Antimonotonicity for Preference Axioms: The Natural Counterpart to Comonotonicity

Giulio Principi* Peter P. Wakker† Ruodu Wang‡

July 17, 2023

Abstract

Comonotonicity ("same variation") of random variables minimizes hedging possibilities and has been widely used in many fields. Comonotonic restrictions of traditional axioms have led to impactful inventions in decision models, including Gilboa and Schmeidler's ambiguity models. This paper investigates antimonotonicity ("opposite variation"), the natural counterpart to comonotonicity, minimizing leveraging possibilities. Surprisingly, antimonotonic restrictions of traditional axioms often do not give new models but, instead, give generalized axiomatizations of existing ones. We, thus, generalize: (a) classical axiomatizations of linear functionals through Cauchy's equation; (b) as-if-risk-neutral pricing through no-arbitrage; (c) subjective probabilities through bookmaking; (d) Anscombe-Aumann expected utility; (e) risk aversion in Savage's subjective expected utility. In each case, our generalizations show where the most critical tests of classical axioms lie: in the antimonotonic cases (maximal hedges). We, finally, present cases where antimonotonic restrictions do weaken axioms and lead to new models, primarily for ambiguity aversion in nonexpected utility.

Keywords: Comonotonicity, bookmaking, hedging, subjective expected utility, ambiguity aversion

*Department of Economics, New York University, USA. ✉ gp2187@nyu.edu

†Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands. ✉ wakker@ese.eur.nl

‡Department of Statistics and Actuarial Science, University of Waterloo, Canada. ✉ wang@uwaterloo.ca

1 Introduction

Comonotonicity has been widely used in mathematics (Hardy, Littlewood & Pólya, 1934, Theorem 236) and in many applied fields, including decision theory.[1] Puccetti & Wang (2015) provided a survey. In particular, comonotonicity was the main tool in Gilboa's (1987) and Schmeidler's (1989) famous ambiguity model, resolving Ellsberg's (1961) paradox. Two variables are comonotonic if they covary in the same direction. For given marginal distributions, comonotonicity maximizes correlation and minimizes hedging possibilities (Hoeffding, 1940).

Antimonotonicity is a natural counterpart to comonotonicity. Two variables X and Y are antimonotonic if they covary in the opposite direction; i.e., if X and $-Y$ are comonotonic. For given marginal distributions, antimonotonicity minimizes correlation and leveraging possibilities. Aouani, Chateauneuf & Ventura (2021) introduced antimonotonic diversification for Choquet integrals. Antimonotonicity turns out to be of interest in its own right, and this paper studies it in general. It will shed new light on many classical results.

Gilboa (1987) and Schmeidler (1989) used comonotonicity to weaken classical independence preference conditions by minimizing hedging effects and maximizing leveraging effects. They, thus, obtained a new preference model, Choquet expected utility, which could accommodate ambiguity aversion in Ellsberg's (1961) paradox. This result, together with Gilboa & Schmeidler (1989), famously opened the new field of decision under ambiguity.

It is very natural to study the counterpart to the Gilboa-Schmeidler approach, where classical independence is weakened using antimonotonicity instead of comonotonicity, now minimizing leveraging effects and maximizing hedging effects. The question then is which new models result this way. Our main result came as a surprise to us: mostly no new models result. That is, we mostly obtain the same classical models as when we do not impose the antimonotonicity restrictions. Whereas these results may at first be taken as negative, it delivers positive implications in the sense that we obtain generalizations of

[1] For comonotonicity in decision theory, see Grabisch (2016). Further examples include fuzzy set theory (Grabisch, Murofushi & Sugeno, 2000), insurance (Dhaene et al., 2002), labor market equilibria (Chade, Eeckhout & Smith, 2017), multiattribute utility theory (Ekeland, Galichon & Henry, 2012), optimal transport (Galichon, 2016), risk allocations (Rüschendorf, 2013), risk attitudes (Yaari, 1969, p. 328), risk measures (Föllmer & Schied, 2016), time preference (Bastianello & Faro, 2023), and welfare theory (Ebert, 2004).

many classical results. Those include the fundamental theorem of asset pricing (as-if risk-neutral pricing through no-arbitrage), and the characterization of linear functionals through Cauchy's equation, subjective probabilities through de Finetti's bookmaking argument, and Anscombe and Aumann's expected utility. In each case, our analysis shows that the most critical versions of the relevant axioms occur in cases of antimonotonicity. We only need to verify the axioms in these cases, which then is sufficient to imply the axioms and corresponding models in full generality. In our proofs, all aforementioned results are derived from our Theorem 1, the antimonotonic generalization of Cauchy's functional equation. Our analysis thus shows the unity in the aforementioned classical results, which were developed independently in their respective fields.

We also study antimonotonic weakenings of convexity conditions for preferences, which involve mathematics of a different nature than the preceding results. We, first, generalize a result of Debreu & Koopmans (1982), providing the most useful axiomatization of risk aversion available in the literature for Savage's (1954) subjective expected utility. This result did not receive the attention that it deserves and we aim to popularize it. We explain its special value, and our contribution to it, in Subsection 4.2.

Whereas in all aforementioned results, antimonotonic restrictions did not lead to new models but instead to generalized axiomatizations, we end with some different cases. For example, antimonotonic convexity is strictly weaker than convexity in nonexpected utility models capturing ambiguity, leading to more general models. This was first shown by Aouani, Chateauneuf & Ventura (2021), and we add some observations and topics for future research.

The outline of this paper is as follows, where in each case below we consider the antimonotonic restrictions of axioms. Section 2 provides axiomatizations of functionals. Subsection 2.1 considers additivity conditions, Subsection 2.2 considers no-arbitrage, and Subsection 2.3 considers affinity conditions. The rest of the paper considers preference axiomatizations. Section 3 considers linear and affine functionals, which involves linear equalities. Subsection 3.1 deals with de Finetti's bookmaking and Subsection 3.2 with Anscombe and Aumann's expected utility. Section 4 turns to convexity conditions, meaning that we now deal with inequalities. After preparations (Subsection 4.1), Subsection 4.2 characterizes concave utility in Savage's subjective expected utility.

Subsection 4.3 considers the Anscombe-Aumann framework for ambiguity, which assumes affine utility. It deviates from all preceding parts in that the antimonotonic weakening of convexity is a true weakening, giving more general ambiguity models. Finally, Section 5 concludes and the Appendix provides proofs.

2 Antimonotonic additivity of functionals

Fix a space $(\Omega, \mathcal{F})$, in which Ω is a *state space* and $\mathcal{F}$ an algebra of subsets of Ω called *events*. We denote by $B(\Omega, \mathcal{F})$ the set of *acts*, i.e., all bounded real-valued measurable functions from Ω to $\mathbb{R}$, equipped with the sup-norm. Images of acts (real numbers) are called *outcomes*. They can be interpreted as monetary. As is common, we identify constant acts with outcomes. Two acts X and Y in $B(\Omega, \mathcal{F})$ are *comonotonic* if

$$\text{for all } \omega, \omega' \in \Omega : (X(\omega) - X(\omega'))\,(Y(\omega) - Y(\omega')) \geq 0. \tag{1}$$

Two acts X and Y in $B(\Omega, \mathcal{F})$ are *antimonotonic* if X and $-Y$ are comonotonic. While comonotonicity captures the idea of two acts covarying positively, antimonotonicity refers to negative covariation between the variables. Each constant act is both comonotonic and antimonotonic with every other act. The next subsections show that antimonotonic additivity implies full-force additivity under each of some additional assumptions, being finiteness of Ω, continuity of I, or monotonicity of I.

