Preferences over rich sets of random variables: On the incompatibility of convexity and semicontinuity in measure*

Alexander Zimper[†] Hirbod Assa[‡]

August 16, 2020

Abstract

This paper considers a decision maker whose preferences are locally upper-or/and lower-semicontinuous 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. Whenever preferences are complete on a rich set of random variables, lower- (resp. upper-) semicontinuity in measure becomes incompatible with convexity of strictly better (resp. worse) sets. We discuss implications for utility representations and risk-measures. In particular, we show that the value-at-risk criterion violates convexity exactly because it is lower-semicontinuous in measure.

Keywords: Continuous Preferences; Utility Representations; Convex Risk Measures; Value-at-Risk
JEL Classification Numbers: D81

*We are grateful to comments and suggestions from Frank Riedel, Alex Ludwig, an anonymous referee 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

We consider a decision maker whose preferences over random variables are locally upper- or/and lower-semicontinuous in the topology of convergence in measure μ. On the one hand, convergence in measure stands for a plausible description of how some decision makers might perceive similarity of random variables. More precisely, we show that this topology can be generated by the k-truncated expectation metric

$$E_{\mu}\left(|X-Y| \wedge k\right) \tag{1}$$

where the truncation value $k > 0$ cuts off any large differences between the random variables X and Y. This topology therefore describes decision makers who care about extremely bad and extremely good events but who also tend to ignore extreme differences in the magnitude of such events. On the other hand, convergence in measure turns out to be particularly interesting because a decision maker whose preferences are lower-semicontinuous in measure is bound to violate in specific choice situations the convexity of strictly better sets. Convexity of strictly better sets, however, is central to standard characterizations of global risk/uncertainty/ambiguity aversion, cf. Cerreira-Vioglio et al. (2011, p.1276):

> "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."

Continuity in measure is thus not merely a purely technical assumption but it comes with strong implications for a decision maker's choice behavior such as, e.g., an aversion against portfolio diversification which would only reduce risk on tail events. To assume that some decision makers have preferences which are continuous in measure might thus be one possible explanation for empirically observed violations of convex choice behavior. For example, experimental studies within the prospect theory framework typically elicit violations of convexity in the form of 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). Similarly, the popularity of the value-at-risk criterion suggests that some decision makers violate convexity when it comes to the risk on tail events.[1]

This paper builds on our previous incompatibility analysis in Assa and Zimper (2018) which was based on the assumption that preferences are complete over the whole set of all

[1] Of course, continuity concepts are not directly testable through finitely many observations of choice behavior. Violations of convexity alone–as, e.g., expressed by the popularity of the value-at-risk criterion–cannot prove that these decision makers' preferences are continuous in measure. We come back to this point when we discuss the behavioral meaning of continuity.

random variables defined on an arbitrary non-atomic measure space. As our main finding we had shown that such preferences cannot be simultaneously (i) globally convex and (ii) globally continuous in the topology of convergence in measure. Our present–local–analysis extends this previous–global–analysis along two dimensions. Firstly, we split up continuity into the two weaker principles of upper- and lower-semicontinuity, respectively, whereby semicontinuity is only required to hold locally, i.e., around selected random variables. 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. Continuity, i.e., the combination of upper-and lower-semicontinuity, then stands for the decision theoretic principle which ensures that $X_n \prec Y_n$ whenever the decision maker's choice set is changed from $\{X, Y, ...\}$ to the set $\{X_n, Y_n, ...\}$ of similar alternatives.

To assume complete preferences over the domain of all random variables is an extremely strong requirement which is bound to be violated by real-life decision makers (see, e.g., Danan et al. 2015 and references therein). As a second generalization, we now require completeness on much smaller domains than the set of all random variables. To this purpose, we introduce the notion of *rich* sets of random variables. Formally, we define a rich set, denoted $R(\mathcal{F})$, through a construction (i.e., partition) procedure that starts out with any given subset of at least two random variables, denoted $\mathcal{F}$. The technical purpose of a rich set is to allow for sequences of random variables which converge in measure but not in alternative topologies that are compatible with convexity such as, e.g., pointwise convergence or convergence in mean. Examples of rich sets–that are constructed from themselves–are the standard vectors space of random variables such as, e.g., all L^p spaces with $0 \leq p \leq \infty$, as well as the vector space of all simple random variables. But much smaller (and non-convex) sets of random variables–where $\mathcal{F}$ only consists of two degenerate random variables–might be rich sets as well (cf. Example 1 below).

