DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

OBJECTIVE LOTTERIES AS AMBIGUOUS OBJECTS: ALLAIS, ELLSBERG, AND HEDGING

Mark Dean
Brown University

Pietro Ortoleva
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 1356

October 2011

Objective Lotteries as Ambiguous Objects: Allais, Ellsberg, and Hedging*

Mark Dean† and Pietro Ortoleva‡

October 19, 2011

Preliminary and Incomplete

Abstract

We derive axiomatically a model in which the Decision Maker can exhibit simultaneously both the Allais and the Ellsberg paradoxes in the standard setup of Anscombe and Aumann (1963). Using the notion of 'subjective', or 'outcome' mixture of Ghirardato et al. (2003), we define a novel form of hedging for objective lotteries, and introduce a novel axiom which is a generalized form of preferences for hedging. We show that this axiom, together with other standard ones, is equivalent to a representation in which the agent reacts to ambiguity using multiple priors like the MaxMin Expected Utility model of Gilboa and Schmeidler (1989), generating an Ellsberg-like behavior, while at the same time, she treats also objective lotteries as 'ambiguous objects,' and use a fixed (and unique) set of priors to evaluate them – generating an Allais-like behavior. We show that this representation is equivalent to one in which the agent evaluates lotteries using a set of concave rank-dependent utility functionals. A comparative notion of preference for hedging is also introduced.

JEL: D81

Keywords: Allais Paradox, Ellsberg Paradox, Hedging, Multiple Priors, Subjective Mixture, Probability Weighting, Rank Dependent Expected Utility

1 Introduction

In last decades a large number of empirical and theoretical work has been devoted to the study of two 'paradoxes' in individual decision making: the Allais paradox in the case of

*We would like to thank Simone Cerreia-Vioglio, 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 University and Northwestern University 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.

objective probabilities, and the Ellsberg paradox in the case of subjective probabilities. Together, they constitute two of the most widely studied and robust phenomena in experimental economics and psychology of individual behavior. Following this overwhelming evidence, two vast literatures have sprung up. The first focuses on generalizing models of (objective) Expected Utility to accommodate Allais-type behavior, while the second aimed at generalizing models of (subjective) Expected Utility to accommodate Ellsberg-type behavior – or 'ambiguity aversion.'

Despite the existence of large number of theoretical models that addresses each of the two phenomena separately, however, far less attention has been devoted to the study of the *relation* between them. Models that study Allais-like behavior usually only look at objective probabilities, thus not allowing for ambiguity. At the same time, models that study Ellsberg-like behavior either do not consider objective probabilities at all or, if they do, they explicitly assume that the agent follows Expected Utility to assess objective lotteries.[1] In particular, to our knowledge there is no model that studies a decision maker who exhibit both tendencies at the same time in a unified way, at least in standard setups. (See Section 3 for more discussion.) This is surprising given the existence of an least one intuitive connection between these two phenomena: loosely speaking, one might expect a decision maker who is 'pessimistic' about the outcome of risky and uncertain events to display both types of behavior at the same time.[2] This conceptual similarity, moreover, is coupled by a technical one: both phenomena can been as violation of some form of linearity/independence of the preferences.

In this paper we develop axiomatically a model which allows for both Allais and Ellsberg-type behavior in the standard Anscombe and Aumann (1963) setup, and we do so by demonstrating a formal link between these two phenomena. Following the intuition of the characterization of ambiguity aversion as a preference for hedging first suggested Schmeidler (1989), we introduce a generalized notion of hedging that can be applied both to objective risk and subjective uncertainty. We do so by using the concept of 'outcome mixtures' rather than 'probability mixtures' to define hedging. This allows to define a unique axiom, a generalized preference for hedging, which leads to both an Allais and an Ellsberg-type behavior. In particular, we characterize a model in which the decision maker evaluates acts using multiple priors, like in the MaxMin Expected Utility model of Gilboa and Schmeidler (1989), but who will also use some 'multiple priors' to evaluate objective lotteries, distorting the original probabilities – therefore treating objective lotteries as if they were 'ambiguous objects.'

We then show how the same preferences can be alternatively represented by using a distortion of probabilities that follows the Rank Dependent Expected Utility (RDEU) model of Quiggin (1982). In particular, our agent will have a *set* of concave probability weighting

[1] Two exceptions are Klibanoff et al. (2005) and Drapeau and Kupper (2010). In the former, 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.

[2] Such a link has been drawn by Wakker (1990), Chew and Wakker (1996), Wakker (1996), Chateauneuf (1999), and Diecidue and Wakker (2001) in the context of rank-dependent utility, and also more recently by Gumen et al. (2011).

functions, and asses objective lotteries using a RDEU model with the worst probability weighting. In Section 2.6.2 we show how this generalizes the RDEU model with concave probability weighting, allowing us the flexibility to model behaviors which, although in line with intuitive notions of 'pessimism,' could not be captured by RDEU.

1.1 Overview of Main Results

One of the central contributions of the paper is to provide a concept of 'preference for hedging' that explains both Allais-style behavior for objective risk, and Ellsberg-style behavior for (subjective) uncertainty. Indeed the link between the latter and hedging in well-known, following the Uncertainty Aversion axiom introduced in Schmeidler (1989). At the same time, however, this is usually done by defining hedging using *probability mixtures:* the mixture between two acts f and g is generally defined as an act that returns in each state a probabilistic mixture between the outcome obtained of f and g in that state. A preference for hedging then implies that, if an agent is indifferent between any two acts, she should rank the probability mixture between them as at least as good. While an identical notion could be applied directly to objective lotteries, it would not capture our intuition of what is driving Allais-like violations of expected utility: a strict preference for hedging would require a strict preference for the probabilistic mixture of a lottery and its certainty-equivalent over the certainty equivalent itself, contradicting the preference for certainty that is often claimed to be driving Allais style behavior.

We therefore define hedging using the concept of *outcome mixtures*: we follow Ghirardato et al. (2003) to define outcome-mixture of consequences, and then introduce a novel notion of outcome-mixture of lotteries. To illustrate, suppose that we have two outcomes, \$0 and \$10, and that we want to 'mix' them with weight 0.5, in the sense of identifying a prospect whose utility lies exactly halfway between that of the two original points. The standard way to do this is to use a probability mixture, i.e. create a lottery p that returns \$0 or \$10 with equal probability: under Expected Utility, the utility of p is exactly half-way between the utility of \$0 and that of \$10. Instead, here we use the idea of an *outcome* mixture, which identifies some element in the space of outcomes that has a utility exactly in the middle of \$0 and \$10. If we knew that the utility is a linear, we could simply use the outcome exactly in the middle – \$5 in the example above. In general, of course, the utility will not be linear, and in the paper we will adapt the methodology developed by Ghirardato et al. (2003) for the case of ambiguity to identify the correct outcome. It turns out that their approach works naturally in our setup as well, and we shall denote $z = \frac{1}{2}x \oplus \frac{1}{2}y$ the outcome mixture thus obtained.