2.1 Antimonotonic additivity implies addivitity

A functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ is *additive* if

$$\text{for all } X, Y : I(X + Y) = I(X) + I(Y). \tag{2}$$

The equation is also known as Cauchy's equation (Aczél, 1966). The functional I satisfies *antimonotonic (am-) additivity* if additivity holds only for all pairs of antimonotonic acts X, Y, while I satisfies *comonotonic additivity* if additivity holds for all pairs of comonotonic acts X, Y. Moreover, I is *linear* if it is *homogeneous*, i.e., $I(\alpha X) = \alpha I(X)$ for all $\alpha \in \mathbb{R}$ and all $X \in B(\Omega, \mathcal{F})$, and additive. *Positive homogeneity* imposes the homogeneity requirement only for $\alpha > 0$. The above definitions are extended to I's defined on subdomains in the obvious manner, imposing the requirements only when all acts involved are contained in the subdomain. We next generalize standard results on Cauchy's equation to antimonotonicity. The following result presents a first situation where an am-condition is equivalent to its full-force version.

Theorem 1. *Suppose that Ω is finite and $\mathcal{F} = 2^\Omega$. For $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$, antimonotonic additivity is equivalent to additivity. In addition, if I is continuous at one point, then it is linear and continuous.*

By standard measure-theoretic techniques, the previous result can be extended to general, possibly infinite, state spaces Ω if continuity is assumed (or monotonicity as in the next subsection). We do not expect the equivalence between antimonotonic additivity and additivity to hold on general state spaces without some extra regularity condition, but this remains to us as an open question.

Proposition 2. *Under continuity, antimonotonic additivity of a functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ is equivalent to linearity and, thus, to additivity.*

2.2 No-arbitrage in finance

We now turn to an application in finance. We interpret acts as financial assets and let I reflect the market price. In finance, additivity of I and even linearity are imposed by common market trade assumptions. Furthermore, I is then assumed *normalized*: $I(0) = 0$ and $I(1) = 1$. This implies $I(k) = k$ for all outcomes k. A central condition in finance is *monotonicity*: $I(X) \geq I(Y)$ whenever $X(\omega) \geq Y(\omega)$ for all $\omega \in \Omega$. By adding this condition, this section differs from the preceding subsection. Any linear and monotone I can be normalized by dividing it by $I(1)$—except in the trivial case where $I(1)$ and the functional are 0. In combination with linearity, monotonicity is equivalent to no-arbitrage in finance. The fundamental theorem of asset pricing entails that no-arbitrage implies as-if risk-neutral pricing: there exists a probability measure[2] P on Ω such that I is the expectation, denoted $\mathbb{E}_P$ or $\mathbb{E}$ for short, under P. We generalize this fundamental theorem

[2]In decision theory, there is much interest in finite additivity. We will, therefore, only require finite additivity of probability measures. A necessary and sufficient condition for countable additivity can readily be added in all our results (Wakker, 1993, Proposition 4.4).

of finance by showing that additivity (and linearity) can be weakened to antimonotonic acts. That is, the critical test of no-arbitrage occurs in cases where leverage possibilities are maximal. This then ensures no-arbitrage everywhere.

Proposition 3. *There exists a probability measure P such that $I = \mathbb{E}_P$ ("as-if risk-neutrality") if and only if I is normalized and satisfies monotonicity and antimonotonic additivity. Here, P is unique.*

In the above proposition, am-additivity again implies additivity, as it did in the preceding subsection. The proposition shows that am-additivity captures the essence of no-arbitrage, given that the other conditions are standard. Gilboa & Samuelson (2022) characterized no-arbitrage for arbitrary sets of acts and discuss its normative status.

We next turn to the case where a probability measure P is given on Ω. A probability space $(\Omega, \mathcal{F}, P)$ results. In the theory of decision making under risk and in the risk management literature, there is much interest in *law-based* functionals I: $I(X) = I(Y)$ whenever X and Y induce the same probability distribution over outcomes. In this case, I depends only on the probability distribution ("law") generated over the outcomes. In decision making under risk, Ω is often suppressed and acts are identified with those probability distributions. Different from the interpretation that we considered before, I does not need to represent a market price here. When P is subjective, Machina & Schmeidler (1992) called law-based functionals probabilistically sophisticated.

Acts that differ only on a null set can be treated the same and can be identified. A set is *null* if it has P value 0. We, therefore, extend the domain of I somewhat and incorporate acts that are bounded except on a null event. We denote this extended domain by $L^\infty(\Omega, \mathcal{F}, P)$. We call P *atomless* if, for each $E \in \mathcal{F}$ and $0 < \mu < P(E)$, there exists $A \subset E$ with $P(A) = \mu$. This is the common definition for finitely additive measures. Under countable additivity, it is equivalent to the absence of atoms. Using Mackenzie's (2019) technique, we could allow for atoms as long as a sufficiently large part of the state space is atomless, but for simplicity we will assume complete absence of atoms.

Proposition 4. *Suppose that Ω is finite with equally-likely states and $\mathcal{F} = 2^\Omega$, or $(\Omega, \mathcal{F}, P)$ is atomless. Then I is normalized, monotone, law-based, and am-additive if and only if it is $\mathbb{E}_P$, the expectation with respect to P.*

In the risk management literature, for a risk measure I, the equality $I(X+Y) = I(X) + I(Y)$ is often interpreted as that no diversification benefit[3] is assigned to the portfolio vector (X, Y); see Wang & Zitikis (2021) in the context of the Basel Accords. In this context, Proposition 4 is very intuitive: If no portfolio of two antimonotonic risks (representing maximum hedging effect) is assigned a diversification benefit, then no portfolio should have any diversification benefit, and hence the risk measure should simply be the expected value. This is in sharp contrast to the idea of assigning no diversification benefit to comonotonic risks, which leads to a large class of risk measures called distortion risk measures; mathematically, they coincide with the dual utility functionals of Yaari (1987). See McNeil, Frey & Embrechts (2015) for using distortion risk measures in risk management.

2.3 Antimonotonic affinity on convex domains

Whereas the preceding result assumed full domains, for many applications, including preference results derived in later sections, it is desirable to have flexibility regarding domains. Hence, this subsection presents affinity conditions that can be used on convex domains D. We call $I : D \to \mathbb{R}$ *affine* if

$$\text{for all } \alpha \in [0,1] \text{ and all } X, Y \in D : I\left(\alpha X + (1-\alpha)Y\right) = \alpha I(X) + (1-\alpha) I(Y). \quad (3)$$

Proposition 5. *Assume $I : D \to \mathbb{R}$, where $D \subset B(\Omega, \mathcal{F})$ is convex and contains a constant act in its interior. Then am-affinity is equivalent to affinity whenever I is continuous or monotonic.*

3 Preference axiomatizations of subjective expected utility

As yet, we took a functional I as primitive. In this section, we instead take a preference relation $\succcurlyeq$ over acts as primitive. This is of interest in economics and decision theory, where choice making, captured through $\succcurlyeq$, is taken as the primitive, and all

[3]The diversification benefit often refers to $I(X)+I(Y)-I(X+Y)$; see McNeil, Frey & Embrechts (2015).

concepts should get their meaning through $\succcurlyeq$. The assumptions in the preceding section are naturally generalized as follows. Still, Ω is a state space endowed with an algebra of events $\mathcal{F}$. The *outcome set* $\mathcal{C}$ is, for now, taken general, with specifications added later. We start with a binary relation $\succcurlyeq$ on $\mathcal{C}$, later extended to nonconstant acts. Acts map Ω to $\mathcal{C}$. Acts are *measurable*, which means that their inverses of preference intervals are measurable. A *preference interval* is a subset of $\mathcal{C}$ that, for each pair of outcomes $x \succcurlyeq z$ contains all outcomes y with $x \succcurlyeq y \succcurlyeq z$. We also assume that acts X are *bounded*, i.e., there exist outcomes $x \succcurlyeq z$ such that $x \succcurlyeq X(\omega) \succcurlyeq z$ for all ω. Thus, formally, *acts* are measurable bounded maps from Ω to $\mathcal{C}$. We again identify constant acts with outcomes. Finally, $\succcurlyeq$ denotes a binary relation on the set of all acts that, for constant acts, agrees with the above relation $\succcurlyeq$ on outcomes, which is why we use the same symbol. As usual, the asymmetric part of $\succcurlyeq$ is $\succ$, its symmetric part is $\sim$, and $\preccurlyeq$ and $\prec$ denote reversed preferences. We call $\succcurlyeq$ *trivial* if $X \sim Y$ for all acts X, Y. We summarize:

Assumption 6 (Structural Assumption)**.** A state space Ω is given with an algebra $\mathcal{F}$ and an outcome set $\mathcal{C}$. Acts are bounded measurable maps from Ω to $\mathcal{C}$, and $\succcurlyeq$ is a binary relation on the set of all acts.

