University of Pretoria
Department of Economics Working Paper Series

Preferences Over Rich Sets of Random Variables: Semicontinuity in Measure versus Convexity

Alexander Zimper
University of Pretoria
Hirbod Assa
University of Liverpool
Working Paper: 2019-40
May 2019

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

Preferences over rich sets of random variables: Semicontinuity in measure versus convexity*

Alexander Zimper† Hirbod Assa‡

May 26, 2019

Abstract

The choice of a continuity concept in decision theoretic models has behavioral meaning because it pins down how the decision maker perceives the similarity of random variables. This paper analyzes the preferences of a decision maker who perceives similarity in accordance with the topology of convergence in measure. As our main insight we show that this decision maker cannot be globally risk- or ambiguity averse whenever her preferences are lower-semicontinuous and complete on a rich set of random variables. Real life decision makers who perceive the similarity of random variables in accordance with convergence in measure might thus account for violations of global convexity as observed in empirical studies. Similarly, the non-convex risk measure *value-at-risk* might be popular among decision makers because it represents preferences that are lower-semicontinuous in measure.

Keywords: Similarity Perceptions; Continuous Preferences; Uncertainty; Ambiguity; Utility Representations; Risk Measures
JEL Classification Numbers: D81

*We are grateful to comments and suggestions from Frank Riedel and Alex Ludwig as well as from seminar participants at the Universities of Bielefeld, Frankfurt, Padova, Udine and the ETH Zürich.

†Department of Economics; University of Pretoria; postal address: Private Bag X20; Hatfield 0028; South Africa; E-mail: alexander.zimper@up.ac.za

‡Institute for Financial and Actuarial Mathematics and Institute for Risk and Uncertainty, University of Liverpool, Center for Doctoral Training, Chadwick Building, G62, Liverpool UK. E-mail: assa@liverpool.ac.uk

1 Introduction

1.1 Continuity and perception of similarity

Continuity of preferences is the fundamental behavioral principle according to which decision makers evaluate similar random variables in a similar way. To be precise, let X and Y be two random variables such that the decision maker strictly prefers Y over X, denoted $X \prec Y$. Upper-semicontinuity below Y implies that $X_n \prec Y$ for any random variables X_n that the decision maker perceives as 'sufficiently similar' to X. On the other hand, lower-semicontinuity above X implies $X \prec Y_n$ for any Y_n 'sufficiently similar' to Y. If the decision maker's choice set is changed from $\{X, Y, ...\}$ to $\{X_n, Y_n, ...\}$, upper- and lower-semicontinuity combined ensure that $X_n \prec Y_n$ whenever these new alternatives are sufficiently similar to the original ones.

The perception of similarity, however, is a subjective notion which might be different for different decision makers. To formally describe similarity perceptions for random variables, we fix throughout this paper the probability space $(\Omega, \mathcal{B}, \mu)$ such that $\Omega = (0, 1)$, $\mathcal{B}$ is the Borel-sigma algebra on the Euclidean interval $(0, 1)$, and μ is the Lebesgue measure. Random variables are functions that map this probability space into the real line. The following example illustrates how similarity of random variables might be perceived differently by different decision makers.

Example 1. Fix the constant random variable X which gives zero in all states of the world, i.e.,

$$X(\omega) = 0 \text{ for all } \omega \in (0, 1).$$

Consider three different sequences of random variables, $\left\{X_n^A\right\}$, $\left\{X_n^B\right\}$, and $\left\{X_n^C\right\}$, which are, respectively, defined for $n \geq 1$ as follows:

$$X_n^A(\omega) = \begin{cases} \frac{1}{n} & \text{if } \omega \in \left(0, \frac{1}{n}\right) & \text{i.e., with prob. } \frac{1}{n} \\ 0 & \text{if } \omega \in \left[\frac{1}{n}, 1\right) & \text{i.e., with prob. } 1 - \frac{1}{n} \end{cases}$$

$$X_n^B(\omega) = \begin{cases} 1 & \text{if } \omega \in \left(0, \frac{1}{n}\right) & \text{i.e., with prob. } \frac{1}{n} \\ 0 & \text{if } \omega \in \left[\frac{1}{n}, 1\right) & \text{i.e., with prob. } 1 - \frac{1}{n} \end{cases}$$

$$X_n^C(\omega) = \begin{cases} n & \text{if } \omega \in \left(0, \frac{1}{n}\right) & \text{i.e., with prob. } \frac{1}{n} \\ 0 & \text{if } \omega \in \left[\frac{1}{n}, 1\right) & \text{i.e., with prob. } 1 - \frac{1}{n} \end{cases}$$

Arguably, all decision makers would agree that the X_n^A become increasingly similar to X with increasing n: the probability of a non-zero outcome converges to zero whereby this non-zero outcome converges itself to zero. There will also be many–but not all–decision makers for whom the X_n^B converge to X since the probability of the (fixed) non-zero outcome converges to zero. Finally, it is also plausible that some decision makers perceive the X_n^C as increasingly similar to X because the probability of the (unboundedly increasing) non-zero outcome converges to zero.□

In mathematical decision theory the similarity of random variables is formally pinned down through the choice of some topology.[1] More specifically, metric topologies capture the similarity between two random variables through a distance function (i.e., metric).

Example 1 revisited. Consider the following three alternative distance functions

$$
\begin{aligned}
d_0(X, X_n) &= \int_\Omega \frac{|X - X_n|}{1 + |X - X_n|} d\mu, \\
d_1(X, X_n) &= \int_\Omega |X - X_n| \, d\mu, \\
d_\infty(X, X_n) &= \inf\{\alpha \in [0, \infty) \mid \mu(|X - X_n| > \alpha) = 0\},
\end{aligned}
$$

generating the topologies of convergence (i) in measure, (ii) in mean, and (iii) in the supremum norm, respectively. The sequence $\{X_n^A\}$ but not the sequences $\{X_n^B\}$ and $\{X_n^C\}$ converges in the supremum norm to X. The sequences $\{X_n^A\}$ and $\{X_n^B\}$ but not the sequence $\{X_n^C\}$ converge in mean to X. All three sequences $\{X_n^A\}$, $\{X_n^B\}$, and $\{X_n^C\}$ converge in measure to X.□

Continuity of preferences is a relative concept that is based on the decision maker's perception of similarity. All three distance functions in the above example correspond to different similarity perceptions that could plausibly be held by different decision makers. If we want to model the continuous preferences of a decision maker for whom all three sequences $\{X_n^A\}$, $\{X_n^B\}$, and $\{X_n^C\}$ converge to X, we can do so in the topology of convergence in measure but not in the topologies of convergence in the supremum norm or in mean, respectively. The choice of a continuity concept in a decision theoretic

[1] A topology imposed on some universal set is the collection of all *open* subsets of this universal set. A random variable Y_n is sufficiently similar to the random variable Y with respect to the chosen topology if Y_n belongs to some sufficiently small open set around Y.

model has thus behavioral meaning because it comes with an assumption about how the decision maker perceives similarity of random variables. Consider for example the following definition of continuous preferences which is standard in axiomatic decision theory.

Standard Continuity Axiom. *For all Z, Z', X, Y, the sets*

$$\{\alpha \in [0,1] \mid \alpha Z + (1-\alpha) Z' \prec Y\} \text{ and } \{\alpha \in [0,1] \mid X \prec \alpha Z + (1-\alpha) Z'\} \quad (1)$$

are open subsets of the Euclidean unit interval.

There is nothing wrong with this standard continuity axiom as long as we are aware that (1) only captures continuity of preferences for a decision maker who perceives the similarity of random variables in accordance with the topology of pointwise convergence. More precisely, let $X \prec Y$ and note that, by (1), the only converging sequences $\{Y_n\}$ that are needed for establishing the lower-semicontinuity of preferences above X are sequences that converge pointwise, i.e.,[2]

$$\{Y_n\} \to Y \text{ if and only if } \lim_{n\to\infty} Y_n(\omega) \to Y(\omega) \text{ for every } \omega \in \Omega.$$

Since the sequences $\{X_n^B\}$ and $\{X_n^C\}$ from Example 1 do not converge pointwise, any decision theoretic model that characterizes continuity of preferences through (1) would, in general, exclude decision makers who have continuous preferences but perceive the similarity of random variables in accordance with, e.g., the convergence in measure or in mean.

1.2 Complete preferences over rich sets

This paper considers a decision maker who perceives similarity of random variables in accordance with the topology of convergence in measure. Whereas our decision maker will, e.g., always satisfy continuity in the topology of pointwise convergence (thereby satisfying the standard continuity axiom) or in the topology of convergence in mean, the converse statement is not true whenever her preferences are complete over some *rich* set of random variables. As a general rule, continuity of preferences over random variables is the harder to establish the more sequences of random variables converge. The analytical

[2] To see this, note that all converging sequences $\{Y_n\} \to Y$ which determine whether the strictly better set at X is open are, by (1), of the form

$$Y_n = \alpha(n) Y + (1-\alpha(n)) Z'$$

such that $\lim_{n\to\infty} \alpha(n) = 1$.

role of a rich set is to allow us the construction of suitable sequences of random variables that converge in measure but not in these alternative topologies.

To be precise, consider the sequence of canonical partitions of the state space Ω into intervals of equal length, denoted $\{\Pi_n\}$, such that

$$\Pi_n = \{\Omega_{1_n}, ..., \Omega_{n_n}\} = \left\{ \left(0, \frac{1}{n}\right), \left[\frac{1}{n}, \frac{2}{n}\right), ..., \left[\frac{n-1}{n}, 1\right) \right\}.$$

Next fix a family $\mathcal{F}$ of reference random variables with the interpretation that we will analyze the semicontinuity of preferences only locally at the reference random variables in $\mathcal{F}$. Define for any pair $X, Y \in \mathcal{F}$ the following random variables for all $\Omega_{i_n} \in \Pi_n$ and all $n \geq 1$:

$$Y_{i_n}(\omega) = \begin{cases} Y(\omega) & \omega \in \Omega \backslash \Omega_{i_n} \\ Y(\omega) + n(X(\omega) - Y(\omega)) & \omega \in \Omega_{i_n}. \end{cases} \tag{2}$$

That is, a random variable Y_{i_n} might differ from Y only on the partition cell Ω_{i_n} which has probability $\mu(\Omega_{i_n}) = \frac{1}{n}$.