Throughout this paper we fix 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. Based on our construction of a rich set $R(\mathcal{F})$, we derive the following fundamental incompatibility results for convexity and semicontinuity in measure μ.

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.*

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. The above incompatibility results for preferences imply the following incompatibility results for these standard utility representations.[2]

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.

[2] 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).

Next consider a decision maker who strictly prefers random variables that are strictly less risky in terms of some fixed risk measure ρ.

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. This paper's descriptive perspective on value-at-risk is different because our analysis offers one possible explanation–in terms of continuity in measure–for the popularity of value-at-risk without any normative judgment. We show that value-at-risk violates convexity exactly because it represents preferences which are lower-semicontinuous in measure. Decision makers with preferences over rich sets that are continuous in measure might therefore feel more comfortable with a lower-semicontinuous risk measure, as, e.g., value-at-risk, than with some convex risk-measure that violates lower-semicontinuity.

Behavioral meaning of continuity: Related literature

Our interpretation of the behavioral meaning of continuous preferences follows Fishburn (1970) who writes:

> "Continuity formalizes the intuitive notion that if two elements in X are not very different then their utilities should be close together. The difference between x and y can be thought of either in terms of their relative proximity under $\prec$ or in terms of a structure for X that is related to $\prec$ in some way." (p.35)

The mathematical 'structure' which captures the 'difference' between random variables is the topology which has to be chosen by the modeler. The choice of a specific topology is the more empirically relevant the better it manages to capture the similarity perceptions of a relevant group of real life decision makers. Although the choice of a topology has thus empirical meaning in the form of similarity perceptions, the problem from a data-analytical perspective is that such similarity perceptions are not reflected in observable choices. Instead, according empirical studies would have to rely on the cognizant answers of people asked about their similarity perceptions or/and about their choices in hypothetical choice situations involving infinite converging sequences.

On the one hand, we are not aware of any empirical studies that investigate real life people's similarity perceptions for random variables. The pragmatic way around this lack of empirical evidence is the modeler's choice of a topology which he/she deems as 'plausible upon introspection'. This is exactly our approach in the present paper where we deem convergence in measure as one–among many others–plausible topology. Wakker (1998) writes: "We find continuity an appealing condition when formulated in a space with a natural metric." When it comes to domains of random variables, alternative 'natural' metrics become available with our metric (1) being, arguably, one of these natural metrics.

On the other hand, behavioral decision theorists in the tradition of Samuelson's (1938) *revealed preference theory* would reject any data in the form of introspective answers as unreliable and only trust data consisting of observed choice behavior (cf. Agner and Loewenstein, 2012). Because continuity concepts are not 'testable' through choice data, i.e., they are not falsifiable through finitely many observed choices, this extremely positivist school of thought would reject any topological or/and continuity assumptions as behaviorally irrelevant. We do not share this view that only choice data is empirically relevant whereas introspective answers must be discarded as unreliable. Going beyond revealed preference theory, questions about subjective survival-, health-, unemployment beliefs within the Health and Retirement Study (HRS) or the Survey of Consumer Finances (SCF) have proved useful in explaining choice behavior with regards to savings and retirement decisions.[3] We hope that empirical studies about similarity perceptions for random variables will also become available in the future so that the choice of a topology could be based on empirical evidence.

Independent of any considerations about empirical relevance, there exists a theoretical literature which investigates behavioral implications of topological assumptions on preferences for infinite spaces used in economic models. Schmeidler (1971) proves that continuity in any connected topological space combined with transitivity implies completeness of the (weak) preference relation. Based on Schmeidler's (1971) finding, Gerasimou (2013) shows that continuity in any connected topological space combined with (some) incompleteness results in a *fragile* preference relation in the sense that $X \prec Y$ implies non-comparability of some X' and Y' belonging to an arbitrarily small neighborhood

[3] For references to the economic literature which uses subjective belief data from the HRS or/and the SCF, see, e.g., Groneck, Ludwig, and Zimper (2016).

around X and Y, respectively. Gerasimou's (2013) argument against the plausibility of such *preference fragility* amounts to an argument in favor of continuity as a behavioral principle:

> "Indeed, one would expect that when decision makers express strict preference for one alternative over another, marginal changes in these two alternatives should not result in them becoming incomparable. If they do, then doubt should perhaps be cast on the validity of the strict-preference comparison between the original alternatives. Finally, introspection and casual empiricism do not seem favorable for the property's descriptive accuracy either." (p.161)

This theoretical approach–which uses continuity on connected spaces in order to derive behavioral implications in the form of completeness versus incompleteness of preferences–is reviewed and extended in the (analytically deep) article by Khan and Uyanık (2019). These authors also make the connection between Schmeidler (1971) and Gerasimou (2013) and earlier contributions in Eilenberg (1941), Sonnenschein (1965), and Sen (1969).

