For any $p, q \in \Delta(X)$, if $p \trianglerighteq_{FOSD} q$ then $p \succeq q$, and if $p \triangleright_{FOSD} q$ then $p \succ q$.*

A.2 (Monotonicity). *For any $f, g \in \mathcal{F}$ if $f(\omega) \succeq g(\omega)$ for all $\omega \in \Omega$, then $f \succeq g$.*

A.3 (Continuity). *$\succeq$ is continuous: the sets $\{g \in \mathcal{F} : g \succeq f\}$ and $\{g \in \mathcal{F} : g \preceq f\}$ are closed for all $f \in \mathcal{F}$.*

Axiom 1 imposes that our preference relation respects FOSD when applied to objective lotteries. Axiom 2 is the standard monotonicity postulate for acts: if an act f returns a consequence which is better than what another act g returns in *every* state of the world, then f must be preferred to

[10]Since our set of consequences X is a generic (compact and connected) set, then the usual definition of FOSD designed for $\mathbb{R}$ would not apply. The definition that follows is a standard generalization which uses, as a ranking for X, the ranking derived from the preferences on degenerate lotteries. It is easy to see that this definition coincides with the standard definition of FOSD in the special case in which $X \subseteq \mathbb{R}$ and $\delta_x \succeq \delta_y$ iff $x \geq y$ for all $x, y \in X$.

g. Axiom 3 is a standard continuity assumption.[11] [12]

In the standard development of subjective Expected Utility theory, to the axioms above one would add the Independence axiom of Anscombe and Aumann (1963).[13] This axiom, together with Axiom 1-3, would have two implications: 1) that the decision maker is a Expected Utility maximizer with respect to objective lotteries; 2) that the decision maker is a Subjective Expected Utility Maximizer with respect to acts. (See Anscombe and Aumann (1963).) As we are interested in violations of both subjective and objective expected utility maximization, this would then be too strong for our analysis. To accommodate for ambiguity aversion, in the literature it is then standard practice to posit a much weaker axiom: Risk Independence, which postulates independence only for constant acts (objective lotteries).[14] This axiom is imposed by virtually all the models defined in the setup of Anscombe and Aumann (1963). However, since we are explicitly aiming to model Allais-style violations of objective expected utility, also this weaker axiom is too strong for our analysis. We will therefore have to depart radically from this approach.

2.2.2 Outcomes-Mixtures of Consequences: the approach of Ghirardato et al. (2003)

One of the goals of this paper is to capture behaviorally the agent's aversion to 'exposure to risk' by extending the notion of a preference for hedging to objective lotteries. Loosely speaking, a preference for hedging means that, if an agent is indifferent between two lotteries, she weakly prefers to both a 'mixture' between them that reduces the overall exposure to risk. To define this, however, we need to define what we mean by a 'mixture' between two lotteries that reduces the exposure to risk – what is usually understood with the idea of 'hedging.' The traditional approach is to define this mixture by creating a more complicated lottery that returns each prize $x \in X$ with a probability which is a convex combination of the probabilities assigned to x by the original lotteries. (We shall refer to these mixtures as *probability-mixtures*.) This approach, however, will not work for our analysis, as the process of probability mixing changes the probabilities of various prizes in a way that might *increase* exposure to risk – thus not providing the 'hedging' that we are looking for. For example, any mixture of this kind between two degenerate lotteries δ_x and δ_y becomes a non-degenerate lottery, introducing some exposure to risk which wasn't there before. And since our agents need not follow Expected Utility and are potentially averse to exposure to risk, then this kind of mixture won't be appropriate for us.

In this paper we will instead introduce the alternative notion of *outcome mixtures* of lotteries that, we will argue, provides the form of 'hedging' that we are looking for. We begin, in this section, by defining the notion of outcome mixture for the consequences in X, following the approach of

[11] We should emphasize that, although Axiom 3 is entirely standard, it is stronger than Archimedean Continuity, often assumed in this literature, which only posits that the sets $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succeq h\}$ and $\{\alpha \in [0,1] : h \succeq \alpha f + (1-\alpha)g\}$ are closed. The main difference is that our axiom above guarantees that the utility function in the representation is also continuous.

[12] Although we assume that the state space Ω is finite, as we mentioned before our analysis could easily be extended to the general case of an infinite state space. To do this, however, we would have to adapt the Continuity axiom by requiring Archimedean continuity on acts, and full continuity on lotteries. That is, we would require: 1) $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succeq h\}$ and $\{\alpha \in [0,1] : h \succeq \alpha f + (1-\alpha)g\}$ are closed; and 2) $\{q \in \Delta(X) : q \succeq p\}$ and $\{q \in \Delta(X) : q \preceq p\}$ are closed for all $p \in \Delta(X)$. We would then obtain representations identical to ours, but in which the measures over Ω are just finitely additive and not necessarily countably additive. If, in addition, we wanted also to obtain countable additivity, we would have to further assume Arrow's Monotone Continuity Axiom (see Chateauneuf et al. (2005)).

[13] The Independence axioms posits that for every $f, g, h \in \mathcal{F}$, and for every $\alpha \in [0,1]$ we have $f \succeq g$ if and only $\alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h$.

[14] The Risk Independence Axiom posits that for every $p, q, r \in \Delta(X)$, and for every $\alpha \in [0,1]$ we have $p \succeq q$ if and only if $\alpha p + (1-\alpha)r \succeq \alpha q + (1-\alpha)r$.

Ghirardato et al. (2003).[15] In the next section we will extend this idea to outcome mixture of lotteries and of acts.

Consider two consequences $x, y \in X$ and suppose that, in the context of some model, the agent assigns utilities to all elements of X, and that we wish to identify the element with a utility precisely in between that of x and y – i.e. the *consequence* z in X that has has a utility which is exactly in the middle between that of x and y. For example, if we knew that the utility function of the agent were linear on $X \subseteq \mathbb{R}$, we could simply take the element $\frac{1}{2}x + \frac{1}{2}y$. Of course this is in general we do not want to restrict ourselves to linear utility. However, if the set of consequences X is connected, and if preferences are well-behaved enough (in a sense that we shall discuss below), then Ghirardato et al. (2003) introduce a technique which allows us to elicit this element for any (continuous) utility function. In what follows we adapt their technique, originally developed for Savage acts, to the case of objective lotteries with weight $\frac{1}{2}$.

Definition 2. For any $x, y \in X$, if $x \succeq y$ we say that $z \in X$ is a $\frac{1}{2}$-*mixture* of x and y, if $\delta_x \succeq \delta_z \succeq \delta_y$ and

$$\frac{1}{2}x + \frac{1}{2}y \sim \tfrac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z} + \tfrac{1}{2}c_{\frac{1}{2}z+\frac{1}{2}y}. \tag{1}$$

We denote z by $\frac{1}{2}x \oplus_{\succeq} \frac{1}{2}y$.

The rationale of the definition above is the following. Consider some $x, y, z \in X$ such that $\delta_x \succeq \delta_z \succeq \delta_y$. Suppose now also that (1) holds. Then, we know that the agent is indifferent between either receiving the probability mixture between x and y, or first taking the probability mixture between x and z, and then the probability mixture between z and y – which would hold if z had a utility exactly half-way between that of x and y. This is trivially true under expected utility. Ghirardato et al. (2003) show that it is also true for all preferences in the much broader class of 'locally bi-separable' preferences – essentially, those for which a cardinally unique utility function can be identified. (See below for a formal definition.) Thanks to our structural assumption, this notion is well-defined: since X is a connected set and preferences are continuous, for any $x, y \in X$ there must exist some $z \in X$ such that $z = \frac{1}{2}x \oplus_{\succeq} \frac{1}{2}y$. We refer to Ghirardato et al. (2001, 2003) for further discussion. We denote $\oplus_{\succeq}$ using the preferences as a subscript to emphasize how such outcome-mixture depends on the original preference relation. However, in most of the following discussion we drop the subscript for simplicity of notation. Once preferences averages between two elements are defined as above, we can then define any other mixture $\lambda x \oplus (1-\lambda)y$ for any dyadic rational $\lambda \in (0, 1)$ simply by applying the definition above iteratively.[16]

Even though the formal concept above is well defined in our setting, without more structure on the preferences there is no sense in which we can guarantee that the utility of an outcome mixture z is in the middle of x and y: for one thing, this presumes that the very notion of utility is well-defined and, in some sense, unique. We now provide a necessary and sufficient condition to guarantee that this is the case. Consider some x, y, z', z'' such that z' and z'' are "in between" x and y. Then, consider the following two lotteries: $\frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z'} + \frac{1}{2}c_{\frac{1}{2}y+\frac{1}{2}z''}$ and $\frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z''} + \frac{1}{2}c_{\frac{1}{2}y+\frac{1}{2}z'}$. The only difference between them is that: in the former x is mixed with z', and y with z'', and

[15] Similar approaches to define mixtures of consequences were used in Wakker (1994), Kobberling and Wakker (2003), and in the many references therein.

[16] Any $\lambda \in [0, 1]$ is dyadic rational if for some finite N, we have $\lambda = \sum_{i=1}^{N} a_i/2^i$, where $a_i \in (0, 1)$ for every i and $a_N = 1$. Then, we use $\lambda x \oplus (1-\lambda)y$ as a short-hand for the iterated preference average $\frac{1}{2}z_1 \oplus \frac{1}{2}(\ldots(\frac{1}{2}z_{N-1} \oplus \frac{1}{2}(\frac{1}{2}z_N \oplus \frac{1}{2}y))\ldots)$, where for every i, $z_i = x$ if $a_i = 1$ and $z_i = y$ otherwise. Alternatively, we could have defined $\lambda x \oplus \lambda y$ for any real number $\lambda \in (0, 1)$, by defining it for dyadic rationals first, and then using continuity of the preferences to define it for the whole $(0, 1)$. The two approaches are clearly identical in our axiomatic structure; we choose to use the most restrictive definition to state the axioms in the weakest form we are aware of.

then they are mixed together; in the latter x is mixed first with z'', and y with z', and then they are mixed together. In both cases, the only weights involved are weights $\frac{1}{2}$, and x is always mixed with some element worse than it, while y is mixed with some element better than it. The only difference is in the 'order' of this mixture. The following axiom imposes that the agent should be indifferent between these two lotteries – she should not care about such 'order.'

A.4 (Objective Tradeoff-Consistency). *For any $x, y, z', z'' \in X$ such that $\delta_x \succeq \delta_{z'} \succeq \delta_y$, $\delta_x \succeq \delta_{z''} \succeq \delta_y$, and $c_{\frac{1}{2}r+\frac{1}{2}s}$ exists for $r = x, y$ and $s = z', z''$. Then, we have*

$$\frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z'} + \frac{1}{2}c_{\frac{1}{2}y+\frac{1}{2}z''} \sim \frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z''} + \frac{1}{2}c_{\frac{1}{2}y+\frac{1}{2}z'}.$$

The axiom above is clearly implied by Risk Independence – following which both lotteries would be indifferent to a lottery that returns each option with probability $\frac{1}{4}$. At the same time, it is much, much weaker than it. For example, it is easy to see that it is compatible with the behavior of an agent who evaluates each lottery of the form $\frac{1}{2}a + \frac{1}{2}b$, where $\delta_a \succeq \delta_b$, by the functional $\gamma(\frac{1}{2})u(a) + (1 - \gamma(\frac{1}{2}))u(b)$, where $\gamma(\frac{1}{2})$ could be any number between 0 and 1 – the elements $\gamma(\frac{1}{2})$ would 'cancel out' leading to the indifference required by the axiom. That is, Axiom 4 does not rule out even extreme forms of probability weighting. It is not hard to see how essentially all generalizations of Expected Utility that have been suggested in the literature satisfy this axiom – from RDEU to Disappointment Aversion.[17] Axioms of this form are not uncommon in the literature: one can easily see Axiom 4 as an adaptation of the E-Substitution Axiom in Ghirardato et al. (2001) for the case of objective lotteries.

Following this literature it is then straightforward to show how Axioms 1-4 are enough to guarantee that the representation that we just hinted to is not only sufficient, but also necessary. The following Lemma is a trivial consequence of (Ghirardato et al., 2001, Lemma 1). (The proof is therefore omitted.)

Lemma 1. *A preference relation satisfies Axioms 1-4 if and only if they are (locally) biseparable with a continuous utility function u, i.e. there exists a cardinally unique continuous utility function $u : X \to \mathbb{R}$ and a parameter $\gamma(\frac{1}{2}) \in (0, 1)$ such that, for any $x, y, z, w \in X$ with $\delta_x \succeq \delta_y$ and $\delta_z \succeq \delta_w$, we have that $\frac{1}{2}x+\frac{1}{2}y \succeq \frac{1}{2}z+\frac{1}{2}w$ if, and only if, $\gamma(\frac{1}{2})u(x)+(1-\gamma(\frac{1}{2}))u(y) \geq \gamma(\frac{1}{2})u(z)+(1-\gamma(\frac{1}{2}))u(w)$.*

The Lemma above shows that Axiom 4 is enough to guarantee that there is a meaningful way in which the outcome mixture of definition 2 locates a z which has a utility half way between that of x and y.

Before we proceed, we note how in this section we have defined, following the literature, one particular way of identifying the outcome mixture of two consequences in X. However, one could think of many other ways of doing so. For example, if for some reason we knew that the utility of an agent is linear on X, we could have simply used a standard convex combination. It is important to note that the analysis that we present below would work with *any* way of defining outcome mixtures. The generalization to mixture of lotteries and acts that we are about to discuss, or the generalized notion of hedging that we will introduce later, are all independent on how the mixture between elements of X is obtained – provided that it does capture the point with a utility precisely 'in the middle.'

[17] One exception to this is the Weighted Utility Model of Chew (1983).

2.2.3 Outcome-mixtures of Lotteries and Acts

One of the key (and, to our knowledge, novel) contributions of our paper is to extend the notion of outcome mixtures to mixture of *lotteries* and of *acts*. This will be an essential step in defining the notion of hedging which is the core of our analysis.

We begin by extending the concept to a mixture of lotteries. First, consider the simplest case: two degenerate lotteries. Their mixture can be easily defined following the notion above: for any two $x, y \in X$, the mixture between δ_x and δ_y is the degenerate lottery $\delta_{\lambda x \oplus (1-\lambda) y}$. We can similarly define the mixture between a degenerate lottery δ_y and a generic lottery $p \in \Delta(X)$: replace what p returns with the $\oplus$-mixture with y, keeping the probabilities constant. That is, for every $p \in \Delta(X)$ with $p = \sum p(x_i)\delta_{x_i}$, we define the mixture as $\alpha p \oplus (1-\alpha)\delta_y = \sum p(x_i)\delta_{\alpha x_i \oplus (1-\alpha) y}$.

Less straightforward, however, is to define outcome-mixture of two non-degenerate lotteries, mainly because there are *many* possible ways to do it. To see why, consider two lotteries $p = \frac{1}{2}x + \frac{1}{2}y$ and $q = \frac{1}{2}z + \frac{1}{2}w$. How can we define a mixture between them? We could combine x with z, and y with w, and we obtain the lottery $\frac{1}{2}(\frac{1}{2}x \oplus \frac{1}{2}z) + \frac{1}{2}(\frac{1}{2}y \oplus \frac{1}{2}w)$. Or, we could combine x with w, and y with z, and then obtain a different lottery. But we can also see p as $p = \frac{1}{4}x + \frac{1}{4}x + \frac{1}{4}y + \frac{1}{4}y$ and q as $q = \frac{1}{4}z + \frac{1}{4}z + \frac{1}{4}w + \frac{1}{4}w$, and combine them in yet many other ways. Or, decompose them differently, to find many other combinations. All of this shows that there is a large number of ways to combine these two lotteries. We denote by $\bigoplus_{p,q}^{\frac{1}{2}}$ the set of all such mixtures.[18]

Alternatively, we could interpret the set $\bigoplus_{p,q}^{\frac{1}{2}}$ as follows.[19] Under the assumption that our agent has a well-defined utility function over the set X, we could see each lottery $p \in \Delta(X)$ as a *random variable* V_p that assigns to each event in $[0,1]$ a certain utility, i.e. $p = V_p : [0,1] \to \mathbb{R}$. Since the $\oplus$-mixture is nothing but a mixture of the utilities, then the mixtures in $\bigoplus_{p,q}^{\frac{1}{2}}$ could be seen as the mixtures between the random variables corresponding to p and q. But of course to define the mixture between two random variables we need to know their *covariance*. The set $\bigoplus_{p,q}^{\frac{1}{2}}$ could then be seen as the set of all mixtures with weight $\frac{1}{2}$ of the random variable corresponding to p and q for *any* possible covariance. The multiplicity of mixtures would then stem from the multiple possible covariances that could be found.

Finally, we define the notion of outcome-mixture for acts. We do so point-wise: an act h is a $\frac{1}{2}$-mixture between two acts f and g if $h(\omega) \in \bigoplus_{f(\omega),g(\omega)}^{\frac{1}{2}}$ for all $\omega \in \Omega$, that is, if for every state it returns a lottery which is a mixture between the lotteries returned by f or g. We denote $\bigoplus_{f,g}^{\frac{1}{2}}$ the set of all such mixtures of two acts.

2.2.4 Main Axioms

Now that we are endowed with the notions of outcome mixtures, we can use them to define the main axiom: *hedging*. We begin with a simple example. Consider some lottery $p = \frac{1}{2}x + \frac{1}{2}y$, and

[18] Formally, $\bigoplus_{p,q}^{\frac{1}{2}}$ is constructed as follows. Consider any lottery p and q, and notice that, because both are simple lotteries, we could always find some $x_1, \dots x_n, y_1, \dots, y_n \in X$, and some $\gamma_1, \dots, \gamma_n \in [0,1]$ such that $p = \sum_{i=1}^n \gamma_i \delta_{x_i}$, $q = \sum_{i=1}^n \gamma_i \delta_{y_i}$. (For example, the lotteries $p = \frac{1}{2}x + \frac{1}{2}y$ and $q = \frac{1}{3}z + \frac{2}{3}w$ could be both written as $p = \frac{1}{3}x + \frac{1}{6}x + \frac{1}{6}y + \frac{1}{3}y$ and $q = \frac{1}{3}z + \frac{1}{6}w + \frac{1}{6}w + \frac{1}{3}w$.) Then, the set $\bigoplus_{p,q}^{\frac{1}{2}}$ will be the set of all combinations r such that $r = \sum_{i=1}^n \gamma_i(\frac{1}{2}x_i \oplus \frac{1}{2}y_i)$. That is, we have $\bigoplus_{p,q}^{\frac{1}{2}} := \{r \in \Delta(X) : \exists x_1, \dots x_n, y_1, \dots, y_n \in X, \exists \gamma_1, \dots, \gamma_n \in [0,1]$ such that $p = \sum_{i=1}^n \gamma_i \delta_{x_i}, q = \sum_{i=1}^n \gamma_i \delta_{y_i}$ and $r = \sum_{i=1}^n \gamma_i \delta_{(\frac{1}{2}x_i \oplus \frac{1}{2}y_i)}\}$.

[19] We thank Fabio Maccheroni for suggesting the following interpretation.

the lottery r that could be obtained by mixing p with itself, i.e. $r \in \bigoplus_{p,p}^{\frac{1}{2}}$. We would argue that an agent who is, in some sense, 'averse to exposure to risk,' should rank r as *at least as good as* p. To wit, notice how r is constructed. At one extreme, it could be constructed by mixing x with x, and y with y, generating $r = p$ – so r is at least as good as p. At the other, r could be obtained by mixing x and y, and y with x, giving us $r = \delta_{\frac{1}{2}x \oplus \frac{1}{2}y}$. That is, r would become a *degenerate lottery* the utility of which is exactly in the middle between that of x and y. An agent who is attracted to certainty, and who dislikes exposure to risk, will then like r at least as much as p. A similar argument would naturally hold for any other way of constructing r: for example, we could have $r = \frac{1}{4}x + \frac{1}{2}\delta_{\frac{1}{2}x \oplus \frac{1}{2}y} + \frac{1}{4}y$, which once again will be at least as good as p for any agent who dislikes exposure to risk. The intuition here is simple: any lottery in $\bigoplus_{p,p}^{\frac{1}{2}}$ has the same *expected* utility as p, but has (weakly) lower variance in utility. In this sense, by mixing good with bad outcomes, the process of hedging reduces the exposure to risk. This means that an agent who is attracted towards such reduction should exhibit a preferences for hedging.

A similar argument naturally applies to hedging between different lotteries. To wit, consider two lotteries $p = \frac{1}{2}x + \frac{1}{2}y$ and $q = \frac{1}{2}z + \frac{1}{2}w$, where $x, z \succ y, w$, and suppose that p is indifferent to q. Take some r which could be obtained by mixing p and q with weight $\frac{1}{2}$, i.e. $r \in \bigoplus_{p,q}^{\frac{1}{2}}$. Again, an agent who is 'averse to exposure to risk' should like r as at least as much as p and q. On the one hand, r can formed by mixing the two 'good' elements with each other (x and z), and the two 'bad' ones (y and w) with each other. But since p and q are indifferent to each other, then these mixture should not be worse for the agent: the expected utility of r will be halfway between that of p and q, but its variance in utility terms must be weakly less than average of the variance of p and q. On the other hand, r could be formed by mixing the good element in p with the bad element in q, and vice-versa, giving us $r = \delta_{\frac{1}{2}x \oplus \frac{1}{2}w} + \frac{1}{2}\delta_{\frac{1}{2}y \oplus \frac{1}{2}z}$. Again, in this case we would have that the process of hedging is similar to 'pulling extremes towards the center', reducing the variability: so an agent who is averse to this variability should not be averse to hedging. This lead us to argue that if we wish to posit an 'aversion to exposure to risk,' we could posit that for any $p, q, r \in \Delta(X)$ such that $p \sim q$ and $r \in \bigoplus_{p,q}^{\frac{1}{2}}$, we should have $r \succeq p \sim q$.