A set $R(\mathcal{F})$ of random variables is a rich set generated by $\mathcal{F}$ if it consists, for any pair $X, Y \in \mathcal{F}$, of all the Y_{i_n} defined by (2). Examples of rich sets that are generated by themselves are all standard vectors space of random variables including, e.g., all L^p, $0 \leq p \leq \infty$, as well as the vector space of all simple random variables. But much smaller sets than these standard vector spaces of random variables can be rich sets as shown in the following example.

Example 2. Let $\mathcal{F} = \{X, Y\}$ such that $X(\omega) = 0$ and $Y(\omega) = 1$ for all $\omega \in \Omega$. The constant reference random variables X and Y generate the rich set $R(\mathcal{F})$ which consists of X, Y and of all X_{i_n}, Y_{i_n}, $n \geq 2$, such that

$$X_{i_n}(\omega) = \begin{cases} 0 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} & \text{i.e., with prob. } 1 - \frac{1}{n} \\ n & \text{if } \omega \in \Omega_{i_n} & \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

and

$$Y_{i_n}(\omega) = \begin{cases} 1 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} & \text{i.e., with prob. } 1 - \frac{1}{n} \\ 1 - n & \text{if } \omega \in \Omega_{i_n} & \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

□

Central to our analysis will be the assumption that the decision maker has complete preferences over *some* rich set of random variables such as, e.g., the rich set of Example 2.

1.3 Main results

The crucial feature of rich sets is that the Y_{i_n} converge in measure to Y while X results, for every n, from a convex combination of the Y_{i_n}, $i_n \in \{1, ..., n\}$. Based on this feature we derive the following fundamental incompatibility results.

Incompatibility results for preferences. *Suppose that a decision maker has complete preferences over a rich set of random variables $R(\mathcal{F})$ such that $X \prec Y$ for some $X, Y \in \mathcal{F}$.*

(i) *Preferences which are upper-semicontinuous in measure μ below $Y \in \mathcal{F}$ violate convexity of the strictly worse set at Y.*

(ii) *Preferences which are lower-semicontinuous in measure μ above $X \in \mathcal{F}$ violate convexity of the strictly better set at X.*

Convexity of strictly better sets is central to standard characterizations of global risk aversion or global uncertainty (i.e., ambiguity) aversion. Cerreira-Vioglio et al. (2011, p.1276) write:

> "Convexity reflects a basic negative attitude of decision makers toward the presence of uncertainty in their choices, an attitude arguably shared by most decision makers and modelled through a preference for hedging/randomization."

Our incompatibility result implies that decision makers who have continuous preferences over a rich set of random variables and who perceive similarity of random variables in accordance with convergence in measure cannot globally share such negative attitude towards uncertainty because their preferences must violate convexity of some strictly better set.

Familiar utility specifications that come with the convexity of strictly better sets are, for example, risk averse expected utility decision makers, rank dependent utility decision makers who are strongly risk averse in the sense of Chew et al. (1985) or Chateauneuf et al. (2005), as well as Choquet expected utility (Gilboa 1987; Schmeidler 1989) and multiple priors decision makers (Gilboa and Schmeidler 1989) who are (simply speaking) jointly risk- and ambiguity averse. Our incompatibility results for preferences imply the following incompatibility results for these familiar utility representations.[3]

[3] For a general class of utility representations with convex strictly better sets–including the variational preferences of Maccheroni et al. (2006)–see Cerreira-Vioglio et al. (2011).

Incompatibility results for utility representations. *Consider a non-trivial utility representation*

$$X \preceq Y \Leftrightarrow U(X) \leq U(Y)$$

for preferences over a rich set of random variables. Suppose that the represented preferences are lower-semicontinuous in measure μ. Then the following incompatibility results apply.

(i) Expected utility. *If*

$$U(Z) = \int_{\Omega} u(Z)\, d\pi$$

for an arbitrary additive probability measure π on $(\Omega, \mathcal{B})$, then the Bernoulli utility function u cannot be concave.

(ii) Choquet expected utility. *If*

$$U(Z) = \int_{\Omega}^{Choquet} u(Z)\, d\nu$$

for an arbitrary non-additive probability measure ν on $(\Omega, \mathcal{B})$, then u cannot be concave while ν is convex.

(iii) Maxmin expected utility. *If*

$$U(Z) = \min_{\pi \in \mathcal{P}} \int_{\Omega} u(Z)\, d\pi$$

for an arbitrary set $\mathcal{P}$ of additive probability measures on $(\Omega, \mathcal{B})$, then u cannot be concave.

The above incompatibility results concern standard economic modeling choices for economic applications because non-convex maximization problems are difficult to work with. Experimental studies within the prospect theory framework, however, typically elicit S-shaped Bernoulli utility (i.e., value) functions defined over gains and losses and inversely S-shaped non-additive probability measures (for an overview on this huge literature see Wakker 2010). This empirical evidence from prospect theory would be consistent with the situation that a relevant proportion of the experimental subjects resembled the present paper's decision maker, i.e., a decision maker with continuous preferences who perceives similarity of random variables in accordance with convergence in measure. Although we do not want to overstretch the empirical relevance of our specific decision maker, we would conjecture that empirically observed deviations from convex preferences might partially be explained through decision makers who do not perceive the similarity of random variables in accordance with, e.g., pointwise convergence or convergence in mean.

Next turn to the concept of risk measures which aim to pin down the riskiness of a random variable through some risk number. Let us consider a decision maker who strictly prefers random variables that are strictly less risky in terms of some fixed risk measure ρ. Then our analysis implies the following incompatibility result for continuity in measure versus convexity of risk measures.

Incompatibility result for risk measures. *Consider a non-trivial risk measure representation*

$$X \preceq Y \Leftrightarrow \rho(Y) \leq \rho(X)$$

for preferences over a rich set of random variables. If ρ is a convex risk measure, then these preferences cannot be lower-semicontinuous in measure μ.

The most popular risk measure is value-at-risk which ranks random variables in accordance with their loss quantile at a fixed confidence level. Because value-at-risk is not a convex risk measure it has been heavily criticized in the axiomatic risk measure literature which imposes convexity as a desirable axiom (cf., Artzner et al. 1997, 1998; Föllmer and Schied 2002, 2010; Delbaen 2007, 2009). This literature argues from a normative perspective according to which any risk measure which is used as a regulatory or/and portfolio management criterion should always reward the diversification of portfolios. Our own descriptive perspective on value-at-risk is very different because we offer a possible explanation for the popularity of value-at-risk without any normative judgment. We argue that value-at-risk violates convexity exactly because it represents preferences which are lower-semicontinuous in measure. Decision makers with continuous preferences over rich sets who perceive similarity in accordance with convergence in measure might therefore feel more comfortable with a lower-semicontinuous risk measure like value-at-risk than with some convex (or coherent) risk-measure that violates lower-semicontinuity. We are not saying that it is necessarily 'good' when banking regulators ignore tail risks by choosing value-at-risk as regulatory criterion for Basel capital requirement regulation. But we consider it as plausible that this choice simply reflects the preferences of such regulators because they really don't care about any tail risks.

The remainder of our analysis proceeds as follows. Section 2 reviews and introduces relevant mathematical concepts. While Section 3 analyzes the incompatibility of continuity in measure and convexity for preferences, Section 4 does so for utility representations. Section 5 discusses implications for risk measures. Section 6 concludes. All formal proofs are relegated to the Mathematical Appendix.

2 Mathematical preliminaries

2.1 Convergence in measure

Fix the probability space $(\Omega, \mathcal{B}, \mu)$ such that $\Omega = (0,1)$, $\mathcal{B}$ is the Borel σ-algebra on $(0,1)$, and μ is the Lebesgue measure. A random variable defined on $(\Omega, \mathcal{B}, \mu)$ is a Borel-measurable function $Z : \Omega \to \mathbb{R}$, i.e., for all $A \in \mathcal{B}(\mathbb{R})$,

$$Z^{-1}(A) \in \mathcal{B}$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. We apply the following identity convention for random variables

$$X = Y \text{ if } X(\omega) = Y(\omega), \ \mu\text{-a.e.}$$

That is, we treat two random variables as identical objects if their outcomes coincide except in states belonging to some subset of Ω with Lebesgue measure zero.

Denote by L_0 the set of all random variables defined on the probability space $(\Omega, \mathcal{B}, \mu)$. The results of this paper will be derived for the random variables in some set $L \subseteq L_0$ with the informal interpretation that the random variables in L are somehow 'relevant' to our decision maker.[4] A sequence of random variables $\{Z_n\} \subset L$ converges to $Z \in L$ in measure μ, denoted $Z_n \to_\mu Z$, if and only if, for all $\epsilon > 0$,

$$\lim_{n \to \infty} \mu(\{\omega \in \Omega \mid |Z_n(\omega) - Z(\omega)| > \epsilon\}) = 0.$$

The topology of convergence in measure is generated by the d_0-metric such that for all $X, Y \in L$

$$d_0(X,Y) = \int_\Omega \frac{|X-Y|}{1+|X-Y|} d\mu$$

(cf. Lemma 13.40 in Aliprantis and Border 2006). That is, we have

$$Z_n \to_\mu Z \text{ if and only if } d_0(Z_n, Z) \to 0.$$

In what follows we denote by (L, d_0) our default metric space such that L is endowed with the topology of convergence in measure.