Although our paper shares with this literature the general motivation that continuity comes with behavioral implications, the details of our analytical approach are rather different. Firstly, we are interested in the implications of continuity on the convexity and not on the completeness of preferences whereby we restrict attention to preferences that are complete on our relevant subdomain of rich sets. Secondly, instead of distinguishing between connected versus not connected topological spaces, we endow different (rich) sets with the specific metric (1) that generates the topology of convergence in measure.[4] Instead of looking into behavioral implications of the abstract property of topological connectedness, we are interested in the behavioral implications of a concrete topology which is, in our opinion, one plausible description of how some decision makers might perceive the similarity of random variables.

The remainder of our analysis proceeds as follows. Section 2 introduces relevant mathematical concepts. Section 3 analyzes the incompatibility of convexity and semicontinuity in measure. Sections 4 and 5 discuss implications for utility representations and risk measures, respectively. Section 6 concludes with a discussion about the relevance of convergence in measure versus pointwise convergence. Formal proofs are relegated to the Mathematical Appendix.

[4] A connected topological space is characterized by the property that only the empty set and the universal set are simultaneously open and closed (cf. Chapter 1.11.1 in Bourbaki 1989). Whereas some of the spaces considered in this paper are connected (e.g., the space of all random variables) others are not (e.g., the 'small' rich set of our Example 1 below).

2 Mathematical preliminaries

2.1 Convergence in measure

Fix the non-atomic probability space $(\Omega, \mathcal{B}, \mu)$ such that $\Omega = (0, 1)$ and μ denotes the Lebesgue measure on the Borel σ-algebra $\mathcal{B}$ on $(0, 1)$. 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.[5] 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 $\epsilon > 0$,

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

For a given constant $k > 0$ introduce the *k-truncated expectation metric* $d_k : L^0 \times L^0 \to [0, \infty)$ such that

$$d_k(X, Y) = \int (|X - Y| \wedge k)\, d\mu = \int_{\omega\in\Omega} \max\{|X(\omega) - Y(\omega)|, k\}\, d\mu. \tag{2}$$

As $|X - Y| \wedge k$ is bounded from below by zero and bounded from above by k for all $X, Y \in L^0$, this metric is well-defined for all $X, Z \in L^0$.

Proposition 1. *The k-truncated expectation metric d_k generates the topology of convergence in measure μ, i.e., $\lim_{n\to\infty} d_k(Z, Z_n) = 0$ iff $Z_n \to_\mu Z$.*

[5] 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.

According to the k-truncated expectation metric (2), large payoff differences are simply cut-off at some value k. We regard this metric as a plausible description of how some real-life decision makers perceive the similarity of random variables, especially when possible payoffs or losses can become arbitrarily large. In what follows we denote by (L, d_k) our default metric space such that L is endowed with the topology of convergence in measure.

Remark. The topology of convergence in measure can be alternatively generated by the metric $d_0 : L^0 \times L^0 \to [0, \infty)$ such that

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

(cf. Lemma 13.40 in Aliprantis and Border 2006). Although the d_k- and d_0-metrics are topologically equivalent, we prefer the straightforward decision-theoretic interpretation of the d_k-metric according to which the modelled decision makers tend to ignore differences on extreme tail events.

2.2 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 $n \geq 1$.

Example 1. A 'small' rich set. 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

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

Note that this rich set is not convex and it only contains random variables with at most two different outcomes in their support.□

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.

Observation 1. *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., $L = R(\mathcal{F}) = \mathcal{F}$.*

Observation 1 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

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

By Observation 1, 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.

2.3 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. *Consider an arbitrary rich set $R(\mathcal{F}) \subseteq L$. If the subset $A \subseteq (L, d_k)$ is open and convex, then*

$$A \cap \mathcal{F} \neq \emptyset \text{ implies } \mathcal{F} \subseteq A.$$

In words: Whenever there exists a random variable Y with $Y \in \mathcal{F}$ and $Y \in A$ such that (i) the rich set $R(\mathcal{F})$ is a subset of L and (ii) A is a convex and open subset of (L, d_k), then all random variables in $\mathcal{F}$ must also belong to A.