We now extend this argument to hedging between *acts*. For simplicity, consider now two non-degenerate acts $f, g \in \mathcal{F}$ such that $f \sim g$ and such that $f(\omega)$ and $g(\omega)$ are *degenerate* lotteries for all ω. Now consider some $h \in \bigoplus_{f,g}^{\frac{1}{2}}$, and notice that $h(\omega) = \delta_{\frac{1}{2}f(\omega) \oplus \frac{1}{2}g(\omega)}$. Since there are no lotteries involved, going from f and g to h does not affect the exposure to *risk* – in either case, there is none. But it will reduce the exposure to *ambiguity*: this is precisely the idea of the original hedging axiom of Schmeidler (1989).[20] An agent who is not ambiguity seeking would then (weakly) prefer hedging, and she will rank h as at least as highly as f and g. Combining the two arguments of attraction towards hedging for lotteries and for acts, we then obtain the following axiom – the main postulate of the paper.

A.5 (Hedging). *For any $f, g \in \mathcal{F}$, and for any $h \in \bigoplus_{f,g}^{\frac{1}{2}}$, if $f \sim g$, then $h \succeq f$.*

This axiom can be seen as capturing *both* pessimism in the form of ambiguity aversion, *and* pessimism in the form of violations of EU in the direction suggested by the Allais paradox.

[20]In fact, if, additionally, we impose Risk Independence, then preference for hedging in outcome mixtures is *identical* to preference for hedging in probabilities. In turns, when applied only to acts that map to degenerate lotteries, this is precisely the axiom suggested in Ghirardato et al. (2003).

Our final axiom is the translation of the idea of the Certainty-Independence axiom of Gilboa and Schmeidler (1989) to our setup:[21] when two acts are mixed with a 'neutral' element, their ranking should not change. As opposed to Certainty-independence, however, the 'neutral element' will not only be a constant acts, but a *degenerate lottery*, which is 'neutral' from the point of view of both risk and ambiguity. Moreover, we will use outcome-mixtures instead of probability ones, because our agent could have non-linear reactions to probability mixtures.[22]

A.6 (Degenerate-Independence (DI)). *For any $f, g \in \mathcal{F}$, dyadic $\alpha \in (0,1)$, and for any $x \in X$,*

$$f \succeq g \quad \Leftrightarrow \quad \alpha f \oplus (1-\alpha)\delta_x \succeq \alpha g \oplus (1-\alpha)\delta_x.$$

2.3 First Representation: subjective view of objective risk

We are now ready to introduce our first representation. To better express it, it will be useful to define the notion of a *measure-preserving* map from lotteries into acts on $[0,1]$. The idea is simple: we can *map* every objective lottery $p \in \Delta(X)$ to an act defined on the space $[0,1]$ that assigns to each state in $[0,1]$ a consequence in X. It is as if the agent imagined that, to determine the prize of the objective lottery p, an imaginary 'urn' of size $[0,1]$ will be used: after assigning to each ball in this imaginary urn a consequence in X, thus creating an act from $[0,1]$ to X, there will be an extraction which will determine the final prize. (A visualization which is rather close to being true in most experimental settings.) For example, the lottery $p = \frac{1}{2}x + \frac{1}{2}y$ could be mapped to the act on $[0,1]$ that returns x after the states $[0, \frac{1}{2})$, and returns y after the states $[\frac{1}{2}, 1]$. Indeed there are many possible such maps; we focus only on those in which each lottery is mapped to an act such that the Lebesgue measure of the states which return a given prize is identical to the probability assigned to that prize by the original lottery. We call these *measure-preserving* maps.[23]

Definition 3. We say that a function $\mu : \Delta(X) \to [0,1]^X$ is *measure-preserving* if for all $p \in \Delta(X)$ and all $x \in X$, $\ell(\mu^{-1}(x)) = p(x)$.

We can then introduce our first representation, the Multiple Priors-Multiple Distortions representation.

Definition 4. Consider a complete and non-degenerate preference relation $\succeq$ on $\mathcal{F}$. We say that $\succeq$ admits a *Multiple Priors and Multiple Distortions Representation (MP-MD)* (u, Π, Φ) if there there exists a continuous *utility function* $u : X \to \mathbb{R}$, a convex and compact set of probability measures Π on Ω, and a convex and weak-compact set of Borel probability measures Φ on $[0,1]$, which contains the Lebesgue measure ℓ and such that every $\phi \in \Phi$ is atomless and mutually absolutely continuous with respect to ℓ, such that $\succeq$ is represented by the functional

$$V(f) := \min_{\pi \in \Pi} \int_{\Omega} \pi(\omega)\ U(f(\omega)) \mathrm{d}\omega.$$

where $U : \Delta(X) \to \mathbb{R}$ is defined as

$$U(p) := \min_{\phi \in \Phi} \int_{[0,1]} \phi(s)\ u(\gamma(p)(s)) \mathrm{d}s$$

[21] A preference relation satisfies *Certainty-Indepedence* if for any $f, g \in \mathcal{F}$, and for any $p \in \Delta(X)$ and $\lambda \in (0,1)$, we have $f \succeq g$ iff $\lambda f + (1-\lambda)p \succeq \lambda g + (1-\lambda)p$.

[22] It is not hard to see that this axiom is actually *strictly weaker* than Certainty-Independence. To wit, notice that the latter implies that the agent satisfies standard independence on constant acts, which in turn implies that probability mixtures and outcome are indifferent for her – we must have $\lambda f + (1-\lambda)\delta_x \sim \lambda f \oplus (1-\lambda)\delta_x$ for all $f \in \mathcal{F}$, $x \in X$, and $\lambda \in (0,1)$. But then, Certainty-Independence would naturally imply the axiom below.

[23] A similar concept is used in Wakker (2001).

for any measure-preserving map $\gamma : \Delta(X) \to [0,1]^X$.

In a Multiple Priors and Multiple Distortions Representation the decision maker is endowed with three elements: a utility function u; a set of priors Π over the states in Ω; and a set of priors Φ over $[0,1]$. With respect to ambiguity, she behaves in a way which is conceptually identical to how she would behave in the MaxMin Expect Utility model of Gilboa and Schmeidler (1989): she has a *set* of priors Π on the states of the world, and she uses the worst one of them to aggregate the utility assigned to the lotteries that the act returns in every state. Where the model above differs from MMEU is in how the evaluation on objective lotteries is done. In particular, our agent need not follow vNM Expected Utility. Instead, first she maps each objective lottery into an act on the space $[0,1]$ (in a measure-preserving fashion). Then she considers a *set* of priors Φ over $[0,1]$, which includes the Lebesgue measure ℓ, and she uses the *worst* one of them to compute the utility of the lottery at hand – much in line with the MMEU model. When $\Phi = \{l\}$ her evaluation of lotteries will be equivalent to vNM Expected Utility, and the model as a whole will coincide with MMEU.[24] But when $\Phi \supset \{l\}$ her ranking of objective lotteries will be different: she will be 'pessimistic' towards them, (weakly) lowering their evaluation by using a prior in Φ which returns a lower expected value that ℓ. Since her valuation of degenerate lotteries will not be affected – as it is independent from the prior in Φ that is used – it is easy to see how this leads to certainty bias and Allais-like behavior. (See Section 2.6.1 for more.)

Let us illustrate this intuition by means of a example, in which we consider an approximation in which our urn contains only 100 balls (instead of a measure 1 of them). Our DM acts as if the outcome of any lottery will be determined by drawing a ball from this urn. However, it is as if the probability of drawing each ball is not necessarily $\frac{1}{100}$, but can be larger or smaller depending on whether that ball is associated with a good or bad prize. Specifically, consider the case where the probability of each ball can be between 0.009 and 0.011, thus leading to a set of priors Φ equal to

$$\Phi = \{p \in \Delta(\{1, 100\}) | p_i \in [0.009, 0.011] \ \forall\ i \in 1..100\}.$$

Consider now the lottery $p = \frac{1}{2}\$10 + \frac{1}{2}\$0$, and say it is mapped to an act such that balls 1-50 give \$10 and balls 51-100 give \$0. In this case, our DM would reduce the probabilities associated with balls 1-50 to 45% thinking that she is 'unlucky' and that the 'good balls' will not come out, while at the same time raising to 55% the probabilities associated with balls 51-100, the ones associated with the 'bad' outcome.

One possible interpretation of the agent's behavior in this representation is that she treats, in some sense, objective lotteries like 'ambiguous objects:' it is as if she didn't quite know how to evaluate them – as if they were 'ambiguous' for her – and she reacted by being 'ambiguity averse' towards them, by mapping each lottery to an act and then following a procedure essentially identical to MMEU. The degree of her aversion is given by the size of the set Φ, and, as we mentioned, if Φ is not a singleton our agent will exhibit violations of expected utility of the type exhibited by the Allais paradox.

2.4 Second representation: minimal of RDEU

The Multiple Priors Multiple Distortions representation describes a form of distortion of objective probabilities that is somewhat different from other established ways to distort objective probabilities, such as the one used by the Rank-Dependent Expected Utility (RDEU) representation, in which a 'weighting function' is applied to the cumulative distribution of the objective lottery. In

[24]More precisely, it concedes with the special case of MMEU in which the utility function on consequences is continuous. This implies that the MP-MD model is a generalization of MMEU with a continuos utility function.

what follows we introduce an alternative representation in which our decision maker follows a more standard procedure to distort probabilities.

We start from recalling the Rank-Dependent Expected Utility model of Quiggin (1982) for preferences over the lotteries in $\Delta(X)$.

Definition 5. We say that a function $\psi : [0,1] \to [0,1]$ is a *probability weighting* if it is increasing and it is such that $\psi(0) = 0$, $\psi(1) = 1$. For every non-constant function u and for every probability weighting ψ, we say that a function is a Rank-Dependent Expected Utility function with utility u and weight ψ, denoted $\text{RDEU}_{u,\psi}$, if, for any enumeration of the elements of the support of p such that $x_{i-1} \preceq x_i$ for $i = 2, \dots, |\text{supp}(p)|$, we have

$$\text{RDEU}_{u,\psi}(p) := \psi(p(x_1))u(x_1) + \sum_{i=2}^{n} \big[\psi(\sum_{j=1}^{i} p(x_j)) - \psi(\sum_{j=1}^{i-1} p(x_j))\big] u(x_i). \tag{2}$$

The main feature of the RDEU model is that the decision maker follows a procedure similar to expected utility, except that she distorts the *cumulative* probability distribution of each lottery using a *probability weighting function*. It is well-known that the RDEU model has many desirable properties, such as preserving continuity (as long at the probability weighting is continuous) and FOSD. Depending on the shape of ϕ, moreover, the model allows for attraction or aversion towards certainty: the former takes place when ϕ is concave – leading to to an Allais-like behavior; the opposite takes place when ϕ is convex; when ϕ is linear it coincides with Expected Utility. (See Quiggin (1982), and also Wakker (1994), Nakamura (1995), Chateauneuf (1999), Starmer (2000), Wakker (2001), Abdellaoui (2002), Kobberling and Wakker (2003) and the many references therein.) The RDEU model is arguably the most well-known representation used to study violations of Expected Utility on objective lotteries. The Cumulative Prospect Theory model of Tversky and Kahneman (1992), for example, is built on its framework.

We are now ready to introduce our next representation, which will be similar to a MP-MD representation, but the decision maker will use an RDEU functional to distort objective probabilities. However, as we are trying to capture Allais-like behavior, such functional will be concave (pessimistic); and since we have a MMEU-like representation, we will have *set* of RDEU distortion the worst of which will be used by the agent.

Definition 6. Consider a complete and non-degenerate preference relation $\succeq$ on $\mathcal{F}$. We say that $\succeq$ admits a *Multiple Priors and Multiple Concave Rank-Dependent Representation (MP-MC-RDEU)* (u, Π, Ψ) if there there exists a continuous *utility function* $u : X \to \mathbb{R}$, a convex and compact set of probability measures Π on Ω, and a convex, (point-wise) compact set of differentiable and concave probability weightings Ψ such that $\succeq$ is represented by the functional

$$V(f) := \min_{\pi \in \Pi} \int_{\Omega} \pi(\omega)\; U(f(\omega)) \mathrm{d}\omega.$$

where $U : \Delta(X) \to \mathbb{R}$ is defined as

$$U(p) := \min_{\psi \in \Psi}\; \text{RDEU}_{u,\psi}(p).$$

Just like in the MP-MD representation, in a Multiple Priors and Multiple Concave Rank-Dependent Representation our agent has a *set* of probabilities which she uses to evaluate acts just like the MMEU model. Here, however, instead of using a set of priors over $[0,1]$, she has a set of probability weightings Ψ, and she uses the worst one of those in a RDEU functional to evaluate objective lotteries. This set has two features. First, it is composed only of *concave* – hence pessimistic – distortions. Second, Ψ could also be a singleton: this means that the RDEU model with concave distortion is a special case of the representation above. (In Section 2.6.3 we discuss a comparison with RDEU more in detail.) At the same time, just like the MP-MD representation, also this representation is a generalization of MMEU – they coincide when Ψ contains only the identity function.

Finally, notice that if a preference relation admits a Multiple Priors and Multiple Concave Rank-Dependent Representation (u, Π, Ψ), then in many cases we can enlarge the set Ψ by adding distortions which are less severe than those already present, and that will therefore leave the behavior unchanged – that is, we can add redundant elements.[25] We therefore define a notion of 'minimal' representation, in which these redundant elements are removed.

Definition 7. We say that a set of probability weightings Ψ included in any representation is *minimal* if there is no $\Psi' \subset \Psi$ such that the same preferences can be represented by a representation of the same form that includes Ψ' instead of Ψ.

2.5 Representation Theorem

We are now ready to introduce our representation theorem.

Theorem 1. *Consider a complete and non-degenerate preference relation $\succeq$ on $\mathcal{F}$. Then, the following are equivalent*

(1) $\succeq$ satisfies Axioms 1-6.

(2) $\succeq$ admits a Multiple Priors and Multiple Distortions Representation (u, Π, Φ).

(3) $\succeq$ admits a Multiple Priors and Multiple Concave RDEU Representation (u, Π, Ψ).

Moreover, u is unique up to a positive affine transformation, Π, Φ are unique, and there exists a unique minimal Ψ.

Theorem 1 shows that the axiomatic structure discussed above is equivalent to both representations. That is, imposing a preference for (generalized) hedging, together with our other more standard axioms, is tantamount to positing that the decision-maker has a MMEU-like representation for her ranking of acts, but also that she has a subjective view of objective risk, as it happens in a MP-MD representation. Moreover, this itself is *equivalent* to the existence of an alternative representation in which the agent evaluates objective lotteries using the *min* of a set of concave RDEU functionals. Finally, all the components of both representations are identified uniquely.

2.6 Discussion

2.6.1 Relation with the Allais Paradox

We now turn to describe how both models above could generate certainty bias and an Allais-like behavior. First of all, our main axiom, Hedging, directly implies a form of certainty bias. Consider $x, y, z \in X$ such that $u(z) = \frac{1}{2}u(x) + \frac{1}{2}u(y)$. This implies that we could obtain δ_z as a outcome mixture of $\frac{1}{2}x + \frac{1}{2}y$ with itself. But then, hedging immediately implies that we have $\delta_z \succeq \frac{1}{2}x + \frac{1}{2}y$,

[25] For example, if the identity function doesn't belong to Ψ, we can add it to the set and leave the behavior unchanged; or, we can add any convex combination of any element of Ψ and the identity function, and again leave the behavior unchanged.

leading to (weak) certainty bias. Our first representation has a similar feature: while the evaluation of degenerate lotteries cannot be not distorted (i.e. we must have $V(\delta_x) = U(\delta_x) = u(x)$), that of non-degenerate ones could be if the decision maker uses a prior ϕ more pessimistic than the Lebesgue measure ℓ. An identical argument applies also to our second representation.[26]

Similarly, it is easy to see how the choice pattern of the Allais experiment can be accommodated by both representations: for example, recall that a special case of our second representation is the RDEU model with concave distortions, which is well-known to allow for such behavior. Importantly, moreover, the two models not only allow for Allais-like behavior, but they rule out the possibility of an *opposite* preference. For brevity, we discuss this in Appendix A.

2.6.2 Relation with the Ellsberg Paradox

In both representations discussed above there is a sense in which the decision maker could be seen as ambiguity averse: when $|\Pi| > 1$, the agent has a set of priors and uses the worst one of those to judge uncertain events – just like the MaxMin Expected Utility model of Gilboa and Schmeidler (1989). Importantly, however, this does not imply that the agent will necessarily exhibit the Ellsberg paradox: both the Ellsberg behavior and its *opposite* are compatible with our representations, even if $|\Pi| > 1$.[27] To wit, consider an urn with 100 balls, which could be Red or Black, in unknown proportions. An experimenter will extract a ball from this urn, and the color of the extracted ball determines the state of the world, R or B. We will analyze the Decision Maker's ranking between three acts: betting on R – i.e. getting \$10 if a red ball is extracted, \$0 otherwise; betting on B; and an objective lottery which pays \$10 or \$0 with equal probability. Let us assume, for simplicity, that the decision maker is indifferent between betting on red or black. The typical Ellsberg behavior is that the decision maker strictly prefers the objective lottery to either of the bets. We say that a decision maker exhibits the opposite behavior is she is indifferent between betting on red or black, but strictly prefers both to the objective lottery. We will now show how both patterns are compatible with our representations. The former case is trivially true in a MP-MD representation when $|\Pi| > 1$ and $|\Phi| = 1$: in this case we know that our model coincides with the MaxMin Expected Utility, which is compatible with the Ellsberg behavior. Consider now the case in which $|\Pi| = 1$ and $|\Phi| > 1$: this agent distorts – pessimistically – objective probabilities but not subjective ones, and therefore will prefer betting on red or black rather than betting on the objective lottery – thus exhibiting the opposite of the Ellsberg paradox. A similar result could of course be obtained also when $|\Pi| > 1$, as long as the distortions of objective probabilities are 'stronger' than those of subjective ones. In fact, this is the case because neither of the representations, nor the axioms, posit that the agent should be more pessimistic for subjective risk than she is for objective one, therefore allowing both for Ellsberg and its opposite. As pointed out in Wakker (2001), Ellsberg-type behavior is a product of relative, rather than absolute pessimism, and while our Hedging axiom regulates the latter, it does not rescrit the former.

In Section 3 we address this issue in more depth, and we provide a novel behavioral axiom which will allow us to characterize axiomatically the special case of our model in which the agent never exhibits the opposite of the Ellsberg behavior.

[26]Note, however, that our model does not necessarily guarantee the property of Negative Certainty Independence of Dillenberger (2010). See Section 4 for more.

[27]Examples of this kind appear already in Epstein (1999) and Wakker (2001).

2.6.3 Comparison with RDEU

As evident from the existence of a Multiple Priors and Multiple Concave RDEU representation, our model is much related to Rank-Dependent Expected Utility. In fact, if we focus on the special case in which $|\Omega| = 1$, our model becomes a model of preferences over (vNM) lotteries in which the agent has a *set* of *concave* probability weightings, and uses the worst of them to evaluate objective lotteries using the RDEU functional form. Since our set of probability weightings could be a singleton, the RDEU model with concave distortions becomes a special case of ours.

There are two important behavioral differences between our model and standard RDEU. First, because each probability weighting used in our model is concave, and because the agent uses the worst one of them, then our agent can never exhibit a behavior that goes 'against Allais:' she is either certainty-biased, or she satisfies Expected Utility – she is never 'certainty-averse.' By contrast, the RDEU model is more flexible, as it also allows for certainty-aversion by allowing convex probability weighting. This is naturally due to our focus on pessimistic agents – the basic goal of our paper – via our main axiom, Hedging.

However, once we focus on concave probability weightings, our model is strictly more general than standard RDEU. While in RDEU the agent has a fixed distortion to be used for every lottery, in our representation she may have multiple distortions, and use a different one depending on the lottery at hand. Importantly, while this flexibility is irrelevant when lotteries have only two prizes in their support, it might play an important role in more general cases. To wit, consider the following 2 lotteries: $p = \frac{1}{3}\ \$0 + \frac{1}{3}\ \$1 + \frac{1}{3}\ \$10,000$; $q = \frac{1}{3}\ \$0 + \frac{1}{3}\ \$9,999 + \frac{1}{3}\ \$10,000$. In the RDEU model the agent must use the *same* probability distortion for both p and q – the rank of the three outcomes is the same, and since in RDEU *only* the relative rank matter, the probability distortion is bound to be the same.[28] So the agent is bound always to distort the intermediate outcome in the same way – despite the fact that in p this intermediate outcome is comparably 'very bad,' while in q it is comparably 'very good'. By contrast, our model could accommodate the situation in which in p both the probabilities of \$0 and of \$1 are much overweighted and the probability of \$10,000 is underweighted; while for q only the probability of \$0 is overweighted, and both that of \$9,999 and of \$10,000 are underweighted – a behavior which, we would argue, is more in line with standard notions of pessimism.

The relation between our models and RDEU becomes more evident once we notice that, as has been noted in the literature, the RDEU representation is formally identical to the Choquet Expected Utility model of Schmeidler (1989), one of the most well-known models used to study ambiguity aversion. In particular, in a setup in with a fixed state space, a given set of outcomes, and an *objective* probability distribution over these states, the axioms of Schmeidler (1989) together with a form of First Order Stochastic Dominance leads exactly to RDEU for acts defined on this space.[29] In a similar spirit, we use a generalized version of Schmeidler (1989)'s *hedging* axiom to obtain a representation which is similar to the MaxMin Expected Utility model for the case of risk: it is not hard to see that, at least in a rather loose sense, our model of decision making under risk compares to RDEU in a similar way to which the Choquet Expected Utility compares to the MaxMin Expected Utility model – hence the differences between RDEU and our model discussed above.

A natural question is then to identify the conditions which guarantee that our set of distortions

[28]In fact, this is one of the fundamental and characterizing features of RDEU: see Diecidue and Wakker (2001).

[29]See, among others, Wakker (1990), Chew and Wakker (1996), Wakker (1996), Chateauneuf (1999), and Diecidue and Wakker (2001) for an in-depth analysis. The key component is the use of Schmeidler (1989)'s axiom of Comonotonic Independence, which posits that if we focus only on acts which 'move together' in the sense of agreeing which are the 'good' and 'bad' states, then independence should be satisfied.