[4]Our preferred interpretation is that the decision maker is 'aware' of the random variables in L. At this point, we do not even require the decision maker to have complete preferences over all random variables in L.

2.2 Vector spaces of random variables and metric topologies

Denote by L^p with $0 < p < \infty$ the vector space of all random variables on $(\Omega, \mathcal{B}, \mu)$ for which the Lebesgue integral

$$\int_\Omega |X|^p \, d\mu$$

is finite. The vector space L^p with $p = 1, 2, 3, ...$ thus consists of all random variables for which the 1st, 2nd, 3rd,... moment exists. Further denote by L^∞ the vector space of all bounded random variables and by L^S the vector space of all simple random variables. The following strict set-inclusions apply

$$L^S \subset L^p \subset L^{p'} \text{ for } 0 \leq p' < p \leq \infty.$$

Fix a pair of random variables and suppose that the alternative distances between X and Y, indexed by p, exist:

$$d_p(X,Y) = \begin{cases} \int_\Omega \frac{|X-Y|}{1+|X-Y|} d\mu & p = 0 \\ \int_\Omega |X-Y|^p \, d\mu & 0 < p < 1 \\ \left(\int_\Omega |X-Y|^p \, d\mu\right)^{\frac{1}{p}} & \text{for } 1 \leq p < \infty \\ \inf\{\alpha \in [0,\infty) \mid \mu(|X-Y| > \alpha) = 0 \text{ a.e.}\} & \text{for } p = \infty \end{cases}$$

The function $d_p : L^{p'} \times L^{p'} \to [0, \infty)$ with $0 \leq p \leq \infty$ is a well-defined metric for any subset L of the vector space $L^{p'}$ whenever $p \leq p'$ (cf. Corollary 13.3 in Aliprantis and Border 2006). That is, we can endow any set of random variables $L \subseteq L^{p'}$ with d_p such that $p \leq p'$ to obtain the metric topology (L, d_p). In particular, the d_0-metric is well defined for all subsets L of random variables including the vector space L^0 of all random variables itself.

Because

$$p' < p \text{ implies } d_{p'}(X,Y) \leq d_p(X,Y)$$

whenever $d_p(X,Y)$ is well-defined (cf. Corollary 13.3 in Aliprantis and Border 2006), we have the following fact.

Fact 1. *Let $0 \leq p' < p \leq \infty$. If a sequence of random variables $\{Z_n\}$ converges to Z in the d_q-metric, then it also converges in the $d_{p'}$-metric. In particular, we have for any d_p-metric*

$$d_p(Z_n, Z) \to 0 \text{ implies } d_0(Z_n, Z) \to 0.$$

The converse statement is not necessarily true (cf. Example 1).

Semicontinuity of preferences is harder to establish on a metric space with more converging sequences than on a metric space with less converging sequences. Preferences that are semicontinuous in our chosen topology (L, d_0) must therefore also be semicontinuous on any metric space (L, d_p) such that $p > 0$.

2.3 Rich sets

Fix a set of references random variables $\mathcal{F}$ whereby we assume that $X \neq Y$ for some $X, Y \in \mathcal{F}$. Recall from the introduction the formal definition of the sequence $\{\Pi_n\}$ of canonical partitions of $\Omega = (0, 1)$ such that, for $n \geq 1$,

$$\Pi_n = \{\Omega_{1_n}, ..., \Omega_{n_n}\} = \left\{ \left(0, \frac{1}{n}\right), \left[\frac{1}{n}, \frac{2}{n}\right), ..., \left[\frac{n-1}{n}, 1\right) \right\}.$$

Denote by $1_{\Omega_{i_n}}$ the indicator function of partition cell Ω_{i_n}, i.e.,

$$1_{\Omega_{i_n}} = \begin{cases} 1 & \omega \in \Omega_{i_n} \\ 0 & \text{else.} \end{cases}$$

Definition. Rich sets. *We say that $R(\mathcal{F})$ is the rich set generated by $\mathcal{F}$ if and only if it consists, for any pair $X, Y \in \mathcal{F}$, of all*

$$Y_{i_n} = Y + n(X - Y) 1_{\Omega_{i_n}}$$

such that $\Omega_{i_n} \in \Pi_n$ for $n \geq 1$.

Let $n = 1$ in the above definition to see that $Y_{1_1} = X$ and $X_{1_1} = Y$ so that it always holds that $\mathcal{F} \subseteq R(\mathcal{F})$. The following fact provides a simple criterion for identifying rich sets that are generated by themselves.

Fact 2. *Suppose that L with $\mathcal{F} \subseteq L$ is a vector space of random variables such that $Z \cdot 1_{\Omega_{i_n}} \in L$ for all $Z \in \mathcal{F}$ and all $\Omega_{i_n} \in \Pi_n$ with $n \geq 1$. Then L is a rich set generated by itself, i.e., $L = R(\mathcal{F}) = \mathcal{F}$.*

Fact 2 follows because $Y + n(X - Y) 1_{\Omega_{i_n}} \in L$ can be constructed from a repeated application of the vector operations addition and scalar multiplication whenever

$$X, Y, X \cdot 1_{\Omega_{i_n}}, Y \cdot 1_{\Omega_{i_n}} \in L.$$

As a consequence of Fact 2, the standard normed vector spaces L^p, $0 \leq p \leq \infty$, as well as L^S of random variables are rich sets generated by themselves because $Z \in L^p$ implies $Z \cdot 1_{\Omega_{i_n}} \in L^p$ for all $\Omega_{i_n} \in \Pi_n$, $n \geq 1$. The same holds for the vector space of all simple random variables, i.e., all random variables with finite support. On the other hand, the vector space of all constant random variables is not a rich set because any rich set must contain some random variables with at least two different outcomes in their supports. Also any set of random variables whose support is on the same bounded subset of the reals cannot be a rich set.

For an example of a 'small' rich set that is not a vector space recall Example 2 from the Introduction where the set of reference random variables consists of two constant random variables. The rich set of Example 2 is not even convex and it only contains random variables with at most two different outcomes in their support.

Remark. In Assa and Zimper (2018) we have considered a decision maker who has complete preferences over the set L^0 of all random variables. Ouf of all the d_p-metrics with $0 \leq p \leq \infty$ only the d_0-metric is well-defined for all random variables in L^0. We thought it therefore natural to conduct the continuity analysis in Assa and Zimper (2018) within the topology of convergence in measure (i.e., for the metric space (L^0, d_0)). To assume complete preferences over all random variables, however, is a strong requirement (cf., e.g., Danan et al. (2015) and references therein for arguments in favor of incomplete preferences). In the present paper we only assume that the decision maker has complete preferences over an arbitrary rich set. Moreover, we now focus on the topology of convergence in measure for the behavioral reason that our decision maker perceives similarity of random variables in accordance with this topology–even in case that other d_p-metrics with $p > 0$ are well-defined for the rich set in question.

2.4 A key mathematical result

A set $A \subseteq L^0$ is *convex* if and only if

$$Y_1, ..., Y_n \in A \text{ implies } \lambda_1 Y_1 + ... + \lambda_n Y_n \in A \text{ for all } \lambda_i \geq 0 \text{ s.t. } \sum_{i=1}^{n} \lambda_i = 1.$$

The following Lemma states the key mathematical result that we will use to prove our incompatibility results.

Lemma 1. *Suppose that $R(\mathcal{F}) \subseteq L$ for some rich set $R(\mathcal{F})$. If $Y \in \mathcal{F}$ belongs to an arbitrary convex and open subset $A \subseteq (L, d_0)$, then all $X \in \mathcal{F}$ must also belong to A.*

Let $\mathcal{F} = R(\mathcal{F}) = L$ be some vector space of random variables. Then Lemma 1 states equivalently that the topological vector space (L, d_0) is *locally non-convex* in the sense that all open balls $B_\varepsilon(Y; d_0)$, $Y \in L$, must be non-convex sets. It is well-known in functional analysis that the null-functional is the only continuous linear functional on a locally non-convex vector space (cf. Theorem 1 in Day 1940). Because the expectation operator is a linear functional, it cannot be continuous on any locally non-convex topological vector space. In particular, we have that the expectation operator is continuous on (L, d_p) for any d_p-metric with $1 \leq p$ but discontinuous for any d_p-metric with $p < 1$ (cf. Section 1.47 in Rudin 1991).[5]

As a generalization of the scope of these familiar results for topological vector spaces, Lemma 1 states that any topological space (L, d_0) is locally non-convex whenever L happens to be some rich set $R(\mathcal{F})$. When we are going to describe in Example 3 a risk-neutral expected utility decision maker with complete preferences over the rich set of Example 2, we will show that these preferences are not continuous as the expectation operator is discontinuous in measure on this rich set.

Remark. In Assa and Zimper (2018) we proved the following result: *The only convex subset of (L^0, d_0) with non-empty interior is the set (L^0, d_0) itself.* Note that this previous result obtains as a special case of Lemma 1 if we specify the set of reference random variables as the whole set all random variables itself, i.e., if we set $\mathcal{F} = R(\mathcal{F}) = L^0$.

3 Incompatibility results for preferences

We consider a preference relation over random variables in $L \subseteq L^0$ with the usual interpretations and conventions. $X \preceq Y$ (weak preference) means: either $X \prec Y$ (strict preference) or $X \sim Y$ (indifference). The strict preference relation $\prec$ is *asymmetric*, i.e., $X \prec Y$ implies not $Y \prec X$. The indifference relation is *symmetric*, i.e., $X \sim Y$ implies $Y \sim X$, as well as *reflexive*, i.e., $X \sim X$. For the results of this section we do not require *transitivity* of $\preceq$. We also do not require *completeness* of $\preceq$ on L but only on some rich set $R(\mathcal{F}) \subseteq L$.