Once outcome mixtures are defined, the key passage of our analysis is to extend this notion of outcome mixture between two prizes to a notion of outcome mixture between two lotteries. First, consider the mixture between a lottery $p = \sum_{x \in X} p(x)x$ and a degenerate lottery that returns the outcome y with probability 1. We can then define their mixture $\frac{1}{2}p \oplus \frac{1}{2}y$ as the lottery that returns, with the same probabilities as p, a mixture of the outcomes of p with y, using the outcome mixture defined above. Next, we define the mixture between two non-degenerate lotteries. The key observation for this case will be that there will be many possible such mixtures. To illustrate, consider for simplicity the lottery p which returns \$10 and \$0 with probability $\frac{1}{2}$, and let us think about the all the mixtures of p with itself, i.e.

with a second lottery that returns \$10 and \$0 with probability $\frac{1}{2}$. Recall that the idea here is to mix the outcomes that this lottery returns. One way to do it would be to mix \$0 with \$0 and \$10 with \$10: we would obtain exactly the lottery p. However, we could also mix \$0 with \$10, and \$10 with \$0, and we would obtain a lottery that returns $\frac{1}{2}\$10 \oplus \frac{1}{2}\$0$ with probability 1 (a degenerate lottery). All of these lotteries could be seen as mixtures between p and itself. We can define the mixture between any lotteries p and q in the same way, and we denote by $\bigoplus_{p,q}^{\frac{1}{2}}$ the set of all such mixtures. Finally, once we have defined the mixture between two lotteries, we can also define the mixture between two acts, simply point-wise: for any two acts f and g, we define $\bigoplus_{f,g}^{\frac{1}{2}}$ as the set of all acts that could be obtained by mixing in every state of the world the lottery returned by f and the lottery returned by g in that state.

Now that we have outcome mixtures of lotteries, we can introduce our main axiom, Hedging, which, we will argue, generates both Allais-type and Ellsberg-type behavior. 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 these include p itself, but also the degenerate lottery that returns $\frac{1}{2}\$10 \oplus \frac{1}{2}\$0$. The key observation is that, loosely speaking, all of these lotteries are (weakly) more 'concentrated towards the mean' (in utility terms) than p, while preserving expected utility. In the extreme case, this gives us the degenerate lottery $\frac{1}{2}\$10 \oplus \frac{1}{2}\$0$. A 'pessimistic' agent should like this reduction of the variance, this 'pulling towards the mean' – and should exhibit a preference for hedging. If we extend this idea to acts, moreover, then hedging acquires also the additional advantage of mixing the outcomes in each state, just like in the case of Uncertainty Aversion in Gilboa and Schmeidler (1989) – a reduction of subjective uncertainty which pessimistic subjects should also like. Thus, our axiom will say that if we have two acts that are indifferent to each other, then any third act that could be obtained as a hedging between the first two should ranked as at last as good as them.

The main theorem of the paper will then show that, in a standard Anscombe-Aumann setup, this axioms, together with other standard ones,[3] is equivalent to the existence of a representation for the preferences of the decision maker with the following components 1) a utility function u on the prizes, 2) a convex and compact set of probability measures Π on the finite set of states of the world Ω, and 3) a convex and compact set of Borel probability measures Φ on $[0,1]$, which contain the Lebesgue measure ℓ.[4] Our agent evaluates acts in a way very similar to the MaxMin Expected Utility (MMEU) model of Gilboa and Schmeidler (1989): she calculates the utility of each act by combining the utility that what this act returns in each state using the *worst* of the priors in Π. That is, $\succeq$ is represented by

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

where $U : \Delta(X) \to \mathbb{R}$ represents her evaluation of lotteries in $\Delta(X)$ – precisely as in the

[3]In particular, we posit Continuity; (standard) Monotonicity; First Order Stochastic Dominance; one axiom that guarantee that the notion of $\oplus$ is well defined; one axiom with the same idea of the Certainty-independence axiom of Gilboa and Schmeidler (1989), but where mixtures defined using outcome mixtures instead of probability mixtures, and using degenerate lotteries instead of constant acts.

[4]Moreover, every $\phi \in \Phi$ is atomless and mutually absolutely continuous with respect to ℓ.

MMEU model. Where our model differs, however, is on how this U is computed for each lottery. In particular, 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,' and in which each ball is associated to a certain prize preserving the correct probabilities. (We call this a *measure-preserving map* γ.) Put another way, our decision maker acts as if the lottery will be actualized by the drawing of a ball from an urn, with the the fraction of balls giving a particular prize equal to the probability of that prize under p. (For example, when she faces the lottery $p = \frac{1}{2}x + \frac{1}{2}y$, she 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.) It is then easy to see that an expected utility agent would use the uniform prior on $[0, 1]$ to evaluate such acts – the Lebesgue measure ℓ. By contrast, we capture the notion of pessimism by considering an agent who has a *set* of priors on $[0, 1]$, which contains ℓ, and who evaluates lotteries using the *worst* one of them. That is, we have

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

This representation as an intuition very similar to MMEU: it amounts to saying that agent treats each objective lottery as an act, and she is ambiguity averse in her dealing with it. Like in MMEU, then, this ambiguity aversion is captured by the multiplicity of priors. We then show how this representation is fully compatible with the Allais paradox: intuitively, degenerate lotteries cannot be distorted, as any prior on $[0, 1]$ must still integrate to 1; on the other hand non-degenerate ones can, and they will receive an utility always (weakly) lower than they would have under Expected Utility, because ℓ must belong to the set of priors, so if the representation uses a different one it must be because it assigns a lower utility.

One feature of the representation above is that the decision maker might distort objective probabilities when she evaluates lotteries. However, she does so following a procedure which is rather different from others suggested in the literature: instead of distorting the distribution of the lottery, she maps it into an imaginary act on the space $[0, 1]$ and uses distorted (i.e. non-uniform) probabilities over this state space. A natural question is then to ask if there existed an alternative, and equivalent representation in which the decision maker followed the more traditional path of distorting the commutative distribution of the lottery. This is the procedure followed by one of the most well-known model used to study violations of Expected Utility on objective lotteries: the Rank Dependent Expected Utility Model (henceforth RDEU) of Quiggin (1982). (See Wakker (1994), Nakamura (1995), Chateauneuf (1999), Starmer (2000), Abdellaoui (2002), Kobberling and Wakker (2003) and the many references therein.) According to this model, the agent uses a rule similar to Expected Utility, but applies a weighting function on the *cumulative* distribution of each lottery. Depending on the shape of this function, the behavior of the agent can be either in line with certainty bias a' la Allais (concave weighting), or the opposite (convex weighting). Our main theorem will then show that we can, in fact, find a representation of this kind: we will show that the representation above (and the axioms) are equivalent to one in which the agent has a utility function and a set of priors over Π , but also a *set* of *concave* weighting functions on the cumulative distributions. She then ranks acts using the worst of the priors in Π, just like the previous representation, but she evaluates objective lotteries using the

RDEU representation with the *worst* of the weighting functions in the set. It turns out that our model is a strict generalization of the RDEU one with a concave distortion; we refer to Section 3 for more discussion.

Finally, we note that the comparative notion ambiguity aversion of Ghirardato and Marinacci (2002) applies naturally in this setup, and leads to a the following representation: an agent who is more attracted to certainty (in the sense of their definition) admits a representation with a weakly larger set of priors Π on Ω, and also a weakly larger set of priors Φ on $[0, 1]$.

The remainder of the paper is organized as follows. Section 2 presents the formal setup, the axioms, and the representation theorem. Section 3 discusses the relevant literature. Section 4 concludes. The proofs appear in the appendix.