Ψ is a singleton – the special case of our model which coincides with concave RDEU on objective lotteries. It is not hard to see that all we need is an axiom that guarantees that the preferences of our agent must be of the RDEU form for objective lotteries. To this end we could use, for example, the Probability trade-off consistency Axiom of Abdellaoui (2002). Or, as the discussion above should make clear, we could simply posit the equivalent of Schmeidler (1989)'s axiom of Comonotonic Independence: we could posit both the Comonotonic Sure-Thing Principle and the Comonotonic Mixture Independence axioms of Chateauneuf (1999), or any other provided by the literature. Together with our other axioms, these will guarantee that we could represent our preferences using a *single* RDEU functional (not necessarily concave) Since from our representation we also know that we can represent as the *min* of a set of *concave* functionals, then we can represent using a unique *concave* RDEU functional.[30]

2.6.4 Maps from lotteries into acts

One of the features of the Multiple Priors Multiple Distortions representation is that our agent maps each objective lottery into an act on $[0,1]$ using a measure-preserving map. This map could take many forms, and the representation guarantees the existence of a set of priors which would work for *any* map, as long as it is measure-preserving. Alternatively, we could have looked for a representation in which the agent uses a *fixed*, specific map, and had a set of priors which would work for this specific map only. For example, we could have focused on the map $\bar{\gamma}$ 'from worst to best:' for any lottery p, enumerate the outcomes in its support from worst to best, i.e. $x_{i-1} \preceq x_i$ for $i = 2, \ldots, |\text{supp}(p)|$, and define $\bar{\gamma}(p)$ as $\bar{\gamma}(p)\big([0, p(x_1)\big) = x_1$ and $\gamma(p)\big([\sum^{i-1} i_{j=1} p(x_i), \sum_{j=1}^{i} p(x_i))\big) = x_i$ for $i = 2, |\text{supp}(p)|$. (Intuitively, γ assigns the worst outcomes to the smallest states in $[0,1]$, and the best outcomes to the higher ones. Then, we have the following observation.) Indeed, if a Multiple Priors Multiple Distortions representation exists, so must a representation with this fixed map alternative representation. At the same time, however, once we focus on a specific map, we can derive additional properties on the set of priors Φ.

Observation 1. Suppose that $\succeq$ admits a MP-MD representation. Then, it must also admit a representation which is identical to an MP-MD representation, but in which: 1) the agent uses the map $\bar{\gamma}$ (defined above) to map lotteries into priors in $[0,1]$; 2) the set of priors on $[0,1]$ used in this representation are all 'decreasing,' i.e. theirs PDFs are all (weakly) decreasing functions. To see why, notice that if $\succeq$ admits a MP-MD representation, then it must also admit a MP-MC-RDEU representation with set of distortions Ψ. For any $\psi \in \Psi$, consider its derivative ψ', and call D the set of all derivatives. Indeed each element of D is a decreasing function, and by construction must integrate to 1 on $[0,1]$. Now consider each member of D as a PDF, and call Φ' the corresponding set of priors on $[0,1]$. It is easy to see that Φ' is in fact the desired set of priors.

2.7 Hedging-Neutrality and Restricted Violations

As we mentioned in the discussion above, one of the features of our representations is that they allow for the simultaneous violations of both Anscombe-Aumann Expected Utility on acts, and of vNM Expected Utility on objective lotteries. We now turn to analyze the behavioral axioms that allow us to restrict violations to only one of these domains. To do this, we can impose various forms of 'hedging neutrality,' i.e. posit that the decision maker has no incentive to hedge in one domain

[30] In turns, this implies that our axioms (esp. Hedging) together with those that characterize RDEU, imply the Attraction for Certainty Axiom of Chateauneuf (1999), or Probabilistic Risk Aversion as defined in Abdellaoui (2002), or the pessimism condition of Wakker (2001), since they are all implied by the existence of a *concave* RDEU representation.

or another. There are three ways in which we could do this: by imposing that the decision maker never has an incentive to hedge; that she never has such incentive between acts that map only to *degenerate* lotteries; that she never has such incentive between degenerate acts. The following axioms formalize this.

A.7 (Hedging Neutrality). *For any $f, g \in \Delta(X)$, and for any $h \in \bigoplus_{p,q}^{\frac{1}{2}}$, if $f \sim g$, then $h \sim f$.*

A.8 (Hedging-Neutrality on Acts). *For any $f, g, h \in \mathcal{F}$ such that $f \sim g$, $h \in \bigoplus_{f,g}^{\frac{1}{2}}$ and such that for all $\omega \in \Omega$ we have $f(\omega) = \delta_x$ and $g(\omega) = \delta_y$ for some $x, y \in X$, we have $h \sim f$.*

A.9 (Hedging-Neutrality on Lotteries). *For any $p, q, r \in \Delta(X)$ such that $p \sim q$ and $r \in \bigoplus_{p,q}^{\frac{1}{2}}$, we have $r \sim p$.*

A different way to capture hedging neutrality is to posit that the agent is indifferent between subjective and objective mixtures. The following axiom imposes this in the weakest possible way: that there exists at least *one* situation in which probability and outcome mixtures coincide. As we shall see below, this will be sufficient to guarantee Independence for all lotteries.

A.10 (Local Neutrality for Subjective and Objective Mixtures). *There exists $x, y \in X$ and $\alpha \in (0,1)$ such that $\delta_x \nsim \delta_y$ and $\alpha x + (1-\alpha) y \sim \delta_{\alpha x \oplus (1-\alpha) y}$.*

The consequences of these axioms, in addition to our previous ones, appear in the following proposition.

Proposition 1. *Consider a non-degenerate preference relation $\succeq$ that admits a MP-MD representation (u, Π, Φ). Then the following holds:*

(a) $|\Pi| = 1$ if, and only if, $\succeq$ satisfies Axiom 8 (Hedging Neutrality on Acts);

(b) The following are equivalent:

- *(1) $\Phi = \{\ell\}$;*
- *(2) $\succeq$ satisfies Axiom 9 (Hedging Neutrality on Lotteries);*
- *(3) $\succeq$ satisfies Risk Independence;*
- *(4) $\succeq$ satisfies Axiom 10 (Local Neutrality for Subjective and Objective Mixtures).*

(c) The following are equivalent:

- *(1) $|\Pi| = |\Phi| = 1$, and $\Phi = \{\ell\}$;*
- *(2) $\succeq$ satisfies Axiom 7 (Hedging Neutrality);*
- *(3) $\succeq$ satisfies Independence.*

We should emphasize point (b) in particular. It shows that to obtains a standard vNM Expected Utility representation under the axiomatic structure of the MP-MD representation, we can either impose standard vNM independence on lotteries (Risk Independence), or simply Hedging Neutrality on Lotteries, or even more simply that there exist at least *one* non-trivial case in which subjective and objective mixtures coincide: all of these postulates are equivalent.

2.8 A comparative notion of attraction towards certainty

We now discuss a comparative notion of *attraction towards certainty*. In particular, we show how the comparative notion of ambiguity aversion introduced in Ghirardato and Marinacci (2002) translates to our setup, and implies both more ambiguity aversion and more probability distortions for objective lotteries. Consider two decision makers, 1 and 2, such that 2 is more attracted to certainty than 1 is: that is, whenever 1 prefers a certain option δ_x to some act f, so does decision maker 2. Such attraction could be interpreted in two ways. First, both agents treat both probabilities and events in the same way, but 2 has a utility function which is more concave than that of 1. Alternatively, the curvature of the utility function could be the same for both agents, but 2 could be 'more pessimistic' than 1 is.[31] The approach of Ghirardato and Marinacci (2002) is to focus on this second case – looking at the relative attraction towards certainty while keeping constant the curvature of the utility function. We therefore need to posit that the curvature of the utility function of the two agents is the same, which in our setup translates to a very simple requirement: we want $\oplus_1 = \oplus_2$, that is, both agents should have the same approach to outcome mixtures. With this in mind, we can then use the following well-known definition introduced by (Ghirardato and Marinacci, 2002, Definition 7):[32]

Definition 8. Let $\succeq_1$ and $\succeq_2$ be two complete and non-degenerate preference relations on $\mathcal{F}$. We say that $\succeq_2$ is *more attracted to certainty than* $\succeq_2$ if the following hold:

1. $\oplus_{\succeq_1} = \oplus_{\succeq_2}$
2. for all $x \in X$ and all $f \in \mathcal{F}$
$$\delta_x \succeq_1 f \;\Rightarrow\; \delta_x \succeq_2 f$$
and
$$\delta_x \succ_1 f \;\Rightarrow\; \delta_x \succ_2 f.$$

This definition has precise consequences in our setup.

Proposition 2. *Let $\succeq_1$ and $\succeq_2$ be two complete and non-degenerate preference relations on $\mathcal{F}$ that admit Multiple Priors and Multiple Distortions Representations (u_1, Π_1, Φ_1) and (u_2, Π_2, Φ_2). Then, the following are equivalent:*

1. *$\succeq_2$ is more attracted to certainty than $\succeq_1$ and $\oplus_{\succeq_1} = \oplus_{\succeq_2}$;*
2. *u_1 is a positive affine transformation of u_2, $\Pi_2 \supseteq \Pi_1$ and $\Phi_2 \supseteq \Phi_1$.*

Proposition 2 shows that if we consider two agents who have the same curvature of the utility function, and such that 2 is more attracted to certainty than 1, then in any MP-MD representation *both* set of priors Π and Φ of 1 are (weakly) smaller than those of 2.[33]

[31] Note that, as X can be an arbitrary compact set, by demanding that the curvature of the utility function is the same, we understand that both agents apply the same utility to each object up to a positive affine transformation.

[32] There are two minor differences between what follows and (Ghirardato and Marinacci, 2002, Definition 7). First, here we require $\oplus_{\succeq_1} = \oplus_{\succeq_2}$, instead of requiring that the two preferences are *cardinally symmetric*, as defined in (Ghirardato and Marinacci, 2002, Definition 5). However, it is not hard to see that these two conditions are equivalent, since both imply that the (unique) utility indexes must be positive affine transformations of each other. The second difference is in the name: they interpret this comparative ranking as higher *ambiguity aversion*, while we interpret it more simply as attraction towards certainty. The reason is, calling this a comparative ambiguity aversion would not be precise here: our agents could be identically ambiguity averse, but have a higher tendency to 'distort probabilities' which lead them to a higher attraction towards certainty.

[33] The fact that the definition in Ghirardato and Marinacci (2002) captures *both* more ambiguity aversion and more

3 The Ellsberg Paradox and Relative Pessimism

In Section 2.6.2 we have shown how both the representations in Theorem 1 allow the agent to violate Savage Expected Utility in a 'pessimistic' fashion, but, *at the same time*, exhibit the opposite of Ellsberg behavior. This happens when the distortions on objective probabilities are more pronounced than those of objective ones – a possibility that has been noted in the literature, for example in Epstein (1999) and Wakker (2001).[34] We now refine our model to rule out this behavior: we characterize axiomatically the special case of our model in which the agent's behavior is *always* (weakly) compatible with the Ellsberg paradox. In particular, we obtain a model in which the decision maker could act in line with both the Allais and the Ellsberg paradoxes, potentially at the same time, while she cannot exhibit the opposite of either.

To better express both the new axiom and the new representation, it will be useful to introduce a notation to represent the reduction from subjective to objective risk. Consider an act $f \in \mathcal{F}$ and suppose that for every state ω it returns a degenerate lottery, i.e. $f(\omega) = \delta_y$ for some $y \in X$. Consider also a prior $\pi \in \Delta(\Omega)$, and notice that we can identify the lottery in $\Delta(X)$ that is the derived from f using probabilities in π: that is, the lottery that returns $f(\omega)$ with probability $\pi(\omega)$. This lottery is simply the reduction of f from *subjective to objective risk using prior* π. We denote it by f^π. (An identical notion is used in Ok et al. (2011).) We can then extend this definition also to acts that return non-degenerate lotteries, preserving the intuition: for any act $f \in \mathcal{F}$ and prior $\pi \in \Delta(\Omega)$, we denote by f^π the lottery that returns, with probability $\pi(\omega)$, the certainty equivalent of $f(\omega)$. That is, f^π denotes the constant act that yield the lottery $\sum \pi(\omega) c_{f(\omega)}$ in every state.

Endowed with this notation, we can define our new axiom and representation.

A.11 (Incomplete Reduction of Uncertainty). *For any $f \in \mathcal{F}$ there exits $\pi \in \Delta(\Omega)$ such that $f^\pi \sim f$ and $g^\pi \succeq g$ for all $g \in \mathcal{F}$.*

The intuition of Axiom 11 is the following. Consider some act f and suppose that the axiom is violated: for every $\pi \in \Delta(\Omega)$ such that $f \sim f^\pi$, we have $g \succ g^\pi$ for some $g \in \mathcal{F}$. This means that however we think the decision maker reduces the subjective uncertainty of f to an objective one, i.e. for any π' such that $f \sim f^{\pi'}$, then there exists some other act g that is evaluated with a prior which is more *optimistic* (for g) than the one used for f – we have that $g \succ g^{\pi'}$. The idea of the axiom is to rule out precisely this case: whatever prior is used (for g), it should be more *pessimistic* for g than the one used to evaluate f. That is, if the decision maker uses different

'probabilistic risk aversion' is well known: see, for example, the discussion in the paper and in Ghirardato (2004). For a comparative definition that captures only ambiguity aversion when the decision maker preferences over objective lotteries might violate Expected Utility, see Epstein (1999).

[34] One might naturally ask whether the term 'ambiguity averse' for such a decision maker would be appropriate. While this is a terminological issue and we abstain from committing to a specific view, we note how this has been subject of a discussion in the literature. On the one hand, intuitively there is a sense in which the agent is ambiguity averse, as she violates Savage Expected Utility precisely in the 'pessimistic' fashion prescribed by ambiguity aversion. In line with this intuition, the definition of Ghirardato and Marinacci (2002) would define her as ambiguity averse. On the other hand, when the agent exhibits the *opposite* of Ellsberg's behavior, then not only she violates the one empirical regularity that has led to the study of ambiguity aversion, but she shows that her pessimism is smaller for subjective than for objective bets – precisely the opposite of the standard intuition of ambiguity aversion. In line with this, the definition in Epstein (1999) would classify her as *not* ambiguity averse. (Wakker, 2001, Section 6) suggests how we should view the Allais paradox as reflecting 'absolute' pessimism, while the Ellsberg paradox as a 'relative' one, because it suggests that there is more pessimism for subjective uncertainty than for objective risk. From this point of view, he suggests how uncertainty could be seen as comprising both ambiguity and risk, while ambiguity aversion should be taken to represent this relative concept. He would therefore classify this agent as uncertainty averse but ambiguity loving. (See also Ghirardato (2004) for more discussion.)

models to evaluate different acts, she should use a model which is pessimistic with the act at hand. It is easy to see that the Incomplete Reduction of Uncertainty is trivially satisfied by the MaxMin Expected Utility representation of Gilboa and Schmeidler (1989). The following theorem shows that it is the necessary and sufficient condition to obtain the special case of our model that we satisfies the Ellsberg behavior.

Theorem 2. *Consider a complete and non-degenerate preference relation $\succeq$ on $\mathcal{F}$. The following are equivalent:*

(1) $\succeq$ satisfies Axioms 1-6 and Axiom 10.

(2) there there exists a continuous utility function *$u : X \to \mathbb{R}$, a convex and compact set of probability measures $\hat{\Pi}$ on Ω, and a convex, (point-wise) compact set of differentiable and concave probability weightings Ψ such that $\succeq$ is represented by the following functional: for any $f \in \mathcal{F}$, and for any enumeration of the states in Ω such that $f(\omega_{i-1}) \preceq f(\omega_i)$ for $i = 2, \ldots, |\mathrm{supp}(p)|$, we have*

$$V(f) := \min_{\pi \in \hat{\Pi}} \; U(f^{\pi})$$

and

$$V(p) = U(p) = \min_{\psi \in \Psi} \; \mathrm{RDEU}_{u,\psi}(p).$$

Moreover, u is unique up to a positive affine transformation, and there exists a unique minimal Ψ.

Theorem 2 shows that if we add Axiom 11 (Incomplete Reduction of Uncertainty) we obtain a representation much reminiscent of the MP-MC-RDEU representation, but with one relevant difference. While objective lotteries are ranked precisely in the same way, the agent has now a new way of evaluating subjective acts. In a MP-MC-RDEU representation the agent simply considers a set of priors Π and uses the worst one of these to evaluate each act. Here, instead, she proceeds in two steps. First, she considers a set of priors $\hat{\Pi}$, and using the most pessimistic one of them she maps the (subjective) act into an objective lottery. Second, once she has found this corresponding objective lottery, she *distorts* it precisely in the same way she distorts the other objective lotteries – using the worst of a set of RDEU distortions. That is, the decision maker distorts subjective acts *twice*.[35]

We should emphasize two features of the representation above. First, while it allows for an Ellsberg-like behavior, it rules out the opposite one: indeed subjective acts are (pessimistically) distorted twice, and by construction they must be distorted weakly ‘more’ than objective lotteries.

Second, notice that the preferences that admit one such representation $(u, \hat{\Pi}, \Psi)$ also admit a MP-MC-RDEU representation (u, Π, Ψ). The key observation is that while the utility u of the two representation is the same (up to affine transformations) and the set of distortions of objective lotteries Ψ is the same (as long as it is minimal), the set $\hat{\Pi}$ is bound to be *smaller* that the set Π. The reason is simple: while in the representation in Theorem 2 acts are distorted twice, first using $\hat{\Pi}$ and then using Ψ, in a MP-MC-RDEU representation the priors in Π are the only distortion that is applied, and must therefore include, in some sense, the combination of the distortions of

[35]While Theorem 2 shows the existence of a representation which is a special case of a MP-MC-RDEU representation, it is not hard to see how it could have instead derived an equivalent one which is instead a special case of a MP-MD representation (following the same steps used to prove the equivalence in Theorem 1). For brevity, we leave this to the reader.

both $\hat{\Pi}$ and Ψ. In fact, Π and $\hat{\Pi}$ will coincide if and only if the agent satisfies Expected Utility on objective lotteries (when Ψ includes only of the identity function). For example, consider the case in which there are two states of the world s_1 and s_2, and suppose that we have $\hat{\Pi} = \{\pi\}$, where $\pi(s_1) = \pi(s_2) = \frac{1}{2}$ – i.e. the agent has a *unique* prior to evaluate them – but at the same time suppose that we have $\Phi = \{\phi\}$ where $\phi(\frac{1}{2}) = \frac{3}{4}$. It is easy to see that the same agent has a MP-MC-RDEU representation in which Π is not a singleton, but rather it must contain both π' and π'' where $\pi'(s_1) = \frac{1}{4}$, $\pi'(s_2) = \frac{3}{4}$ and $\pi''(s_1) = \frac{3}{4}$ and $\pi''(s_2) = \frac{1}{4}$.

Given that we have two sets of priors, Π and $\hat{\Pi}$, and given that these sets are often given a specific interpretation, one might ask which one of them should be seen as the 'correct' one to use. In particular, in the example above which one is the correct set of priors? Is it a singleton or not? If we follow the interpretation often implicitly or explicitly suggested in the literature that the set of priors is the set of the possible 'models of the world' used by the agent, which are employed by the decision maker to asses subjective uncertainty, and therefore to reduce it to an objective one, then from this point of view it would then seem that the 'true' set of priors is $\hat{\Pi}$, and not Π. In fact, if the two sets differ from each other, i.e. when the agent violate vNM EU on objective lotteries, then considering the set of Π might actually be misleading, as it includes both the 'true' priors and the distortions applied to objective lotteries. For instance, in the example above the decision maker effectively has a *unique* way of reducing subjective uncertainty to risk – something that would not emerge were we to only look at the set Π.

We should also emphasize how identifying the 'true' set of priors along these lines might actually be important from a modeling perspective. A natural example concerns *updating*: when new information about the state of the world is revealed, it is reasonable to expect the agent to update her models of the world, i.e. $\hat{\Pi}$, but not her distortions of objective lotteries, i.e. Ψ; therefore, she should not update the full Π. Being able to identify the correct set of priors might then be relevant if we wish to formalize how the agent actually reacts to new information and updates her priors.

This discussion emphasizes one final point. To properly identify the set $\hat{\Pi}$ we need to be able to observe not only the preferences of the agent over (Savage) acts, but also her preferences over objective lotteries, as is the case here. Observing only the former could in fact allow us to pin down the set Π, but we wouldn't be able to separate it from $\hat{\Pi}$. This means that even if we are *only* interested in modeling acts and we are not concerned with the agent's approach to objective lotteries at all – because, for example, we don't believe many of these exist in the real world – we might at the same time be interested in observing how the agent reacts to them so that we can properly identify how she actually approaches subjective uncertainty. In turn, this means that a setup in which both subjective and objective uncertainty are involved, like the one of Anscombe and Aumann (1963), although originally introduced for mathematical convenience, is instead important to properly identify the behavior that we are interested in.[36]

4 Overview of the related literature

A large literature, much too large to be surveyed here, has been devoted to developing models that allow either for Allais or for Ellsberg-type behavior. However, far fewer models exist that allow for *both* features at the same time in setups, such as that of Anscombe and Aumann (1963), where both phenomena could appear independently. On the one hand, the majority of models meant to study Allais-like behavior do not allow for the presence of subjective probabilities at all, thus

[36]As we mention in Section 2.1, in our analysis we don't need the full setup of Anscombe and Aumann (1963): we also simply observe the preferences of the agent over the union of Savage acts and objective lotteries over the same prize space.

ruling out the presence of ambiguity aversion. On the other hand, the vast majority of models meant to capture ambiguity aversion either do not consider objective lotteries at all, operating in the setup of Savage (1954); or do consider them, operating in the setup of Anscombe and Aumann (1963), but also assume that agents satisfy vNM independence on objective lotteries, thus ruling out the possibility of Allais-like behavior.[37] The relevant axiom for this assumption, which posits that the agent satisfies vNM independence on constant acts, is usually called Risk Independence, and it is implied by almost all the weakening of independence suggested in the literature for this setup.[38] From the point of view of the literature on Ambiguity Aversion, therefore, one can see our paper as taking the standard setup of Anscombe and Aumann (1963), and generalizing MMEU with a representation that coincides with it on acts which do not involve objective lotteries, while at the same time weakening the assumption of Risk Independence and allowing for Allais-type behavior for objective lotteries. Indeed ours is not the first paper to relax Risk Independence in this setup. First of all, Ghirardato et al. (2001, 2003) show that one can obtain exactly a MaxMin Expected Utility representation by considering outcome mixtures, while at the same time disregarding objective lotteries – thus not restricting, but also not modeling, how the agent reacts to them. On the other hand, Drapeau and Kupper (2010) considers a model which corresponds to one in which agents exhibit uncertainty averse preferences a' la Cerreia et al. (2010) on acts that do not involve objective lotteries, while modeling her reaction to objective risk in a way similar to the model of Cerreia-Vioglio (2010).[39] As we shall see in our discussion of the latter, however, while this allows for violations of vNM independence, these are not necessarily in the direction suggested by the Allais paradox. By contrast, in our model agents always violate vNM independence for objective lotteries in the direction suggested by Allais paradox – this is precisely our goal. Finally, in Klibanoff et al. (2005) a corollary to the main theorem generalizes the representation to the case of non-EU preferences on objective lotteries; this case, however, is not fully axiomatized, and does not model jointly the attitude towards risk and uncertainty.[40]

From a procedural point of view, our paper considers the notion of outcome mixtures, which we denote by the symbol $\oplus$, instead of probability mixtures. Procedures of this kind are indeed not new: we refer to Ghirardato et al. (2001, 2003), Wakker (1994), Kobberling and Wakker (2003), and to the many references therein. More precisely, one could see our approach as the translation of the one of Ghirardato et al. (2003) to the case of objective probabilities. We then use this approach to introduce the novel notion of outcome mixture of lotteries and of acts, a central step in our analysis.