Weak ordering holds if *completeness* ($X \succcurlyeq Y$ or $Y \succcurlyeq X$ for all X, Y) and *transitivity* ($X \succcurlyeq Y$ and $Y \succcurlyeq Z$ implies $X \succcurlyeq Z$) hold. The following definitions for preferences will naturally generalize the preceding definitions for functionals, so that we use the same terms. An outcome x is a *certainty equivalent* (*CE*) of an act X if $x \sim X$. In general, it need not always exist and need not be unique. *Monotonicity* holds if $X \succcurlyeq Y$ whenever $X(\omega) \succcurlyeq Y(\omega)$ for all $\omega \in \Omega$. Two acts X, Y are *comonotonic* if there are no states ω, ω' such that $X(\omega) \succ X(\omega')$ and $Y(\omega) \prec Y(\omega')$. They are *antimonotonic* if there are no states ω, ω' such that $X(\omega) \succ X(\omega')$ and $Y(\omega) \succ Y(\omega')$. A real-valued functional I *represents* $\succcurlyeq$, or $\succcurlyeq$ *maximizes* I, if the preference domain is contained in the domain of I and $X \succcurlyeq Y \Leftrightarrow I(X) \geq I(Y)$.

Subjective expected utility or *expected utility*, or *EU* for short, holds if there exist a probability measure P on $\mathcal{F}$ and a *utility function* $U : \mathcal{C} \to \mathbb{R}$ such that $\succcurlyeq$ maximizes *expected utility* $\int_\Omega U(X)\mathrm{d}P$, where this integral, called the *EU* of X, is assumed to be well-defined and finite.

3.1 de Finetti's bookmaking for subjective probabilities

We here assume $\mathcal{C} = \mathbb{R}$, as in Section 2. Further, we assume that the preference relation $\succcurlyeq$ on outcomes is the natural ordering $\geq$. Therefore, the preceding definitions of monotonicity, comonotonicity, and antimonotonicity for preferences agree with those for representing functionals in the first part of the paper. *Subjective expected value*, or *expected value* or *EV* for short, holds if EU holds with U the identity function.

The following condition is commonly used to axiomatize EV maximization. *Additivity* holds for $\succcurlyeq$ if

$$\text{for all acts } X, Y, Z : X \sim Y \Longrightarrow X + Z \sim Y + Z. \tag{4}$$

If a certainty equivalent exists for every act, as is the case in all results in this paper, then a convenient restriction is:

$$\text{for all acts } X, Z \text{ and outcomes } x : X \sim x \Longrightarrow X + Z \sim x + Z. \tag{5}$$

The condition is weaker than Eq. 4 in the sense of restricting to constant $Y = x$. However, it readily implies Eq. 5 by two-fold application with the (same) CE for X and Y and transitivity. The condition is appealing because we can then use "ironing out" (Li, 2020): in any sum of acts we can replace any act by its CE (any if several exist), readily giving a quantitative evaluation of the sum. The condition is well-suited for our purposes because the constant act x is automatically antimonotonic with the other acts.

Definition 7. *Am-additivity* holds for $\succcurlyeq$ if the implication of Eq. 5 is imposed only if X and Z are antimonotonic.

The following result generalizes de Finetti's (1931) bookmaking theorem by adding an antimonotonic restriction.

Proposition 8. *Assume Structural Assumption 6 with $\succcurlyeq$ on $\mathcal{C} = \mathbb{R}$ the natural ordering $\geq$. There exists a probability measure P such that expected value holds if and only if there exists a certainty equivalent for every act and weak ordering, monotonicity, and am-additivity hold.*

de Finetti and many other authors who wrote on bookmaking, assumed additivity more or less implicitly[4] but emphasized the importance of monotonicity. They used the above result, without the antimonotonicity restriction, and several variations, to argue that it is rational to use subjective probabilities in the context of uncertainty. Linearity of utility, as implied here, is reasonable for moderate stakes (l'Haridon & Vieider, 2019, p. 189; Savage, 1954, p. 91). de Finetti's result was historically important as a foundation of Bayesianism. Our result shows, again, that the most critical case of "bookmaking" occurs when there are maximal possibilities of hedging (antimonotonicity). That is, de Finetti needed to defend his condition only for antimonotonic cases.

3.2 Anscombe and Aumann's subjective expected utility

Anscombe & Aumann (1963) presented a famous preference axiomatization of expected utility. Fishburn (1970, §13.1) proposed a simplified version of their framework that is most popular today and we will use that. It is important because it is mostly used as point of departure for developing ambiguity theories (§4.3). The outcome set $\mathcal{C}$ is nowadays often chosen to be a mixture space, and so will we do. Examples of mixture spaces include money intervals, convex subsets of linear spaces, or, the most common case in the literature on ambiguity, probability distributions over general "prizes." For simplicity, readers not familiar with general mixture spaces may take in mind any such example.

We call $\mathcal{C}$ a *mixture space* if it is endowed with a mixture operation. A *mixture operation* generalizes convex combinations on linear spaces. It maps $\mathcal{C} \times [0,1] \times \mathcal{C}$ to $\mathcal{C}$, and is denoted $\alpha x + (1-\alpha)y$. It is required to satisfy the following conditions:

(i) $1x + 0y = x$ (identity);

(ii) $\alpha x + (1-\alpha)y = (1-\alpha)y + \alpha x$ (commutativity);

(iii) $\alpha(\beta x + (1-\beta)y) + (1-\alpha)y = \alpha\beta x + (1-\alpha\beta)y$ (associativity).

A function $U : \mathcal{C} \to \mathbb{R}$ is *affine* if

$$\text{for all outcomes } x, y, \text{ and } 0 < \alpha < 1: \; U(\alpha x + (1-\alpha)y) = \alpha U(x) + (1-\alpha)U(y). \tag{6}$$

[4] The bookmaking argument usually makes yet stronger assumptions by also incorporating positive scalar multiplications and, thus, linear combinations. We showed that such assumptions are not needed because they are implied by the other conditions.

It is an *interval scale* if it is unique up to multiplication by a positive factor and addition of a constant.

The mixture operation naturally extends to acts by applying it statewise, which again gives a mixture operation. *Mixture continuity* holds for $\succcurlyeq$ if the sets

$$\{\alpha \in [0,1] : \alpha X + (1-\alpha)Z \succcurlyeq Y\} \text{ and } \{\alpha \in [0,1] : Y \succcurlyeq \alpha X + (1-\alpha)Z\}$$

are closed for all acts X, Y, Z. Together with some other conditions, mixture continuity implies the existence of a certainty equivalent for each act.