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.[5] 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 will also 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 $(1 - \alpha)$. Moreover, 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 $\mathcal{F}$ such that $p(\omega) = p$ for all $\omega \in \Omega$. Finally, given some $p, q \in \Delta(X)$ and some $E \in \Sigma$, xEy denotes the acts that yield outcome x if E is realized, and y otherwise.

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

[5]It is standard practice to generalize our analysis to the more general case in which X is a connected and compact topological space with topology τ. 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, specifically the discussion after Axiom 3.

2.2 Axioms and Subjective and Objective Mixtures

2.2.1 Basic Axioms

We start with imposing some basic axioms on our preference relation. First of all, we postulate that it respects First Order Stochastic Dominance (FOSD) on *objective* lotteries. In particular, we use the following standard definition.[6]

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 then impose that our preference relation respects FOSD.

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 $p \succ q$.*

We next impose 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 g.

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

Next, since we are after a representation theorem we need to posit a form of continuity of the preferences.[7] [8]

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 $f \in \mathcal{F}$.*

In the standard development of subjective Expected Utility theory, to the axioms above one would add the Independence axiom of Anscombe and Aumann (1963).[9] This axiom,

[6]Indeed, since 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 we had $\delta_x \succeq \delta_y$ iff $x \geq y$ for all $x, y \in X$.

[7]We should emphasize that, although the following axiom is entirely standard, it stronger than the Archemedean Continuity usually 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.

[8]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)).

[9]The Independence axioms posits that for every $f, g, h \in \mathcal{F}$, and for every $\alpha \in [0,1]$

$$f \succeq g \Leftrightarrow \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h.$$

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 an Ellsberg-type behavior, in the literature on ambiguity aversion, an alternative much weaker axiom has been suggested: Risk Independence, which postulates independence only for objective lotteries.[10] 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, and from the use of probability mixtures, and rather use the alternative notion of *outcome* mixtures as we shall define below.

Before we proceed, however, we posit a much more basic form of coherence of the preferences, which will not rule out Allais-like behavior while at the same time guaranteeing that minimum amount of coherence that we will need in our analysis. Consider some x, y, z', z'' such that z' and z'' are "in between" x and y. Then, consider the following two lotteries:

$$\frac{1}{2}ce_{\frac{1}{2}x+\frac{1}{2}z'} + \frac{1}{2}ce_{\frac{1}{2}y+\frac{1}{2}z''}$$

and

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

The only difference between them is that: in the former x is mixed with z', and y with z'', and 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, however, the only weights involved are weights $\frac{1}{2}$, and x is always mixed with some element worst than it, while y is mixed with some element better than it. The only difference is then in the 'order' of this mixture. Our axiom then imposes that these two lotteries should be indifferent for the agent, who 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 $ce_{\frac{1}{2}r+\frac{1}{2}s}$ exists for $r = x, y$ and $s = z', z''$. Then, we have*

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

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 weaker than it. For example, it is easy to see that it is compatible with the

[10] The Risk Independence Axiom posits that for every $f, g, h \in \mathcal{F}$, and for every $\alpha \in [0, 1]$

$$f \succeq g \;\Leftrightarrow \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h.$$

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 forms of probability weighting.

Axioms of this form are not uncommon in the literature: in particular, this is an adaptation of the E-Substitution Axiom in Ghirardato et al. (2001) for the case of objective lotteries. In fact, following this literature it is straightforward to show how Axioms 1-4 are enough to guarantee that the representation hinted to in the previous paragraph 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 there exists a cardinally unique 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,* $\gamma(\frac{1}{2})u(x) + (1-\gamma(\frac{1}{2}))u(y) \geq \gamma(\frac{1}{2})u(z) + (1-\gamma(\frac{1}{2}))u(w)$.

2.2.2 Probability-Mixtures and Outcomes-Mixtures

We now turn to discuss one of the central the notions of our analysis: *outcome-mixtures* of lotteries. The intuition is the following. To be able to define behaviorally a notion of aversion to 'exposure to risk,' it will be useful to define a notion of hedging between lotteries: a 'mixture' between what the two lotteries return constructed in such a way that the utility of this mixture is 'in between' that of original lotteries. The traditional approach to do this 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 lottery. (We shall refer to these mixtures as *probability-mixtures*.) Unfortunately, however, this approach would not work for our analysis: the mixture is created by changing the probabilities, and our agent might not be neutral to it. 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.

Instead of probability mixtures, in this paper we will use the alternative notion of *outcome-mixtures* of lotteries. We will first of all follow the literature and define the simpler notion of outcome-mixture for consequences;[11] then, we extend this idea to a mixture over lotteries. To illustrate, consider first two outcomes $x, y \in X$. A probability-mixture with weight α between them is the lottery that returns x with probability α and y with probability $(1-\alpha)$. Under Expected Utility, such probability mixture has a utility which is exactly half-way between that of x and y. By contrast, an outcome-mixture between x and y is an *outcome* z in X that has has a utility which is exactly in the middle between that of x and y. The whole point now is to understand: 1) if such z exists, and 2) how to identify it. Indeed if we knew that the utility function of the agent were linear on $X \subseteq \mathbb{R}$, we could

[11]Similar approaches to define mixtures of consequences were used in Wakker (1994), Kobberling and Wakker (2003), Ghirardato et al. (2001, 2003), and in the many references therein. More precisely, in what follows we adapt the approach Ghirardato et al. (2003) to the case of objective probabilities.

simply take the element $\frac{1}{2}x + \frac{1}{2}y$. But of course this is in general not true. However, if the set of consequences X is connected, then Ghirardato et al. (2003) propose a technique which allows us to elicit this element anyway. In what follows we adapt their technique, originally developed for Savage acts, to the case 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 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 mixture between x and y, or first taking the mixture between x and z, and then the mixture between z and y – which would hold if z had a utility exactly half-way between that of x and y. This is even more explicit if we consider the representation in Lemma 1: it is immediate to see that we must have that we have that $u(z) = \frac{1}{2}u(x) + \frac{1}{2}u(z)$ if and only if we had $\delta_x \succeq \delta_z \succeq \delta_y$ and (1) holds. Thanks for our structural assumption, this notion is well-defined: since X is a connected set and u is continuous, this means that, under Axioms 1-4, 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, except in Section 2.8, 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.[12]

Definition 2 defines the outcome-mixture between two consequences in X. One of the key (and, to our knowledge, novel) contributions of our paper is to extend this notion to mixture of *lotteries* and of *acts*. Let us start from the simplest case: two degenerate lotteries. Indeed 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 will be the degenerate lottery $\delta_{\lambda x+(1-\lambda)y}$. We can similarly define the notion of mixture between a degenerate lottery δ_x and a generic lottery $p \in \Delta(X)$: replace what p returns with the $\oplus$-mixture with x, keeping the probabilities constant. That is, we have that, for ant $p \in \Delta(X)$ $y \in X$ and α :

$$\alpha(\sum p(x_i)\delta_{x_i}) \oplus (1-\alpha)\delta_y = \sum p(x_i)\delta_{\alpha x_i \oplus (1-\alpha)y}.$$

Slightly more complicated, however, is to define outcome-mixture of two lotteries, mainly because there are many possible ways to do it. To see why, consider two lotteries $p = \frac{1}{2}x + \frac{1}{2}y$