[5] It is possible to extend our incompatibility analysis from (L, d_0) to (L, d_p) spaces with $0 < p < 1$, which are also locally non-convex. However, such extension would require sequences that are generated by different partitions of Ω than just the canonical partitions that we are using for the construction of rich sets.

Introduce the *strictly better* set at X

$$S^*(X) = \{Z \in L | X \prec Z\}.$$

as well as the *strictly worse* set at Y

$$s^*(Y) = \{Z \in L | Z \prec Y\}.$$

Definitions. Semicontinuity of $\preceq$ in measure μ

(i) $\preceq$ *is lower-semicontinous in measure μ above X if and only if the strictly better set $S^*(X)$ is open in (L, d_0).*

(ii) $\preceq$ *is upper-semicontinuous in measure μ below Y if and only if the strictly worse set $s^*(Y)$ is open in (L, d_0).*

Let us give behavioral interpretations of both concepts of semicontinuous preferences whereby we fix $X \prec Y$. Upper-semicontinuity in measure μ below Y means that $X_n \prec Y$ for sufficiently large n whenever the X_n converge in measure μ to X. A decision maker with upper-semicontinuous preferences will thus keep preferring Y over the X_n whenever the X_n are sufficiently similar to X whereby we pin down similarity by convergence in measure. Conversely, a violation of upper-semicontinuity in measure μ below Y implies the existence of some sequence $\{X_n\}$ that converges in measure μ to X such that

$$X \prec Y \preceq X_n$$

for all $n \geq M$ with M being sufficiently large.

Analogously, lower-semicontinuity in measure μ above X means that $X \prec Y_n$ for sufficiently large n whenever the Y_n converge in measure μ to Y. A decision maker with lower-semicontinuous preferences will keep preferring the Y_n over X whenever the Y_n are sufficiently similar in measure μ to Y. A violation of lower-semicontinuity in measure μ above X implies the existence of some sequence $\{Y_n\}$ that converges in measure μ to Y such that

$$Y_n \preceq X \prec Y$$

for all $n \geq M$ with M being sufficiently large.

Theorem 1. *Consider a preference relation $\preceq$ on L which is complete on some rich set $R(\mathcal{F}) \subseteq L$ such that $X \prec Y$ for some $X, Y \in \mathcal{F}$.*

(i) *If $\preceq$ is lower-semicontinuous in measure μ above X, the strictly better set $S^*(X)$ cannot be convex.*

(ii) *If $\preceq$ is upper-semicontinuous in measure μ below Y, the strictly worse set $s^*(Y)$ cannot be convex.*

Sketch of the proof. Based on a straightforward application of Lemma 1 we formally prove Theorem 1 in the Appendix. To see the intuition behind these fundamental incompatibility results, let us briefly sketch the basic proof idea of Theorem 1. We use two facts about the

$$Y_{i_n} = Y + n(X - Y) 1_{\Omega_{i_n}}.$$

Firstly, the Y_{i_n} converge in measure to Y whereby for any $\varepsilon > 0$ there is some large enough n such that $d_0(Y_{i_n}, Y) < \varepsilon$ for all $i_n \in \{1, ..., n\}$. Secondly, by construction,

$$\frac{1}{n}\sum_{i_n=1}^{n} Y_{i_n} = X \text{ for all } n.$$

Consequently, by lower-semicontinuity above X, we have $X \prec Y_{i_n}$ for all Y_{i_n} that are sufficiently similar to Y in the topology of convergence in measure. If the strictly better set $S^*(X)$ was convex, we would then obtain for large enough n the contradiction

$$X \prec \frac{1}{n}\sum_{i_n=1}^{n} Y_{i_n} = X.$$

This proves part (i).

Part (ii) follows because upper-semicontinuity below Y implies $X_{i_n} \prec Y$ for all X_{i_n} for sufficiently large n. If the strictly worse set at Y was convex, we would then obtain the contradiction

$$Y = \frac{1}{n}\sum_{i_n=1}^{n} X_{i_n} \prec Y.$$

□

4 Incompatibility results for utility representations

Suppose now that there exists an utility representation for given preferences. That is, there exists some $U : L \to \mathbb{R}$ such that, for all $X, Y \in L$,

$$\begin{aligned} X &\prec Y \Leftrightarrow U(X) < U(Y), \\ X &\sim Y \Leftrightarrow U(X) = U(Y). \end{aligned}$$

The utility function U is continuous in measure μ at $Z \in L$ if and only if, for every $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$Z_n \in B_\delta (Z; d_0) \text{ implies } |U(Z) - U(Z_n)| < \varepsilon.$$

The corresponding definitions of upper- and lower-semicontinuity of U are given as follows.

Definition: Lower- and upper-semicontinuity of U.

(i) *U is lower-semicontinuous in measure μ at Y if and only if, for every $\varepsilon > 0$, there exists some $\delta > 0$ such that*

$$Y_n \in B_\delta (Y; d_0) \text{ implies } U(Y_n) > U(Y) - \varepsilon. \quad (3)$$

(ii) *U is upper-semicontinuous in measure μ at X if and only if, for every $\varepsilon > 0$, there exists some $\delta > 0$ such that*

$$X_n \in B_\delta (X; d_0) \text{ implies } U(X_n) < U(X) + \varepsilon.$$

We say that U is lower-semicontinuous (resp. upper-semicontinuous) whenever U is lower-semicontinuous (resp. upper-semicontinuous) at all $Z \in L$. The following proposition (proved in the Appendix) establishes that any utility representation U which is lower-semicontinuous (resp. upper-semicontinuous) must represent preferences that are lower-semicontinuous above all $Z \in L$ (resp. upper-semicontinuous below all $Z \in L$).

Proposition 1.

(i) *Suppose that $\preceq$ violates lower-semicontinuity in measure μ above X. Then U violates lower-semicontinuity in measure μ at some Y such that $X \prec Y$.*[6]

(ii) *Suppose that $\preceq$ violates upper-semicontinuity in measure μ below some Y. Then U violates upper-semicontinuity in measure μ at some X such that $X \prec Y$.*

[6] The converse statement is, in general, not true. A violation of lower-semicontinuity of U at some $Y \in L$ implies that the strictly better set $S^*(c) = \{Z \in L \mid c < U(Z)\}$ cannot be open for some $c \in \mathbb{R}$ (cf., Theorem 1, p.76 in Berge 1996). However, we do not always have that $c = U(X)$ for some $X \in \mathcal{F}$.

Next we extend the familiar definitions of concave versus convex functions whose domains are convex subsets of the real line to utility functions whose domains are convex sets of random variables.

Definitions: Concavity versus convexity of U. *Let L be a convex set.*

(i) *U is concave on L if and only if, for all $X, Y \in L$ and all $\lambda \in (0, 1)$,*

$$U(\lambda X + (1 - \lambda) Y) \geq \lambda U(X) + (1 - \lambda) U(Y).$$

(ii) *U is convex on L if and only if, for all $X, Y \in L$ and all $\lambda \in (0, 1)$,*

$$U(\lambda X + (1 - \lambda) Y) \leq \lambda U(X) + (1 - \lambda) U(Y).$$

Proposition 2. *Let L be a convex set.*

(i) *If U is concave on L, then the strictly better set $S^*(Z)$ is convex for all $Z \in L$.*

(ii) *If U is convex on L, then the strictly worse set $s^*(Z)$ is convex for all $Z \in L$.*

Combining Theorem 1 with Propositions 1 and 2 gives us the following incompatibility results for utility representations.

Theorem 2. *Consider a preference relation $\preceq$ on a convex set L which is complete on some rich set $R(\mathcal{F}) \subseteq L$ such that $X \prec Y$ for some $X, Y \in \mathcal{F}$.*

(i) *Suppose that U is concave on L. Then U cannot be lower-semicontinuous in measure μ at Y. More precisely, there must exist some sequence $\{Y_{i_n}\} \to_\mu Y$ on $\mathcal{F}$ such that*

$$\lim_{n \to \infty} U(Y_{i_n}) \leq U(X) < U(Y).$$

(ii) *Suppose that U is convex on L. Then U cannot be upper-semicontinuous in measure μ at X. More precisely, there must exist some sequence $\{X_{i_n}\} \to_\mu X$ on $\mathcal{F}$ such that*

$$\lim_{n \to \infty} U(X_{i_n}) \geq U(Y) > U(X).$$

4.1 Operators for utility random variables

Fix some increasing Bernoulli utility function $u : \mathbb{R} \to \mathbb{R}$. Recall that u is concave if and only if, for all $x, y \in \mathbb{R}$ and all $\lambda \in (0,1)$,

$$u(\lambda x + (1-\lambda) y) \geq \lambda u(x) + (1-\lambda) u(y). \tag{4}$$

For convex u the inequality in (4) is reversed.

Let $Z \in L$ and note that $u(Z) : \Omega \to \mathbb{R}$ such that

$$u(Z)(\omega) = u(Z(\omega))$$

is itself a random variable defined on $(\Omega, \mathcal{B})$. We refer to $u(Z)$ as utility random variable. For a given set L of random variables, introduce the following set of utility random variables

$$L_u = \{u(Z) \mid Z \in L\}$$

and denote by $co(L_u)$ the convex hull of L_u, i.e., the set of all utility random variables that are convex combinations of the utility random variables in L_u.

An operator on $co(L_u)$, denoted I, is a mapping $I : co(L_u) \to \mathbb{R}$. The operator I satisfies monotonicity on $co(L_u)$ if and only if, for all $u(Z), u(Z') \in co(L_u)$,

$$u(Z(\omega)) \leq u(Z'(\omega)) \text{ for all } \omega \text{ implies } I(u(Z)) \leq I(u(Z')).$$

Definitions: Concavity and convexity of I.