Consider, for example, the relevant special case such that $\mathcal{F} = R(\mathcal{F}) = L$ is some vector space of random variables. For such vector spaces Lemma 1 implies that the universal set L itself is the only non-empty, convex and open subset in the topology of convergence in measure. Put equivalently, the topological vector space (L, d_k) is *locally non-convex*. To see this, note that all open balls $B_\varepsilon(Y; d_k) \subset L$ must be non-convex sets as L is the only non-empty, convex and open subset of (L, d_k).

Remark. 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) such that these d_p-metrics are defined as follows for $p > 0$:

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

It is possible to extend our subsequent incompatibility analysis from (L, d_k) 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.

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 $R(\mathcal{F}) \subseteq L$.

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-semicontinuous in measure μ above X if and only if the strictly better set $S^*(X)$ is open in (L, d_k).*

(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_k).*

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.*

4 Incompatibility results for utility representations

4.1 General analysis

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

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

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*

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

(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*

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

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 $Z \in L$).

Observation 2.

(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 $X \prec Y$.*

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 $\lambda \in (0,1)$,*

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

[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 $X \in \mathcal{F}$.

Observation 3. *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 Observations 2 and 3 gives us the following incompatibility results for utility representations.

Proposition 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 $\{Y_{i_n}\} \rightarrow_\mu Y$ 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 $\{X_{i_n}\} \rightarrow_\mu X$ on $\mathcal{F}$ such that*

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

4.2 Operators for utility random variables

Fix some increasing Bernoulli utility function $u : \mathbb{R} \rightarrow \mathbb{R}$. 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 $u(Z) : \Omega \rightarrow \mathbb{R}$ 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 $\lambda \in (0,1)$,*

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

Observation 4. *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 Proposition 2 with Observation 4 gives us the following results.

Proposition 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 Proposition 3 to the standard utility representations *expected utility*, *Choquet expected utility*, and *multiple priors expected utility*, respectively.

4.3 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 $co(L_u)$. In what follows, we derive a string of corollaries to Proposition 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 $R(\mathcal{F})$ 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 $X, Y \in \mathcal{F}$).

4.3.1 Expected utility

Suppose 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 $Z \in L$,*

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

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

(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 2. Consider the rich set $R(\mathcal{F})$ of Example 1 which consists of $X = 0$, $Y = 1$ 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

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

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

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

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 2 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 $\mu_{Z_n} \to \mu_Z$, if and only if

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

for all bounded and μ-almost everywhere continuous functions $u : \mathbb{R} \to \mathbb{R}$ (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 inverse 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.

Discussion: Bounded Bernoulli utility. A bounded Bernoulli utility function guarantees continuity in measure μ of expected utility preferences on any rich set. However, such bounded Bernoulli utility of an expected utility decision maker is subject to the same criticism that Azevedo and Gottlieb (2012) formulate for inverse S-shaped value functions of prospect theory in their article "Risk-neutral firms can extract unbounded profits from consumers with prospect theory preferences". The following example illustrates the basic idea of these authors' criticism for our formal framework.

Example 3. Consider a decision maker who has expected utility preferences with respect to μ over the (rich) domain of all random variables with either one or two outcomes in their support such that the increasing Bernoulli utility function satisfies

$$u(x) = -u(-x) \text{ for all } x \in \mathbb{R} \text{ and } \lim_{x \to \infty} u(x) = c > 0.$$

Because the Bernoulli utility function is bounded from below by $-c$ and from above by c, this expected utility representation must be continuous in measure μ. Define, at first, the random variable

$$X^n(\omega) = \begin{cases} 0 & \text{if } \omega \in \Omega \backslash \Omega_{i_n} & \text{i.e., with prob. } 1 - \frac{1}{n} \\ -\frac{1}{2}n^2 & \text{if } \omega \in \Omega_{i_n} & \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

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

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

Next, define for any given X^n, a corresponding random variable Y^n such that

$$Y^n(\omega) = \begin{cases} \varepsilon_n & \text{if } \omega \in \Omega \backslash \Omega_{i_n} \quad \text{i.e., with prob. } 1 - \frac{1}{n} \\ -n^2 & \text{if } \omega \in \Omega_{i_n} \quad \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

where $\varepsilon_n \geq 0$ is implicitly defined through the expected utility indifference equation

$$\begin{aligned} \int_\Omega u(X^n)\, d\mu &= \int_\Omega u(Y^n)\, d\mu \\ &\Leftrightarrow \\ u(\varepsilon_n)\left(1 - \frac{1}{n}\right) &= \left[u\left(n^2\right) - u\left(\frac{1}{2}n^2\right)\right]\frac{1}{n}. \end{aligned}$$