[12] 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 axioms structure; we choose to use the most restrictive definition to state the axioms in the weakest form we are aware of.

and $q = \frac{1}{2}z + \frac{1}{2}w$. How can define a mixture between p and q? 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 and q as $p = \frac{1}{4}x + \frac{1}{4}x + \frac{1}{4}y + \frac{1}{4}y$ and $p = \frac{1}{4}z + \frac{1}{4}z + \frac{1}{4}w + \frac{1}{4}w$, and combine them in yet other ways. Or, decompose them in different ways, to find yet many other combinations. All of this shows that there is a large number of ways to combine these lotteries, and for any two lotteries $p, q \in \Delta(X)$ we denote by $\bigoplus_{p,q}^{\frac{1}{2}}$ the set of all such mixtures between p and q and weight $\frac{1}{2}$. 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 $x_1, \ldots x_n, y_1, \ldots, y_n \in X$, and some $\gamma_1, \ldots, \gamma_n \in [0,1]$ such that $p = \sum_{i=1}^n \gamma_i x_i$, $q = \sum_{i=1}^n \gamma_i y_i$.[13] Then, the set $\bigoplus_{p,q}^{\frac{1}{2}}$ will be the set of all combinations r such that $r = \sum_{i=1}^n \gamma_i(\frac{1}{2}x_i \oplus \frac{1}{2}y_i)$.[14]

Alternatively, we could interpret the set $\bigoplus_{p,q}^{\frac{1}{2}}$ as follows.[15] 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 variable we need to know their *covariance*. The set $\bigoplus_{p,q}^{\frac{1}{2}}$ could then be seen as the set of all $\frac{1}{2}$-mixtures of the random variable corresponding to p and q for *any* possible covariance. The multiplicity of mixtures could then be seen as stemming from the multiple possible covariances that could be found.

Finally, we define the notion of outcome-mixture for acts. As standard, we do it point-wise: an act h is a $\frac{1}{2}$-mixture between two acts f and g if, and only if, $h(\omega) \in \bigoplus_{f(\omega),g(\omega)}^{\frac{1}{2}}$ for all $\omega \in \Omega$. We denote $\bigoplus_{f,g}^{\frac{1}{2}}$ the set of all outcome-mixtures of two acts.

2.2.3 Main Axioms

Now that we are endowed with the notions of outcome mixtures, we can use them to define the two central axioms of the paper. We start from a form of positive homogeneity and translation invariance much in the spirit of the Certainty-Independence axiom of Gilboa and Schmeidler (1989):[16] 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

[13]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 $q = \frac{1}{3}z + \frac{1}{6}w + \frac{1}{6}w + \frac{1}{3}w$.

[14]That is, we have $\bigoplus_{p,q}^{\frac{1}{2}} := \{r \in \Delta(X) : \exists x_1, \ldots x_n, y_1, \ldots, y_n \in X, \exists \gamma_1, \ldots, \gamma_n \in [0,1]$ such that $p = \sum_{i=1}^n \gamma_i x_i$, $q = \sum_{i=1}^n \gamma_i y_i$, and $r = \sum_{i=1}^n \gamma_i(\frac{1}{2}x_i \oplus \frac{1}{2}y_i)\}$.

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

[16]Recall that 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 $\lambda f + (1-\lambda)p \succeq \lambda g + (1-\lambda)p$.

ones, because our agent could have non-linear reactions to probability mixtures.[17]

A.5 (Degenerate-Independence (DI)). *For any $f, g \in \mathcal{F}$, dyadic $\alpha \in (0,1)$, and for any $x \in X$,*

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

Finally, we posit our last and most important axiom: *preference for hedging.* Let us develop the intuition for it with a simple example. Consider some lottery $p = \frac{1}{2}x + \frac{1}{2}y$, and the lottery r that could be obtained by mixing p with itself, i.e. $r \in \bigoplus_{p,p}^{\frac{1}{2}}$. We would like to argue that any 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. On the one hand, we could have that x is mixed with x, and y is mixed with y, generating $r = p$ – so r is at least as good as p. On the other hand, 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* centered exactly at the prize 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: by the process of hedging, the agent could mix some good and some bad outcomes, 'pulling them towards the center,' reducing the exposure to risk and the variability that she is subject to (in terms of utility). This means that an agent who is attracted towards such reduction should be attracted towards 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, we would like to argue that an agent who is 'averse to exposure to risk,' should like r as at least as much as p and q. In particular, r can formed by 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. However, 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 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* – thus far we have only discussed hedging between lotteries (degenerate acts). For simplicity, consider now two non-degenerate

[17] It is not hard to see that this axiom is actually *striclty 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.

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 $h \in \bigoplus^{\frac{1}{2}}_{f,g}$, and notice that we must have $h(\omega) = \delta_{\frac{1}{2}f(\omega)\oplus\frac{1}{2}g(\omega)}$. Since there are no lotteries involved, going from f and g to h does not reduce the exposure to *risk*, as the hedging above did. At the same time, however, it will reduce the exposure to *ambiguity* – this is identical in spirit to the hedging axiom of Gilboa and Schmeidler (1989).[18] 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.6 (Hedging). *For any $f, g \in \mathcal{F}$), and for any $h \in \bigoplus^{\frac{1}{2}}_{f,g}$, if $f \sim g$, then $h \succeq f$.*

2.3 First Representation: subjective view of objective risk

We are now ready to introduce our first and main representation. To better express it, however, it will be useful to define the notion of a *measure-preserving* map from lotteries into acts on $[0, 1]$. The idea is simple: for every objective lottery $p \in \Delta(X)$, we can *map* it 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.[19] 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: the lottery above could just as easily be mapped to the act that returns y to the states $[0, \frac{1}{2})$, and x to the states $[\frac{1}{2}, 1]$. Of all the possible maps, however, we focus only on those in which each lottery in $\Delta(X)$ 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.

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

[18] In fact, if, additionally, we apply risk independence, then preference for hedging in outcome mixtures is identical to preference for hedging in probabilities.

[19] A visualization which is rather close to being true in most experimental settings.

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

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

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 (continuous) utility function u; a set of priors Π over the states in Ω, which is convex and (weak) compact; and a set of (Borel) priors Φ over $[0,1]$, which is convex and (weak) compact, contains the Lebesgue measure ℓ, and contains only priors which are atom less and mutually absolutely continuous w.r.t. it. She then ranks acts as follows. 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 each of the (objective) lotteries that the act return in every state.

Where the model above differs from MMEU is in how the evaluation on objective lotteries is done. Our agent, in fact, need not follow vNM Expected Utility to compute it. Rather, first she maps each objective lottery into an act on the space $[0,1]$ (in a measure-preserving fashion). It is easy to see that if she followed Expected Utility, she would then evaluates each of these imaginary acts using the Lebesgue measure ℓ. Instead, our agent has a *set* of priors Φ over $[0,1]$, which includes ℓ, and she uses the *worst* one of them to compute her utility – much in line with the MMEU model. This will lead her to give a (weakly) lower evaluation to each non-degenerate lottery, while keeping her evaluation of degenerate ones the same. In fact, her behavior will coincide with Expected Utility only in the special case in which Φ is a singleton – which, by construction, means that it contains only ℓ.