Our model is naturally related also to the models that study violations of Expected Utility in the case of risk (and not ambiguity) using representations and techniques reminiscent of the ones developed to study ambiguity aversion. First of all, our work is conceptually closely related to that of Maccheroni (2002) and Cerreia-Vioglio (2010): both provide representation in which the decision

[37]In addition, a few papers consider objective lotteries together with subjective uncertainty while using Savage acts: for example, Klibanoff et al. (2005). These papers as well add the additional assumption that the agent satisfies vNM Expected Utility on lotteries.

[38]This is true for the models in Gilboa and Schmeidler (1989) and Maccheroni et al. (2006), since both Centainty Independence and Weak-Certainty Independence imply the much weaker Risk Independence. And it is also true in the much more general models of Cerreia et al. (2010), Cerreia-Vioglio et al. (2011), and Ghirardato and Siniscalchi (2010). See Gilboa and Marinacci (2011) for a survey.

[39]More precisely, since Drapeau and Kupper (2010) studies a preorder which corresponds to the risk perception instead of studying the agent's preferences, as standard in their literature, their results are formally equivalent but 'inverted:' instead of positing quasi-concavity, they posit quasi-convexity, and instead of obtaining the *inf* over a set of measures, they obtain the *sup*.

[40]In addition, Chew and Sagi (2008) suggest how using the notion of 'conditional small worlds' that they introduce could generate a behavior which is consistent with both the Ellsberg and the Allais paradoxes.

maker treats objective lotteries as 'ambiguous objects,' as we do. Neither model, however, studies ambiguity aversion – they both work in the setup of vNM. Even focusing on this setup, moreover, both models have a fundamental difference with ours. While in our representation the decision maker has a fixed utility function, but has multiple probability distortions to evaluate lotteries, in both the papers above the opposite holds: the agent uses the *correct* probabilities, which are fixed, but at the same time she acts as if she had ambiguity over her utility. In particular, Maccheroni (2002) assumes that preferences are continuous, satisfy a weakening of vNM independence, as well as traditional convexity,[41] and obtains a representation such that the agent has a *set* of utility functions, and evaluates each lottery according to the *worst* of these utilities for that lotteries – a representation which is much the counterpart of ours, with multiple utilities instead of multiple probability distortions. Then, Cerreia-Vioglio (2010) generalizes this model by dropping independence entirely, and only requiring a weaker form of convexity, quasi-convexity.[42] He then derives a representation which generalizes that in Maccheroni (2002) in a similar way in which Cerreia et al. (2010) generalizes the one in Gilboa and Schmeidler (1989). The conceptual difference in the representation between these models and ours entails an important difference in behavior: while our model is designed to address the Allais paradox, and, more in general, attraction towards certainty, the ones in Maccheroni (2002) and Cerreia-Vioglio (2010) have a different goal, and agents in both of their models may exhibit certainty *aversion* – the opposite of Allais. This is particularly easy to see in the model of Maccheroni (2002): since there are multiple utilities and the agent is considering the worst one of them, then she would rather not face a certain outcome, where the worst utility can be chosen by the malevolent nature, but rather face a lottery, where nature needs to choose the worst utility for all elements in the support, and therefore cannot make the agent 'too worse off.'[43] From this point of view, therefore, one can see Maccheroni (2002) and Cerreia-Vioglio (2010) as exploring the consequences of convexity or quasi-convexity, while we aim to study a notion of pessimism.

Although there are few models that allows for both the Allais and the Ellsberg paradox to be present at the same time in the setup of Anscombe and Aumann (1963), as we mentioned in the introduction the idea of a connection between these behaviors is not new. We have already discussed (Section 2.6.3) how previous authors have noted that the RDEU representation is formally identical to the Choquet Expected Utility model of Schmeidler (1989), and how, in a similar spirit, we use a generalized version of Schmeidler (1989)'s *hedging* axiom to obtain a representation which is similar to the MaxMin Expected Utility model for the case of risk. In addition to deriving a different model, as opposed to this literature we also study the case of *simultaneous* presence of ambiguity aversion and Allais-like behavior, instead of focusing on a specific one of these. This is possible because we operate in the standard setup of Anscombe and Aumann (1963), where both features could be present at the same time and independently. By contrast, the approach followed by most of these papers would not apply in such setup, and, in general, would not apply to the case in which lotteries are elements of the simplex, as in von-Neumann Morgenstern or as in the questions of the Allais experiment.[44]

Perhaps the paper most closely related to ours is Wakker (2001). This paper focuses on the

[41] A preference relation $\succeq$ on a convex set is *convex* if for all p, q, r, if $p \succeq r$ and $q \succeq r$, then $\alpha p + (1-\alpha)q \succeq r$.

[42] More precisely, he only requires that, for any to lotteries p, q such that $p \sim q$ we have $\alpha p + (1-\alpha)q \succeq p$ for all $\alpha \in (0,1)$.

[43] Consider, for example, an agent whose preferences are represented a' la Maccheroni (2002) with the following utilities: $u_1(x) = 0$, $u_1(y) = 1$, $u_2(x) = 1$, $u_2(y) = 0$. Indeed this agent would rank $x \sim y$, but she would also rank $\frac{1}{2}x + \frac{1}{2}y \succ x$, in violation of attraction towards certainty.

[44] Most of this literature studies a setup in which the object of choice are Savage acts defined on a given set of states of world with an objective probability distributions over them – a setup where it is much easier to posit Schmeidler (1989)'s Comonotonic Independence.

case in which the preferences of the agent are of the Choquet Expected Utility form – of which RDEU is the special case that applies when lotteries are objective – and shows that a generalization of the common consequence effect can be used to characterize pessimism (convexity) in both the objective and subjective domains. Our generalized notion of preference for hedging could be seen as an assumption with a similar spirit – providing a generalized notion of pessimism that applies to both objective risk and subjective uncertainty. This notion is applicable to a broader class of preferences than those considered in Wakker (2001): on subjective uncertainty, it is well known that the multiple priors model is more general that the Choquet expected utility model; and we show in Section 2.6.3 that our representation generalizes concave RDEU. On the other hand, to define our notion we use outcome mixtures, which forces us to impose a richer structure on the space of consequences (connectedness). Wakker (2001) also shows that the conditions that imply pessimism for subjective uncertainty do not guarantee Ellsberg-type behavior in the presence of non-EU behavior over objective risk. We obtain a similar result, but we include a novel axiom that allows us to characterize a model in which this is guaranteed.

Starting from Segal (1987, 1990), a different channel to connect the Allais and Ellsberg paradox has been suggested: both could be seen as stemming from a failure of reduction of compound lotteries. In particular, Segal (1990) shows how RDEU can be derived precisely from such postulate; Dillenberger (2010) then links preference for one shot resolution of uncertainty with an axiom called Negative Certainty Independence, which is strongly linked to Allais-like behavior. At the same time, Segal (1987) argues how the Ellsberg paradox could be seen in a similar light: he argues that "the ambiguous lottery $(x, S; 0, S)$ (ambiguous in the sense that the decision maker does not know the probability of S) should be considered a two-stage lottery, where the first, imaginary, stage is over the possible values of the probability of S" (Segal, 1987, pg. 177). From this point of view, then, the two paradoxes are linked. This connection, however, is based on a specific interpretation of the Ellsberg paradox; as opposed to our analysis, moreover, this approach is based on the richer setup in which two-stage lotteries are observable.[45] Halevy (2007) tests in a lab experiment whether subjects who exhibit Ellsberg-like behavior also fail to reduce compound lotteries, and finds that about half of them do (while many of the others have a behaviors compatible with the models of Halevy and Feltkamp (2005), Klibanoff et al. (2005), and Seo (2009)). Dean and Ortoleva (2012) also tests the relation between ambiguity aversion and failure to reduce compound lotteries, and also documents an extremely strong relation.

Our work is also related to the recent Gumen et al. (2011), which is also built on the intuition of subjective evaluations of objective lotteries. In particular, they introduce a framework where they can analyze subjective distortions of objective probabilities: they study the preferences of a decision maker on space of pairs composed of 1) a probability measure over a state space and of 2) an Anscombe-Aumann act (over the same state space) – an object that they call a 'info-act.' The idea is that an info-act captures either a situation of objective uncertainty, or that of subjective risk. Using this framework, they are then able to define a behavioral notion of 'pessimism' for risky prospects in a way reminiscent of uncertainty aversion. After defining the natural mappings between these preferences and the more standard preferences over lotteries, they then show how their behavioral notion of pessimism in the info-act world implies that the corresponding preference over lotteries exhibits a form of pessimism consistent with the Allais paradox — for example, if they admitted a RDEU representation, it would have a concave probability weighting. Their paper has therefore a different focus from ours: while we derive a characterization theorem in the standard

[45]Indeed one could also see the setup of Anscombe and Aumann (1963) as 'rich,' as it entails both objective and subjective uncertainty with an implicit assumption about the timing of resolution of each of them. As we argued in Section 2.1, however, this feature is entirely irrelevant for us: we could have carried out our analysis even if we simply observed the agent's preference over the union of vNM lotteries and of Savage acts.

setup of Anscombe and Aumann (1963), they introduce a novel space that allows them to define a more general notion of pessimism, but do not look for a representation theorem.

Finally, our model is also naturally related to other generalizations of vNM Expected Utility. We have already discussed (Section 2.6.3) how our model is much related to the RDEU model of Quiggin (1982): while the restriction of our model to objective lotteries (i.e. when $|\Omega| = 1$) is not nested with the general formulation of RDEU, it is a strict generalization of RDEU with concave probability distortions. Yaari (1987) suggests a 'dual theory' of choice under risk, in which, instead of imposing linearity with respect to probability mixtures, we have linearity with respect to direct mixing of payments of risky prospects. Our approach differs as we do not need to impose even the latter linearity, as we use the notion of outcome-mixtures. Our results are also related to those of Dillenberger (2010), which shows the equivalence, under some basic assumptions, between Negative Certainty Independence (NCI), and PORU, which is preference for one-shot resolution of uncertainty. Moreover, he also shows that NCI is not satisfied by RDEU unless it is Expected Utility. On the one hand, it is easy to construct examples of our model which might violate NCI. On the other hand, whether the only model in our class of preferences that satisfies NCI is Expected Utility is still an open question. A second strand of literature aims to capture Allais-type behavior by weakening the requirement of independence to that of betweenness:[46] see, among others, Chew (1983), Dekel (1986), and the disappointment aversion model of Gul (1991). It is well known that this class of model is distinct from the RDEU class. Similarly, it is also easy to construct an example of our model which violates the betweenness axiom.

5 Conclusion

In this paper, we have introduced a novel link between two of the most discussed paradoxes in decision theory: ambiguity aversion and the Allais paradox. We have demonstrated that a preference for hedging, properly defined, can lead to both behaviors, and we have derived a representation which generalizes the Gilboa and Schmeidler (1989) multiple priors model by allowing the agent to treat objective probabilities like subjective objects, with 'multiple priors' of their own. The resulting model of choice under risk is a generalization of the RDEU model with concave distortions.

While our model does not require an agent who exhibits Allais-type behavior to be ambiguity averse, or *vice versa*, our result on the existence of a similar channel to capture both tendencies – preference for hedging – suggests that this might be the case. In this light, we emphasize the recent Dean and Ortoleva (2012), which tests the existence of an *empirical* relation between these behaviors. They show not only the significant presence of each individual paradox, but also the presence of a *significant* positive relationship between the propensity to exhibit each of them: subjects who exhibit one behavior are significantly more likely to exhibit the other.

We conclude by discussing possible extensions to the model. In the present paper we have derived a MMEU-like representation for non-Expected Utility by deriving the equivalents of preferences for hedging and certainty independence. A natural extension would then be to generalize this latter model following the generalizations of MMEU. For example, one could relax Axiom 6 and look for a representation along the lines of the Variational Preferences of Maccheroni et al. (2006) or the Uncertainty Averse preferences of Cerreia et al. (2010). We should point out, however, that this might not be straightforward. The reason is, our approach was constructed adapting the notion of outcome mixtures of consequences introduced in Ghirardato et al. (2003), which requires a form of bi-separability that need not be satisfied by the generalizations above – in particular, they might

[46] A preference relation satisfied betweenness if, for any $p, q \in \Delta(X)$, $p \sim q$ implies $\alpha p + (1-\alpha)q \sim p$ for all $\alpha \in [0,1]$.

not satisfy the conditions of Lemma 1. For this reason, a different approach to outcome-mixtures in X would be required. If this were found, however, then one could immediately use our notion of mixture of lotteries and acts – which, as we argued, is irrespective of how mixtures on X are constructed – and define a preference for hedging, paving the way for the generalizations hinted above.

Appendix A: Relation with the Allais Paradox

In Section 2.6.1 we have argued how both the MP-MD and the MP-MC-RDEU can allow for the behavior observed in the Allais paradox. In what follows, we show that both models also rule out the possibility of an *opposite* preference. To wit, consider the following four lotteries: $p_1 = \$1$, $p_2 = .01 \cdot \$0 + .89 \cdot \$1 + .1 \cdot \$x$, $p_3 = .89 \cdot 0\$ + .11 \cdot \$1$, and $p_4 = .9 \cdot \$0 + .1 \cdot \$y$. Recall that the Allais experiment asked to compare the lotteries above assuming $x = y = \$5$, and then observed the first preferred to the second, but the fourth preferred to the third. Let us now instead choose x and y in such a way to make $p_1 \sim p_2$ and $p_3 \sim p_4$. Then, we have a choice pattern which conforms with 'Allais' if and only if $x \geq y$. We will now prove that this must be the case for any MP-MD representation (u, Π, Φ); for simplicity we assume that u is linear (the argument could be easily generalized). Let us define the following three events on the unit interval: $E_1 = [0, 0.89), E_2 = [0.89, 090), E_3 = (0.90, 1]$. Then, consider the (measure preserving) map from lotteries into acts on $[0, 1]$ defined by the following table:

	E_1	E_2	E_3
p_1	\$1	\$1	\$1
p_2	\$1	\$0	\$x
p_3	\$0	\$1	\$1
p_4	\$0	\$0	\$y

Let α be the smallest weight put on E_3 by any prior in Φ, and β be the smallest weight put on E_2 by one of the priors for which $\phi(E_2) = \alpha$. Notice first of all that we must have $u(p_2) \leq (1 - \alpha - \beta) + \alpha x$, since p_2 could be evaluated using the prior above or a worse one, so $1 = u(p_1) = u(p_2) \leq (1 - \alpha - \beta) + \alpha x$, hence $\frac{\alpha+\beta}{\alpha} \leq x$. Notice also that we must have $u(p_4) = \alpha y$, and $u(p_3) \leq \min(0.11, \alpha + \beta)$.[47] Suppose first that we have $\alpha + \beta \leq 0.11$. Then, we have $\alpha y = u(p_4) = u(p_3) \leq \alpha + \beta$, hence $y \leq \frac{\alpha+\beta}{\alpha}$, which means $x \geq y$ as desired. Suppose instead that $\alpha + \beta > 0.11$. This means that we have $\alpha y = u(p_4) = u(p_3) \leq 0.11$, so $y \leq \frac{0.11}{\alpha}$. Since $x \geq \frac{\alpha+\beta}{\alpha}$ and $\alpha + \beta > 0.11$ we have $x > \frac{0.11}{\alpha}$, so $x > y$ as sought.

Appendix B: Proofs

Proof of Theorem 1

Proof of (1) $\Rightarrow$ (2). The proof will proceed with the following 6 steps: 1) we construct a derived preference relation on the Savage space with consequences X and set of states $\Omega \times [0, 1]$; 2) we prove that the continuity properties of the original preference relation imply some continuity property of the derived preference relation. 3) we prove that this derived relation is *locally* bi-separable (in the sense of Ghirardato and Marinacci (2001)) for some event in the space $\Omega \times [0, 1]$; 4) we prove that this derived relation admits a representation remininscent MaxMin Expected Utility in the larger Savage space; 5) we use this result to provide a representation for the restriction of $\succeq$ to constant acts; 6) we merge the two representations to obtain the desired representation for the acts in $\mathcal{F}'$.

Step 1. Denote by Σ^* the Borel σ-algebra on $[0, 1]$, and consider a state space $\Omega' := \Omega \times [0, 1]$ with the appropriate sigma-algebra $\Sigma' := \Sigma \times \Sigma^*$. Define $\mathcal{F}'$ the set of simple Savage acts on Ω', i.e. Σ'-measurable, finite valued functions $f' : \Omega' \to X$. To avoid confusion, we use $f', g', \ldots$ to denote generic elements of this space.[48] Define $\oplus$ on $\mathcal{F}'$ like we did in $\mathcal{F}$: once we have $\oplus$ defined on X, for any $f', g' \in \mathcal{F}'$ and $\alpha \in (0, 1)$, $\alpha f' \oplus (1 - \alpha) g'$ is the act in $\mathcal{F}'$ such that

[47] We know that $u(p_3) \leq 0.11$ since Φ contains the Lebesgue measure.

[48] Following the same abuses of notation of the main setup, for any $x \in X$ we also refer to the constant act $x \in \mathcal{F}'$ which returns x in every state.

$(\alpha f' \oplus (1-\alpha)g')(\omega') = \alpha f'(\omega') \oplus (1-\alpha)g'(\omega')$ for all $\omega' \in \Omega'$. (Moreover, since each act in $\mathcal{F}'$ is a function from $\Omega \times [0,1]$ into X, for all $f' \in \mathcal{F}'$ and for all $\omega \in \Omega$ abusing notation we can denote $f'(\omega, \cdot) : [0,1] \to X$ as the act that is constant in the first componente (Ω) but not on the second component ($[0,1]$).)

We now define two maps, one from $\mathcal{F}$ to $\mathcal{F}'$, and the other from $\mathcal{F}'$ to $\mathcal{F}$. Define first of all $\gamma^{-1} : \mathcal{F}' \to \mathcal{F}$ as

$$\gamma^{-1}(f')(\omega)(x) = \ell(f'(\omega, \cdot)^{-1}(x))$$

where $\ell(\cdot)$ denotes the Lebesgue measure. It is easy to see that $\gamma^{-}1(f)$ is well defined. Now define $\gamma : \mathcal{F} \to 2^{\mathcal{F}'}$ as

$$\gamma(f) = \{f' \in \mathcal{F}' : f = \gamma^{-1}(f')\}.$$

Notice that, by construction, we must have $\gamma(f) \cap \gamma(g) = \emptyset$ for all $f, g \in \mathcal{F}$ such that $f \neq g$. (Otherwise, we would have some $f' \in \mathcal{F}'$ such that $\gamma^{-1}(f') = f$ and $\gamma^{-1}(f') = g$, which is not possible since $f \neq g$.) Moreover, notice that we must have that $\gamma(\delta_x) = \{x\}$. Finally, notice that $\gamma\mathcal{F} := \cup_{f \in \mathcal{F}} \gamma(f) = \mathcal{F}'$ by construction.

Define now $\succeq'$ on $\mathcal{F}'$ as follows: $f' \succeq' g'$ if, and only if, $f \succeq g$ for some $f, g \in \mathcal{F}$ such that $f = \gamma^{-1}(f')$ and $g = \gamma^{-1}(g')$. Define by $\sim'$ and $\succ'$ is symmetric and asymmetric parts. (Notice that this implies $f' \sim' g'$ if $f', g' \in \gamma(f)$ for some $f \in \mathcal{F}$.)

We will now claim that $\succeq'$ is a complete preference relation on $\mathcal{F}'$.

Claim 1. $\succeq^*$ is a complete preference relation.

Proof. The completeness of $\succeq'$ is a trivial consequence of the completeness of $\succeq$ and the fact that $\gamma(\mathcal{F}) = \mathcal{F}'$. Similarly, the reflexivity follows from the reflexivity of $\succeq'$. To prove that $\succeq'$ is transitive, consider some $f', g', h' \in \mathcal{F}'$ such that $f' \succeq' g'$ and $g' \succeq' h'$. By construction, we must have some $f, g, h \in \mathcal{F}$ such that $f = \gamma^{-1}(f')$, $g = \gamma^{-1}(g')$, and $h = \gamma^{-1}(h')$ such that $f \succeq g$ and $g \succeq h$. By transitivity of $\succeq$, we also have $f \succeq h$, hence $f' \succeq' h'$ as sought. □

Step 2. We now prove that the continuity properties of $\succeq$ are inherited by $\succeq^*$. For any sequence $(f'_n) \in (\mathcal{F}'^{\infty}$, and any $f' \in \mathcal{F}'$, we say that $f'_n \to f'$ pointwise if $f_n(\omega) \to f$ (in the relevant topology) for all $\omega \in \Omega'$.