In analogy to Azevedo and Gottlieb (2012), the economic interpretation of these random variables is as follows: The decision maker is indifferent between (i) the random variable Y^n that results in the small positive payout ε_n with large probability $1 - \frac{1}{n}$ and the large loss $-n^2$ with small probability $\frac{1}{n}$ and (ii) the random variable X^n that gives the moderate loss $-\frac{1}{2}n^2$ with large probability $1 - \frac{1}{n}$ and nothing else. If a risk-neutral firm offers to trade Y^n to the decision maker in exchange for X^n, the decision maker–who is indifferent between X^n and Y^n by construction–has no problem to accept this trade. The firm gains from this trade the random payoff $Z^n = X^n - Y^n$, i.e.,

$$Z^n(\omega) = \begin{cases} -\varepsilon_n & \text{if } \omega \in \Omega \backslash \Omega_{i_n} \quad \text{i.e., with prob. } 1 - \frac{1}{n} \\ \frac{1}{2}n^2 & \text{if } \omega \in \Omega_{i_n} \quad \text{i.e., with prob. } \frac{1}{n} \end{cases}$$

so that her expected profit becomes

$$\int_\Omega Z^n d\mu = \frac{1}{2}n - \varepsilon_n\left(1 - \frac{1}{n}\right).$$

Finally, note that

$$\lim_{n\to\infty} \int_\Omega Z^n d\mu = \infty$$

because boundedness, i.e., $\lim_{x\to\infty} u(x) = c$, implies

$$\begin{aligned} \lim_{n\to\infty} u(\varepsilon_n)\left(1 - \frac{1}{n}\right) &= \lim_{n\to\infty} \left[u\left(n^2\right) - u\left(\frac{1}{2}n^2\right)\right]\frac{1}{n} = 0 \\ &\Rightarrow \\ \lim_{n\to\infty} \varepsilon_n &= 0. \end{aligned}$$

That is, the firm's expected profit from trading Y^n for X^n goes to infinity if n gets large. But this would result in an existence problem for a trade equilibrium whenever the firm could endogenously choose an arbitrarily large number n (as assumed by Azevedo and Gottlieb (2012)).□

The structure of the above example–which pitches ever smaller utility differences of an inverse S-shaped Bernoulli utility function against linearly increasing expected value differences–is at the heart of the formal argument in Azevedo and Gottlieb (2012) who look at joint limit conditions of probability weights and inverse S-shaped value functions from prospect theory. These authors use the fact that a risk-neutral firm might construct hypothetical gambles in order to extract large expected profits (at diminishing $-\varepsilon_n$ risk to itself) to argue against the suitability of prospect theory in economic applications:

> "Despite the relative success in explaining empirical regularities, prospect theory is rarely applied to strategic and market environments. This paper presents problems that one necessarily faces when attempting to incorporate prospect theory in those environments.
>
> We show that, under conditions satisfied by virtually all functional forms used in the literature, individuals with prospect theory preferences accept gambles with arbitrarily large negative expected values. This result severely limits the applicability of prospect theory when the supply side of the market is endogenous." Azevedo and Gottlieb (2012, p.1292)

As our example shows, the criticism of Azevedo and Gottlieb (2012) is not restricted to prospect theory with inverse S-shaped value functions but equally applies to expected utility preferences with inverse S-shaped Bernoulli utility functions when the domain includes random variables with arbitrarily large negative outcomes. As these authors admit, inverse S-shaped Bernoulli utility or/and value functions explain well empirically observed choice data. Additionally, bounded Bernoulli utility functions–which imply an inverse S-shape–have the two conceptual advantages that they ensure (i) continuity in measure on a rich domain as well as (ii) existence of an expected utility representation over large sets of random variables such as, e.g., the set of all random variables.[8] Giving up bounded Bernoulli utility functions might thus come with more severe disadvantages for relevant economic applications than the problem described in Azevedo and Gottlieb (2012). After all, this problem only applies to hypothetical market situations in which (i)