The classical independence axiom is as follows: *strong independence* holds if

$$\text{for all } X, Y, Z \in D \text{ and } 0 < \alpha < 1 : X \sim Y \Longrightarrow \alpha X + (1-\alpha)Z \sim \alpha Y + (1-\alpha)Z. \quad (7)$$

As in the preceding subsection, the following reformulation is convenient:

$$\text{for all } X, Z \in D,\ x \in \mathcal{C},\ \text{and } 0 < \alpha < 1 : X \sim x \Longrightarrow \alpha X + (1-\alpha)Z \sim \alpha x + (1-\alpha)Z. \quad (8)$$

Similarly to Eq. 5, the condition in Eq. 8 is again weaker than strong independence in the sense of restricting to constant $Y = x$, but readily implies Eq. 7 by two-fold application with the (same) CE for X and Y and transitivity. The condition is appealing because of its central role in backward induction: in a lottery we can replace any conditional sublottery by its CE, and do this repeatedly until a certainty equivalent of the entire lottery results (ironing out in Li, 2020).

Definition 9. *Am-independence* holds for $\succcurlyeq$ if the implication of Eq. 8 is imposed only if X and Z are antimonotonic.

Theorem 10. *Assume Structural Assumption 6 with $\mathcal{C}$ a mixture space. The following statements are equivalent.*

(i) Weak ordering, monotonicity, mixture continuity, and strong independence hold.

(ii) Weak ordering, monotonicity, mixture continuity, and am-independence hold.

(iii) Subjective expected utility holds with U affine.

In (iii), U is an interval scale.

4 Antimonotonic convexity and concavity

We have so far investigated linear-type representations and axioms that in fact dealt with linear equalities and their antimonotonic generalizations, processing both outcomes and events linearly. This section turns to antimonotonic generalizations of convexity and concavity axioms, which means that we relax linearity and that we now deal with inequalities. Subsection 4.2 maintains linearity in events/probabilities by assuming expected utility. However, unlike preceding sections, it allows for nonlinear utility. It will provide an antimonotonic axiomatization of concave utility. Subsection 4.3 will take a dual approach. It assumes linear or, equivalently, affine utility of outcomes. But now, unlike preceding sections, it allows for nonlinear event weighting. That is, it investigates deviations from expected utility, in particular, implications of antimonotonic convexity for ambiguity models.

4.1 Preparations for convexity

We continue to work with preferences as primitives, as in Structural Assumption 6. We also assume that the outcome set $\mathcal{C}$ is a mixture set, so that we can handle both monetary outcomes and more general outcome sets such as commodity bundles or sets of probability distributions over prizes. This section investigates the following conditions.

Definition 11 (Convexity)**.** Preferences are *convex* if

$$\text{for all acts } X, Y \text{ and } 0 < \alpha < 1 : X \sim Y \Longrightarrow \alpha X + (1-\alpha)Y \succcurlyeq X. \tag{9}$$

Preferences are *am-convex* if the above implication is imposed only for antimonotonic X, Y.

Convexity of preference is a common assumption in consumer theory (Mas-Colell, Whinston, & Green, 1995). It is also called quasiconvexity or, sometimes, quasiconcavity because it is equivalent to quasiconcavity of any representing function. It may reflect preference for smoothing, diversification, and hedging, in models discussed next. It is remarkable that the same mathematical condition, while capturing utility of commodity bundles in consumer theory, also provides the best characterization of risk aversion in subjective expected utility and best captures ambiguity aversion in the currently most popular ambiguity models, as the following subsections will show.

4.2 Convexity of preference for expected utility and utility curvature

This subsection allows for nonlinear utility. We will maintain continuity:

Definition 12. Utility U on the mixture space $\mathcal{C}$ is *mixture continuous* if, for all outcomes x, y, $U(\alpha x + (1-\alpha)y)$ is continuous in α.

The condition is implied by affinity and also by common continuity conditions on convex subsets of linear spaces. Hence, assuming it is less restrictive than most other continuity conditions.

We next present an appealing implication of am-convexity in Savage's expected utility, where utility is not assumed to be affine in outcomes, and utility curvature captures different risk (or uncertainty) attitudes. To avoid triviality, we assume *non-degenerateness*, i.e., there exists an event A with $0 < P(A) < 1$.

Theorem 13. *Assume Structural Assumption 6 with non-degenerate expected utility and a mixture continuous utility function U. The following three statements are equivalent.*

(i) $\succcurlyeq$ *satisfies convexity.*

(ii) $\succcurlyeq$ *satisfies am-convexity.*

(iii) U *is concave.*

Debreu & Koopmans (1982) showed that (i) and (iii) in the theorem are equivalent, more generally, even without continuity of utility. Wakker & Yang (2019, 2021) generalized Debreu & Koopmans' result to Choquet expected utility. The characterizations used in these results, through convexity of preference with respect to outcome mixing, are appealing because they make risk aversion directly testable for subjective probabilities. To explain this point, we first note that concave utility captures risk aversion under expected utility. In decision under risk, where probabilities are objective and known beforehand, the conditions most commonly used to characterize risk aversion involve preference for expected value or aversion to mean-preserving spreads. These conditions use probabilities as inputs. This use is problematic for decision under uncertainty because then probabilities are subjective and not directly observable, as in Savage (1954). The main purpose of preference axiomatizations is to make theoretical properties directly observable. Therefore, the aforementioned common conditions for risk aversion, using probabilities as input, are not well suited for the context of uncertainty. Theorem 13 and its predecessors make risk aversion directly observable and testable for subjective probabilities. They are the only way to do so known to us.

The contribution of our Theorem 13 to its predecessors by Debreu & Koopmans (1982) and Wakker & Yang (2019, 2021) concerns, again, the central topic of this paper: we only need to inspect the most critical cases with maximal hedging possibilities. If risk aversion (and preference for diversification) passes those tests, then it holds everywhere.

4.3 Convexity of preference and affine utility

In §3.2 we presented a version of Anscombe and Aumann's axiomatization of expected utility. However, their framework has proved extremely useful for developing deviations from expected utility. It serves, for instance, to accommodate Ellsberg's (1961) ambiguity aversion. Famous contributions are Gilboa & Schmeidler's (1989) axiomatization of multiple priors, and Schmeidler's (1989) axiomatization of Choquet expected utility. These contributions initiated the field of ambiguity theory, a large and important field today (Gilboa & Marinacci, 2016; Trautmann & van de Kuilen, 2015). Section 4 of Lu (2021) provided an extension to random choice.

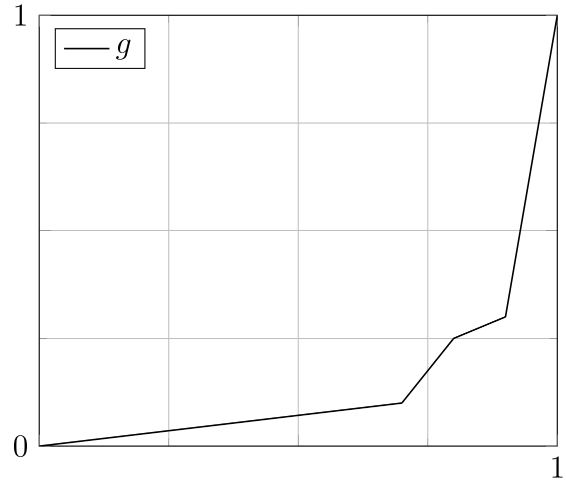
We define the *Anscombe-Aumann framework* formally, as it is used today. It is in fact the simplified framework proposed by Fishburn (1970). It assumes: (i) Structural Assumption 6; (ii) the outcome space $\mathcal{C}$ is a mixture space; (iii) an affine utility function $U : \mathcal{C} \to \mathbb{R}$ is given that represents preferences $\succcurlyeq$ on $\mathcal{C}$; (iv) monotonicity holds. The primary interpretation of outcomes is now that they are probability distributions over "prizes," in which case the mixture operation concerns probabilistic mixing. Affine utility U is then justified by assuming expected utility maximization over the probability distributions over prizes, which indeed gives linearity with respect to probabilistic mixtures. Affine utility facilitates mathematical analyses so much that the Anscombe-Aumann framework has propelled the field of ambiguity over the last decades.