Let us illustrate this intuition by means of a simple 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 – in other words, out agent acts as if she were *ambiguity averse* over these balls. Thus, the lottery $p = \frac{1}{2}\$10 + \frac{1}{2}\$0$ might be 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, thinking that she is 'unlucky' and that the 'good balls' will not come out, while at the same time raising the probabilities associated with balls 51-100, the ones associated with the 'bad' outcome.

This main feature of the model above, which sets is apart from the rest of the literature on ambiguity, is that our agent treats objective probabilities like 'ambiguous objects:' it is as if she mapped each lottery into an acts, and she were *ambiguity averse* over such acts, following a procedure essentially identical to the MMEU model to evaluate them. That is, our decision maker is, in some sense, ambiguity averse towards objective lotteries, where the degree of her aversion is given by the size of the set Φ. In turns, it is not hard to see how, whenever Π and Φ are not singletons, our decision maker will have an 'attraction' towards certainty, and towards degenerate lotteries in particular. In fact, she discounts both uncertainty and risk by using the worst prior in Π and, at the same time, the worst prior in Φ. The objects which we can be sure are left unaffected are degenerate lotteries: we must have that $V(\delta_x) = U(\delta_x) = u(x)$ by construction. As we will see in Section 2.6.1, this will allow our model to represent both the Ellsberg and the Allais paradoxes.

2.4 Second representation: minimal of RDEU

The Multiple Priors Multiple Distortions representation discussed in the previous paragraph describes an agent who distorts objective probabilities by mapping lotteries into acts, and treating such acts as ambiguous objects that she evaluates using multiple priors. This form of distortion of probabilities seems very different from other established way to distort objective probabilities, such as the ones part of the Rank-Dependent Expected Utility (RDEU) representation, in which a 'weighting function' is applied to the cumulative distribution of the objective lottery. One could therefore ask if an alternative representation could be found, in which our decision maker follows a more standard procedure to distort probabilities.

We start from recall 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 $RDEU_{u,\psi}$, if, for any enumeration of the elements of the support of p such that $x_{i-1} \preceq x_i$ for $i = 2, \ldots, |\text{supp}(p)|$, we have

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

The main feature of the RDEU model is that the decision maker follows a procedure similar to expected utility, except that she uses distorted probabilities. In particular, the distortion comes from a probability weighting function which modifies the way the agent treats the *cumulative* probability distribution of the lottery at hand. 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. Moreover, when the probability weighing is linear, RDEU coincides with Expected Utility. Depending on the shape of ϕ, the model allows both for an attraction towards certainty: when ϕ is concave it leads to an Allais-like behavior; the opposite takes place when ϕ is convex. (We refer to Quiggin (1982), and also Wakker (1994), Starmer (2000), 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 can then write a representation similar to our previous one, but in which the Decision Maker uses an RDEU functional to distort probabilities. Indeed since we have an model of Allais-like behavior, such functional will be concave, 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

$$U(p) := \min_{\psi \in \Psi}\ RDEU_{u,\psi}(p).$$

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 following the MMEU model. Here, however, instead of having a set of priors over $[0, 1]$, she has a *set* of probability weights, and she will use the worst one of those in a RDEU functional to evaluate objective risk. That is, our agent treats objective lotteries using the *min* of a set of RDEU distortions. Moreover, notice two features of this set. 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 2.6.2 we discuss a comparison with RDEU more in detail.)

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 and Φ are unique.

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, the components of which are identified uniquely. 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.

2.6 Discussion

2.6.1 Our representations and the Allais Paradox

We now turn to describe how our model could generate certainty bias and an Allais-like behavior. Intuitively, it is easy to see how both derive from the our main axiom, Axiom 6 (Hedging): for any $x, y, z \in X$ such that $u(z) = \frac{1}{2}u(x) + \frac{1}{2}u(y)$, we must have that $\delta_z \succeq \frac{1}{2}x + \frac{1}{2}y$, leading to (weak) certainty bias. In turns, as we have seen our first representation has a similar feature: while the evaluation of degenerate lotteries is not distorted (i.e. we must have $V(\delta_x) = U(\delta_x) = u(x)$), that of non-degenerate ones could be reduced when the decision maker uses some prior ϕ which returns a lower amount than the Lebesgue measure ℓ. An identical argument applies also to our second representation.

Similarly, it is easy to see how the choice pattern of the Allais experiment could be accommodated by our model: 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 it. Importantly, moreover, we will now show how our model not only is able to accommodate for Allais-like behavior, but that it rules 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 lotteries like the ones above but in which x and y were equal to each other and set to \$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.

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 $u(p_3) \leq \min(0.11, \alpha + \beta)$.[20] 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.

[20]Notice that we know that $u(p_3) \leq 0.11$ since Φ contains the Lebesgue measure.

2.6.2 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, and it can be seen as one 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, moreover, then the RDEU model with concave distortions becomes a special case of ours.

There are however, 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 in 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.

On the other hand, however, once we focus on concave probability weightings, while in RDEU the agent has a fixed distortion to be used for every lottery, in our representation she could have multiple ones, and use a different one depending on the lottery at hand. Importantly, we would argue that this additional flexibility could prove to be desirable from a modeling prospective. Let us illustrate this with an example. 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 lotteries 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.[21] 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, while 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.

A natural question is then to identify the conditions which guarantee that our set of distortions Ψ is a singleton – the special case of our model which coincides with concave RDEU on objective lotteries. It turns out that to characterize such special case all we need is an axiom that guarantees that the preferences of our agent must be of the RDEU form for objective lotteries. This means that we could simply impose the Probability trade-off consistency Axiom of Abdellaoui (2002), or both the Comonotonic Sure-Thing Principle and the Comonotonic Mixture Independence axioms of Chateauneuf (1999), as either of these axioms, as proved in the respective papers, together with our monotonicity and continuity imply 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 that we can represent using a unique *concave* RDEU functional.[22]

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

[22] In turns, this implies that our axioms together, and Hedging in particular, must imply both the Attraction

2.6.3 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 *fixed*, specific map, and had a set of priors which would work for this specific map only. Indeed, if a Multiple Priors Multiple Distortions representation exists, so must this alternative representation – it simply posits a weaker requirement. At the same time, however, once we focus on a specific map, we might be able to derive additional properties on the set of priors Φ on $[0,1]$. We will now argue that this is indeed the case.

Let us focus a specific map $\bar{\gamma}$ from $\Delta(X)$ into acts on $[0,1]$, the one that goes '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 $i = 2, \ldots, |\text{supp}(p)|$. Now define $\bar{\gamma}(p)$ as $\bar{\gamma}(p)\big([0, p(x_1)\big) = x_1$ and $\gamma(p)\big([\sum^{i-1} {}_{j=1} p(x_i), \sum_{j=1}^{i} p(x_i))\big) = x_i$ 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.

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 a fixed map $\bar{\gamma}$ 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. Too 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 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.

Observation 1 shows that if we focus on a specific map, the one from 'worst to best,' then not only we could find a set of priors which work for that specific map, which is trivial, but in particular every prior in this set must be 'decreasing,' i.e. assign a higher weight to 'earlier' states: in particular, it must overweight 'early' states, and underweight 'late' ones. And since for this map early states are the ones to which the worst outcomes are assigned, while the good ones are assigned to late, underweighted states, then we have a representation in which each of the priors used by the agent, each of her distortions, must be 'pessimistic.'