(i) *I is concave on $co(L_u)$ if and only if, for all $X, Y \in L$ and all $\lambda \in (0,1)$,*

$$I(\lambda u(X) + (1-\lambda) u(Y)) \geq \lambda I(u(X)) + (1-\lambda) I(u(Y)).$$

(ii) *I is convex on $co(L_u)$ if and only if, for all $X, Y \in L$ and all $\lambda \in (0,1)$,*

$$I(\lambda u(X) + (1-\lambda) u(Y)) \leq \lambda I(u(X)) + (1-\lambda) I(u(Y)).$$

Proposition 3. *Let L be a convex set and assume that, for all $Z \in L$,*

$$U(Z) = I(u(Z)) \tag{5}$$

for some operator I on $co(L_u)$.

(i) *Suppose that u is concave, I satisfies monotonicity and concavity on co* (L_u)*. Then* U *is concave on* L*.*

(ii) *Suppose that u is convex, I satisfies monotonicity and convexity on co* (L_u)*. Then* U *is convex on* L*.*

Combining Theorem 2 with Proposition 3 gives us the following results.

Theorem 3. *Consider a preference relation* $\preceq$ *on a convex set* L *which is complete on some rich set* $R(\mathcal{F}) \subseteq L$ *such that* $X \prec Y$ *for some* $X, Y \in \mathcal{F}$*. Suppose that these preferences have a utility representation which is of the operator form (5).*

(i) *If* u *is concave,* I *satisfies monotonicity and concavity on co* (L_u)*, then* U *cannot be lower-semicontinuous in measure* μ *at* Y*.*

(ii) *If* u *is convex,* I *satisfies monotonicity and convexity on co* (L_u)*, then* U *cannot be upper-semicontinuous in measure* μ *at* X*.*

The next subsection applies Theorem 3 to the standard utility representations *expected utility*, *Choquet expected utility*, and *multiple priors expected utility*, respectively.

4.2 Standard utility representations

Suppose now that I is an operator defined on an arbitrary vector space V. The operator I is called *superlinear* on V if it satisfies the following two properties:

(i) Positive Homogeneity: for all $\alpha \geq 0$ and all $v \in V$,

$$I(av) = aI(v),$$

(ii) Superadditivity: for all $v, v' \in V$,

$$I(v + v') \geq I(v) + I(v').$$

The operator I is *sublinear* on V if superadditivity is replaced with *subadditivity* (i.e., for all $v, v' \in V$, $I(v + v') \leq I(v) + I(v')$).

Standard utility representations for preferences over random variables are of the form (5) such that I stands for a specific concept of an expectation operator defined on a suitable vector space of utility random variables V that includes $co\left(L_u\right)$. In what follows, we derive a string of corollaries to Theorem 3 under the assumption that the operator I in (5) takes on specific functional forms discussed in the literature. All these corollaries assume complete preferences on an arbitrary rich set $R\left(\mathcal{F}\right)$ such that $X \prec Y$ for some $X, Y \in \mathcal{F}$; (in particular, we rule out trivial preferences according to which $X \sim Y$ for all $X, Y \in \mathcal{F}$).

Suppose, at first, that I is the standard expectations operator with respect to some additive probability measure π. Because the expectation operator satisfies monotonicity and is super- as well as sublinear (i.e. linear), we obtain the following result.

Corollary 1. *Suppose that U is of the expected utility form, i.e., for all $Z \in L$,*

$$U\left(Z\right) = I\left(u\left(Z\right)\right) = \int_{\Omega} u\left(Z\right) d\pi$$

for an arbitrary additive probability measure π defined on $\left(\Omega, \mathcal{B}\right)$.

(i) *If U is lower-semicontinuous in measure μ, the Bernoulli utility function u cannot be concave.*

(ii) *If U is upper-semicontinuous in measure μ, the Bernoulli utility function u cannot be convex.*

(iii) *If U is continuous in measure μ, the Bernoulli utility function u cannot be linear.*

Let us illustrate Corollary 1 for the simplest example of a rich set we can think of.

Example 3. Consider the rich set $R\left(\mathcal{F}\right)$ of Example 2 which consists of $X = 0$, $Y = 1$ and of all X_{i_n}, Y_{i_n}, $n \geq 2$, such that

$$X_{i_n}\left(\omega\right) = \begin{cases} 0 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} \quad \text{i.e., with prob. } 1 - \frac{1}{n} \\ n & \text{if } \omega \in \Omega_{i_n} \quad \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

and

$$Y_{i_n}\left(\omega\right) = \begin{cases} 1 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} \quad \text{i.e., with prob. } 1 - \frac{1}{n} \\ 1 - n & \text{if } \omega \in \Omega_{i_n} \quad \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

If the Bernoulli utility function u was concave, U cannot be lower-semicontinuous in measure μ at Y; that is, for some converging sequence $\left\{Y_{i_n}\right\} \rightarrow_{\mu} Y$ we must have that

$$\lim_{n \rightarrow \infty} \int_{\Omega} u\left(Y_{i_n}\right) d\pi \leq \int_{\Omega} u\left(X\right) d\pi < \int_{\Omega} u\left(Y\right) d\pi.$$

Conversely, for a convex u there must exist some converging sequence $\{X_{i_n}\} \rightarrow_\mu X$ such that

$$\lim_{n\to\infty} \int_\Omega u(X_{i_n})\, d\pi \geq \int_\Omega u(Y)\, d\pi > \int_\Omega u(X)\, d\pi.$$

To see this for the special case $u(x) = x$ and $\pi = \mu$, observe that for all converging sequences

$$\lim_{n\to\infty} \int_\Omega Y_{i_n} d\mu = 0 \text{ and } \lim_{n\to\infty} \int_\Omega X_{i_n} d\mu = 1$$

while

$$\int_\Omega X d\mu = 0 \text{ and } \int_\Omega Y d\mu = 1.$$

□

An immediate consequence of Corollary 1 is that there cannot exist a risk-neutral expected utility decision maker whose preferences are continuous in measure μ on some rich set of random variables (cf. Example 3 where $u(x) = x$ is equivalent to risk-neutrality). Next recall that a sequence of random variables $\{Z_n\}$ converges to Z in distribution[7], denoted $\mu_{Z_n} \rightarrow \mu_Z$, if and only if

$$\int_{\mathbb{R}} u(x)\, d\mu_{Z_n} \rightarrow \int_{\mathbb{R}} u(x)\, d\mu_Z$$

for all bounded and μ-almost everywhere continuous functions $u : \mathbb{R} \rightarrow \mathbb{R}$ (cf. Theorem 25.8. in Billingsley 1996). Because convergence in measure of the expected utility representation implies convergence in distribution for this utility representation, a Bernoulli utility function that is bounded from above and from below (μ-almost everywhere) guarantees that expected utility preferences are continuous in measure. In other words, a bounded Bernoulli utility function, which takes on an S-shape over the real line, is thus always a sufficient condition for continuity in measure μ of the expected utility representation of complete preferences on a rich set.

Turn now to the concept of Choquet expected utility for which I in (5) becomes the Choquet expectation operator with respect to some non-additive probability measure ν

[7] The distribution μ_Z of random variable Z is the probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\mu_Z(A) = \mu(\{\omega \in \Omega \mid Z(\omega) \in A\}) \text{ for all } A \in \mathcal{B}(\mathbb{R}).$$

(Schmeidler 1989). The Choquet expectation operator satisfies monotonicity. Moreover, it is superlinear for any convex ν, i.e., for any ν such that, for all $A, B \in \mathcal{B}$,

$$\nu (A \cup B) + \nu (A \cap B) \geq \nu (A) + \nu (B) \tag{6}$$

(cf. Corollary and Proposition 3 in Schmeidler 1986). In contrast, it is sublinear for any concave ν (i.e., for any ν such that inequality (6) is reversed).

Corollary 2. *Suppose that U is of the Choquet expected utility form, i.e., for all $Z \in L$,*

$$\begin{aligned} U(Z) &= I(u(Z)) = \int_{\Omega}^{Choquet} u(Z)\, d\nu \\ &= \int_0^{\infty} \nu (u(Z) \geq x)\, dx - \int_{-\infty}^{0} (1 - \nu (u(Z) \geq x))\, dx \end{aligned}$$

for an arbitrary non-additive probability measure ν defined on $(\Omega, \mathcal{B})$.

(i) *If U is lower-semicontinuous in measure μ, we cannot simultaneously have that u is concave while ν is convex.*

(ii) *If U is upper-semicontinuous in measure μ, we cannot simultaneously have that u is convex while ν is concave.*

Choquet expected utility (CEU) theory uses convex non-additive probability measures to describe ambiguity averse decision makers. To express a behavioral relevant combination of ambiguity aversion with (standard) risk aversion, the typical modeling choice for an CEU decision maker combines a convex non-additive probability measure with a concave Bernoulli utility function. By Corollary 2(i) such CEU decision maker cannot have non-trivial preferences on a rich set of random variables that are lower-semicontinuous in measure μ.

Finally, turn to the concept of multiple priors expected utility where the expectation operator I is defined with respect to a set of additive probability measures (i.e., multiple priors). Recall that I satisfies monotonicity and superlinearity if I is the minimal expectation operator whereas I satisfies monotonicity and sublinearity if I is the maximal expectation operator (cf. Lemma 3.3. in Gilboa and Schmeidler 1989).

Corollary 3.

(i) *Suppose that U is of the maxmin expected utility form, i.e., for all $Z \in L$,*

$$U(Z) = I(u(Z)) = \min_{\pi \in \mathcal{P}} \int_{\Omega} u(Z)\, d\pi$$

for some set $\mathcal{P}$ of additive probability measures on $(\Omega, \mathcal{B})$. If U is lower-semicontinuous in measure μ, the Bernoulli utility function u cannot be concave.