We again study convexity of preference. Any utility effect as in Theorem 13 has now been ruled out by the affinity assumption. Hence, as follows from Theorem 13, strict convexities must now in fact speak to deviations from EU. In the first axiomatized ambiguity models (Gilboa & Schmeidler, 1989; Schmeidler, 1989), and in many that followed later, convexity was found to be equivalent to ambiguity aversion, explaining Ellsberg's (1961) famous paradox. Hence, convexity has as yet been the most central condition in ambiguity theories. This subsection deviates from preceding sections not only by considering inequalities and by allowing for deviations from EU, but also because antimonotonic restrictions now do lead to strict weakenings of preference conditions and, thus, to new models.

The following example shows that antimonotonic convexity is strictly weaker than convexity. To prepare, a *weighting function* W maps events to $[0,1]$ and satisfies $W(\emptyset)=0$, $W(S)=1$, and $A \supset B \Rightarrow W(A) \geq W(B)$. We call W *convex* if $W(A \cup B) + W(A \cap B) \geq W(A) + W(B)$ for all events. This implies *pseudo-convexity*: $W(A) \leq W(A \cup B) - W(B) \leq 1 - W(A^c)$ for all disjoint events A, B. We assume, for simplicity, that $U \geq 0$. *Choquet expected utility* holds if there exists a weighting function W such that the preference relation maximizes $X \mapsto \int_{[0,\infty)} W(U(X) \geq x)\mathrm{d}x$. Aouani, Chateauneuf & Ventura (2021) proved that for such preferences pseudo-convexity is equivalent to am-convexity (their Theorem 1 and Corollary 1).

Example 14. Assume Structural Assumption 6 with $S = [0,1]$, $\mathcal{F}$ the usual Borel sigma-algebra, P the usual Lebesgue measure (uniform distribution), and $\mathcal{C} = [0,\infty)$. The function g satisfies $g(0) = 0$, $g(0.7) = 0.1$, $g(0.8) = 0.25$, $g(0.9) = 0.3$, $g(1) = 1$, and is linear in between (see Figure 1). The preference relation $\succcurlyeq$ maximizes Choquet expected utility with $W(E) = g(P(E))$. We here in fact have the special case of Choquet expected utility of Yaari (1987) with law invariance and linear utility. Pseudo-convexity of W is equivalent to g being superadditive (i.e., $g(x+y) \geq g(x) + g(y)$ for $x, y \in [0,1]$ with $x + y \leq 1$) and satisfying $g(x) + g(y) \leq 1 + g(x+y-1)$ for $x, y \in [0,1]$ with $x + y \geq 1$. Therefore, W here is pseudo-convex, and by the aforementioned result of Aouani, Chateauneuf & Ventura (2021) antimonotonic convexity holds. However, W is not convex and, hence, neither is $\succcurlyeq$ (Wakker & Yang, 2019 Corollary 7; Wang et al., 2020 Theorem 3).

Figure 1: The weighting function g



Example 14 shows that antimonotonic convexity is strictly weaker than convexity. Thus, here, preference for hedging only when the hedging possibility is maximal (antimonotonicity) does not imply a general preference for hedging, unlike what we found in Theorem 13.

For Choquet expected utility with affine utility, Aouani, Chateauneuf & Ventura (2021) fully characterized antimonotonic convexity by pseudo-convexity of W (their Theorem 1 and Corollary 1). They provided many related results for superadditivity, supermodularity, and other properties, with interesting implications for uncertainty attitudes and diversification. Beissner & Werner (2022) provided optimization techniques for nonexpected utility models that neither are differentiable nor satisfy convexity of preference.

We end with some preliminary observations on the role of antimonotonicity in other ambiguity theories. To prepare, we define *uncertainty reduction*:

$$\text{for all acts } X, \text{ outcomes } x, \text{ and } 0 < \lambda < 1 : X \sim x \Longrightarrow \lambda X + (1 - \lambda)x \succcurlyeq X. \tag{10}$$

Uncertainty reduction is weaker than antimonotonic convexity because every constant act x is antimonotonic with every other act. Thus, antimonotonic convexity is between convexity and uncertainty reduction. Cerreia-Vioglio et al. (2011) characterized convexity in full generality in the Anscombe-Aumann framework. Under some further assumptions, mainly weak certainty independence, Maccheroni, Marinacci & Rustichini (2006) axiomatized convexity through their variational functionals. Similarly, under some further assumptions, Castagnoli et al. (2022) axiomatized uncertainty reduction through their star-shaped representing functionals. Hence, given the other assumptions, antimonotonic convex functionals will be in the "middle" between convex and star-shaped functionals. Interestingly, Castagnoli et al. (2022) showed that their star-shaped functionals are maxima of concave functionals. This result can be applied to am-convex preferences and functionals. We leave the further study of antimonotonicity in ambiguity models to future work.

5 Conclusion

This paper has provided a systematic study of antimonotonic restrictions of axioms for preference relations and functionals. Antimonotonicity is the natural counterpart to the well-known comonotonicity. We obtained many generalizations of classical theorems, for each showing where the most critical tests of axioms are. Those tests concern the cases with maximal possibilities for hedging. Our results highlight the asymmetry between antimonotonicity and comonotonicity. For ambiguity theories, antimonotonicity provides a tool for developing new theories.

Acknowledgements

We thank Fabio Maccheroni and Efe Ok for helpful comments and discussions. Giulio Principi gratefully acknowledges the support of the MacCracken Fellowship. Ruodu Wang acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03823) and Canada Research Chars (CRC-2022-00141).

Appendix: Proofs

Proof of Theorem 1. We assume am-additivity and derive additivity. Write $\Omega = \{\omega_1, \ldots, \omega_n\}$. Because X and $-X$ are antimonotonic,

$$I(0) = I(0+0) = I(0) + I(0) = 0$$

and

$$0 = I(0) = I(X - X) = I(X) + I(-X),$$

implying $I(-X) = -I(X)$.

For any comonotonic X and Y, $X+Y$ and Y are comonotonic so that $X+Y$ and $-Y$ are antimonotonic. Hence,

$$I(X) = I(X+Y-Y) = I(X+Y) + I(-Y) = I(X+Y) - I(Y).$$

Comonotonic additivity follows.

Consider two general X, Y. With $\Omega = \{\omega_1, \ldots, \omega_n\}$, we can write $X = X^{\uparrow} + X^{\downarrow}$ with $X^{\uparrow}(\omega_i)$ weakly increasing and $X^{\downarrow}(\omega_i)$ weakly decreasing in i, and $Y = Y^{\uparrow} + Y^{\downarrow}$ similar. By comonotonic additivity (CA) and am-additivity (AA):

$$\begin{aligned} I(X+Y) &\overset{\text{(def)}}{=\!=} I(X^{\uparrow} + X^{\downarrow} + Y^{\uparrow} + Y^{\downarrow}) \\ &\overset{\text{(AA)}}{=\!=} I(X^{\uparrow} + Y^{\uparrow}) + I(X^{\downarrow} + Y^{\downarrow}) \\ &\overset{\text{(CA)}}{=\!=} I(X^{\uparrow}) + I(Y^{\uparrow}) + I(X^{\downarrow}) + I(Y^{\downarrow}) \\ &\overset{\text{(AA)}}{=\!=} I(X^{\uparrow} + X^{\downarrow}) + I(Y^{\uparrow} + Y^{\downarrow}) = I(X) + I(Y). \end{aligned}$$

This shows that I is additive. Theorem 5.1.1 in Aczél (1966) shows that linearity follows under mild extra conditions such as continuity at one point or monotonicity (defined later).