Claim 2. For any $(f_n) \in (\mathcal{F})^{\infty}$, $f \in \mathcal{F}$, if there exists $(f'_n) \in (\mathcal{F}'^{\infty}$, $f' \in \mathcal{F}'$ such that $f_n = \gamma^{-1}(f'_n)$ for all n, $f = \gamma^{-1}(f')$, and such that $f'_n \to f'$ pointwise, then we must have that $f_n \to f$.

Proof. We will prove the claim for the case in which f_n and f are constant acts, i.e. $f_n, f \in \Delta(X)$. The extension to the general case follows trivially. Assume that f'_n and f' as above exist: we will now prove that if $p_n = \gamma^{-1}(f'_n)$ for all n, and if $p = \gamma^{-1}(f)$, then $p_n \to p$ (weakly). Consider now some continuous v, and notice that we must have that $\int_X v(u) \mathrm{d}p_n = \int_{[0,1]} v(f'_n) \mathrm{d}\ell$ by contruction of γ. (Recall that ℓ is the Lebesgue measure.) Moreover, since v is continuous and since f'_n *pointwise* converges to f', we must then have that $\int_{[0,1]} v(f'_n) \mathrm{d}\ell \to \int_{[0,1]} v(f') \mathrm{d}\ell = \int_X v(u) \mathrm{d}p$: in turns, this means $\int_X v(u) \mathrm{d}p_n = \int_X v(u) \mathrm{d}p$. Since this was proved for a generic continuous v, we must have $p_n \to p$ (in weak convergence).

□

Step 3. We now prove that $\succeq'$ is locally biseparable for some event $A \in \Sigma'$. Consider the event $A = \Omega \times [0, \frac{1}{2}]$. Define Σ_A as the algebra generated by A, i.e. $\Sigma_A := \{\emptyset, A, A^C, \Omega'\}$, and by $\mathcal{F}'_A$ the corresponding set of acts, which is a subset of $\mathcal{F}'$. We will now prove that the restriction of $\succeq'$ on acts measurable under A is biseparable in the sense of Ghirardato and Marinacci (2001). We procced by a sequence of Claims.

Claim 3. There exist $x, y \in X$ such that $\delta_x \succ \delta_y$.

Proof. Suppose, by means of contradiction, that $\delta_x \sim \delta_y$ for all $x, y \in X$. Then, we would have that $p \trianglerighteq_{FOSD} q$ for all $p, q \in \Delta(X)$. By Axiom 1 (FOSD), therefore, we would have $p \sim q$ for all $p, q \in \Delta(X)$. In turns, by Axiom 2 (Monotonicity) we must have $f \sim g$ for all $f, g \in \mathcal{F}$, but this contradicts the assumption that $\succeq$ is non-degenerate. □

Claim 4 (Dominance)**.** For every $f', g' \in \mathcal{F}'$, if $f'(\omega') \succeq' g'(\omega')$ for every $\omega' \in \Omega'$, then $f' \succeq g'$.

Proof. Consider some $f', g' \in \mathcal{F}'$ such that $f'(\omega') \succeq' g'(\omega')$ for every $\omega' \in \Omega'$. Now consider $f'(\omega, \cdot)$ and $g'(\omega, \cdot)$ for some $\omega \in \Omega$, and notice that we have that both $\gamma^{-1}(f'(\omega, \cdot))$ and $\gamma^{-1}(g'(\omega, \cdot))$ are constant acts (in $\mathcal{F}$). Since we have $f'(\omega, A) \succeq' g'(\omega, A)$ for all $A \in \Sigma^*$, and since $x \succeq' y$ if and only if $\delta_x \succeq \delta_y$, then we must also have that $\gamma^{-1}(f'(\omega, \cdot)) \trianglerighteq_{FOSD} \gamma^{-1}(g'(\omega, \cdot))$ by construction. By Axiom 1 (FOSD), then, we must have $\gamma^{-1}(f'(\omega, \cdot)) \succeq \gamma^{-1}(g'(\omega, \cdot))$ for all $\omega \in \Omega$. In turns, this means that, for the acts $\hat{f}, \hat{g} \in \mathcal{F}$ defined by $\hat{f}(\omega) := \gamma^{-1}(f'(\omega, \cdot))$ and $\hat{g}(\omega) := \gamma^{-1}(g'(\omega, \cdot))$ for all $\omega \in \Omega$, we have $\hat{f} \succeq \hat{g}$ by Axiom 2 (Monotonicity). But then, notice that we must have that $f' \in \gamma(\hat{f})$ and $g' \in \gamma(\hat{g})$ by construction. But this means that we have $f' \succeq' g'$ as sought. $\square$

Claim 5. For any $x, y \in X$, $\gamma^{-1}(xAy) = \frac{1}{2}x + \frac{1}{2}y$.

Proof. Notice first of all that, since $xAy \in \mathcal{F}'$ is a constant act, then so much be $\gamma^{-1}(xAy)$. Moreover, notice that by definition of γ^{-1} we must have that for all $\omega \in \Omega$, $\gamma^{-1}(xAy)(\omega)(x) = \frac{1}{2}$; similarly, for all $\omega \in \Omega$ $\gamma^{-1}(xAy)(\omega)(y) = \frac{1}{2}$. This implies that we have $\gamma^{-1}(xAy(\omega)(x) = \frac{1}{2}x + \frac{1}{2}y$ as sought. $\square$

Claim 6. For every $x, y \in X$, there exists $z \in X$ such that $z \sim' xAy$.

Proof. Consider $x, y \in X$, and notice that $\gamma^{-1}(xAy) = \frac{1}{2}x + \frac{1}{2}y$ by claim 5. Now notice that, by Axiom 3 (Continuity) and 1 (FOSD), there must exist $z \in X$ such that $\frac{1}{2}x + \frac{1}{2}y \sim \delta_z$. We have previously observed that $\gamma^{-1}(z) = \delta_z$, which implies $\gamma^{-1}(z) \sim \gamma^{-1}(xAy)$, which implies $xAy \sim' z$ as sought. $\square$

Given Claim 6, for any $x, y \in X$, define $ce'(xAy) := z$ for some $z \in X$ such that $xAy \sim' z$.

Claim 7 (Essentiality)**.** A is an essential event for $\succeq'$.[49]

Proof. Consider any $x, y \in X$ such that $\delta_x \succ \delta_y$ – Claim 3 guarantee that they exist. Now consider the $p = \frac{1}{2}x + \frac{1}{2}y$. By Axiom 1 (FOSD) we must have $\delta_x \succ p \succ \delta_y$. Now consider the act $xAy \in \mathcal{F}'$. Notice that we have $xAy(\omega \times [0, 5]) = x$ and $xAy(\omega \times [0.5, 1]) = \delta_y$ for all $\omega \in \Omega$. By construction, therefore, we must have $xAy \in \gamma(p)$, $x \in \gamma(\delta_x)$ and $y \in \gamma(\delta_y)$. By definition of $\succ'$, then, we have $x \succ' xAy \succ' y$ as sought. $\square$

Claim 8 (A-Monotonicity)**.** For any non-null event $B \in \Sigma_A$, and $x, y, z \in X$ such that $x, y \succ z$ we have

$$x \succ' y \Leftarrow xBz \succ' yBz.$$

Moreover, for any non-universal[50] $B \in \Sigma_A$, $x, y, z \in X$ s.t. $x, y \succeq z$

$$x \succ' y \Leftarrow zBx \succ' zBy$$

Proof. Consider an event $B \in \Sigma_A$, and $x, y, z \in X$ such that $x \succ' y$. Notice that by construction this implies $\delta_x \succ \delta_y$. Notice also that the non-null events in Σ_A are A, A^C, and Ω'. In the case of $B = \Omega'$ we have $xBz = x$ and $yBz = y$, which guarantees that $x \succ' y$. Now consider the case in which $B = A$. By Claim 5, $\gamma^{-1}(xAz) = \frac{1}{2}x + \frac{1}{2}z$, and $\gamma^{-1}(yAz) = \frac{1}{2}y + \frac{1}{2}z$. Since $\delta_x \succ \delta_y$, then $\frac{1}{2}x + \frac{1}{2}z \vartriangleright_{FOSD} \frac{1}{2}y + \frac{1}{2}z$, which, by Axiom 1 (FOSD), implies $\frac{1}{2}x + \frac{1}{2}z \succ \frac{1}{2}y + \frac{1}{2}z$, hence $\gamma^{-1}(xAz) \succ \gamma^{-1}(yAz)$. By construction of $\succeq'$ this implies $xAz \succ' yAz$. Now consider the case in which $B = A^c$. This implies that we have $xA^Cz = zAx$ and $yA^Cz = zAy$. Notice, however, that by construction we must have $zAx \in \gamma(\frac{1}{2}x + \frac{1}{2}z)$. Since we also have $xAz \in \gamma(\frac{1}{2}x + \frac{1}{2}z)$, by construction of $\succeq'$ we must have $zAx \sim' xAz$. Similarly, we have $zAy \sim' yAz$. We have already proved that we must have $xAz \succ yAz$, and this, by transitivity, implies $zAA \succ' zAy$ as sought.

Now consider some $B \in \Sigma_A$ which is non-universal. If $B = \emptyset$, we trivially have that $x \succ' y \Leftarrow zBx \succ' zBy$. Now consider the case in which $B = A$. In this case we have $x \succ' y$ and we need to show $zAx \succ' zAy$: but this is exactly what we have showed above. Similarly, when $B = A^C$, we need to show that if $x \succ' y$ then $xAz \succ' yAx$ – which is again exactly what we have shown before. $\square$

[49]We recall that an event E is essential if we have $x \succ' xAy \succ' y$ for some $x, y \in X$.
[50]An event is *universal* if $y \sim xAy$ for all $x, y \in X$ such that $x \succ y$.

Claim 9 (A-Continuity)**.** Let $\{g'_\alpha\}_{\alpha \in D} \subseteq \mathcal{F}'_A$ be a net that pointwise converges to g'. For every $f' \in \mathcal{F}'$, if $g'_\alpha \succeq' f$ (resp. $f \succeq' g'_\alpha$) for all $\alpha \in D$, then $g' \succeq' f'$ (resp. $f' \succeq' g'$).

Proof. This claim is a trivial consequence of the continuity of $\succeq$ and of Claim 2. To see why, consider $f', g' \in \mathcal{F}'$ and a net $\{g'_\alpha\}_{\alpha \in D} \subseteq \mathcal{F}'_A$ that pointwise converges to g' such that $g'_\alpha \succeq' f'$ for all $\alpha \in D$. By contruction we must have $\gamma^{-1}(g'_\alpha) \succeq \gamma^{-1}(f')$. Now, notice that, if g'_α pointwise converges to some g', then we must have that $\gamma^{-1}(g'_\alpha)$ converges to $\gamma^{-1}(g')$ by Claim 2. But then, by continuity of $\succeq$ (Axiom 3), we must have $\gamma^{-1}(g') \succeq \gamma^{-1}(f')$, and therefore $g' \succeq' f'$ as sought. The proof of the opposite case ($f \succeq' g'_\alpha$ for all $\alpha \in D$) is analogous. $\square$

Claim 10 (A-Substitution)**.** For any $x, y, z', z'' \in X$ and $B, C \in \Sigma_A$ such that $x \succeq' z' \succeq' y$ and $x \succeq' z'' \succeq' y$, we have

$$ce'_{xBz'}Cce'_{z''By} \sim' ce'_{xCz''}Bce'_{z'Cy}.$$

Proof. Consider first the case in which $B = \emptyset$. In this case, the claim becomes $ce'_{z'}Cce'_{y} \sim' ce'_{z'Cy}$, which is trivially true. The case $C = \emptyset$ is analogous. Now consider the case $B = \Omega'$. The claim becomes $ce'_{x}Cce'_{z''} \sim' ce'_{xCz''}$ which again is trivially true. The case in which $C = \Omega'$ is again analogous.

We are left with the case in which $B = A$ and $C = A^C$. (The case $B = A^C$ and $C = A$ is again analogous.) In this case the claim becomes $ce'_{xAz'}A^C ce'_{z''Ay} \sim' ce'_{xA^Cz''}Ace'_{z'^Cy}$, which is equivalent to $ce'_{z''Ay}Ace'_{xAz'} \sim' ce'_{z''Ax}Ace'_{yAz'}$. Now notice that since $ce'_{xAy} \in X$ for all $x, y \in X$, by claim 5, we have that $\gamma^{-1}(ce'_{z''Ay}Ace'_{xAz'}) = \frac{1}{2}ce'_{z''Ay} + \frac{1}{2}ce'_{xAz'}$. At the same time, consider some $r, s \in X$, and notice that, since $ce'_{rAs} \sim' rAs$ by contruction, then we must have $\gamma^{-1}(ce'_{rAs}) \sim \gamma^{-1}(rAs)$. Since $\gamma^{-1}(rAs) = \frac{1}{2}r + \frac{1}{2}s$ again by claim 5, then we have that $\gamma^{-1}(ce'_{rAs}) \sim \frac{1}{2}r + \frac{1}{2}s$. Moreover, since $ce'_{rAs} \in X$, then we must have that $\delta_{ce'_{rAs}} \sim \delta_{c_{\frac{1}{2}z+\frac{1}{2}s}}$. Since this is true for all $r, s \in X$, then by Axiom 1 (FOSD) we must have $\frac{1}{2}c_{\frac{1}{2}z''+\frac{1}{2}y} + \frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z'} \sim \frac{1}{2}ce'_{z''Ay} + \frac{1}{2}ce'_{xAz'}$, hence $\gamma^{-1}(ce'_{z''Ay}Ace'_{xAz'}) \sim \frac{1}{2}c_{\frac{1}{2}z''+\frac{1}{2}y} + \frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z'}$. By analogous arguments, we must have $\gamma^{-1}(ce'_{z''Ax}Ace'_{yAz'}) \sim \frac{1}{2}c_{\frac{1}{2}z''+\frac{1}{2}x} + \frac{1}{2}c_{\frac{1}{2}y+\frac{1}{2}z'}$. At the same time, Axiom 4 we must have $c_{\frac{1}{2}z''+\frac{1}{2}y} + \frac{1}{2}c_{\frac{1}{2}x+\frac{1}{2}z'} \sim \frac{1}{2}c_{\frac{1}{2}z''+\frac{1}{2}x} + \frac{1}{2}c_{\frac{1}{2}y+\frac{1}{2}z'}$, which by transitivity implies $\gamma^{-1}(ce'_{z''Ay}Ace'_{xAz'}) \sim \gamma^{-1}(ce'_{z''Ax}Ace'_{yAz'})$, hence $ce'_{z''Ay}Ace'_{xAz'} \sim ce'_{z''Ax}Ace'_{yAz'}$ as sought. $\square$

Notice that these claims above prove that $\succeq'$ is locally-biseparable in the sense of Ghirardato and Marinacci (2001).

Step 4. We now prove that $\succeq'$ admits a representation similar to MMEU. We proceed again by claims.

Claim 11 (C-Independence)**.** For any $f', g' \in \mathcal{F}'$, $x \in X$ and $\alpha \in (0, 1)$

$$f' \sim' g' \Rightarrow \alpha f' \oplus (1-\alpha)x \sim' \alpha g' \oplus (1-\alpha)x.$$

Proof. Consider $f', g' \in \mathcal{F}'$ such that $f' \sim' g'$. Notice that we could have $f' \sim' g'$ in two possible cases: 1) $\gamma^{-1}(f') = \gamma^{-1}(g')$; 2) $\gamma^{-1}(f') \neq \gamma^{-1}(g')$ but $\gamma^{-1}(f') \sim \gamma^{-1}(g')$. In either case, we must have $\gamma^{-1}(f') \sim \gamma^{-1}(g')$. By Axiom 6, then, we must have that for any $x \in X$ and $\alpha \in (0, 1)$, $\alpha\gamma^{-1}(f') \oplus (1-\alpha)\delta_x \sim \alpha\gamma^{-1}(g') \oplus (1-\alpha)\delta_x$. Let us now consider $\alpha\gamma^{-1}(f') \oplus (1-\alpha)\delta_x$, and notice that, by construction, we must have that $f' \oplus (1-\alpha)x \in \gamma(\alpha\gamma^{-1}(f') \oplus (1-\alpha)\delta_x)$: in fact, we must have that for every $\omega \in \Omega$ and every $y \in X$, $(\alpha\gamma^{-1}(f') \oplus (1-\alpha)\delta_x)(\omega)(\alpha y \oplus (1-\alpha)x) = \ell(f'(\omega)^{-1}(\alpha y \oplus (1-\alpha)x)$. In turns, this means that $\gamma^{-1}(f' \oplus (1-\alpha)x) = \alpha\gamma^{-1}(f') \oplus (1-\alpha)\delta_x$. Similarly, $g' \oplus (1-\alpha)x \in \gamma(\alpha\gamma^{-1}(g') \oplus (1-\alpha)\delta_x)$ and $\gamma^{-1}(g' \oplus (1-\alpha)x) = \alpha\gamma^{-1}(g') \oplus (1-\alpha)\delta_x$. Since we have $\alpha\gamma^{-1}(f') \oplus (1-\alpha)\delta_x \sim \alpha\gamma^{-1}(g') \oplus (1-\alpha)\delta_x$, then by transitivity $\gamma^{-1}(f' \oplus (1-\alpha)x) \sim \gamma^{-1}(g' \oplus (1-\alpha)x)$, hence $f' \oplus (1-\alpha)x \sim' g' \oplus (1-\alpha)x$ as sought. $\square$

Claim 12 (Hedging)**.** For any $f', g' \in \mathcal{F}'$ such that $f' \sim' g'$

$$\tfrac{1}{2}f' \oplus \tfrac{1}{2}g' \succeq' f'.$$

Proof. Consider $f', g' \in \mathcal{F}'$ such that $f' \sim' g'$. Notice that we could have $f' \sim' g'$ in two possible cases: 1) $\gamma^{-1}(f') = \gamma^{-1}(g')$; 2) $\gamma^{-1}(f') \neq \gamma^{-1}(g')$ but $\gamma^{-1}(f') \sim \gamma^{-1}(g')$. In either case, we must have $\gamma^{-1}(f') \sim \gamma^{-1}(g')$. Now consider the act $\frac{1}{2}f' \oplus \frac{1}{2}g'$: we will now prove that, for all $\omega \in \Omega$, $\gamma^{-1}(\frac{1}{2}f'(\omega, \cdot) \oplus \frac{1}{2}g'(\omega, \cdot)) \in \bigoplus^{\frac{1}{2}}_{\gamma^{-1}(f'(\omega,\cdot)),\gamma^{-1}(g'(\omega,\cdot))}$. To see why, notice that for all $\omega \in \Omega$, $(\frac{1}{2}f'(\omega, \cdot) \oplus \frac{1}{2}g'(\omega, \cdot))(A) = \frac{1}{2}f'(\omega, A) \oplus \frac{1}{2}g'(\omega, A)$ for all $A \in \Sigma^*$: that is, for every event in $[0,1]$ is assigns an $x \in X$ which is the $\oplus$-$\frac{1}{2}$-mixtures of what is assigned by $f'(\omega,\cdot)$ and $g'(\omega,\cdot)$. But this means that $\gamma^{-1}(\frac{1}{2}f'(\omega,\cdot) \oplus \frac{1}{2}g'(\omega,\cdot))$ must be a constant act (lottery in $\Delta(X)$) such that, if $H_x^{f',g'} := \{A \in \Sigma^* : x = \frac{1}{2}f'(\omega,A)] \oplus \frac{1}{2}g'(\omega,A)\}$, then $\gamma^{-1}(\frac{1}{2}f'(\omega,\cdot) \oplus \frac{1}{2}g'(\omega,\cdot))(x) = \ell(\underset{A\in H_x^{f',g'}}{\cup} A)$. But then, we must have that $\gamma^{-1}(\frac{1}{2}f'(\omega,\cdot) \oplus \frac{1}{2}g'(\omega,\cdot)) \in \bigoplus^{\frac{1}{2}}_{\gamma^{-1}(f'(\omega,\cdot)),\gamma^{-1}(g'(\omega,\cdot))}$. By construction of $\oplus$ in the space $\mathcal{F}'$, then, we must have that $\gamma^{-1}(\frac{1}{2}f' \oplus \frac{1}{2}g'^{\frac{1}{2}}_{\gamma^{-1}(f'),\gamma^{-1}(g')}$. But then, since we have already enstablished that we have $\gamma^{-1}(f') \sim \gamma^{-1}(g')$, by Axiom 5 (Hedging) we must have that $\gamma^{-1}(\frac{1}{2}f' \oplus \frac{1}{2}g') \succeq \gamma^{-1}(f')$, which implies $\frac{1}{2}f' \oplus \frac{1}{2}g' \succeq' f'$ as sought. $\square$

Claim 13. There exists a continuous non-constant function $u : X \to \mathbb{R}$ and a non-empty, weak* compact and convex set P of finitely additive probabilities of Σ' such that $\succeq'$ is represented by the functional

$$V'(f') := \min_{p\in P} \int_{\Omega'} u(f')\mathrm{d}p.$$

Moreover, u is unique up to a positive affine transformation and P is unique. Moreover, $|P| = 1$ if and only if $\succeq'$ is such that for any $f', g' \in \mathcal{F}'$ such that $f' \sim' g'$ we have $\frac{1}{2}f' \oplus \frac{1}{2}g' \sim' f'$.

Proof. This Claim follows directly from Theorem 5 in (Ghirardato et al., 2001, page 12), where the essential event for which axioms are defined is the event A defined above. (It should be noted that weak*-compactness of P follows as well.) The last part of the Theorem, which characterizes the case in which $|P| = 1$, is a well-known property of MMEU representations. (See Gilboa and Schmeidler (1989).) $\square$

Step 5. We now use the result above to provide a representation of the restriction of $\succeq$ to constant acts. To this end, let us first look at the restriction of $\succeq'$ to acts in $\mathcal{F}'$ that are constant in their first component: define $\mathcal{F}^* \subset \mathcal{F}'$ as $\mathcal{F}^* := \{f' \in \mathcal{F}' : f'(\omega,\cdot) = f'(\omega',\cdot) \text{ for all } \omega,\omega' \in \Omega\}$. Define by $\succeq^*$ the restriction of $\succeq'$ to $\mathcal{F}^*$.