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 standard Anscombe-Aumann Expected Utility on acts, and of vNM Expected Utility on objective lotteries. We now turn to analyze

for Certainty Axiom of Chateauneuf (1999), as well as that our preferences on objective lotteries exhibits Probabilistic Risk Aversion as defined in Abdellaoui (2002), since both are implied by the existence of a *concave* RDEU representation.

which would be 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. 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 it 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 $r \sim p$.*

The consequences of these axioms, in addition to our previous ones, appear in the following proposition. (The proof follows standard arguments and it is therefore omitted.)

Proposition 1. *Consider a 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) $|\Phi| = 1$ and $\Phi = \{\ell\}$ if, and only if, $\succeq$ satisfies Axiom 9 (Hedging Neutrality on Lotteries), which, in turns, holds if and only if $\succeq$ satisfies Risk Independence.

(c) $|\Pi| = |\Phi| = 1$, and $\Phi = \{\ell\}$ if and only if $\succeq$ satisfies Axiom 7 (Hedging Neutrality).

2.8 A comparative notion of attraction towards certainty

We now introduce a comparative notion of *attraction towards certainty.* We take an approach which follows the notions of comparative ambiguity aversion in Epstein (1999) and Ghirardato and Marinacci (2002), extend it to our more general setup, and show that the same intuition of attraction towards certainty implies both more ambiguity aversion and more probability distortions.

Ghirardato and Marinacci (2002) present the following notion. Consider two decision maker, 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. The approach of Ghirardato and Marinacci (2002) focuses on this second case – look at the relative attraction towards certainty while keeping constant the curvature of the utility function. It turns out that in our setup imposing that the utility function has the same curvature translates in a very simple requirement: we want $\oplus_1 = \oplus_2$, that is, both agents should have the same approach to outcome mixtures – which, as our derivation of $\oplus$ argues, implies that they have the same curvature of the utility function. With this in mind, we can then use the following definition introduced by (Ghirardato and Marinacci, 2002, Definition 7):[23]

Definition 7. 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
$$\delta_x \succ_1 f \;\Rightarrow\; \delta_x \succ_2 f.$$

It turns out that this definition, which was developed for a standard setup in which agent's preferences are vNM on objective lotteries, works in out setup as well.

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 $\Phi_2 \supseteq \Phi_1$.*

That is, if we consider two agents who have the same outcome-mixtures, derived from $\oplus$, and such that 2 is more attracted to certainty than 1, then *both* set of priors Π and Φ of 1 are smaller than those of 2.

3 Overview of the related literature

A large literature has been devoted to developing models that allow for Allais or Ellsberg-type behavior. To the best of our knowledge, however, almost no models in the literature allow for *both* features at the same time in standard setups such as that of Anscombe and Aumann (1963). On the one hand, most models meant to study Allais-like behavior do not consider the presence of subjective probabilities at all, thus ruling out Ellsberg-type behavior. On

[23]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.

the other hand, models meant to capture ambiguity aversion either do not consider objective lotteries at all, operating in the setup of Savage; 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.[24] 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 essentially all the weakening of independence suggested in the literature for this setup.[25] (See Gilboa and Marinacci (2011) for a survey.) 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 obtaining a representation that coincides with MaxMin Expected Utility 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 an 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).[26] As we shall see in our discussion of the latter, however, while the latter allows 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.

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, 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). First of all, our work is conceptually very closely related to that of Maccheroni (2002) and Cerreia-Vioglio (2010): both papers provide representation in which the decision maker treats objective lotteries as 'ambiguous objects,' as we do, and find representation reminiscent of the ones developed to study ambiguity aversion.

[24]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.

[25]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).

[26]More precisely, since Drapeau and Kupper (2010) studies a preorder which corresponds to the risk perception instead of standard 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*.

Neither, however, studies Ellsberg-like behavior – they both work in the setup of vNM. Even focusing in the case of risk, 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,[27] 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.[28] 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). This conceptual difference in the representation 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.'[29] 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.

We have already discussed (Section 2.6.2) 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. Our results are also related to those of Dillenberger (2010). Dillenberger (2010) studies two properties: Negative Certainty Independence (NCI), and PORU, which is preference for one-shot resolution of uncertainty. He shows that they are equivalent under some basic assumptions. 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[30]: see, among others, Chew (1983),

[27] 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 $\alpha p + (1-\alpha)q \succeq r$.

[28] 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 $\alpha \in (0, 1)$.

[29] 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.

[30] 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 $\alpha \in [0, 1]$.

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.

Although there are few models that allows for both the Allais and the Ellsberg paradox to be present at the same time in the standard setup of Anscombe and Aumann (1963), the idea of a connection between these behaviors is not new. Previous authors have noted that 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. (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. 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. There are, however, some conceptual differences between our approach and the one followed by the literature. The first, most important difference is that these papers identify a connection between the functional forms of RDEU and Choquet EU, without discussing common restriction about their approach to risk or ambiguity; people could be 'pessimistic' or 'optimistic' in both domains, or indeed pessimistic on one domain and optimistic in another. By contrast, one of the goals of our paper is to identify a unique axiom which implies a forms of pessimism in both aspects. Put differently, while we argue that Hedging implies both Allais and Ellsberg-like behavior, this literature shows that comonotonic independence implies a similar representation in both dimensions, although without restrictions as to whether either Allais or Ellsberg-like behavior would emerge. And indeed one cannot impose Schmeidler (1989)'s hedging axiom to both environments to derive pessimism, because, as we've discussed, such axiom would not imply pessimism for risk.[31] By contrast, one could do this precisely using the axiom that we introduce, which generalizes Schmeidler (1989)'s intuition to capture pessimism on objective risk as well. A secondary difference is that we derive a model with the *simultaneous* presence of Ellsberg and Allais-like behavior, instead of focusing on a specific one of these. This possible because our approach applies to the standard setup of Anscombe and Aumann (1963), where both features could be present at the same time. 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.[32]

[31]We should emphasize, moreover, that this axiom could in general not even be defined in the environment in which most of this literature operates, as it does not allow the notion of mixture required to define the axiom.

[32]Most of this literature studies a setup in which we have a given set of states of world with an objective

Segal (1987) and Segal (1990) suggest a different channel to connect the Allais and Ellsberg paradox: he argues that 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, while in 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). As opposed to our analysis, however, his approach is based on a setup much richer than the standard one of Anscombe and Aumann (1963) – to be able to discuss reduction of compound lotteries, two-stage lotteries need to be observable.

Finally, 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. 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 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.

4 Conclusion

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

While our model does not require an agent who exhibits Allais-type behavior to also exhibit Ellsberg-type behavior, or *vice versa*, intuitively our result on the existence of a similar channel to capture both tendencies would seem to make this more likely to be the case. In this light, we highlight a recent experimental result in Dean and Ortoleva (2011), which tests the existence of an *empirical* relation between these behaviors. Preliminary results suggest not only the significant presence of each individual bias, but also the presence of a

probability distributions – a setup where it is much easier to posit comonotonic independence.

significant, although small, positive relationship between the two. We interpret this as further evidence of the fact that not only these two phenomena exists and could be economically relevant, at least as much as could be suggested by an experiment on undergraduate students, but also that they are related to each other.