[8] Point (i) is explained in detail in the present paper. For point (ii) see our analysis in Assa and Zimper (2018) and references therein. For example, Savage's (1954) subjective expected utility axiomatization requires a bounded (Bernoulli) utility function as he considers complete preferences over the set of all random variables. Compare Wakker (1993) who writes:

> "Ever since, the extension of Savage's theorem to unbounded utility has been an open question, and with that the question "what is wrong with Savage's axioms?". [...] I think that "what is wrong with Savage's axioms", is primarily his requirement of completeness of the preference relation on the set of all (alternatives=) acts [...]." (p.448)

one trading partner is risk-neutral and (ii) both trading partners are endogenously able to construct arbitrary gambles which they can trade with each other.

4.3.2 Choquet expected utility and maxmin expected utility

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 (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.[9] 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

5.1 General analysis

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

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

[9] 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 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 speak of a decision maker with ρ-preferences.

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.[10]

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.[11]

Proposition 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 show that the value-at-risk criterion must violate convexity on rich sets because it represents preferences that are lower-semicontinuous in measure.

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

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

5.2 Value-at-risk

The value-at-risk of random variable $Z \in L^0$ at confidence level α corresponds to some α-quantile of Z (cf. Chapter 4.4 in Föllmer and Schied 2016) whereby the α-quantiles of Z are the members of the interval $\left[q_Z^-(\alpha), q_Z^+(\alpha)\right]$ such that

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

Let us restrict attention to the lower α-quantile, $q_Z^-(\alpha)$, and the upper α-quantile, $q_Z^+(\alpha)$, respectively, to define

$$\mathrm{VaR}_\alpha^-(Z) = -q_Z^-(\alpha) \text{ and } \mathrm{VaR}_\alpha^+(Z) = -q_Z^+(\alpha).$$

For any Z with a continuous and strictly increasing distribution function both value-at-risk definitions coincide. In general, we have that

$$\mathrm{VaR}_\alpha^-(Z) = \mathrm{VaR}_\alpha^+(Z) \tag{10}$$

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 might not reduce to a single value (cf. Lemma A.19. in Föllmer and Schied 2016). Whenever the equality (10) holds for a random variable Z, both value-at-risk definitions VaR_α^- and VaR_α^+ are continuous in measure μ at Z.[12] If (10) is violated at some Z, however, VaR_α^- will result for some converging sequences in an 'upward jump' to a higher value-at-risk whereas VaR_α^+ will result for some sequences in a 'downward jump' to a lower value-at-risk at Z. The following example illustrates these possible discontinuities for both value-at-risk definitions for the non-generic case in which (10) is violated.

Example 4. Consider the random variable

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

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

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

[12] In Assa and Zimper (2018) we had incorrectly claimed that value-at-risk preferences are continuous in measure without the qualifying condition (10).

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

$$\mathrm{VaR}_\alpha^-(Z) = 1 \text{ and } \mathrm{VaR}_\alpha^+(Z) = 0$$

while, for all n,

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

implying

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

□

In the non-generic case that (10) is violated for a given Z, VaR_α^- is an upper-semicontinuous function in measure μ at Z whereas VaR_α^+ 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, VaR_α^--preferences are lower-semicontinuous at every Z whereas VaR_α^+-preferences are lower-semicontinuous at Z if and only if the generic case (10) 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 VaR_α^- preferences over $R(\mathcal{F})$, then the strictly better set at X cannot be convex.*

(ii) *If the decision maker has complete VaR_α^+ preferences over $R(\mathcal{F})$ such that the generic case*

$$\mathrm{VaR}_\alpha^-(Y) = \mathrm{VaR}_\alpha^+(Y)$$

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

We illustrate Proposition 4 through a simple example.[13]

[13] We write $\mathrm{VaR}_\alpha(Z)$ whenever (10) holds for Z.

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

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

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

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

$\square$

By Proposition 4, value-at-risk cannot be a convex risk measure on any rich set because it represents preferences that are 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}

$$Y_{i_n} \sim Y \text{ if } n > \frac{1}{\alpha}. \tag{11}$$

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_k-metric, i.e.,

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

is never zero. On the other hand, these random variables have become so similar in the d_k-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.