Claim 14. There exists a unique nonempty, closed and convex set Φ of finitely additive probabilities over Σ^* sucht that $\succeq^*$ is represented by

$$\hat{V}^*(f^*) := \min_{p\in\Phi} \int_{[0,1]} u(f^*(s))\ \mathrm{d}p$$

Proof. This result follows trivially from Claim 13 once we define Φ as projection of P on $[0,1]$. $\square$

Claim 15. There exists a unique nonempty, closed and convex set Φ of finitely additive probabilities over Σ^* such that, for any enumeration of the support of $\{x_1,\ldots,x_{|\mathrm{supp}(p)|}\}$, the restriction of $\succeq$ to $\Delta(X)$ is represented by the functional

$$V^*(p) := \min_{\phi\in\Phi} \sum_{i=1}^{|\mathrm{supp}(p)|} \phi([\sum_{j=1}^{i-1} p(x_j), \sum_{j=1}^{i} p(x_j)])u(x_i)$$

Proof. Construct the set Φ of closed and convex finitely additive probabilities over Σ^* following Claim 14, and define $\hat{V}^*$ accordingly. Notice first of all that, by construction of γ and by definition of $\mathcal{F}^*$, we must have that $\gamma(p) \subseteq \mathcal{F}^*$ for all $p \in \Delta(X)$. We will now argue that, for all $p, q \in \Delta(X)$, we have $p \succeq q$ if and only if $f^* \succeq^* g^*$ for some $f^*, g^* \in \mathcal{F}^*$ such that $\gamma^{-1}(f^*) = p$ and $\gamma^{-1}(g^*) = q$. To see why, notice that if $p \succeq q$, then we must have $f^* \succeq'^*$, hence $f^* \succeq^* g^*$. Conversely, suppose that we have $f^* \succeq^* g^*$ for some $f^*, g^* \in \mathcal{F}^*$ such that $\gamma^{-1}(f^*) = p$ and $\gamma^{-1}(g^*) = q$, but $q \succ p$. But then, by definition of $\succeq'$ we should have $g^* \succ'^*$, a contradiction.

Notice now that for every $p \in \Delta(X)$, if $f^*, g^* \in \gamma(p)$, then we must have $\hat{V}(f^*) = \hat{V}(g^*)$: the reason is, by construction of $\succeq'$ we must have $f^* \sim'^*$, hence $f^* \sim^* g^*$, hence $\hat{V}(f^*) = \hat{V}(g^*)$. Define now $V : \Delta(X) \to \mathbb{R}$ as $V^*(p) := \hat{V}^*(f^*)$ for some $f^* \in \gamma(p)$. By the previous observation this is well defined. Now notice that we have $p \succeq q$ if and only if $f^* \succeq^* g^*$ for some $f^*, g^* \in \mathcal{F}^*$ such that $\gamma^{-1}(f^*) = p$ and $\gamma^{-1}(g^*) = q$, which holds if and only if $\hat{V}^*(f^*) \geq \hat{V}^*(g^*)$, which in turns hold if and only if $V^*(p) \geq V^*(q)$, which means that V^* represents the restriction of $\succeq$ on $\Delta(X)$ as sought. $\square$

Claim 16. $\succeq^*$ satisfies Arrow's Monotone Continuity axiom. That is, for any $f, g \in \mathcal{F}^*$ such that $f \succ^* g$, and for any $x \in X$ and sequence of events in Σ^* $E_1, \dots, E_n$ with $E_1 \subseteq E_2 \subseteq \dots$ and $\cap_{n\geq 1} E_n = \emptyset$, there exists $\bar{n} \geq 1$ such that

$$xE_{\bar{n}}f \succ^* g \;\text{ and }\; f \succ^* xE_{\bar{n}}g.$$

Proof. Consider f, g, x, and $E_1, \dots$ as in the claim above. Notice first of all that for any $s \in \Omega'$, there must exist some $\hat{n}$ such that for all $n \geq \hat{n}$ we have $s \notin E_n$: otherwise, if this was not true for some $s \in \Omega'$, we would have $s \in \cap_{n\geq 1} E_n$, a contradiction. In turn, this means that we have $xE_n f \to f$ pointwise: for any $s \in \Omega'$, there must exist some n such that $s \notin E_n$, and therefore $xE_n f(s) = f(s)$ as sought. Notice then that by Claim 2, we must therefore have that $\gamma^{-1}(xE_n f) \to \gamma^{-1}(f)$. We now show that we must have some $\bar{n}_1 \geq 1$ such that $xE_{\bar{n}_1} f \succ^* g$ for all $n \geq \bar{n}_1$. Assume, by means of contradiction, that this is not the case: for every $n \geq 1$, there exists some $n' \geq n$ such that $g \succeq^* xE'_n f$. Construct now the subsequence of $E_1, \dots$ which includes these events, i.e. the events such that $g \succeq^* xE^{'}_{n'} f$: by the previous argument it must be a subsequence of $E_1, \dots$ and we must have that $E'_1 \subseteq E'_2 \subseteq \dots$ and $\cap_{n\geq 1} E'_n = \emptyset$. This means that we have $g \succeq^* xE'_n f$ for all n. By contruction this then means that we have $\gamma^{-1}(g) \succeq \gamma^{-1}(xE'_n f)$. Now consider $\gamma^{-1}(xE'_n f)$, and notice that we have proved above that $\gamma^{-1}(xE'_n f) \to \gamma^{-1}(f)$ as $n \to \infty$. By Axiom 3 (Continuity), then, we must have that $\gamma^{-1}(g) \succeq \gamma^{-1}(f)$, which in turns means that $g \succeq^* f$, a contradiction. An identical argument shows that there must exist $\bar{n}_2 \geq 1$ such that $f \succ^* xE_{\bar{n}_2} g$ for all $n \geq \bar{n}_2$. Any $n \geq \max\{n_1, n_2\}$ will therefore give us the desired rankings. □

Claim 17. The measures in Φ are countably additive.

Proof. In Claim 16 we have showed that $\succeq^*$ satisfies Arrow's Monotone Continuity Axioms. Using Theorem 1 in Chateauneuf et al. (2005) we can then show that Φ must be countably additive. □

Claim 18. The measures in Φ are atomless.

Proof. We will first of all follow a standard approach and define the likelihood ranking induced by the $\succeq^*$. In particular, define $\succeq_L$ on Σ^* as

$$A \succeq_L B \;\Leftrightarrow\; \min_{\phi\in\Phi} \phi(A) \geq \min_{\phi\in\Phi} \phi(B).$$

Theorem 2 in Chateauneuf et al. (2005) show that every $\phi \in \Phi$ is atomless if and only if for all $A \in \Sigma^*$ such that $A \succ_L \emptyset$, there exists $B \subseteq A$ such that $A \succ_L B \succ_L \emptyset$. $A \in \Sigma^*$ such that $A \succ_L \emptyset$, and notice that this implies that we have $\min_{\phi\in\Phi} \phi(A) > 0$, hence $\phi(A) > 0$ for all $\phi \in \Phi$. Since every $\phi \in \Phi$ is mutually absolutely continuous with respect to the Lebesgue measure, this implies $\ell(A) > 0$. Since ℓ is atomless, then there exists $B \subseteq A$ such that $\ell(A) > \ell(B) > 0$. Notice that this implies $\ell(A\backslash B) > 0$. Again since all $\phi \in \Phi$ are mutually absolutely continuous with respect to the Lebesgue measure, we must therefore have $\phi(A) > 0$, $\phi(A\backslash B) > 0$ and $\phi(B) > 0$ for all $\phi \in \Phi$. But this means that we have $\phi(A) = \phi(B) + \phi(A\backslash B) > \phi(B) > 0$ for all $\phi \in \Phi$. But this implies $\min_{\phi\in\Phi} \phi(A) > \min_{\phi\in\Phi} \phi(B) > 0$, hence $A \succ_L B \succ_L \emptyset$ as sought. □

Claim 19. The Lebesgue measure ℓ belongs to Φ.

Proof. Assume by means of contradiction that $\ell \notin \Phi$. By the uniqueness of Φ, we know that there must therefore exist some $f \in \mathcal{F}^*$ such that $\hat{V}^*(f) := \min_{p\in\Phi} \int_{[0,1]} u(f(s))\, \mathrm{d}p > \int_{[0,1]} u(f(s))\, \mathrm{d}\ell$. Call p_1 a generic element of $\arg\min_{p\in\Phi} \int_{[0,1]} u(f(s))\, \mathrm{d}p$. Notice that, since $\int_{[0,1]} u(f^*(s))\, \mathrm{d}p_1 > \int_{[0,1]} u(f(s))\, \mathrm{d}\ell$, it must be the case that $p_1(A) > \ell(A)$ for some $A \subset [0,1]$ such that $u(f(A)) > \int_{[0,1]} u(f^*(s))\, \mathrm{d}\ell$, and that $p_1(B) < \ell(B)$ for some $B \subset [0,1]$ such that $u(f(B)) < u(f(A))$.

Suppose first of all that $\ell(A) \geq \ell(B)$. Now consider some $f' \in \mathcal{F}^*$ constructed as follows. Consider any $C \subseteq A$ such that $\ell(C) = \ell(B)$ and $p_1(C) > \ell(C)$. (This must be possible since $p_1(A) > \ell(A)$.) Notice that we must therefore have $p_1(C) > p_1(B)$ since $p_1(C) > \ell(C) = \ell(B) > p_1(B)$. Now construct the act f' as: $f'(s) = f(s)$ if $s \notin C \cup B$; $f'(C) = f(B)$; and $f'(B) = f(A)$. (Notice that what we have done is that we have moved the 'bad' outcomes to some events to which p_1 assigns a likelihood above the Lebesgue measure, while we have moved the 'good' outcomes to some event to which p_1 assigns a likelihood below the Lebesgue measure.) Notice now that, by construction, we must have that $f, f' \in \gamma(p)$ for some $p \in \Delta(X)$, hence we must have $f \sim^* f'$. At the same time, since $p_1(C) > p_1(B)$ and since $u(f(B)) < u(f(A)) = u(f(C))$, we must also have $\hat{V}^*(f) = \int_{[0,1]} u(f(s))\ \mathrm{d}p_1 > \int_{[0,1]} u(f'(s))\ \mathrm{d}p \geq \min_{p\in\Phi} \int_{[0,1]} u(f'(s))\ \mathrm{d}p = \hat{V}^*(f')$. But this means that we have $\hat{V}^*(f) > \hat{V}^*(f')$, hence $f \succ f'$, contradicting $f \sim^* f'$. The proof for the case in which $\ell(A) < \ell(B)$ is specular.

□

Claim 20. All measures in Φ are mutually absolutely continuous, and, in particular, they are all mutually absolutely continuous with respect to the Lebesgue measure ℓ.

Proof. To prove this, we will prove that for every event E in $[0,1]$, if E is null for $\succeq^*$ if and only if $\ell(E) = 0$.[51] In turns this means that all measures are mutually absolutely continuous with respect to each other.

Consider some measurable $E \subset [0,1]$ such that $\ell(E) = 0$. Suppose, by means of contradiction, that $\{\phi \in \Phi : \phi(E) > 0\} \neq \emptyset$. Then, consider any $x, y \in X$ such that $\delta_x \succ \delta_y$ (which must exist by non-triviality), and construct the act $yEx \in \mathcal{F}^*$. Since $\{\phi \in \Phi : \phi(E) > 0\} \neq \emptyset$, then we must have that $\min_{\phi\in\Phi} \phi(E)u(y) + (1-\phi(E))u(x) < u(x)$, which in turns means that $yEx \prec^* x$ (by Claim 14), hence $yEx \prec' x$. However, notice that, since $\ell(E) = 0$, we must have that $\gamma^{-1}(yEx) = \delta_x = \gamma^{-1}(x)$. By construction of $\succeq'$, then, we must have $yEx \sim' x$, contradicting $yEx \prec' x$.

Consider now some measurable $E \subset [0,1]$ such that $\ell(E) > 0$. We now want to show that $\phi(E) > 0$ for all $\phi \in \Phi$. Suppose, by means of contradiction, that $\{\phi \in \Phi : \phi(E) = 0\} \neq \emptyset$. Then, consider any $x, y \in X$ such that $\delta_x \succ \delta_y$ (which must exist by non-triviality), and construct the act $xEy \in \mathcal{F}^*$. Since $\{\phi \in \Phi : \phi(E) = 0\} \neq \emptyset$, then we must have that $\min_{\phi\in\Phi} \phi(E)u(y) + (1-\phi(E))u(x) = u(y)$, which in turns means that $xEy \sim^* y$ (by Claim 14), hence $xEy \sim' y$. However, notice that, since $\ell(E) > 0$, then $\gamma^{-1}(xEy) \rhd_{FOSD} \gamma^{-1}(y)$, which implies that we must have $\gamma^{-1}(xEy) \succ \gamma^{-1}(y)$ by Axiom 2 (Monotonicity), which implies $xEy \succ' y$ by construction of $\succeq'$, contradicting $xEy \sim' y$. □

Claim 21. Φ is weak compact.

Proof. We already know that Φ is weak* compact. At the same time, we also know that every element in Φ is countably-additive: we can then apply Lemma 3 in Chateauneuf et al. (2005) to prove the desired result. (Notice that this argument could be also derived from standard Banach lattice techniques: as cited by Chateauneuf et al. (2005) one could follow Aliprantis and Burkinshaw (2006), especially Section 4.2.) □

Step 6. We now derive the main representations. First of all, define as $\hat{\mathcal{F}}$ the subset of acts in $\mathcal{F}'$ that are constant in the second component: $\hat{\mathcal{F}} := \{f' \in \mathcal{F}' : f'(\omega, [0,1]) = x \text{ for some } x \in X\}$. Define $\hat{\succeq}$ the restriction of $\succeq$ to $\hat{\mathcal{F}}$. Now notice that there exists a convex and compact set of finitely additive probability measures Π on Σ, such that $\hat{\succeq}$ is represented by the functional

$$\hat{V}(\hat{f}) := \min_{\pi\in\Pi} \int_{\Omega} \pi(\omega) u(\hat{f}(\omega, [0,1]))\mathrm{d}\omega.$$

Moreover, Π is unique. Again, this trivially follows from Claim 13, where Π is the project of P on Ω.

Claim 22. For any $p \in \Delta(X)$ there exists one $x \in X$ such that $\delta_x \sim p$.

Proof. The claim trivially follows from Axiom 3 (Continuity) and Axiom 1 (FOSD). □

By Claim 22 we know that $ce(p)$ is well defined for all $p \in \Delta(X)$. Now, for any act f, construct the act $\bar{f} \in \mathcal{F}$ as $\bar{f}(\omega) := \delta_{c_{f(\omega)}}$. Notice that for any $f, g \in \mathcal{F}$, we must have $f \succeq g$ if and only if $\bar{f} \succeq \bar{g}$ by Axiom 2 (Monotonicity). At the same time, notice that, by construction of γ, for every $f \in \mathcal{F}$, $|\gamma(\bar{f})| = 1$ and $\gamma(\bar{f})(\omega, [0,1]) = \delta_{c_{f(\omega)}}$. This means also that $\gamma(\bar{f}) \in \hat{\mathcal{F}}$ for all $f, g \in \mathcal{F}$. In turns, we must have that for all $f, g \in \mathcal{F}$, $f \succeq g$ if and only if $\gamma(\bar{f}) \succeq'$

[51] Recall that in this case we can define null events by saying that an event E is null if and only if $\phi(E) = 0$ for some $\phi \in \Phi$.

$\gamma(\bar{g})$, which is equivalent to $\gamma(\bar{f})\hat{\succsim}\gamma(\bar{g})$, which we know is true if and only if $\min_{\pi\in\Pi} \int_\Omega \pi(\omega)u(\gamma(\bar{f})(\omega,[0,1]))\mathrm{d}\omega \geq \min_{\pi\in\Pi} \int_\Omega \pi(\omega)u(\gamma(\bar{f})(\omega,[0,1]))\mathrm{d}\omega$. At the same time, we know that for each $f \in \mathcal{F}$, we have that $\gamma(\bar{f})(\omega,[0,1]) = \delta_{c_{f(\omega)}}$. In turns, this means that we have

$$f \succsim g \Leftrightarrow \min_{\pi\in\Pi} \int_\Omega \pi(\omega)u(\delta_{c_{f(\omega)}})\mathrm{d}\omega \geq \min_{\pi\in\Pi} \int_\Omega \pi(\omega)u(\delta_{c_{g(\omega)}})\mathrm{d}\omega.$$

At the same time, from Claim 15, we know that for all $p \in \Delta(X)$, $u(c_p) = V^*(\delta_{ce_p}) = V^*(p)$, where the first equality holds by construction of V^*, while the second equality holds because V^* represents the restriction of $\succsim$ to $\Delta(X)$ and because $ce_p \sim p$ for all $p \in \Delta(X)$. Given the definition of V^* above, therefore, we obtain that $\succsim$ is represented by the functional

$$V(f) := \min_{\pi\in\Pi} \int_\Omega \pi(\omega) \min_{\phi\in\Phi} \sum_{i=1}^{|\mathrm{supp}(p)|} \phi([\sum_{j=1}^{i-1} p(x_j), \sum_{j=1}^{i} p(x_j)])u(x_i)\mathrm{d}\omega,$$

which is the desired representation. (The uniqueness properties have been proved in the various steps.) Finally, notice that, if $\succsim$ satisfies Axiom 7, then we must have that $\succsim'$ is such that for any $f', g' \in \mathcal{F}'$ such that $f' \sim' g'$ we have $\frac{1}{2}f' \oplus \frac{1}{2}g' \sim' f'$. But then, by Claim 13 we have that $|P| = 1$, which implies $|\Pi| = |\Phi| = 1$. Moreover, since $\ell \in \Phi$, we must therefore have $\Phi = \{\ell\}$.

Proof of (2) $\Rightarrow$ (3). Consider a preference relation that admits a Multiple-Priors and Multiple Distortions representation (u, Π, Φ). The proof will proceed with the following three steps: 1) starting from a MP-MD representation, we will fix a measure-preserving function $\mu : \Delta(X) \to [0,1]^X$ such that it is, in some sense that we shall define below, *monotone* (in the sense that it assigns better outcomes to higher states in $[0,1]$); 2) we will prove that we can find an alternative representation of $\succsim$ which is similar to a Multiple-Priors and Multiple Distortions representation with (u, Π, Φ'), but which holds only for the measure-preserving map defined-above, and in which the set of priors Φ' on $[0,1]$ is made only of 'decreasing' priors (they assign a higher value to earlier states); 3) we will prove that this representation implies the existence of a MP-MC-RDEU representation.

Step 1. Let us consider a measure-preserving function $\mu : \Delta(X) \to [0,1]^X$ with the following two properties: for any $p \in \Delta(X)$ and for any $x \in X$, $\mu^{-1}(x)$ is convex; for any $p \in \Delta(X)$ and $x, y \in \mathrm{supp}(p)$, if $\delta_x \succ \delta_y$, then for any $r \in \mu^{-1}(x)$ and $s \in \mu^{-1}(y)$, we have $r > s$. The idea is that μ maps lotteries into acts which in which the set of states that return a given outcome is convex (first property), and such that the best outcomes are returned always by higher states (in $[0,1]$).

We now define a binary relation B on Φ as follows: for any $\phi, \phi' \in \Phi$, we have $\phi B \phi'$ if, and only if, $\int_{[0,1]} u(\mu(p))\mathrm{d}\phi \leq \int_{[0,1]} u(\mu(p))\mathrm{d}\phi'$ for *all* $p \in \Delta(X)$. Notice that the relation B depends on both u and μ; notice, moreover, that we have $\phi B \phi'$ and $\phi' B \phi$ off $\phi = \phi'$, which means that B is reflexive. Finally, notice that B is also transitive by construction.

Claim 23. *B is upper-semicontinuous when B is metrized using the weak metric. That is, for any $(\phi_m) \in \Phi^\infty$ and $\phi, \phi' \in \Phi$, if $\phi_m \to \phi'$ weakly and $\phi_m B \phi$ for all m, then $\phi' B \phi$.*

Proof. Suppose that we have ϕ_m, ϕ, and ϕ' as in the statement of the claim. This means that for any $p \in \Delta(X)$ we have $\int_{[0,1]} u(\mu(p))\mathrm{d}\phi_m \leq \int_{[0,1]} u(\mu(p))\mathrm{d}\phi$. Notice moreover that, by construction of μ, there must exist $x_1, \ldots, x_n \in X$ and $y_0, \ldots, y_n \in [0,1]$, where $y_0 = 0$ and $y_n = 1$, such that $\mu(p)(y) = x_i$ off $y \in [y_{i-1}, y_i]$ for $i = 1, n$. In turns, this means that for any $\bar{\phi} \in \Phi$, we have $\int_{[0,1]} u(\mu(p))\mathrm{d}\bar{\phi} = \sum_{i=1}^n u(x_i)\bar{\phi}([y_{i-1}, y_i])$. This means that we have $\sum_{i=1}^n u(x_i)\phi_m([y_{i-1}, y_i]) \leq \sum_{i=1}^n u(x_i)\phi([y_{i-1}, y_i])$. At the same time, recall that ϕ' is absolutely continuous with respect to the Lebesgue measure: this means that, by Portmanteau Theorem[52], since $\phi_m \to \phi'$ weakly, then we must have that $\phi_m([y_{i-1}, y_i]) \to \phi'([y_{i-1}, y_i])$ for $i = 1, \ldots, n$. But this means that we have $\sum_{i=1}^n u(x_i)\phi_m([y_{i-1}, y_i]) \to \sum_{i=1}^n u(x_i)\phi'([y_{i-1}, y_i])$, hence $\sum_{i=1}^n u(x_i)\phi'([y_{i-1}, y_i]) \leq \sum_{i=1}^n u(x_i)\phi([y_{i-1}, y_i])$, so $\int_{[0,1]} u(\mu(p))\mathrm{d}\phi' \leq \int_{[0,1]} u(\mu(p))\mathrm{d}\phi$. Since this must be true for any $p \in \Delta(X)$, we therefore have $\phi' B \phi$ as sought. □

[52] See (Billingsley, 1995, Chapter 5).

Step 2. Now define the set

$$\text{MAX}(\Phi, B) := \{\phi \in \Phi : \nexists \phi' \in \Phi \text{ s.t. } \phi' B \phi \text{ and } \phi' \neq \phi\}.$$

Claim 24. $\text{MAX}(\Phi, B) \neq \emptyset$.