Proof of Proposition 2. It is direct that linearity implies additivity, which implies am-additivity. We, therefore, assume the latter and derive linearity. For any fixed finite partition, am-additivity implies linearity for the simple acts defined on that partition by

Theorem 1. Linearity follows for all simple acts because any pair of simple acts is measurable w.r.t. a joint simple partition. Finally, by standard integration techniques, linearity extends to all bounded acts: each can be "sandwiched" between dominating and dominated simple functions. By continuity, its I value then is the limit of the I values of the limiting simple acts.

Proof of Proposition 3. It is direct that $I = \mathbb{E}_P$ implies the conditions of I. We, therefore, assume those conditions and derive $I = \mathbb{E}_P$.

For any fixed finite partition, am-additivity implies linearity for the simple acts defined on that partition. This follows because in Theorem 1 we can replace continuity at one point by monotonicity, as explained at the end of its proof. Linearity follows for all simple acts because any pair of simple acts is measurable w.r.t. a joint finite partition. To obtain the $\mathbb{E}_P$ representation for all simple acts, define $P(E) = I(1_E)$ for all E, which is nonnegative by monotonicity. It uniquely determines P. We have $P(\Omega) = 1$ because I is normalized. Linearity implies additivity of P, and $I = \mathbb{E}_P$ for all simple functions.

Next, by standard integration techniques, the expectation is extended to all bounded acts: Each can be "sandwiched" between dominating and dominated simple acts. Its I value is the limit of the I values of the limiting simple acts, that is, $\mathbb{E}_P$, as we show in the remainder of this proof. For some $\varepsilon > 0$ and simple acts X and Y, assume $|X(\omega) - Y(\omega)| \leq \varepsilon$ for all ω. Then

$$|I(X) - I(Y)| = |I(X - Y)| \leq I(|X - Y|) \leq I(\varepsilon)$$

by monotonicity, and the latter tends to 0 for ε tending to 0 by linearity of I on simple (including constant) acts.

Proof of Proposition 4. That $\mathbb{E}_P$ satisfies all conditions is immediate. We, hence, show that $I = \mathbb{E}_P$ follows from the other conditions. By Proposition 3, $I = \mathbb{E}_Q$ for an "as-if risk-neutral" Q.[5] For any P-equally-likely partition $E_1, \ldots, E_n$ of Ω, all 1_{E_j} have the same I value by law-basedness. Hence, all E_j have the same Q value, being $1/n$, and $Q = P$ for

[5]It was denoted by P in Proposition 3 but here we denote it differently to distinguish it from the P given beforehand.

all these events and their unions. We are done if Ω is finite. If Ω is infinite and atomless, we have the above partitions for every n. Then $Q = P$ for every E with rational probability $P(E) = j/n$. By monotonicity, this follows for all E, also with irrational P value, by enclosing them with P-rational-probability sub- and supersets.

Proof of Proposition 5. We assume am-affinity and derive affinity under continuity or monotonicity. The reversed implication is trivial.

We may assume $0 \in \mathrm{int}D$ and $I(0) = 0$. To see this point, take a constant $k \in \mathrm{int}D$. Define $D' = D - k$: D' contains all acts resulting from subtracting k from acts in D. Next define I' on D' correspondingly: $I'(X) = I(X + k) - I(k)$. These I' and D' share all relevant properties, including antimonotonicity, with I and D. It suffices to prove our results for I' and D'. We may omit primes.

The functional I is positively homogeneous: For each $0 < \alpha < 1$ and $X \in D$,

$$I(\alpha X) = I(\alpha X + (1 - \alpha)0) = \alpha I(X) + (1 - \alpha)I(0) = \alpha I(X),$$

using antimonotonicity of X and 0.

We next extend I to I^* defined on the whole vector space $B(\Omega, \mathcal{F})$ using positive homogeneity. That is, for each $X \in B(\Omega, \mathcal{F})$ we can find $\lambda > 0$ so small that $\lambda X \in D$, and then define $I^*(X) = I(\lambda X)/\lambda$. By associativity of scalar multiplication, I^* is well-defined (independent of the particular λ chosen) and positively homogeneous. Further, I^* is am-affine because antimonotonicity and am-affinity are compatible with multiplication by a common scalar. We next derive am-additivity of I^*. Consider antimonotonic $X, Y \in B(\Omega, \mathcal{F})$. Using positive homogeneity:

$$I^*(X + Y) = 2I^*(X/2 + Y/2) = 2(I^*(X)/2 + I^*(Y)/2) = I^*(X) + I^*(Y).$$

Am-additivity holds for I^*.

Continuity of I on D implies continuity of I^* on $B(\Omega, \mathcal{F})$, and monotonicity of I similarly extends to I^*. Hence, under continuity, I^* is linear by Proposition 2. Under monotonicity, I^* is linear by Proposition 3 applied to the normalization of I^* (dividing it by $I^*(1)$). Affinity of I^* and I follows.

Proof of Proposition 8. EV directly implies the other conditions. We next assume the other conditions and derive EV. To derive am-additivity of the certainty equivalent (CE) functional (uniquely defined given that $\succcurlyeq$ coincides with $\geq$ on outcomes) , assume X, Y antimonotonic. Then $X \sim \mathrm{CE}(X)$ implies $X + Y \sim \mathrm{CE}(X) + Y$ and $Y \sim \mathrm{CE}(Y)$ implies $Y + \mathrm{CE}(X) \sim \mathrm{CE}(Y) + \mathrm{CE}(X)$. By transitivity, $X + Y \sim \mathrm{CE}(X) + \mathrm{CE}(Y)$. Thus, CE is am-additive. Further, it is monotonic and normalized. By Proposition 3, it is $\mathbb{E}_P$. It represents $\succcurlyeq$.

Proof of Theorem 10. That (iii) implies (i), and (i) implies (ii), is direct. We, therefore, assume (ii) and derive (iii). If all outcomes are indifferent then so are, by monotonicity, all acts and, hence, the result is trivial, with U constant. So, we assume nontriviality. On the outcome set standard mixture independence axioms hold because antimonotonicity does not impose any restriction. By Herstein & Milnor (1953), there exists an affine representation on outcomes. We, until further notice, fix two outcomes $M \succ m$, and consider only acts X with $M \succcurlyeq X(\omega) \succcurlyeq m$ for all ω. By monotonicity and mixture continuity, for each such act there exists a $0 \leq p \leq 1$ such that $pM + (1-p)m \sim X$. By Theorem 4 of Herstein & Milnor (1953), p is uniquely determined and represents preferences over acts. We denote it by $\mathrm{MP}(X)$, the *matching probability* of X. We next show that MP is an expectation representation for all acts.

We write $p^* = pM + (1-p)m$ for all $p \in [0,1]$. The idea of the proof is to replace all outcomes by their equivalent p^*, which by monotonicity does not affect preference, and then by isomorphisms everything follows from preceding results. The switches between isomorphic spaces below involve some notational burden.

We first show that MP is am-affine. Assume X and Y antimonotone and $\alpha \in (0,1)$. Write $p = \mathrm{MP}(X)$ and $q = \mathrm{MP}(Y)$. Now

$$\alpha X + (1-\alpha)Y \sim \alpha p^* + (1-\alpha)Y \sim \alpha p^* + (1-\alpha)q^* = (\alpha p + (1-\alpha)q)^*,$$

where the first two equivalences follow from am-independence and the last equality from affinity of MP on outcomes (also readily and more basically from associativity in mixture spaces). The equality

$$\mathrm{MP}(\alpha X + (1-\alpha)Y) = \alpha p + (1-\alpha)q$$

follows: MP is am-affine.