According to Dekel (1989), a transitive preference relation on L 'exhibits 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 5. 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

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

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 $\frac{1}{m}\sum_{i=1}^{m} Y_{i_n}$.□

Why does the portfolio manager of Example 6 violate diversification of her portfolio? By combining the Y_{i_n}, $i = 1, ..., m$, through the convex combination $\frac{1}{m}\sum_{i=1}^{m} Y_{i_n}$, the manager 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. Below this confidence level, however, the portfolio manager perceives the Y_{i_n} as sufficiently similar to a zero-loss portfolio so that her indifference between such tail-risk portfolios and the zero-loss portfolio is perfectly in accordance with preferences that are lower-semicontinuous in measure.

5.3 Average value-at-risk

We conclude this section with a simple example that illustrates the incompatibility result of Theorem 4. To this purpose, we consider the convex risk-measure *average value-at-risk* which is defined as the average of the random variable's value-at-risk over the confidence levels in $(0, \beta)$.[14]

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*

$$\mathrm{AVaR}_\beta (Z) = \frac{1}{\beta}\int_0^\beta \mathrm{VaR}_\alpha (Z) \, d\alpha.$$

In contrast to value-at-risk, a decision maker with AVaR_β-preferences 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.

[14] 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 Schied (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. Fix any $\beta \in (0,1]$ and consider a decision maker with AVaR_β-preferences over the rich set from Example 5. The average value-at-risk of the two constant random variables is trivially given as

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

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

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

implying

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

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 discussion

We have defined rich sets of random variables with the property that complete preferences on such rich sets cannot be simultaneously convex and continuous in measure. More specifically, we have shown that upper-semicontinuity in measure is incompatible with convexity of strictly worse sets whereas lower-semicontinuity is incompatible with convexity of strictly better sets.

Convexity of strictly better sets, however, is central to standard characterizations of global risk aversion or global uncertainty (i.e., ambiguity) aversion and it is also central to the definition of convex or/and coherent risk measures. Decision theoretic models that impose convexity (e.g., global ambiguity aversion) as a behavioral axiom–and which must therefore violate lower-semicontinuity in measure–work instead with the following continuity axiom (cf., Maccheroni et al. 2006; Cerreira-Vioglio et al. 2011):

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

$$\{\alpha \in [0,1] \mid \alpha Z + (1-\alpha) Z' \prec Y\} \text{ and } \{\alpha \in [0,1] \mid X \prec \alpha Z + (1-\alpha) Z'\} \quad (12)$$

are open subsets of the Euclidean unit interval.

Translated into our topological framework for subsets of random variables, this standard continuity axiom captures preferences that are continuous in the topology of pointwise convergence. More precisely, let $X \prec Y$ and note that, by (12), 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.,[15]

$$\{Y_n\} \rightarrow Y \text{ if and only if } \lim_{n \rightarrow \infty} Y_n(\omega) \rightarrow Y(\omega) \text{ for every } \omega \in \Omega. \tag{13}$$

We regard it as perfectly plausible that many, or even the majority of, decision makers perceives similarity of random variables in accordance with pointwise convergence (13). For such decision makers the standard continuity axiom (12) is the appropriate continuity concept so that these decision makers might, e.g., express global uncertainty aversion.

Unfortunately, we are not aware of any empirical studies about real life people's similarity perceptions of random variables. Based on anecdotal evidence, however, we would like to argue that there exists a non-negligible subset of decision makers whose similarity perceptions of random variables are better captured by convergence in measure than by pointwise convergence. This anecdotal evidence concerns the popularity of the value-at-risk criterion for the evaluation of portfolios which expresses preferences that are non-convex but lower-semicontinuous in measure. 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 $\alpha = .001$. Decision makers with VaR_α-preferences would 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 $\frac{1}{1001}$. That is, these decision makers do not really care about tail-risks that only happen with a chance of $\frac{1}{1001}$ or less. Such ignorance towards the magnitude of catastrophic tail events is exactly what continuity in measure–in contrast to continuity with respect to pointwise convergence–can capture.

[15] To see this, note that all converging sequences $\{Y_n\} \rightarrow Y$ which determine whether the strictly better set at X is open are, by (12), of the form

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

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

Appendix: Formal proofs

Proof of Proposition 1. The if-part. We show that $\lim_{n\to\infty} d_k(X, X_n) = 0$ implies $X_n \to_\mu X$. For an arbitrary ϵ with $0 < \epsilon < k$ define the event

$$A_n = \{\omega \in \Omega \mid \mu(|X_n(\omega) - X(\omega)| > \epsilon)\}.$$

Because of

$$\begin{aligned} d_k(X, X_n) &= \int_\Omega (|X - Y| \wedge k) 1_{A_n} + |X - X_n| 1_{A_n^c} d\mu \\ &\geq \int_\Omega \epsilon \cdot 1_{A_n} + 0 \cdot 1_{A_n^c} d\mu \\ &= \epsilon\mu(A_n), \end{aligned}$$

$\lim_{n\to\infty} d_k(X, X_n) = 0$ implies $\lim_{n\to\infty} \mu(A_n) = 0$.