(ii) *Suppose that U is of the maxmax expected utility form, i.e., for all $Z \in L$,*

$$U(Z) = I(u(Z)) = \max_{\pi \in \mathcal{P}} \int_{\Omega} u(Z)\, d\pi$$

for some set $\mathcal{P}$ of additive probability measures on $(\Omega, \mathcal{B})$. If U is upper-semicontinuous in measure μ, the Bernoulli utility function u cannot be convex.

Multiple priors models express ambiguity aversion through maxmin expected utility.[8] By Corollary 3(i), the typical modeling choice, which combines maxmin expected utility with a concave Bernoulli utility function, cannot describe a decision maker have non-trivial preferences on a rich set of random variables that are lower-semicontinuous in measure μ.

5 Incompatibility results for risk measures

This section considers a decision maker who ranks random variables in accordance with some risk measure $\rho : L \to \mathbb{R}$ such that

$$\begin{aligned} X &\prec Y \Leftrightarrow \rho(Y) < \rho(X), \\ X &\sim Y \Leftrightarrow \rho(Y) = \rho(X). \end{aligned} \tag{7}$$

The interpretation is that the decision maker prefers less risky to more risky random variables perceives whereby she perceives the riskiness of random variables in accordance with ρ. Whenever (7) holds for some risk measure ρ, we will speak of a decision maker with ρ-preferences.

[8]To see the formal relationship between the CEU- and the mutliple priors representation of ambiguity aversion, observe that for a convex ν

$$\int_{\Omega}^{Choquet} u(Z)\, d\nu = \min_{\pi \in \mathcal{P}} \int_{\Omega} u(Z)\, d\pi$$

where $\mathcal{P}$ is defined as the core of ν.

The most fundamental property that any risk measure should satisfy is monotonicity, i.e., for all $Z, Z' \in L$,

$$Z(\omega) \leq Z'(\omega) \text{ for all } \omega \text{ implies } \rho(Z) \geq \rho(Z').$$

The axiomatic literature on risk measures additionally imposes convexity as another fundamental property to ensure that the diversification of a portfolio can never increase risk.[9]

Definition: Convexity of risk measures. *Let L be a convex set. The risk measure ρ is convex on L if and only if, for all $X, Y \in L$ and all $\lambda \in (0,1)$,*

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y).$$

Obviously, any ρ-preferences (7) could be equivalently represented by the utility function $U : L \to \mathbb{R}$ such that, for all $Z \in L$,

$$U(Z) = -\rho(Z). \tag{8}$$

Because of (8) our incompatibility analysis for utility representations carries immediately over to convex risk measures.[10]

Theorem 4. *Assume that L is a convex set that contains an arbitrary rich set $R(\mathcal{F})$ such that $X \prec Y$ for some $X, Y \in \mathcal{F}$. If ρ is convex on L, ρ-preferences cannot be lower-semicontinuous in measure μ at Y. More precisely, there must exist some sequence $\{Y_{i_n}\} \to_\mu Y$ on $\mathcal{F}$ such that*

$$\lim_{n\to\infty} \rho(Y_{i_n}) \geq \rho(X) > \rho(Y). \tag{9}$$

The most prominent risk measure used by financial practitioners is the value-at-risk criterion which happens to violate convexity. In what follows, we argue that the value-at-risk criterion must violate convexity on rich sets because it represents preferences that are lower-semicontinuous in measure.

[9] Coherent risk measures, defined on some vector space $L \subseteq L^0$, have to satisfy positive homogeneity and subadditivity which implies convexity.

[10] Note that lower-semicontinuity of U becomes, by (8), upper-semicontinuity of ρ.

5.1 Value-at-risk

The value-at-risk of random variable $Z \in L^0$ at confidence level α is an α-quantile of Z. Recall that the α-quantiles of Z are the members of the interval

$$\left[q_Z^-(\alpha), q_Z^+(\alpha)\right] \tag{10}$$

such that

$$\begin{aligned} q_Z^-(\alpha) &= \sup\{x \in \mathbb{R} \mid \mu(Z < x) < \alpha\}, \\ q_Z^+(\alpha) &= \sup\{x \in \mathbb{R} \mid \mu(Z < x) \leq \alpha\}. \end{aligned}$$

For interpretational reasons it is convenient to define the value-at-risk of Z in terms of the distribution function of the corresponding loss random variable $-Z$ (whose positive outcomes stand for the losses of Z).

Definitions. Value-at-risk. *Consider the following two alternative definitions for the value-at-risk of random variable Z at confidence level $\alpha \in (0,1)$.*

(i)

$$\begin{aligned} \mathrm{VaR}_{<\alpha}(Z) &= -q_Z^-(\alpha) \\ &= \inf\{x \in \mathbb{R} \mid \mu(x < -Z) < \alpha\} \\ &= \inf\{x \in \mathbb{R} \mid F_{-Z}(x) > 1-\alpha\}. \end{aligned}$$

(ii)

$$\begin{aligned} \mathrm{VaR}_{\leq\alpha}(Z) &= -q_Z^+(\alpha) \\ &= \inf\{x \in \mathbb{R} \mid \mu(x < -Z) \leq \alpha\} \\ &= \min\{x \in \mathbb{R} \mid F_{-Z}(x) \geq 1-\alpha\}. \end{aligned}$$

For any Z with a continuous distribution function both value-at-risk definitions coincide, i.e.,

$$\mathrm{VaR}_{<\alpha}(Z) = \mathrm{VaR}_{\leq\alpha}(Z) = \min\{x \in \mathbb{R} \mid F_{-Z}(x) \geq 1-\alpha\}. \tag{11}$$

In general, (11) holds for any given Z for almost all confidence levels $\alpha \in (0,1)$ because there are at most countably many discontinuity points in the distribution function at which the quantile interval (10) might not reduce to a single value (cf. Lemma A.19. in Föllmer and Schiedt 2016). Whenever the equality (11) holds for a random variable Z, both value-at-risk definitions $\text{VaR}_{<\alpha}$ and $\text{VaR}_{\leq\alpha}$ are continuous in measure μ at Z. If (11) is violated at some Z, however, $\text{VaR}_{<\alpha}$ will result for some converging sequences in an 'upward jump' whereas $\text{VaR}_{\leq\alpha}$ will result for some sequences in a 'downward jump' at Z. The following example illustrates these possible discontinuities for both value-at-risk definitions for the non-generic case in which (11) is violated.

Example 4. Consider Z such that

$$Z(\omega) = \begin{cases} -1 & \text{if } \omega \in (0, \alpha] \quad \text{i.e., with prob. } \alpha \\ 0 & \text{if } \omega \in (\alpha, 1) \quad \text{i.e., with prob. } 1 - \alpha \end{cases}$$

and the following two sequences $\{Z_n^+\}$, $\{Z_n^-\}$ such that

$$\begin{aligned} Z_n^+(\omega) &= \begin{cases} -1 & \text{if } \omega \in \left(0, \alpha + \frac{1}{n}\right] \quad \text{i.e., with prob. } \alpha + \frac{1}{n} \\ 0 & \text{if } \omega \in \left(\alpha + \frac{1}{n}, 1\right) \quad \text{i.e., with prob. } 1 - \left(\alpha + \frac{1}{n}\right) \end{cases} \\ Z_n^-(\omega) &= \begin{cases} -1 & \text{if } \omega \in \left(0, \alpha - \frac{1}{n}\right] \quad \text{i.e., with prob. } \alpha - \frac{1}{n} \\ 0 & \text{if } \omega \in \left(\alpha - \frac{1}{n}, 1\right) \quad \text{i.e., with prob. } 1 - \left(\alpha - \frac{1}{n}\right) \end{cases} \end{aligned}$$

Both sequences $\{Z_n^+\}$ and $\{Z_n^-\}$ converge in measure μ to Z. Note that

$$\text{VaR}_{<\alpha}(Z) = 1 \text{ and } \text{VaR}_{\leq\alpha}(Z) = 0$$

while, for all n,

$$\begin{aligned} \text{VaR}_{<\alpha}\left(Z_n^+\right) &= \text{VaR}_{\leq\alpha}\left(Z_n^+\right) = 1, \\ \text{VaR}_{<\alpha}\left(Z_n^-\right) &= \text{VaR}_{\leq\alpha}\left(Z_n^-\right) = 0, \end{aligned}$$

implying

$$\begin{aligned} \lim_{n\to\infty} \text{VaR}_{<\alpha}\left(Z_n^-\right) &< \text{VaR}_{<\alpha}(Z), \\ \lim_{n\to\infty} \text{VaR}_{\leq\alpha}\left(Z_n^+\right) &> \text{VaR}_{\leq\alpha}(Z). \end{aligned}$$

□

In the non-generic case that (11) is violated for a given Z, $\text{VaR}_{<\alpha}$ is an upper-semicontinuous function in measure μ at Z whereas $\text{VaR}_{\leq\alpha}$ is a lower-semicontinuous function in measure μ at Z. Because an upper-(lower)semicontinuous risk measure function corresponds, by (8), to a lower-(upper)semicontinuous utility representation, $\mathrm{VaR}_{<\alpha}$-preferences are lower-semicontinuous at every Z whereas $\mathrm{VaR}_{\leq\alpha}$-preferences are lower-semicontinuous at Z if and only if the generic case (11) holds for Z.

Proposition 4. *Consider a rich set $R(\mathcal{F})$ such that $X \prec Y$ for some $X, Y \in \mathcal{F}$.*

(i) *If the decision maker has complete $\mathrm{VaR}_{<\alpha}$ preferences over $R(\mathcal{F})$, then the strictly better set at X cannot be convex.*

(ii) *If the decision maker has complete $\mathrm{VaR}_{\leq\alpha}$ preferences over $R(\mathcal{F})$ such that the generic case*

$$\mathrm{VaR}_{<\alpha}(Y) = \mathrm{VaR}_{\leq\alpha}(Y)$$

holds, then the strictly better set at X cannot be convex.