Proof. Since Φ is weak compact and B is upper-semi-continuous (in the weak metric), then standard results in order theory show that $\text{MAX}(\Phi, B) \neq \emptyset$: see for example Theorem 3.2.1 in Ok (2011). $\square$

Claim 25. For any $p \in \Delta(X)$ we have $\min\limits_{\phi \in \text{MAX}(\Phi,B)} \int_{[0,1]} u(\mu(p))\mathrm{d}\phi = \min\limits_{\phi \in \Phi} \int_{[0,1]} u(\mu(p))\mathrm{d}\phi$

Proof. Since by construction $\text{MAX}(\Phi, B) \subseteq \Phi$, it trivially follows that the right hand side of the equation is smaller or equal than the left hand side for all $p \in \Delta(X)$. We are left to prove the converse. To this end, say by means of contradiction that there exists some $p \in \Delta(X)$ and some $\hat{\phi} \in \Phi \backslash \text{MAX}(\Phi, B)$ such that $\int_{[0,1]} u(\mu(p))\mathrm{d}\hat{\phi} < \min\limits_{\phi \in \text{MAX}(\Phi,B)} \int_{[0,1]} u(\mu(p))\mathrm{d}\phi$. This means that we cannot have $\phi' B \hat{\phi}$ for any $\phi' \in \text{MAX}(\Phi, B)$. Since B is transitive, we must therefore have that $\hat{\phi} \in \text{MAX}(\Phi, B)$, a contradiction. $\square$

Finally, define the set

$$\Phi' := \{\phi \in \text{MAX}(\Phi, B) : \phi \in \operatorname*{arg\,min}_{\phi \in \text{MAX}(\Phi,B)} \int_{[0,1]} u(\mu(p))\mathrm{d}\phi \text{ for some } p \in \Delta(X)\}.$$

We now define the notion of state-decreasing priors.

Definition 9. A prior ϕ on $[0, 1]$ is *state-decreasing* if there are do not exist any x_1, x_2, x_3, x_4 s.t. $x_1 < x_2 < x_3 < x_4$, $\ell([x_1, x_2]) = \ell([x_3, x_4])$ and $\pi([x_1, x_2]) < \pi([x_3, x_4])$.

Claim 26. Every prior $\phi \in \Phi'$ is state-decreasing.

Proof. Suppose by means of contradiction that there exists $\phi' \in \Phi'$ which is not state-decreasing. This means that there exist x_1, x_2, x_3, x_4 s.t. $x_1 < x_2 < x_3 < x_4$, $\ell([x_1, x_2]) = \ell([x_3, x_4])$ and $\phi'([x_1, x_2]) < \phi'([x_3, x_4])$. Now notice the following. If we have a MP-MD representation, then for any measure preserving map $\mu' : \Delta(X) \to [0, 1]^X$ we must have

$$\min_{\phi \in \Phi} \int_{[0,1]} u(\mu(p))\mathrm{d}\phi = u(ce_p) = \min_{\phi \in \Phi} \int_{[0,1]} u(\mu'(p))\mathrm{d}\phi$$

for all $p \in \Delta(X)$, for any $ce_p \in X$ such that $\delta_{ce_p} \sim p$. Since this must be true for every measure preserving μ' and for every p, then there must exist some $\hat{\phi} \in \Phi$ such that $\phi'(A) = \hat{\phi}(A)$ for all $A \subset [0, 1]$ such that $A \cap ([x_1, x_2] \cup [x_3, x_4] = \emptyset$, and $\hat{\phi}([x_1, x_2]) = \phi'([x_3, x_4])$ and $p\hat{h}i([x_3, x_4]) = \phi'([x_1, x_2])$: the reason is, if we take a measure preserving map μ' which is identical to μ except that it maps to $[x_3, x_4]$ whatever μ maps to $[x_1, x_2]$, and vice-versa, then there must exist a prior which minimizes the utility when μ' is used, and which returns exactly the same utility. Now notice that we must have that, by construction, $\phi'([x_1, x_2]) < \phi'([x_3, x_4])$, and hence $\hat{\phi}([x_1, x_2]) > \hat{\phi}([x_3, x_4])$, where $x_1 < x_2 < x_3 < x_4$, $\ell([x_1, x_2]) = \ell([x_3, x_4])$. (The two priors are otherwise the same.) But since μ assigns prizes with a higher utility to higher states, then this means that we have

$$\int_{[0,1]} u(\mu(p))\mathrm{d}\hat{\phi} \leq \int_{[0,1]} u(\mu(p))\mathrm{d}\phi'$$

for all $p \in \Delta(X)$. Since $\hat{\phi} \neq \phi'$, therefore, we have that $\hat{\phi} B \phi'$, which contradicts the fact that $\phi' \in \Phi' \subseteq \text{MAX}(\Phi, B)$. $\square$

This analysis leads us to the following claim:

Claim 27. There exists a closed, weak compact subset Φ'' of priors on $[0,1]$ such that every $\phi \in \Phi''$ is state-decreasing, atomless, mutually absolutely continuous with respect to ℓ such that $\succeq$ is represented by

$$V(f) := \min_{\pi \in \Pi} \int_{\Omega} \pi(\omega)\bar{U}(f(\omega))\mathrm{d}\omega$$

where $\bar{U} : \Delta(X) \to \mathbb{R}$ is defined as

$$\bar{U}(p) = \min_{\phi \in \Phi''} \int_{[0,1]} \phi(s)u(\mu(p))\mathrm{d}s.$$

(Here u, Π *and* μ are defined above.)

Proof. Simply define the set Φ'' as the closed convex hull of Φ'. Notice that this operation maintains the property that every ϕ in it is state-decreasing, and that it represents the preferences. Therefore, the result follows from Claim 25 and 26 . ☐

The set Φ'' in Claim 27 might not be unique.[53] However, we will now argue that there exists a unique *minimal* Φ'', where by minimal we understand a representation with a set Φ'' such that there is no $\hat{\Phi}'' \subset \Phi''$ which represents the same preferences and satisfies all the properties required by Claim 27 . Consider any minimal representation of the form above with set of priors $\bar{\Phi}'' \subseteq \Phi''$. To prove its uniqueness, assume by contradiction that there exists another minimal representation of the same preferences with a set of priors $\hat{\Phi}'' \neq \bar{\Phi}''$. Now construct the set H as the closed convex hull of $(\Phi \backslash \bar{\Phi}'') \cup \hat{\Phi}''$.[54] It is easy to see that we must have $H \neq \Phi$, since $\bar{\Phi}'' \neq \hat{\Phi}''$ and by the fact that $\bar{\Phi}''$ is minimal. The key observation is then to notice that (u, Π, H) is *also* a MP-MD representation of the same preferences. (By the uniqueness properties of it we can assume that the utility function is the same). To see why, consider first a measure-preserving map μ' which maps better outcomes to higher states in $[0,1]$, as the map μ defined above. For any such map, for each lottery at least one of the minimizing priors must belong to $\bar{\Phi}''$ in the first representation, by construction. At the same time, the value of these acts computed using the worst prior in $\bar{\Phi}''$ must be equivalent to the value computed using the worst prior in $\hat{\Phi}''$, because both represent the same preferences in the representation in Claim 27, hence must have the same certainty equivalents for each lottery. But then, the minimizing priors in the second representation must belong to $\hat{\Phi}''$, and thus for any increasing map both are representations of the same preferences.

Let us now consider a map μ'' which is not 'increasing,' i.e. which need not map better outcomes to better states. Now notice that for any such map in the first representation we cannot have a lottery for which the all minimizing prior belongs to $\bar{\Phi}''$. To see why, notice that if this was the case, we could also construct a lottery for which the minimizing priors also belongs to $\bar{\Phi}''$ (for the map at hand), but for which the value computed using a prior in $\bar{\Phi}''$ is strictly lower if we used an 'increasing' map (as μ' above) instead of μ''. The reason is, we can simply consider a lottery which is 'fine enough', i.e. returns different outcomes with small probability, so that the fact that the μ'' is not 'increasing' matters. (Recall that any prior in $\bar{\Phi}''$ assigns higher weight to lower states, which means that using it we obtain lower values for maps that assign better outcomes to higher states.) But this means that for this lottery we would obtain a strictly lower utility when we use a map like μ' as opposed to when we use μ'', which is impossible because the MP-MD representation should be independent of the map used (the certainty equivalents must be the same). This proves that for any map μ'', we cannot have that the unique minimizing prior belongs to $\bar{\Phi}''$. We now turn to argue that also for such map the two MP-MD representations must represent the same preferences. Given our last result, the only possibility for this not to be the case is that, for this map μ'', there exists a lottery p for which all the minimizing priors in the second representation belong to $\hat{\Phi}''$. We will now argue that this cannot be the case. If it were, then value of p computed using map μ'' and a prior in $\hat{\Phi}''$ must be strictly below the value of p computed using the same map and the worst prior in $(\Phi \backslash \bar{\Phi}'')$. And since we have proved that the we cannot have that the unique minimizing prior for the first representation belongs to $\bar{\Phi}''$, then this means that the value of p

[53] For example, if it doesn't already include it, one could add the identity function to the set, or any convex combination of the identity function with any member of the set, and leave the representation unchanged.

[54] Recall that Φ the set of distortions of the MP-MD representation of the same preferences.

computed using map μ'' and a prior in $\hat{\Phi}''$ is *strictly* below that computed using map μ'' and the worst prior in Φ. At the same time, notice that the value of p computed using the worst prior in $\hat{\Phi}''$ and map μ'' is weakly above that computed using the worst prior in $\hat{\Phi}''$ and an 'increasing' map like μ' above. In turns, however, we have proved that for any such map, this must be equal to the value computed using the worst prior in $\bar{\Phi}''$; by construction, this must be weakly higher than the value assigned by the first representation when using map μ'. But this means that we have some p such that the value assigned by the first representation when using map μ' is *strictly lower* than the one assigned by the same representation when using map μ''. But this contradicts the fact that a MP-MD representation represents the same preferences regardless of the map, as these lotteries would have a different certainty equivalent depending on which map we use.

Step 3. Consider now the representation in Claim 27, and for every $\phi \in \Phi''$ construct first the corresponding probability density function (PDF), pdf_ϕ. Notice that pdf_ϕ is well-defined since every ϕ is mutually absolutely continuous with respect to the Lebesgue measure (this follows from the Radon-Nikodym Theorem, (Aliprantis and Border, 2005, Theorem 13.18)). Moreover, notice that since every $\phi \in \Phi''$ is state-decreasing, then every pdf_ϕ is a decreasing function in $[0,1]$. Moreover, since every $\phi \in \Phi$ is mutually absolutely continuous with respect to the Lebesgue measure, then pdf_ϕ is never flat at zero. For each $\phi \in \phi''$, construct now the corresponding cumulative distribution function, and call the set of them Ψ. Notice that every $\psi \in \Psi$ must be *concave*, strictly increasing, and differentiable functions – because the corresponding PDFs exist, are decreasing, and never flat at zero. We are left to show that Ψ is point-wise compact: but this follows trivially from the standard result that for any two distributions ϕ, ϕ' on $[0,1]$ with corresponding CDFs ψ and ψ' such that both are continuous on $[0,1]$, we have that $\phi \to \phi'$ weakly if, and only if, $\psi \to \psi'$ pointwise.[55] The desired representation then follows trivially, as does the existence of a minimal representation. Finally, the unique properties of the minimal representation follow trivially from the uniqueness properties of the representation in Claim 27, discussed above.

Proof of (3) $\Rightarrow$ (1). We start by proving the necessity of Axiom 3 (Continuity). For brevity in what follows we will only prove that if $\succeq$ admits the representation in (3), then for any $(p_n) \in (\Delta(X))^\infty$, and for any $p, q \in \Delta(X)$, if $p_n \succeq q$ for all n and if $p_n \to p$ (in the topology of weak convergence), then $p \succeq q$. The proof for the specular case in which $p_n \preceq q$ for all n is identical, while the extension to non-constant acts follows by standard arguments once the convergence for constant acts is enstablished. To avoid confusion, we denote $p_n \to^w p$ to indicate weak convergence, $f_n \to^p f$ to denote point-wise convergence, and $\to$ to indicate convergence in $\mathbb{R}$.

Claim 28. Consider $\psi_n \in \Psi^\infty$, $\psi \in \Psi$, $p_n \in \Delta(X)^\infty$, $p \in \Delta(X)$ such that $\psi_n \to^p \psi$ and $p_n \to^w p$. Then $\mathrm{RDEU}_{u,\psi_n}(p_n) \to \mathrm{RDEU}_{u,\psi}(p)$.

Proof. Consider $\psi_n \in \Psi^\infty$, $\psi \in \Psi$, $p_n \in \Delta(X)^\infty$, and $p \in \Delta(X)$ as in the statement of the Claim. (What follows is an adaptation of the Proofs in (Chateauneuf, 1999, Remark 9) to our case.) Notice that since X is a connected and compact set, and since u is continuous, we can assume wlog $u(X) = [0,1]$. Also, for any $t \in [0,1]$, define $A_t := \{x \in X : u(x) > t\}$. Then, notice that for any $p \in \Delta(X)$ and $\psi \in \Psi$ we have

$$\mathrm{RDEU}_{u,\psi}(p) = \int_0^1 \psi(p(A_t))\mathrm{d}t.$$

Define now $H_n, H : [0,1] \to [0,1]$ by $H_n(t) = \psi_n(p_n(A_t))$ and $H(t) = \psi(p(A_t))$. We then have $\mathrm{RDEU}_{u,\psi_n}(p_n) = \int_0^1 H_n(t)\mathrm{d}t$ and $\mathrm{RDEU}_{u,\psi}(p) = \int_0^1 H(t)\mathrm{d}t$. Since $|H_n(t)| \leq 1$ for all $t \in [0,1]$ and for all n, then by the Dominated Convergence Theorem (see (Aliprantis and Border, 2005, Theorem 11.21)) to prove that $\mathrm{RDEU}_{u,\psi_n}(p_n) \to \mathrm{RDEU}_{u,\psi}(p)$ we only need to show that $H_n(t) \to H(t)$ for almost all $t \in [0,1]$. To do this, we denote $M_p := \{r \in [0,1] : \exists x \in \mathrm{supp}(p)$ such that $u(x) = r\}$, and we will show that we have $H_n(t) \to H(t)$ for all $t \in [0,1] \backslash M_p$: since p is a simple lottery (with therefore finite support), this will be enough.

Consider some $t \in [0,1] \backslash M_p$, and notice that we must have that A_t is a continuity set of p. To see why, notice that, since u is continuous, A_t must be open, and we have that $\delta A_t = \{x \in X : u(x) = t\}$; and since $t \notin M_t$, then we must

[55] This is a standard result. See, for example, the discussion in (Billingsley, 1995, Chapter 5).

have $p(\delta A_t) = 0$. By Portmanteau Theorem[56] we then have $p_n(A_t) \to p(A_t)$. We will now argue that for any such t we must also have $H_n(t) \to H(t)$, which will conclude the argument. To see why, consider any $t \in [0,1] \backslash M_p$, and notice that we must have $|H_n(t) - H(t)| = |\psi_n(p_n(A_t)) - \psi(p(A_t))| < |\psi_n(p_n(A_t)) - \psi_n(p(A_t))| + |\psi_n(p(A_t)) - \psi(p(A_t))|$. At the same time: $|\psi_n(p_n(A_t)) - \psi_n(p(A_t))|$ can be made arbitrarily small since $p_n(A_t) \to p(A_t)$ and ψ_n is continuous; and $|\psi_n(p(A_t)) - \psi(p(A_t))|$ can be made arbitrarily small since $\psi_n \to^p \psi$. But then, we must have $H_n(t) \to H(t)$ as sought.

□

Notice, therefore, that we can apply standard generalizations of Berge's Theorem of the maximum, such as (Aliprantis and Border, 2005, Theorem 17.13),[57] and therefore prove that Axiom 3 (Continuity).

Next, we turn to prove the necessity of Axiom 5 (Hedging). To this end, let us define the notion of enumeration.

Definition 10. A *simple enumeratio*n of a lottery q is a step function $x : [0,1] \to X$ such that $l(\{z \in [0,1] | f(z) = w\} = q(w)\ \forall\ w \in \text{supp}(q)$.

Let $N(x) \in \mathbb{N}$ be the number of steps in x, and x_n be the value of $f(x)$ at each step, and $p^x(x_n)$ be the Lesbegue measure of each step x_n.

Claim 29. Let p be some lottery, and x, y be two simple enumerations of p such that $x_{i-1} \preceq x_i$ for all $2 \leq i \leq n$. Then, if ψ is a concave RDU functional and u is a utility function that represents $\succeq$, we have $W(x) = \psi(p^x(x_1))u(x_1) + \sum_{i=2}^{N(x)}(\psi(\sum_{j=1}^{i} p^x(x_j)) - \psi(\sum_{j=1}^{i-1} p^x(x_j)))u(x_i)$, which must be smaller or equal to $\psi(p^y(y_1))u(y_1) + \sum_{i=2}^{N(y)}(\psi(\sum_{j=1}^{i} p^y(y_j)) - \psi(\sum_{j=1}^{i-1} p^y(y_j)))u(y_i) = W(y)$.

Proof. We begin by proving the claim for cases in which p^y map to rational numbers, then extend the claim using the continuity of W. As $p^y(y_i)$ is rational, for all.$i \in 1..N(y)$ we can write each $p^y(y_i) = \frac{m_i}{n_i}$ for some set of integers $\{m_i\}$ and $\{n_i\}$. This means that there are a set of natural numbers $\{k_i\}$ such that $p^y(y_i) = \frac{k_i}{\prod n_i}$. Notice that we can rewrite the step function y as a different step function $\bar{y}$ defined by the intervals $\left\{[\frac{j}{\prod n_i}, \frac{j+1}{\prod n_i})\right\}_{j=0}^{\prod n_i - 1}$, where the value of the function in the interval $[\frac{j}{\prod n_i}, \frac{j+1}{\prod n_i})$ is equal to the value of y in the interval $[p^y(y_l), p^y(y_m))$ where $l = \max\left\{t \in \mathbb{N} | p^y(y_t) \leq \frac{j}{\prod n_i}\right\}$ and $\min\left\{t \in \mathbb{N} | p^y(y_t) \geq \frac{j+1}{\prod n_i}\right\}$. In other words, we have split the original step function y up into a finite number of equally spaced steps, while preserving the value of the original function (again we can do this because the original function had steps defined by rational number). We can therefore now think of $\bar{y}$ as consisting of a finite numer of equally lengthed elements that can be interchanged using the procedure we discuss below Note that redefining y in this way does not change the function - i.e. $y(t) = \bar{y}(t)\ \forall\ t$, and nor does it affect its utility - i.e. $W(y) = W(\bar{y})$.

Now order the steps of $\bar{y}$ using $\succeq$, breaking ties arbitrarily. Let $\bar{y}^1$ denote the worst step of $\bar{y}$, $\bar{y}^2$ the next worst element and so on. We next define a sequence of enumerations and functions recursively:

1. Let ${}^1\bar{y} = \bar{y}$. Define the function ${}^1r : \{1...N(y)\} \to \mathbb{N}$ such that ${}^1r(j)$ is the original position of $\bar{y}^j$ for all j (i.e. ${}^1r(j) = \{n \in \mathbb{N} | {}^1\bar{y}_{{}^1r(j)} = \bar{y}^j\}$).

2. Define ${}^i\bar{y}$ as ${}^i\bar{y}(t) = \bar{y}^i$ for $t \in [\frac{i-1}{\prod n_i}, \frac{i}{\prod n_i})$; ${}^i\bar{y}(t) =^{i-1} \bar{y}_i$ for $t \in [\frac{{}^{i-1}r(i)-1}{\prod n_i}, \frac{{}^{i-1}r(i)}{\prod n_i})$; and ${}^i\bar{y}(t) =^{i=1} \bar{y}(t)$ otherwise.

3. Define ${}^ir(j)$ as the position of y^j in ${}^i\bar{y}$ for all j (i.e. ${}^ir(j) = \{n \in \mathbb{N} | {}^i\bar{y}_{{}^ir(j)} = \bar{y}^j\}$)

[56] See (Billingsley, 1995, Chapter 5).

[57] In particular, in our case the correspondence ρ in the statement of the theorem would be constant and equal to Ψ, which is non-empty and compact, while the function f in the statement of the theorem would correspond to the function $\text{RDEU}_{u,\psi}(p)$ seen as a function of both ψ and p – which, as we have seen, is continuous.

So, at each stage, this procedure takes the previous function, looks for the ith worst step of $\bar{y}$ and switches it into the ith position in the enumeration (while moving whatever was in that slot back to where the worst element came from. The function ${}^{i}r$ keeps track of the location of each of the steps of $\bar{y}$ in each iteration i. The first thing to note is that the final element in this sequence, ${}^{\prod n_i}\bar{y}$, is equivalent to x, in the sense that $W(x) = W({}^{\prod n_i}\bar{y})$: Clearly, each of these switches preserve the Lesbegue measure associated to each prize, thus ${}^{\prod n_i}\bar{y}$ is an enumeration of p. Furthermore ${}^{\prod n_i}\bar{y}_{i-1} \preceq^{\prod n_i} \bar{y}_i$ for all i by construction, meaning that $u({}^{\prod n_i}\bar{y}(t)) = u(x(t))$ for all t.

Next, we show that $W({}^{i}y) \leq W({}^{i-1}y)$ for all $i \in \{2, .. \prod n_i\}$ First, note that it must be the case that ${}^{i-1}\bar{y}_i \succeq \bar{y}^i$: in words, the ith worst element of $\bar{y}$ must be weakly worse than whatever is in the ith slot in ${}^{i-1}\bar{y}$. To see this, note that, if this were not the case, then it must be the case that ${}^{i-1}\bar{y}_i = \bar{y}^j$ for some $j < i$. But, by the iterative procedure, $\bar{y}^j$ must be in slot ${}^{i-1}\bar{y}_j \neq^{i-1} \bar{y}_i$. Next, note that it must be the case that ${}^{i-1}r(i) \geq i$. By the iterative procedure, for all $j < i$, ${}^{i-1}\bar{y}_j = \bar{y}^j \neq \bar{y}^i$. Thus, as ${}^{i-1}r(i)$ is the location of $\bar{y}^i$ in ${}^{i-1}\bar{y}_j$, it must be the case that ${}^{i-1}r(i) \geq i$.