To invoke Proposition 5, we adjust the domain of MP to become a subset of $B(\Omega, \mathcal{F})$. For each act X, we define $X' : \Omega \to [0,1]$ by $X'(\omega) = \text{MP}(X(\omega))$ for all ω. This X' is measurable because every inverse of a preference interval is an event, and X' is also bounded. Define I by $I(X') = \text{MP}(X)$. This I is well-defined because all X with the same X' are indifferent by monotonicity. This I inherits monotonicity from MP. It is also am-affine: Consider antimonotonic X', Y' and $0 < \alpha < 1$. We take underlying X, Y with $X(\omega) = X'(\omega)^*$ and $Y(\omega) = Y'(\omega)^*$; they are also antimonotonic. For every ω,

$$(\alpha X(\omega) + (1-\alpha)Y(\omega))' = \alpha X'(\omega) + (1-\alpha)Y'(\omega)$$

because MP is affine on outcomes. Hence,

$$I(\alpha X' + (1-\alpha)Y') = \text{MP}(\alpha X + (1-\alpha)Y).$$

By am-affinity of MP, this is $\alpha\text{MP}(X) + (1-\alpha)\text{MP}(Y) = \alpha I(X') + (1-\alpha)I(Y')$; I is am-affine. It is affine by Proposition 5. It is normalized.

By standard techniques (e.g., I's affinity implies strong independence) I is $\mathbb{E}_P$ for a probability measure P, first for all indicator functions, then for all simple X', and then, by monotonicity, for all X'. Because $\text{MP}(X) = I(X')$, MP is the EU functional with MP on outcomes as affine utility function U. We have obtained the desired representation for all acts with outcomes between m and M.

We now turn to acts with outcomes not between m and M. For any other outcomes $M^* \succcurlyeq M \succcurlyeq m \succcurlyeq m^*$ we can similarly obtain an expectation representation. We can rescale all these to take value 0 at m and value 1 at M. They then all agree on common domain and are all part of one expectation functional defined on the whole domain.

Proof of Theorem 13. It is clear that (iii) implies (i), and (i) implies (ii). We, therefore, assume (ii), and derive (iii). Assume, for contradiction, that U is not concave. Then there are outcomes M', m' and $0 < \lambda' < 1$ such that

$$U(\lambda' M' + (1-\lambda')m') < \lambda' U(M') + (1-\lambda')U(m').$$

By mixture continuity, we can find the largest $0 \leq \sigma < \lambda'$ such that $m = \sigma M' + (1-\sigma)m'$ satisfies $U(m) = \sigma U(M') + (1-\sigma)U(m')$ and the smallest $1 \geq \tau > \lambda'$ such that $M = \tau M' + (1-\tau)m'$ satisfies $U(M) = \tau U(M') + (1-\tau)U(m')$. We have

$$\text{for all } 0 < \lambda < 1 : U(\lambda M + (1-\lambda)m) < \lambda U(M) + (1-\lambda)U(m). \tag{11}$$

By non-degenerateness, we can take $A \in \mathcal{F}$ with $0 < P(A) = p < 1$. We write (x, y) for $x_A y$. First assume $(M, m) \sim (m, M)$ and notice that they are antimonotonic. This occurs if $P(A) = 0.5$ or $U(m) = U(M)$. Then, by antimonotonic convexity,

$$((m+M)/2, (m+M)/2) \succcurlyeq (m, M),$$

implying

$$U((m+M)/2) \geq (U(m) + U(M))/2,$$

contradicting Eq. 11. From now on we may assume $U(M) > U(m)$ and $p = P(A) > 0.5$. Otherwise we would interchange M and m, and/or A and A^c. We have $(M, m) \succ (m, M)$. In the remainder of this proof, we will only use outcomes x of the form $x = \lambda M + (1-\lambda)m$ for some λ. We assume without further mention that outcomes are of this form. Mapping x to $\lambda(x)$ provides an isomorphism of the outcome space with the interval $[0, 1]$, used for defining average increases below.

We define $x_0 = m$. By mixture continuity, there exists $m \prec x_1 \prec M$ with $(x_1, m) \sim (x_0, M)$. If there are several such, we take the one closest to m, i.e., we take $x_1 = \lambda M + (1-\lambda)m$ with λ minimal. By mixture continuity we can inductively define a "standard sequence" $m = x_0, x_1, x_2, \ldots, x_n$ such that $(x_{j+1}, m) \sim (x_j, M)$ for all $j < n$ each x_j closest to m so that $\lambda(x_{j+1}) > \lambda(x_j)$. and $(M, m) \prec (x_n, M)$. We have

$$\text{for all } j : U(x_{j+1}) - U(x_j) = \frac{(1-p)(U(M) - U(m))}{p}. \tag{12}$$

We first consider the case $x_n \prec M$. We then similarly define a "standard sequence" $M = y_{n+1}, y_n, y_{n-1} \ldots, y_1$ such that $(y_{j-1}, M) \sim (y_j, m)$ and $\lambda(y_{j-1}) < \lambda(x_{j-1}) < \lambda(y_j)$ and y_j closest to m for all j. We have $(m, M) \succ (y_1, m)$ and

$$\text{for all } j : U(y_j) - U(y_{j-1}) = \frac{(1-p)(U(M) - U(m))}{p}. \tag{13}$$

For every j we have $x_{j-1} \prec y_j \prec x_j$ and, further, there exists a $0 < \lambda < 1$, dependent on j, such that $y_j = \lambda x_j + (1-\lambda)x_{j-1}$. By $m \preccurlyeq x_{j-1} \preccurlyeq x_j \prec M$, we have antimonotonicity of (x_{j-1}, M) and (x_j, m). Antimonotonic convexity and $(x_{j-1}, M) \sim (x_j, m)$ imply

$$\lambda(x_j, m) + (1-\lambda)(x_{j-1}, M) \succcurlyeq (x_j, m) \sim (x_{j-1}, M). \tag{14}$$

We next show that, because the triple of outcomes m, $\lambda M + (1-\lambda)m$, M on A^c in Eq. 14 bring in a kind of strict convexity, the triple x_{j-1}, $y_j (= \lambda x_j + (1-\lambda)x_{j-1})$, x_j on A must bring in a kind of concavity, and enough so to maintain the aforementioned antimonotonic convexity. This point is elaborated on next.

The U value of the left act in Eq. 14 exceeds the U value of the other two acts and, therefore, also the $\lambda/1-\lambda$ convex combination of the latter two U values. That is,

$$\begin{aligned} &pU(\lambda x_j + (1-\lambda)x_{j-1}) + (1-p)U(\lambda m + (1-\lambda)M) \\ &\geq p(\lambda U(x_j) + (1-\lambda)U(x_{j-1})) + (1-p)(\lambda U(m) + (1-\lambda)U(M)). \end{aligned}$$

This and

$$(1-p)U(\lambda m + (1-\lambda)M) < (1-p)(\lambda U(m) + (1-\lambda)U(M))$$

(implied by Eq. 11) imply (dropping p)

$$U(\lambda x_j + (1-\lambda)x_{j-1}) > \lambda U(x_j) + (1-\lambda)U(x_{j-1}).$$

The triple x_{j-1}, y_j (which equals $\lambda x_j + (1-\lambda)x_{j-1}$), and x_j exhibit a kind of concavity.

Using the above isomorphism with $[0,1]$, the aforementioned "concavity" means that the average increase of U over $[x_{j-1}, y_j]$ exceeds that over $[y_j, x_j]$:

$$(U(y_j) - U(x_{j-1}))/\lambda > (U(x_j) - U(y_j))/(1-\lambda).$$

A similar proof shows that the average increase of U over $[y_j, x_j]$ exceeds that over $[x_j, y_{j+1}]$. In this proof, write $x_j = \lambda' y_j + (1-\lambda')y_{j+1}$ and proceed as above with y_j for

x_{j-1}, x_j for y_j, y_{j+1} for x_j, Eq. 13 for Eq. 12, and λ' for λ. The two results together imply that the average increase over an interval decreases as we move from m to M from $[y_j, x_j]$ to $x_j, y_{j+1}]$, to $[y_{j+1}, x_{j+1}]$, and so on.