We illustrate Proposition 4 through a simple example.[11]

Example 5. Let $X = -1$ and $Y = 0$, i.e., X gives a constant loss of one while Y gives a constant loss of zero. The rich set $R(\mathcal{F})$ generated by $\mathcal{F} = \{X, Y\}$ consists of X, Y and of all X_{i_n} and Y_{i_n}, $n \geq 2$, such that

$$X_{i_n}(\omega) = \begin{cases} -1 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} \quad \text{i.e., with prob. } 1 - \frac{1}{n} \\ n-1 & \text{if } \omega \in \Omega_{i_n} \quad \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

and

$$Y_{i_n}(\omega) = \begin{cases} 0 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} \quad \text{i.e., with prob. } 1 - \frac{1}{n} \\ -n & \text{if } \omega \in \Omega_{i_n} \quad \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

Fix some $\alpha \in (0, 1)$ and observe that

$$\mathrm{VaR}_\alpha(X) = 1 \text{ and } \mathrm{VaR}_\alpha(Y) = 0$$

as well as

$$\mathrm{VaR}_\alpha(Y_{i_n}) = \begin{cases} 0 & \text{if } n > \frac{1}{\alpha} \\ n & \text{if } n \leq \frac{1}{\alpha} \end{cases}$$

In accordance with Proposition 4, these VaR_α-preferences must satisfy lower-semicontinuity in measure μ above X. To see this, note that, for all $\{Y_{i_n}\} \to_\mu Y$,

$$\mathrm{VaR}_\alpha(Y_{i_n}) = \mathrm{VaR}_\alpha(Y) = 0 \text{ for } n > \frac{1}{\alpha}.$$

[11] We write $\mathrm{VaR}_\alpha(Z)$ whenever (11) holds for Z.

Consequently, the strictly better set at X cannot be convex. To verify this directly, observe that convexity of the strictly better set at X would result, by

$$\begin{array}{ccc} \mathrm{VaR}_\alpha(X) & > & \mathrm{VaR}_\alpha(Y_{i_n}) \\ & \Leftrightarrow & \\ X & \prec & Y_{i_n} \end{array}$$

for all $i_n \in \{1, ..., n\}$ with $n > \frac{1}{\alpha}$, in the contradiction

$$X \prec \left(\frac{1}{n} \sum_{i_n=1}^{n} Y_{i_n} \right) = X.$$

□

Following Dekel (1989) we say that transitive preferences on L 'exhibit diversification' if

$$Z^1 \sim \cdots \sim Z^n$$

implies

$$Z^1 \preceq \sum_{i=1}^{n} \lambda_i Z^i \text{ for all } \lambda_i \geq 0 \text{ such that } \sum_{i_n=1}^{n} \lambda_i = 1.$$

Let us revisit Example 5 to illustrate how value-at-risk preferences may result in the choice of non-diversified portfolios.

Example 6. Portfolio choice. Consider a portfolio manager with VaR_α-preferences over the rich set from Example 4. For any $m \in \mathbb{N}$ such that

$$\alpha < \frac{m}{n} \leq 1$$

construct the mixed portfolio $\frac{1}{m} \sum_{i=1}^{m} Y_{i_n}$ and observe that

$$\mathrm{VaR}_\alpha \left(\frac{1}{m} \sum_{i=1}^{m} Y_{i_n} \right) = n.$$

On the other hand, we have for all $n > \frac{1}{\alpha}$ that

$$\mathrm{VaR}_\alpha(Y_{i_n}) = 0 \text{ for all } i,$$

implying

$$\mathrm{VaR}_\alpha\left(Y_{i_n}\right) < \mathrm{VaR}_\alpha\left(\frac{1}{m}\sum_{i=1}^{m} Y_{i_n}\right) \text{ for all } i.$$

These VaR_α-preferences do not 'exhibit diversification' to the effect that the portfolio manager would always choose any non-diversified Y_{i_n} over the diversified portfolio $\frac{1}{m}\sum_{i=1}^{m} Y_{i_n}$.□

5.2 Discussion

By Proposition 4, value-at-risk cannot be a convex risk measure on any rich set because it represents preferences that are (generically) lower-semicontinuous in measure μ. In Example 5 this lower-semicontinuity of VaR_α-preferences is expressed through the decision maker's indifference between all Y_{i_n} for sufficiently large n since we have for all Y_{i_n}

$$Y_{i_n} \sim Y \text{ if } n > \frac{1}{\alpha}. \tag{12}$$

What happens here is that the decision maker with VaR_α-preferences ignores the difference between the tail of the Y_{i_n}, for which the bad loss of n happens with probability strictly smaller than α, and the tail of Y, for which no loss happens at all. On the one hand, this decision maker still perceives some difference between the Y_{i_n} and the Y because their distance in the d_0-metric, i.e.,

$$d_0\left(Y_{i_n}, Y\right) = \frac{1}{1+n},$$

is never zero. On the other hand, these random variables have become so similar in the d_0-metric for sufficiently large n that the decision maker who perceives similarity in accordance with convergence in measure stops caring about these differences in the sense that she becomes indifferent between these random variables.

For the same reason this decision maker does not like to diversify her portfolio in Example 6. By combining the Y_{i_n}, $i = 1, ..., m$, through the convex combination $\frac{1}{m}\sum_{i=1}^{m} Y_{i_n}$, the decision maker starts to care about the prospect of some non-zero loss because the probability of such loss has been lifted over the likelihood threshold (i.e., the confidence level α) for which losses matter to her.

Do such VaR_α-preferences plausibly describe some decision makers? We would like to argue 'yes' because the popularity of value-at-risk amounts to indirect evidence that value-at-risk resembles the preferences of some people pretty well. Recall that the Basel value-at-risk regulation for bank capital requires banks to absorb losses with a 99.9 per cent probability which corresponds to a confidence level of $\alpha = .001$. Some decision makers might then be indifferent between a random variable giving always a zero loss and a random variable that gives a zero loss with probability $\frac{1000}{1001}$ and a substantial loss with probability $\frac{1}{1001}$. These decision maker would be, e.g., the regulating policy makers who fixed the confidence level at $\alpha = .001$ for exactly the reason that they do not care about a bad outcome that only happens with a chance of $\frac{1}{1001}$.

The critique of the axiomatic risk measure literature against the use of the value-at-risk criterion for banking regulation is normative in nature in that it demands that policy makers (or risk managers) should actually care about catastrophic events regardless of the fact that they are highly unlikely! This paper looks at preferences exclusively from a descriptive perspective which takes a decision maker's preferences as a primitive. This descriptive perspective offers a possible explanation for the popularity of value-at-risk. Namely, value-at-risk may work for decision maker pretty well who (i) perceive the similarity of random variables in accordance with the topology of convergence in measure and (ii) who have lower-continuous preferences on some rich set. In contrast, such decision makers might feel uncomfortable with convex risk measures because these measures must violate, by Theorem 4, the lower-semicontinuity of their preferences.

To illustrate this last point, let us look at *average value-at-risk* as an example of a convex risk measure that is based on the value-at-risk criterion (whereby we can equivalently use $\mathrm{VaR}_{\leq\alpha}$ or $\mathrm{VaR}_{<\alpha}$ in the following definition).

Definition. Average value-at-risk. *Fix some $\beta \in (0, 1]$. The average value-at-risk of the random variable Z for the confidence level interval $(0, \beta)$ is*

$$\mathrm{AVaR}_{\beta}(Z) = \frac{1}{\beta}\int_{0}^{\beta} \mathrm{VaR}_{\leq\alpha}(Z)\, d\alpha.$$

That is, the average value-at-risk of a random variable is the average of the random variable's value-at-risk over the confidence levels in $(0, \beta)$.[12] In contrast to value-at-risk, a decision maker with AVaR_{β}-preferences therefore also takes all tail losses into account. Because average value-at-risk is a convex risk measure, AVaR_{β}-preferences must violate lower-semicontinuity in measure on rich sets.

[12] Our formal definition of the average value-at-risk also appears in the literature as the definition of the "conditional value-at-risk" or of the "expected shortfall". We follow here Föllmer and Schiedt (2016, p.233) who argue that the notion of the average over the interval $(0, \beta)$ is more precise as it clarifies that the conditional distribution in question is the uniform distribution.

Example 7. Average value-at-risk. Consider the rich set $R(\mathcal{F})$ of Example 5. Fix any $\beta \in (0,1]$. The average value-at-risk of the two constant random variables is trivially given as

$$\text{AVaR}_\beta(X) = 1 \text{ and } \text{AVaR}_\beta(Y) = 0.$$

Focus now on the Y_{i_n} and observe that, for $n > \frac{1}{\beta}$,

$$\text{AVaR}_\beta(Y_{i_n}) = \frac{1}{\beta}\left(\int_0^{\frac{1}{n}} n d\alpha + \int_{\frac{1}{n}}^{\beta} 0 d\alpha\right) = \frac{1}{\beta},$$

implying

$$\lim_{n\to\infty} \text{AVaR}_\beta(Y_{i_n}) \geq \text{AVaR}_\beta(X) > \text{AVaR}_\beta(Y)$$

in accordance with (9). Consequently, AVaR_β-preferences violate lower-semicontinuity in measure μ above X on this rich set because we have for sufficiently large n that

$$Y_n \preceq X \prec Y.$$

□

6 Concluding remarks and outlook

Continuity and convexity of preferences are both fundamental behavioral principles whereby continuity is a relative concept which depends on how a decision maker perceives similarity. This paper considers a decision maker who perceives similarity of random variables in accordance with the topology of convergence in measure. We introduce the notion of a rich set which encompasses any standard vector space of random variables but also much smaller sets containing only random variables with at most two different outcomes in their support. If our decision maker has complete preferences over a rich set of random variables, lower-semicontinuity of preferences becomes incompatible with the convexity of strictly better sets. As one implication, utility representations that express risk- or ambiguity aversion cannot describe decision makers whose preferences are lower-semicontinuous in measure. As another implication, value-at-risk is a non-convex risk measure exactly because it represents preferences that are lower-semicontinuous in measure.