Next, note that ${}^{i}y$ and ${}^{i-1}y$ differ only on the intervals $[\frac{i-1}{\prod n_i}, \frac{i}{\prod n_i})$ and $[\frac{{}^{i-1}r(i)-1}{\prod n_i}, \frac{{}^{i-1}r(i)}{\prod n_i})$. Thus, we can write the difference between $W({}^{i}y)$ and $W({}^{i-1}y)$ as $(\psi(\sum_{j=1}^{i} p({}^{i}y_j) - \psi(\sum_{j=1}^{i-1} p({}^{i}y_j)))\ (u({}^{i-1}\bar{y}_i) - u(\bar{y}^i)) + (\psi(\sum_{j=1}^{r(i)} p({}^{i}y_j) - \psi(\sum_{j=1}^{r(i)-1} p({}^{i}y_j)))(u(\bar{y}_i) - u({}^{i-1}\bar{y}_i))$. This is equal to $\left(\left(\psi\left(\frac{i}{\prod n_i}\right) - \psi\left(\frac{i-1}{\prod n_i}\right)\right) - \left(\psi\left(\frac{{}^{i-1}r(i)}{\prod n_i}\right) - \psi\left(\frac{{}^{i-1}r(i)-1}{\prod n_i}\right)\right)\right)(u({}^{i-1}\bar{y}_i) - u(\bar{y}^i))$.

Now, as ${}^{i-1}\bar{y}_i \succeq \bar{y}^i$, it must be the case that $u({}^{i-1}\bar{y}_i) \geq u(\bar{y}^i)$, and so $(u({}^{i-1}\bar{y}_i) - u(\bar{y}^i)) \geq 0$. Furthermore, it must be the case that the term in the first parentheses is also weakly positive by the concavity of ψ. To see this, define the function $\bar{\psi}(x) = \psi\left(x + \frac{i-1}{\prod n_i}\right) - \psi\left(\frac{i-1}{\prod n_i}\right)$. This is a concave function with $\bar{\psi} \geq 0$, and so is subadditive. This means that we have $\bar{\psi}(\frac{{}^{i-1}r(i)}{\prod n_i} - \frac{i-1}{\prod n_i}) \leq \bar{\psi}(\frac{{}^{i-1}r(i)-1}{\prod n_i} - \frac{i-1}{\prod n_i}) + \bar{\psi}((\frac{{}^{i-1}r(i)}{\prod n_i} - \frac{i-1}{\prod n_i}) - (\frac{{}^{i-1}r(i)-1}{\prod n_i} - \frac{i-1}{\prod n_i}))$. In turns, this is equal to $\bar{\psi}(\frac{{}^{i-1}r(i)-1}{\prod n_i} - \frac{i-1}{\prod n_i}) + \bar{\psi}(\frac{1}{\prod n_i})$. If we then substitute the original function we get $\psi(\frac{{}^{i-1}r(i)}{\prod n_i}) \leq \psi(\frac{{}^{i-1}r(i)-1}{\prod n_i}) + \psi(\frac{i}{\prod n_i})$-$\psi(\frac{i-1}{\prod n_i})$. This means $\psi(\frac{{}^{i-1}r(i)}{\prod n_i}) - \psi(\frac{{}^{i-1}r(i)-1}{\prod n_i}) \leq \psi(\frac{i}{\prod n_i}) - \psi(\frac{i-1}{\prod n_i})$. Thus, by iteration we have $W(y) = W(\bar{y}) = W({}^{1}\bar{y}) \geq W\left({}^{\prod n_i}\bar{y}(t)\right) = W(x)$ and we are done.

To extend the proof to enumerations with irrational p functions, take such a function y, and associated x that is the rank order enumeration of y, whereby $p^y(y_i)$ is not guaranteed to be rational for all $i \in 1...N(y)$. Now note that p^y is a vector in $\mathbb{R}^{N(y)}$. Note that I can construct a sequence of vectors $q^i \in \mathbb{Q}^{N(y)}$ such that $\{q^i\} \to p^y$. Define the simple enumeration y^i as the step function whereby $y^i(t) = y_n$ for $t \in [\sum_{j=0}^{n-1} q_{i-1}^i, \sum_{j=0}^{n} q_{i-1}^i)$. The utility of the enumeration y^i is given by

$$W(y^i) = \psi(q_1^i)u(y_1) + \sum_{k=2}^{N(y)} \left(\psi(\sum_{j=0}^{k} q_{i-1}^i)) - \psi(\sum_{j=0}^{k-1} q_{i-1}^i)\right) u(y_k).$$

As $q_j^i \to p^y(y_j)$, and as ψ is continuous, then it must be the case that $W(y^i) \to W(y)$. Similarly, if we let x^i be the rank enumeration of y^i, then it must be the case that $W(x^i) \to W((x)$. Thus, if it were the case that $W(y) > W(x)$, then there would be some i such that $W(y^i) > W\left(x^i\right)$. But as y^i is rational, this contradicts the above result. □

We now turn to prove that the Axiom 5 is satisfied. Again we will prove this only for degenerate acts – the extension to the general case being trivial. Let p, q be two lotteries such that $p \sim q$ and $r \in \bigoplus_{p,q}^{\frac{1}{2}}$. Let x be the enumeration of r, then there must be two enumerations z^x and z^y such that: 1) $z_i = \frac{1}{2}z_i^x \oplus \frac{1}{2}z_i^y$ for all i; 2) for every x_i, $\sum_{i|z_i^x = x_i} r(z_i) = p(x_i)$ and $\sum_{i|z_i^y = y_i} r(z_i) = p(y_i)$. Now, the utility of r is given by $U(r) = \min_{\pi\in\Pi} \sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z_i)$ which is equal to $\min_{\pi\in\Pi} \sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))(\frac{1}{2}(u(z_i^x) + u(z_i^y)))$, which is in turns equal to $\min_{\pi\in\Pi}[\frac{1}{2}(\sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z_i^x)) + \frac{1}{2}(\sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z_i^y))]$. This must be larger or equal than $\frac{1}{2}\min_{\pi\in\Pi}(\sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z_i^x)) + \frac{1}{2}\min_{\pi\in\Pi}(\sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z^y))$.

Note that the enumerations are not in rank order, but, by Claim 29, reordering can only decrease the utility of the enumeration by shuffling them into the rank order for every $\pi \in \Pi$. Let $\bar{z}^x$ and $\bar{z}^y$ be the rank order enumerations of z^x. We must then have that $U(r)$ is larger or equal than $\frac{1}{2}\min_{\pi\in\Pi}(\sum_i (\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z_i^x)) +$

$\frac{1}{2}\min_{\pi\in\Pi}(\sum_i(\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(z^y))$. This is larger or equal than $\frac{1}{2}\min_{\pi\in\Pi}(\sum_i(\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(\bar{z}_i^x)) + \frac{1}{2}\min_{\pi\in\Pi}(\sum_i(\pi(\sum_{j=0}^{i-1} r(z_j)) - \pi(\sum_{j=0}^{i} r(z_j)))u(\bar{z}_i^y))$, which is then equal to $\frac{1}{2}U(p) + \frac{1}{2}U(q)$ as sought.

We now turn to Axiom 1 (FOSD). Let π be a continuous RDEU functional. We know (e.g. (Wakker, 1994, Theorem 12,)) that it respects FOSD. Thus, suppose that p first order stochastically dominates q, and let $\pi^* \in \Pi$ be the functional that minimizes the utility of p. We know that the utility of q under this functional has to be lower than the utility of p, thus the utility of q (which is assessed under the functional that minimazes the utility of q) is lower than that of p

Finally, Axiom 2 (Monotonicity) and Axiom 6 (Degenerate Independence) follow form standard arguments, while Axiom 4 follows from Lemma 1.

Proof of Proposition 1

The proof of all the steps except the equivalence between (b).(1) and (b).(4) follows standard arguments and it is therefore omitted. If $\Phi = \{\ell\}$ it is also trivial to see that Axiom 10 is satisfied. Assume now that Axiom 10 holds, and that $\succeq$admits a MP-MD representation (u, Π, Φ). By Theorem 1 we know that it will also admit a Minimal Multiple Priors and Multiple Concave RDEU Representation (u, Π, Ψ). We now argue that we must have $|\Psi| = 1$ and that it contains only the identify function. Suppose this is not the case, and say $\psi \in \Psi$ where ψ is not the identity function. Since ψ must be concave and it must be a probability weighting function (increasing, $\psi(0) = 0$, $\psi(1) = 1$), then we must also have that $\psi(x) > x$ for all $x \in (0, 1)$. But this implies that we have $RDEU_{u,\psi}(\alpha x + (1-\alpha)y) < \alpha u(x) + (1-\alpha)u(y)$. Since $\psi \in \Psi$, we must then have that for all $x, y \in X$ such that $u(x) \neq u(y)$, and for all $\alpha \in (0, 1)$, we must have $\alpha x + (1-\alpha)y \prec \delta_{\alpha x \oplus (1-\alpha)y}$, in direct violation of Axiom 10. We must therefore have that $|\Psi| = 1$ and that it contains only the identify function, which in turn implies $\Phi = \{\ell\}$ as sought.

Proof of Proposition 2

Notice first of all that both $\succeq_1$ and $\succeq_2$ we can follow Steps from 1 to 4 of the proof of Theorem 1, and obtain two preference relation $\succeq_1'$ and $\succeq_2'$ on $\mathcal{F}'$, both of which admit a representation as in Claim 13 of the form (u_1', P_1') and (u_2', P_2'). Notice, moreover, that we must have $u_1' = u_1$ and $u_2' = u_2$, and we must also have, by construction, $P_1' = \Pi_1 \times \Phi_1$ and $P_2' = \Pi_2 \times \Phi_2$.

Suppose now that we have that $\succeq_2$ is more attracted to certainty than $\succeq_1$. Then, we must have $\oplus_{\succeq_1} = \oplus_{\succeq_2}$, which implies that u_1 is a positive affine transformation of u_2. But this means that u_1' is a positive affine transformation of u_2', which means that, since both $\succeq_1'$ and $\succeq_2'$ are biseparable and have essential events (as proved in the steps from the proof of Theorem 1), then by (Ghirardato and Marinacci, 2002, Proposition 6) $\succeq_1'$ and $\succeq_2'$ are cardinally symmetric. Moreover, since $\succeq_2$ is more attracted to certainty than $\succeq_1$, it is easy to see that we must have that $\succeq_2'$ is more uncertainty averse than $\succeq_1'$ in the sense of (Ghirardato and Marinacci, 2002, Definition 4). We can then apply (Ghirardato and Marinacci, 2002, Theorem 17), and obtain that we must have $P_2' \supseteq P_1'$. Since $P_1' = \Pi_1 \times \Phi_1$ and $P_2' = \Pi_2 \times \Phi_2$, this implies $\Pi_2 \supseteq \Pi_1$ and $\Phi_2 \supseteq \Phi_1$.

Now suppose that we have $\Pi_2 \supseteq \Pi_1$, $\Phi_2 \supseteq \Phi_1$, and that u_1 is a positive affine transformation of u_2. This first of all implies $\oplus_{\succeq_1} = \oplus_{\succeq_2}$. Moreover, it also implies that u_1' is a positive affine transformation of u_2', and we must have $P_2' \supseteq P_1'$. Again by (Ghirardato and Marinacci, 2002, Theorem 17) we then have that $\succeq_2'$ is more uncertainty averse than $\succeq_1'$ in the sense of (Ghirardato and Marinacci, 2002, Definition 4), which implies that $\succeq_2$ is more attracted to certainty than $\succeq_1$, as sought.

Proof of Theorem 2

The proof of the necessity of the axioms is trivial (in light of Theorem 1) and therefore left to the reader. We now prove the sufficiency of the axioms. Since $\succeq$ satisfies axioms 1-6 we know it admits a Multiple Priors - Multiple

Concave RDEU representation (u, Π, Ψ). Define $V : \mathcal{F} \to \mathbb{R}$ as in the definition of a Multiple Priors and Multiple Concave Rank-Dependent Representation. For any $g \in \mathcal{F}$, construct the set $\Pi_g := \{\pi \in \Delta(\Omega) : g^\pi \sim g$ and $f^\pi \succeq f$ for all $f \in \mathcal{F}\}$. By Axiom 11, $\Pi_g \neq \emptyset$ for all $g \in \mathcal{F}$. Now define $\Pi' := \cup_{g \in \mathcal{F}} \Pi_g$. Notice the following three properties of the set Π'. First, $\Pi' \neq \emptyset$; second, for all $f \in \mathcal{F}$ we must have $f^\pi \succeq f$ for all $\pi \in \Pi'$; third for all $f \in \mathcal{F}$ there exists $\pi \in \Pi'$ such that $f^\pi \sim f$. Define $\hat{\Pi}$ as the closed convex hull of Π'.

Now consider any $f \in \mathcal{F}$, and any $\pi_f \in \hat{\Pi}$ such that $f^{\pi_f} \sim f$. This means that we have, for any enumeration of the states in Ω such that $f(\omega_{i-1}) \preceq f(\omega_i)$ for $i = 2, \ldots, |\text{supp}(p)|$, we have

$$V(f) = V(f^{\pi_f}) = \min_{\psi \in \Psi} \ \psi(\pi_f(\omega_1))U(f(\omega_1)) + \sum_{i=2}^{n} \big[\psi(\sum_{j=1}^{i} \pi_f(\omega_j)) - \psi(\sum_{j=1}^{i-1} \pi_f(\omega_j))\big]U(f(\omega_i)).$$

In turns, this means that we have $V(f) = V(f^{\pi_f}) = \min_{\psi \in \Psi} \text{RDEU}_{u,\psi}(f^{\pi_f})$. For simplicity of notation, define $H_f : \Delta(\Omega) \to \mathbb{R}$ as $H_f(\pi) := \min_{\psi \in \Psi} \text{RDEU}_{u,\psi}(f^\pi)$. Notice that we have that $H_f(\pi) = V(f^\pi)$ by construction. Moreover, notice that H_f is continuous by construction as well. We now prove $V(f) \leq \min_{\pi \in \hat{\Pi}} H_f(\pi)$. We proceed in steps. First of all, notice that $V(f) \leq \inf_{\pi \in \Pi'} H_f(\pi)$.[58] This is the case because, by the properties of Π' discussed above, we must have $f^\pi \succeq f$ for all $\pi \in \Pi'$. Now notice that $V(f) \leq \min_{\pi \in \bar{\Pi}'} H(\pi)$, where by $\bar{\Pi}'$ denotes the closure of Π'. To see why, say by means of contradiction that there exists $\pi \in \bar{\Pi}'$ such that $H(\pi') < V(f)$. Since we know that $V(f) \leq \inf_{\pi \in \Pi'} H(\pi)$, however, this means that there exists a sequence (π_n) in Π' such that $\pi_n \to \pi'$, $V(f) \leq H(\pi_n)$, but $H(\pi') < V(f)$. But this clearly contradicts the continuity of H. This proves that $V(f) \leq \min_{\pi \in \bar{\Pi}'} H(\pi)$.

Claim 30. For any $\pi, \pi \in \Delta(\Omega)$ and $\lambda \in (0,1)$, if $V(f) \leq H(\pi)$ and $V(f) \leq H(\pi')$ then $V(f) \leq H(\lambda\pi + (1-\lambda)\pi')$.

Proof. Notice that we have $V(f) \leq \lambda(H(\pi)) + (1-\lambda)(H(\pi')) \leq \min_{\psi \in \Psi} (\lambda \text{RDEU}_{u,\psi}(f^\pi) + (1-\lambda)\text{RDEU}_{u,\psi}(f^{\pi'}))$, where the last inequality is due to the fact that we are taking the min only once. In turns, by concavity of ψ the latter is smaller or equal to $\min_{\psi \in \Psi} \text{RDEU}_{u,\psi}(f^{\lambda\pi + (1-\lambda)\pi'}) = H(\lambda\pi + (1-\lambda)\pi')$ as sought. $\square$

Our previous results, together with Claim 30, imply $V(f) \leq \min_{\pi \in \hat{\Pi}} H(\pi)$ as sought. At the same time, we know that $V(f) \geq H(\pi^f)$ where $\pi^f \in \hat{\Pi}$, which means that we have $V(f) \geq \min_{\pi \in \hat{\Pi}} H(\pi)$, and hence $V(f) = \min_{\pi \in \hat{\Pi}} H(\pi)$ as sought.

References

Abdellaoui, M. (2002): "A Genuine Rank-Dependent Generalization of the Von Neumann-Morgenstern Expected Utility Theorem," *Econometrica*, 70, 717–736.

Aliprantis, C. D. and K. C. Border (2005): *Infinite Dimensional Analysis*, Springer, Berlin Heidelberg New York.

Aliprantis, C. D. and O. Burkinshaw (2006): *Positive operators*, Springer-Verlag New York Inc.

Anscombe, F. J. and R. J. Aumann (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199–205.

Billingsley, P. (1995): *Probability and measure*, Wiley (New York).

Cerreia, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2010): "Uncertainty Averse Preferences," *Journal of Economic Theory (forthcoming).*

Cerreia-Vioglio, S. (2010): "Maxmin Expected Utility on a Subjective State Space: Convex Preferences under Risk," Mimeo Bocconi University.

Cerreia-Vioglio, S., P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi (2011): "Rational Preferences under Ambiguity," *Economic Theory This issue.*

[58] Notice that this is not what we need to prove as we are considering only Π' and not the larger $\hat{\Pi}$.

Chateauneuf, A. (1999): "Comonotonicity axioms and rank-dependent expected utility theory for arbitrary consequences," *Journal of Mathematical Economics*, 32, 21–45.

Chateauneuf, A., F. Maccheroni, M. Marinacci, and J. M. Tallon (2005): "Monotone continuous multiple priors," *Economic Theory*, 26, 973–982.

Chew, S. and P. Wakker (1996): "The Comonotonic Sure-Thing Principle," *Journal of Risk and Uncertainty*, 12, 5–26.

Chew, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51, 1065–1092.

Chew, S. H. and J. S. Sagi (2008): "Small worlds: Modeling attitudes toward sources of uncertainty," *Journal of Economic Theory*, 139, 1–24.

Dean, M. and P. Ortoleva (2012): "Estimating the Relationship between Economic Preferences: A Testing Ground for Unified Theories of Behavior," Mimeo, California Insitute of Technology.

Dekel, E. (1986): "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom* 1," *Journal of Economic Theory*, 40, 304–318.

Diecidue, E. and P. Wakker (2001): "On the intuition of rank-dependent utility," *Journal of Risk and Uncertainty*, 23, 281–298.

Dillenberger, D. (2010): "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, 78, 1973–2004.

Drapeau, S. and M. Kupper (2010): "Risk Preferences and their Robust Representation," Mimeo Humboldt University Berlin.

Epstein, L. G. (1999): "A definition of uncertainty aversion," *Review of Economic Studies*, 66, 579–608.

Ghirardato, P. (2004): "Defining Ambiguity and Ambiguity Attitude," in *Uncertainty in economic theory: essays in honor of David Schmeidler's 65th birthday*, ed. by I. Gilboa, London: Routledge, vol. 63.

Ghirardato, P., F. Maccheroni, M. Marinacci, and M. Siniscalchi (2001): "A subjective spin on roulette wheels," California Insitute of Technology, Social Science Working Paper 1127.

——— (2003): "A subjective spin on roulette wheels," *Econometrica*, 71, 1897–1908.

Ghirardato, P. and M. Marinacci (2001): "Risk, ambiguity, and the separation of utility and beliefs," *Mathematics of Operations Research*, 864–890.

——— (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251–289.

Ghirardato, P. and M. Siniscalchi (2010): "A more robust definition of multiple priors," Mimeo Northwestern University.

Gilboa, I. and M. Marinacci (2011): "Ambiguity and the Bayesian paradigm," Mimeo Bocconi University.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with a non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.

Gul, F. (1991): "A theory of disappointment aversion," *Econometrica*, 59, 667–686.

Gumen, A., E. A. Ok, and A. Savochkin (2011): "Decision-Making under Subjective Risk," Mimeo New York University.

Halevy, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536.

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

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73, 1849–1892.

Kobberling, V. and P. Wakker (2003): "Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique," *Mathematics of Operations Research*, 28, 395–423.

Maccheroni, F. (2002): "Maxmin under risk," *Economic Theory*, 19, 823–831.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

Nakamura, Y. (1995): "Rank dependent utility for arbitrary consequence spaces," *Mathematical Social Sciences*, 29, 103–129.

Ok, E. (2011): "Elements of Order Theory," Mimeo, New York University.

Ok, E., P. Ortoleva, and G. Riella (2011): "Incomplete Preferences under Uncertainty: Indecisiveness in Beliefs vs. Tastes," *Econometrica (forthcoming)*.

Quiggin, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3, 323–343.

Savage, L. J. (1954): *The foundations of statistics*, Wiley Publications in Statistics.

Schmeidler, D. (1989): "Subjective probability and expected utility without additivity," *Econometrica*, 57, 571–587.

Segal, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202.

——— (1990): "Two-stage lotteries without the reduction axiom," *Econometrica: Journal of the Econometric Society*, 349–377.

Seo, K. (2009): "Ambiguity and second-order belief," *Econometrica*, 77, 1575–1605.

Starmer, C. (2000): "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk," *Journal of Economic Literature*, 38, 332–382.

Tversky, A. and D. Kahneman (1992): "Advances in prospect theory: cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P. (1990): "Characterizing optimism and pessimism directly through comonotonicity," *Journal of economic theory*, 52, 453–463.

——— (1994): "Separating marginal utility and probabilistic risk aversion," *Theory and decision*, 36, 1–44.

——— (1996): "The sure-thing principle and the comonotonic sure-thing principle: An axiomatic analysis* 1," *Journal of Mathematical Economics*, 25, 213–227.

——— (2001): "Testing and Characterizing Properties of Nonadditive Measures Through Violations of the Sure-Thing Principle," *Econometrica*, 69, 1039–1059.

Yaari, M. E. (1987): "The dual theory of choice under risk," *Econometrica: Journal of the Econometric Society*, 95–115.