A Appendix A: Proofs

A.1 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 $f', g', \dots$ to denote generic elements of this space.[33] 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 $(\alpha f' \oplus (1-\alpha)g')(\omega') = \alpha f'(\omega') \oplus (1-\alpha)g'(\omega')$ 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 $f \in \mathcal{F}$.)

[33] 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.

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. $\square$

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).

$\square$

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. $\square$

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. □

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. □

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. □

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 $\succeq'$.[34]

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. □

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[35] $B \in \Sigma_A$, $x, y, z \in X$ 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 $\frac{1}{2}x + \frac{1}{2}z \rhd_{FOSD} \frac{1}{2}y + \frac{1}{2}z$, 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

[34]We recall that an event E is essential if we have $x \succ' xAy \succ' y$ for some $x, y \in X$.

[35]An event is *universal* if $y \sim xAy$ for all $x, y \in X$ such that $x \succ y$.

$zAx \in \gamma(\frac{1}{2}x + \frac{1}{2}z)$. 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. $\square$

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

$$ce'_{xBz'} C ce'_{z''By} \sim' ce'_{xCz''} B ce'_{z'Cy}.$$

Proof. Consider first the case in which $B = \emptyset$. In this case, the claim becomes $ce'_{z'} C ce'_y \sim' ce'_{z'Cy}$, which is trivially true. The case $C = \emptyset$ is analogous. Now consider the case $B = \Omega'$. The claim becomes $ce'_x C ce'_{z''} \sim' ce'_{xCz''}$ 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 $ce'_{xAz'} A^C ce'_{z''Ay} \sim' ce'_{xA^C z''} A ce'_{z'^C y}$, which is equivalent to $ce'_{z''Ay} A ce'_{xAz'} \sim' ce'_{z''Ax} A ce'_{yAz'}$. Now notice that since $ce'_{xAy} \in X$ for all $x, y \in X$, by claim 5, we have that $\gamma^{-1}(ce'_{z''Ay} A ce'_{xAz'}) = \frac{1}{2} ce'_{z''Ay} + \frac{1}{2} ce'_{xAz'}$. 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 $\delta_{ce'_{rAs}} \sim \delta_{ce_{\frac{1}{2}z + \frac{1}{2}s}}$. Since this is true for all $r, s \in X$, then by Axiom 1 (FOSD) we must have $\frac{1}{2} ce_{\frac{1}{2}z'' + \frac{1}{2}y} + \frac{1}{2} ce_{\frac{1}{2}x + \frac{1}{2}z'} \sim \frac{1}{2} ce'_{z''Ay} + \frac{1}{2} ce'_{xAz'}$, hence $\gamma^{-1}(ce'_{z''Ay} A ce'_{xAz'}) \sim \frac{1}{2} ce_{\frac{1}{2}z'' + \frac{1}{2}y} + \frac{1}{2} ce_{\frac{1}{2}x + \frac{1}{2}z'}$. By analogous arguments, we must have $\gamma^{-1}(ce'_{z''Ax} A ce'_{yAz'}) \sim \frac{1}{2} ce_{\frac{1}{2}z'' + \frac{1}{2}x} + \frac{1}{2} ce_{\frac{1}{2}y + \frac{1}{2}z'}$. At the same time, Axiom 4 we must have $ce_{\frac{1}{2}z'' + \frac{1}{2}y} + \frac{1}{2} ce_{\frac{1}{2}x + \frac{1}{2}z'} \sim \frac{1}{2} ce_{\frac{1}{2}z'' + \frac{1}{2}x} + \frac{1}{2} ce_{\frac{1}{2}y + \frac{1}{2}z'}$, which by transitivity implies $\gamma^{-1}(ce'_{z''Ay} A ce'_{xAz'}) \sim \gamma^{-1}(ce'_{z''Ax} A ce'_{yAz'})$, hence $ce'_{z''Ay} A ce'_{xAz'} \sim ce'_{z''Ax} A ce'_{yAz'}$ 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 5, 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. □

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 6 (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. □

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).) □

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 $\mathcal{F}^* := \{f' \in \mathcal{F}' : f'(\omega, \cdot) = f'(\omega', \cdot)$ 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 $[0, 1]$. □

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 $\{x_1, \ldots, x_{|\text{supp}(p)|}\}$, the restriction of $\succeq$ to $\Delta(X)$ is represented by the functional

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

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. □

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^* $E_1, \ldots, E_n$ with $E_1 \subseteq E_2 \subseteq \ldots$ 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 $\ell(E) = 0$.[36] 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.

[36]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 $\phi \in \Phi$.

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: $\hat{\mathcal{F}} := \{f' \in \mathcal{F}' : f'(\omega, [0,1]) = x$ 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 $\bar{f}(\omega) := \delta_{ce_{f(\omega)}}$. 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 $\gamma(\bar{f})(\omega, [0,1]) = \delta_{ce_{f(\omega)}}$. 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 $\gamma(\bar{f}) \succeq' \gamma(\bar{g})$, which is equivalent to $\gamma(\bar{f}) \hat{\succeq} \gamma(\bar{g})$, 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 $\gamma(\bar{f})(\omega, [0,1]) = \delta_{ce_{f(\omega)}}$. In turns, this means that we have

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

At the same time, from Claim 15, we know that for all $p \in \Delta(X)$, $u(ce_p) = V^*(\delta_{ce_p}) = V^*(p)$, where the first equality holds by construction of V^*, while the second equality holds because V^* represents the restriction of $\succeq$ 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 $\succeq$ 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. Finally, notice that, if $\succeq$ satisfies Axiom 7, then we must have that $\succeq'$ is such that for any $f', g' \in \mathcal{F}'$ such that $f' \sim' g'$ we have $\frac{1}{2} f' \oplus \frac{1}{2} g' \sim' f'$. 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) ⇒ (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 $\succeq$ 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 $x, y \in \text{supp}(p)$, 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 $x_1, \dots, x_n \in X$ and $y_0, \dots, y_n \in [0,1]$, 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[37], 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 $i = 1, \dots, n$. 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. □

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

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

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 $\min_{\phi \in \text{MAX}(\Phi,B)} \int_{[0,1]} u(\mu(p))\text{d}\phi = \min_{\phi \in \Phi} \int_{[0,1]} u(\mu(p))\text{d}\phi$

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 $\int_{[0,1]} u(\mu(p))\text{d}\hat{\phi} < \min_{\phi \in \text{MAX}(\Phi,B)} \int_{[0,1]} u(\mu(p))\text{d}\phi$. 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

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

We now define the notion of state-decreasing priors.

Definition 8. 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

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

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 $\hat{phi}([x_3, x_4]) = \phi'([x_1, x_2])$: 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

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

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 Π'' of priors on $[0,1]$ such that every $\phi \in \Phi$ is state-decreasing, atom less, 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 . □

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.[38] The desired representation then follows trivially.

Proof of (3) ⇒ (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 $RDEU_{u,\psi_n}(p_n) \to RDEU_{u,\psi}(p)$.

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

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

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

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 $RDEU_{u,\psi_n}(p_n) = \int_0^1 H_n(t)\mathrm{d}t$ and $RDEU_{u,\psi}(p) = \int_0^1 H(t)\mathrm{d}t$. 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 $RDEU_{u,\psi_n}(p_n) \to RDEU_{u,\psi}(p)$ 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 have $p(\delta A_t) = 0$. By Portmanteau Theorem[39] 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),[40] and therefore prove that Axiom 3 (Continuity).