By Eq. 11, the average U increase over $[m, y_1]$ is strictly below that of $[m, M]$. But we have, just, partitioned that interval $[m, M]$ into $2n + 1$ intervals that all have a strictly smaller average increase than $[m, y_1]$. A contradiction has resulted.

We, finally, turn to the case of $x_n \sim M$. We take $z_0, \ldots, z_{2n}$ such that $z_j = \lambda_j M + (1 - \lambda_j)m$, $z_{2j} = x_j$, $U(z_{2j+1}) = (U(x_j) + U(x_{j+1}))/2$, z_j closest to m, $\lambda_{j+1} > \lambda_j$ for all j. We also define $m' = \lambda M + (1 - \lambda)m$ such that $U(m') = (U(M) + U(m))/2$. By Eq. 11, $\lambda > 0.5$. We have $(z_j, M) \sim (z_{j+1}, m') \sim (z_{j+2}, m)$ for all j. By am-convexity, $\lambda(z_j, M) + (1 - \lambda)(z_{j+2}, m) \succcurlyeq (z_j, M)$ and, hence, $\lambda(z_j, M) + (1 - \lambda)(z_{j+2}, m) \succcurlyeq (z_{j+1}, m')$. Hence, $U(\lambda z_j + (1 - \lambda)(z_{j+2}) \geq (U(z_j) + U(z_{j+2}))/2$ whereas $\lambda < 0.5$. Given that z_{j+1} was chosen closest to m, $\lambda_{j+1} < 0.5\lambda_j + 0.5\lambda_{j+2}$. The average increase of U over $[z_j, z_{j+1}]$ strictly exceeds that over $[z_{j+1}, z_{j+2}]$. This holds for all j. It is in contradiction with the average increase of U over $[z_0, z_1]$ strictly being below that over $[m, M]$ (remember: $z_0 = m$) as follows from Eq. 11.

References

Aczél, J. (1966). *Lectures on functional equations and their applications.* Academic Press, New York.

Anscombe, F. J. and Aumann, R. J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, **34**(1), 199–205.

Aouani, Z., Chateauneuf, A., and Ventura, C. (2021). Propensity for hedging and ambiguity aversion. *Journal of Mathematical Economics*, **97**, 102543.

Bastianello, L. and Faro, J. H. (2023). Choquet expected discounted utility. *Economic Theory*, **75**, 1071–1098.

Beissner, P. and Werner, J. (2022). Optimal allocations with α-MaxMin utilities, Choquet expected utilities, and prospect theory. *Theoretical Economics*, forthcoming.

Castagnoli, E., Cattelan, G., Maccheroni, F., Tebaldi, C. and Wang, R. (2022). Star-shaped risk measures. *Operations Research*, **70**(5), 2637–2654.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Montrucchio, L. (2011). Uncertainty averse preferences. *Journal of Economic Theory*, **146**, 1275–1330.

Chade, H., Eeckhout, J. and Smith, L. (2017). Sorting through search and matching models in economics. *Journal of Economic Literature*, **55**(2), 493–544.

de Finetti, B. (1931). Sul significato soggettivo della probabilitá. *Fundamenta Mathematicae*, **17**, 298–329. Translated into English by Khale, M. as *On the Subjective Meaning of Probability*. In Monari, P. & Cocchi, D. (1993, eds.). *Probabilitá e Induzione*, 291–321. Clueb, Bologna.

Debreu, G. and Koopmans, T. C. (1982). Additively decomposed quasiconvex functions. *Mathematical Programming*, **24**, 1–38.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R. and Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, **31**, 3–33.

Ebert, U. (2004). Social welfare, inequality, and poverty when needs differ. *Social Choice and Welfare*, **23**, 415–448.

Ekeland, I., Galichon, A. and Henry, M. (2012). Comonotonic measures of multivariate risks. *Mathematical Finance*, **22**, 109–132.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, **75**, 643–669.

Fishburn, P. C. (1970). *Utility Theory for Decision Making*. Wiley, New York.

Föllmer, H. and Schied, A. (2016). *Stochastic Finance. An Introduction in Discrete Time*. Fourth Edition. Walter de Gruyter, Berlin.

Galichon, A. (2016). *Optimal Transport Methods in Economics*. Princeton University Press.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, **16**, 65–88.

Gilboa, I. and Marinacci, M. (2016). *Ambiguity and the Bayesian paradigm*. In Arló-Costa, H., Hendricks, V. F., van Benthem, J. F. A. K. (2016). *Readings in Formal Epistemology*, 385–439. Springer, Berlin.

Gilboa, I. and Samuelson, L. (2022). What were you thinking? Decision theory as coherence test. *Theoretical Economics*, **17**, 507–519.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, **18**(2), 141–153.

Grabisch, M. (2016). *Set Functions, Games and Capacities in Decision Making*. Springer, Berlin.

Grabisch, M., Murofushi, T. and Sugeno, M. (2000). *Fuzzy Measures and Integrals: Theory and*

Applications. Physica-Verlag, Berlin.

Hardy, G. H., Littlewood J. E. and Pólya, G. (1934). *Inequalities*. Second edition 1952, reprinted 1978. Cambridge University Press, Cambridge.

Herstein, I. N. and Milnor, J. (1953). An axiomatic approach to measurable utility. *Econometrica*, **21**(2), 291–297.

Hoeffding, W. (1940). Masstabinvariante Korrelationstheorie. *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin*, **5**, 179–233.

l'Haridon, O. and Vieider, F. (2019). All over the map: a worldwide comparison of risk preferences. *Quantitative Economics*, **10**, 185–215.

Li, J. (2020). Preferences for partial information and ambiguity. *Theoretical Economics*, **15**, 1059–1094.

Lu, J. (2021). Random ambiguity. *Theoretical Economics*, **16**, 539–570.

Maccheroni, F., Marinacci, M. and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, **74**(6), 1447–1498.

Machina, M. J. and Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica*, **60**, 745–780.

Mackenzie, A. (2019). A foundation for probabilistic beliefs with or without atoms. *Theoretical Economics*, **14**, 709–778.

Mas-Colell, A., Whinston, D. M., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, New York.

McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton University Press, Princeton.

Puccetti, G. and Wang R. (2015). Extremal dependence concepts. *Statistical Science*, **30**(4), 485–517.

Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons, New York. (Second edition 1972, Dover Publications, New York.)

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, **57**(3), 571–587.

Trautmann, S. T. and van de Kuilen, G. (2015). *Ambiguity attitudes*. In Keren, G. and Wu, G. *The Wiley Blackwell Handbook of Judgment and Decision Making*. (Ch. 3), 89–116. Blackwell, Oxford.

Wakker, P. P. (1993). Unbounded utility for Savage's "Foundations of Statistics," and other models. *Mathematics of Operations Research*, **18**, 446–485.

Wakker, P. P. and Yang, J. (2019). A powerful tool for analyzing concave/convex utility and weighting functions. *Journal of Economic Theory*, **181**, 143–159.

Wakker, P. P. and Yang, J. (2021). Concave/convex weighting and utility functions for risk: a new light on classical theorems. *Insurance: Mathematics and Economics*, **100**, 429–435.

Wang, R., Wei, Y. and Willmot, G. E. (2020). Characterization, robustness and aggregation of signed Choquet integrals. *Mathematics of Operations Research*, **45**(3), 993–1015.

Wang, R. and Zitikis, R. (2021). An axiomatic foundation for the Expected Shortfall. *Management Science*, **67**, 1413–1429.

Yaari, M. E. (1969). Some remarks on measures of risk aversion and on their uses. *Journal of Economic Theory*, **1**, 315–329.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, **55**(1), 95–115.