The only-if part. We show that $X_n \to_\mu X$ implies $\lim_{n\to\infty} d_k(X, X_n) = 0$. For an arbitrary ϵ with $0 < \epsilon < k$ pick some sufficiently fine partition such that $\mu(A_n) < \epsilon$. Note that

$$\begin{aligned} d_k(X, X_n) &= \int_\Omega \left((|X - Y| \wedge k) 1_{A_n} + (|X - Y|) 1_{A_n^c}\right) d\mu \\ &\leq \int_\Omega \left(k \cdot 1_{A_n} + \epsilon \cdot 1_{A_n^c}\right) d\mu \\ &= k\mu(A_n) + \epsilon[1 - \mu(A_n)]. \end{aligned}$$

For every $\epsilon > 0$ we have that $X_n \to_\mu X$ implies $\lim_{n\to\infty} \mu(A_n) = 0$. Consequently, $\lim_{n\to\infty} d_k(X, X_n) \leq \epsilon$ for all $\epsilon > 0$ which proves the only-if part.□□

Proof of Lemma 1. Consider a rich set $R(\mathcal{F}) \subseteq L$ such that $Y \in \mathcal{F}$ as well as $Y \in A$ for some open and convex $A \subseteq (L, d_k)$. To prove the Lemma, we have to show that $\mathcal{F} \subseteq A$.

Step 1. Fix some $k > 0$. Next, 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

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

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

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

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

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

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

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

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 $X \in A$ for any $X \in \mathcal{F}$.□□

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 Observation 2. 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

$$Z \in B_\delta (Y; d_k) \text{ but } Z \notin S^* (X). \tag{14}$$

Let $\delta_n = \frac{1}{n}$ and pick $Y_n \in B_{\delta_n} (Y; d_k)$ such that $Y_n \notin S^* (X)$. By (14), such Y_n exist for all $n \geq 1$. This constructs a converging sequence $\{Y_n\}$, $d_k (Y_n, Y) \to 0$, 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 Observation 3(i). 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 Observation 4. 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

Agner, E., Loewenstein, G. (2012) "Behavioral Economics" In *Handbook of the Philosophy of Science: Philosophy of Economics* Chapter **13**, pp. 641-690, Editor: Uskali Mäki, Amsterdam: Elsevier.

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.

Azevedo, E.M., Gottlieb, D. (2012) "Risk-Neutral Firms Can Extract Unbounded Profits from Consumers with Prospect Theory Preferences" *Journal of Economic Theory* **147**, 1291-1299.

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.

Bourbaki, N. (1989) *Elements of Mathematics. General Topology. Chapters 1-4.* Berlin: Springer.

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.

Eilenberg, S. (1941) "Ordered Topological Spaces" *American Journal of Mathematics* **63**, 39-45.

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

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

Groneck, M., Ludwig, A., Zimper, A. (2016) "A Life-cycle Model with Ambiguous Survival Beliefs" *Journal of Economic Theory* **162**, 137-180.

Kahn, M.A., Uyanık, M. (2019) "Topological Connectedness and Behavioral Assumptions on Preferences: A Two-way Relationship" *Economic Theory* (forthcoming) https://doi.org/10.1007/s00199-019-01206-7

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.

Savage, L.J. (1954) *The Foundations of Statistics*. New York: John Wiley.

Samuelson, P.A. (1938) "A Note on the Pure Theory of Consumer's Behaviour" *Economica* **5**, 61-71.

Schmeidler, D. (1971) "A Condition for the Completeness of Partial Preference Relations" *Econometrica* **39**, 403-404.

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.

Sen, A. (1969) "Quasi-transitivity, Rational Choice and Collective Decisions" Review of Economic Studies **36**, 381-393.

Sonnenschein, H. (1965) "The Relationship between Transitive Preference and the Structure of the Choice Space" *Econometrica* **33**, 624-634.

Wakker, P.P. (1988) "The Algebraic versus the Topological Approach to Additive Representations" Journal of Mathematical Psychology **32**, 421-435.

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