Next, we turn to prove the necessity of Axiom 6 (Hedging). To this end, let us define the notion of enumeration.

Definition 9. A *simple enumeratio*n of a lottery q is a step function $x : [0, 1] \to X$ such that $l(\{z \in [0, 1]|f(z) = w\} = q(w)\ \forall\ w \in \mathrm{supp}(q)$.

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

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

[40] 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 $RDEU_{u,\psi}(p)$ seen as a function of both ψ and p – which, as we have seen, is continuous.

function that represents $\succeq$, we have

$$W(x) = \psi(p^x(x_1))u(x_1) + \sum_{i=2}^{N(x)} \left(\psi(\sum_{j=1}^{i} p^x(x_j)) - \psi(\sum_{j=1}^{i-1} p^x(x_j)) \right) u(x_i)$$
$$\leq \psi(p^y(y_1))u(y_1) + \sum_{i=2}^{N(y)} \left(\psi(\sum_{j=1}^{i} p^y(y_j)) - \psi(\sum_{j=1}^{i-1} p^y(y_j)) \right) u(y_i) = W(y).$$

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. ${}^1r(j) = \left\{ n \in \mathbb{N} | {}^1\bar{y}_{{}^1r(j)} = \bar{y}^j \right\}$)

2. Define ${}^i\bar{y}$ as

$$\begin{aligned} {}^i\bar{y}(t) &= \bar{y}^i \text{ for } t \in [\frac{i-1}{\prod n_i}, \frac{i}{\prod n_i}) \\ &= {}^{i-1}\bar{y}_i \text{ for } t \in [\frac{{}^{i-1}r(i) - 1}{\prod n_i}, \frac{{}^{i-1}r(i)}{\prod n_i}) \\ &= {}^{i=1}\bar{y}(t) \text{ otherwise} \end{aligned}$$

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

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 ${}^ir$ 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

$$\begin{aligned}
W({}^i y) - W({}^{i-1}y) &= \left(\psi(\sum_{j=1}^{i} p({}^i y_j) - \psi(\sum_{j=1}^{i-1} p({}^i y_j))\right)(u({}^{i-1}\bar{y}_i) - u(\bar{y}^i)) \\
&\quad + \left(\psi(\sum_{j=1}^{r(i)} p({}^i y_j) - \psi(\sum_{j=1}^{r(i)-1} p({}^i y_j))\right)(u(\bar{y}_i) - u({}^{i-1}\bar{y}_i)) \\
&= \left(\psi\left(\frac{i}{\prod n_i}\right) - \psi\left(\frac{i-1}{\prod n_i}\right)\right)(u({}^{i-1}\bar{y}_i) - u(\bar{y}^i)) + \\
&\quad \left(\psi\left(\frac{{}^{i-1}r(i)}{\prod n_i}\right) - \psi\left(\frac{{}^{i-1}r(i)-1}{\prod n_i}\right)\right)(u(\bar{y}_i) - u({}^{i-1}\bar{y}_i)) \\
&= \left(\left(\psi\left(\frac{i}{\prod n_i}\right) - \psi\left(\frac{i-1}{\prod n_i}\right)\right) - \left(\psi\left(\frac{{}^{i-1}r(i)}{\prod n_i}\right) - \psi\left(\frac{{}^{i-1}r(i)-1}{\prod n_i}\right)\right)\right) \\
&\quad \times(u({}^{i-1}\bar{y}_i) - u(\bar{y}^i)).
\end{aligned}$$

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 a new 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

$$\begin{aligned}
\bar{\psi}\left(\frac{{}^{i-1}r(i)}{\prod n_i} - \frac{i-1}{\prod n_i}\right) &\leq \bar{\psi}\left(\frac{{}^{i-1}r(i)-1}{\prod n_i} - \frac{i-1}{\prod n_i}\right) \\
&\quad + \bar{\psi}\left(\left(\frac{{}^{i-1}r(i)}{\prod n_i} - \frac{i-1}{\prod n_i}\right) - \left(\frac{{}^{i-1}r(i)-1}{\prod n_i} - \frac{i-1}{\prod n_i}\right)\right) \\
&= \bar{\psi}\left(\frac{{}^{i-1}r(i)-1}{\prod n_i} - \frac{i-1}{\prod n_i}\right) + \bar{\psi}\left(\frac{1}{\prod n_i}\right).
\end{aligned}$$

By substituting back for the original function gives

$$\begin{aligned}\psi\left(\frac{{}^{i-1}r(i)}{\prod n_i}\right) &\leq \psi\left(\frac{{}^{i-1}r(i)-1}{\prod n_i}\right)+\psi\left(\frac{i}{\prod n_i}\right)-\psi\left(\frac{i-1}{\prod n_i}\right) \\ &\Rightarrow \psi\left(\frac{{}^{i-1}r(i)}{\prod n_i}\right)-\psi\left(\frac{{}^{i-1}r(i)-1}{\prod n_i}\right)\leq\psi\left(\frac{i}{\prod n_i}\right)-\psi\left(\frac{i-1}{\prod n_i}\right).\end{aligned}$$

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 $t \in [\sum_{j=0}^{n-1} q^i_{i-1}, \sum_{j=0}^{n} q^i_{i-1})$. The utility of the enumeration y^i is given by

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

As $q^i_j \to p^y(y_j)$, 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 $W(y^i) > W(x^i)$. But as y^i is rational, this contradicts the above result. □

We now turn to prove that the Axiom 6 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

$$\begin{aligned}
U(r) &= \min_{\pi\in\Pi}\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z_i)\\
&= \min_{\pi\in\Pi}\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)\left(\frac{1}{2}(u(z_i^x)+u(z_i^y))\right)\\
&= \min_{\pi\in\Pi}[\frac{1}{2}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z_i^x)\right)\\
&\quad+\frac{1}{2}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z_i^y)\right)]\\
&\geq \frac{1}{2}\min_{\pi\in\Pi}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z_i^x)\right)\\
&\quad+\frac{1}{2}\min_{\pi\in\Pi}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z^y)\right)
\end{aligned}$$

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

$$\begin{aligned}
U(r) &\geq \frac{1}{2}\min_{\pi\in\Pi}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z_i^x)\right)\\
&\quad+\frac{1}{2}\min_{\pi\in\Pi}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(z^y)\right)\\
&\geq \frac{1}{2}\min_{\pi\in\Pi}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(\bar{z}_i^x)\right)\\
&\quad+\frac{1}{2}\min_{\pi\in\Pi}\left(\sum_i\left(\pi\left(\sum_{j=0}^{i-1}r(z_j)\right)-\pi\left(\sum_{j=0}^{i}r(z_j)\right)\right)u(\bar{z}_i^y)\right)\\
&= \frac{1}{2}U(p)+\frac{1}{2}U(q)
\end{aligned}$$

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 5 (Degenerate Independence) follow form standard arguments, while Axiom 4 follows from Lemma 1.

A.2 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.

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.*

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.

Dean, M. and P. Ortoleva (2011): "Allais and Ellsberg: an experimental test of their relationship," Mimeo California Institute 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., 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.

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.

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

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.

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.