Empirical evidence from experimental studies within the prospect theory framework suggest the prevalence of S-shaped value functions and inverse S-shaped non-additive probability measures. These findings stand for violations of global convexity. This paper's incompatibility analysis shows that lower-semicontinuity of preferences in the topology of convergence in measure would require such shapes whenever Choquet expected utility decision makers have complete preferences over some rich set of random variables. That is, the reason for this empirical evidence might be real life decision makers who have continuous preferences whereby their perceptions of similarity between random variables are well described by the topology of convergence of measure. Under the behavioral assumption of continuous preferences, our analysis predicts that such decision makers will tend to violate convexity in more choice situations than, e.g., decision makers who perceive similarity of random variables in accordance with convergence in mean. To investigate the empirical relevance of this paper's theoretical analysis, future empirical research should therefore look into the relationship between different similarity perceptions of random variables, on the one hand, and non-convexity of preferences, on the other hand.

Appendix: Formal proofs

Proof of Lemma 1. Step 1. Fix $Y \in \mathcal{F}$ and consider an arbitrary $X \in \mathcal{F}$. Since $R(\mathcal{F})$ is rich, we have

$$Y_{i_n} = Y + n(X - Y)\, 1_{\Omega_{i_n}} \in R(\mathcal{F})$$

for $n \geq 1$. Next note that

$$\begin{aligned} d_0(Y, Y_{i_n}) &= \int_\Omega \frac{|Y - Y_{i_n}|}{1 + |Y - Y_{i_n}|} d\mu = \int_\Omega \frac{\left|n(X - Y_{i_n})1_{\Omega_{i_n}}\right|}{1 + \left|n(X - Y_{i_n})1_{\Omega_{i_n}}\right|} d\mu \\ &= \int_\Omega \frac{|n(X - Y)|}{1 + |n(X - Y)|} 1_{\Omega_{i_n}} d\mu \\ &< \int_\Omega 1_{\Omega_{i_n}} d\mu = \frac{1}{n}. \end{aligned}$$

Consequently, we have for all Y_{i_n}, $i_n \in \{1, ..., n\}$, that

$$Y_{i_n} \in B_\varepsilon(Y; d_0) \text{ for } \varepsilon \geq \frac{1}{n}$$

where $B_\varepsilon(Y; d_0)$ denotes an open ball around Y in (L, d_0) with radius ε.

Step 2. Fix any open set $A \subseteq (L, d_0)$ such that $Y \in A$. By definition, there must exist some sufficiently small $\varepsilon > 0$ such that

$$B_\varepsilon(Y; d_0) \subseteq A$$

so that, by Step 1, $Y_{i_n} \in A$ for all $i_n \in \{1, ..., n\}$ with $n \geq \frac{1}{\varepsilon}$.

Step 3. Finally, note that convexity of A implies

$$\begin{aligned} \frac{1}{n}\sum_{i_n=1}^{n} Y_{i_n} &= \frac{1}{n}\sum_{i_n=1}^{n} Y + n(X - Y)\, 1_{\Omega_{i_n}} \\ &= Y + \sum_{i_n=1}^{n} (X - Y) 1_{\Omega_{i_n}} \\ &= Y + X - Y = X, \end{aligned}$$

which gives the desired result $X \in A$.□□

Proof of Theorem 1. Ad (i). Assume that $S^*(X)$ is convex. Because of $X \prec Y$, we have $Y \in S^*(X)$. If $S^*(X)$ was open, Lemma 1 implies that $X \in S^*(X)$, a contradiction. Consequently, $\preceq$ cannot be lower-semicontinuous in measure above X. □

Ad (ii). Assume now that $s^*(Y)$ is convex. Because of $X \prec Y$, we have $X \in s^*(Y)$. An open $s^*(Y)$ would, by Lemma 1, imply the contradiction $Y \in s^*(Y)$. Consequently, $\preceq$ cannot be upper-semicontinuous in measure below Y. □□

Proof of Proposition 1. If $\preceq$ is not lower-semicontinuous above X, the strictly better set $S^*(X)$ is not open. That is, there must exist some $Y \in S^*(X)$ which is not an interior point of $S^*(X)$, i.e., for all $\delta > 0$, there are Z such that

$$Z \in B_\delta(Y; d_0) \text{ but } Z \notin S^*(X). \tag{13}$$

Let $\delta_n = \frac{1}{n}$ and pick $Y_n \in B_{\delta_n}(Y; d_0)$ such that $Y_n \notin S^*(X)$. By (13), such Y_n exist for all $n \geq 1$. This constructs a converging sequence $\{Y_n\}$, $d_0(Y_n, Y) \to 0$, such that $Y_n \notin S^*(X)$ for all n, implying

$$U(Y_n) \leq U(X) \text{ for all } n \text{ whereas } U(X) < U(Y).$$

Let

$$\varepsilon = U(Y) - U(X) > 0$$

to see that, for all n,

$$U(Y_n) \leq U(Y) - \varepsilon.$$

But this violates (3) so that U is not lower-semicontinuous at Y. The argument for upper-semicontinuity proceeds analogously.□□

Proof of Proposition 2. If $S^*(X)$ is empty, the result obtains trivially. Suppose therefore that $Y, Z \in S^*(X)$ so that $U(X) < U(Y)$ as well as $U(X) < U(Z)$. If U is concave on a convex L, we have for any $\lambda \in (0,1)$

$$\begin{aligned} U(\lambda Y + (1-\lambda) Z) &\geq \lambda U(Y) + (1-\lambda) U(Z) \\ &> U(X), \end{aligned}$$

implying

$$\lambda Y + (1-\lambda) Z \in S^*(X).$$

□□

Proof of Proposition 3. If u is concave we have, for all ω,

$$u(\lambda X(\omega) + (1-\lambda) Y(\omega)) \geq \lambda u(X(\omega)) + (1-\lambda) u(Y(\omega)).$$

Next observe that

$$\begin{aligned} U(\lambda X + (1-\lambda) Y) &= I(u(\lambda X + (1-\lambda) Y)) \\ &\geq I(\lambda u(X) + (1-\lambda) u(Y)), \text{ by monotonicity} \\ &\geq \lambda I(u(X)) + (1-\lambda) I(u(Y)), \text{ by concavity} \\ &= \lambda U(X) + (1-\lambda) U(Y), \end{aligned}$$

which proves part (i). Part (ii) is proved analogously.□□

References

Aliprantis, D.C., Border, K. (2006) *Infinite Dimensional Analysis.* 2nd edition. Berlin: Springer.

Artzner, P., Delbaen, F., Eber, J., Heath, D. (1997) "Thinking Coherently" *Risk* **10**, 68-71.

Artzner, P., Delbaen, F., Eber, J., Heath, D. (1999) "Coherent Measures of Risk" *Mathematical Finance* **9**, 203-228.

Assa, H., Zimper, A. (2018) "Preferences over all Random Variables: Incompatibility of Convexity and Continuity" *Journal of Mathematical Economics* **75**, 71-83.

Berge, C. (1997) *Topological Spaces.* New York: Dover Publications. (An "unabridged and unaltered republication" of the original English publication from 1963.)

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

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Montrucchio, L. (2011) "Uncertainty Averse Preferences" *Journal of Economic Theory* **146**, 1275-1330.

Chateauneuf, A., Cohen, M., Meilijson, I. (2005) "More Pessimism than Greediness: A Characterization of Monotone Risk Aversion in the Rank-dependent Expected Utility Model" *Economic Theory* **25**, 649-667.

Chew, S., Karni, E., Safra, Z. (1987) "Risk Aversion in the Theory of Expected Utility with Rank Dependent Preferences" *Journal of Economic Theory* **42**, 370-381.

Day, M.M. (1940) "The Spaces L^p with $0 < p < 1$" *Bull. Amer. Math. Soc.* **46**, 816-823.

Danan, E., Gajdos, T., Tallon, J.-M. (2015) "Harsanyi's Aggregation Theorem with Incomplete Preferences" *American Economic Journal: Microeconomics* **7**, 61-69.

Dekel, E. (1989) "Asset Demands without the Independence Axiom" *Econometrica* **57**, 163-169.

Delbaen, F. (2002) "Coherent Risk Measures on General Probability Spaces" In *Advances in Finance and Stochastics*, pp. 1-37. Berlin: Springer.

Delbaen, F. (2009) "Risk Measures for Non-integrable Random Variables" *Mathematical Finance* **19**, 329-333.

Föllmer, H., Schied, A. (2002) "Convex Measures of Risk and Trading Constraints" *Finance and Stochastics* **6**, 429-447.

Föllmer, H., Schied, A. (2016) *Stochastic Finance. An Introduction in Discrete Time.* 4th edition. Berlin: Walter de Gruyter GmbH.

Gilboa, I. (1987) "Expected Utility with Purely Subjective Non-Additive Probabilities" *Journal of Mathematical Economics* **16**, 65-88.

Gilboa, I., and D. Schmeidler (1989) "Maxmin Expected Utility with Non-Unique Priors" *Journal of Mathematical Economics* **18**, 141-153.

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

Rudin, W. (1991) *Functional Analysis*. 2nd edition. New York: McGraw-Hill.

Schmeidler, D. (1986) "Integral Representation Without Additivity" *Proceedings of the American Mathematical Society* **97**, 255-261.

Schmeidler, D. (1989) "Subjective Probability and Expected Utility Without Additivity" *Econometrica* **57**, 571-587.

Wakker, P.P. (2010) *Prospect Theory for Risk and Ambiguity.* Cambridge University Press.