University of Pretoria
Department of Economics Working Paper Series

Preferences Over all Random Variables: Incompatibility of Convexity and Continuity

Hirbod Assa
University of Liverpool
Alexander Zimper
University of Pretoria and Kiel Institute for the World Economy
Working Paper: 2017-14
March 2017

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

Preferences over all random variables: Incompatibility of convexity and continuity*

Hirbod Assa[†] Alexander Zimper[‡]

March 7, 2017

Abstract

We consider preferences over all random variables on a given nonatomic probability space. We show that non-trivial and complete preferences cannot simultaneously satisfy the two fundamental principles of convexity and continuity. As an implication of this incompatibility result there cannot exist any non-trivial continuous utility representations over all random variables that are either quasi-concave or quasi-convex. This rules out risk-averse (or seeking) expected utility representations and, more generally, risk- and uncertainty-averse (or seeking) Choquet expected utility representations for this large space of random variables.

Keywords: Large Space; Preference for Diversification; Utility Representation; Risk Measures
JEL Classification Numbers: D81

*We are grateful to comments and suggestions from Keith Conrad, Tzachi Gilboa, Ehud Lehrer, Frank Riedel, the math.stackexchange user Soma, and Peter Wakker.

[†]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

[‡]Department of Economics; University of Pretoria; Kiel Institute for the World Economy; postal address: Private Bag X20; Hatfield 0028; South Africa; E-mail: alexander.zimper@up.ac.za

1 Introduction

Fix an arbitrary nonatomic probability space (Ω, Σ, μ). The set of all random variables defined over this space, denoted $L^0(\mu)$, consists of all Σ-measurable functions $X : \Omega \to \mathbb{R}$. Most of the literature on preferences over random variables restricts attention to rather small subsets of random variables such as, e.g., random variables with finite support. Whenever larger classes of random variables are considered they typically belong to a technically convenient space $L^p(\mu) \subset L^0(\mu)$, with $1 \leq p \leq \infty$, such that $X \in L^p(\mu)$ with $p < \infty$ if, and only if, the integral

$$\int_\Omega |X|^p \, d\mu \tag{1}$$

exists. For example, $L^1(\mu)$ collects all random variables with finite expected value; $L^2(\mu)$ collects all random variables with finite variance; and $L^\infty(\mu)$ denotes the set of all bounded random variables.

This paper takes an extreme stand and considers preferences over ALL random variables whereby we endow $L^0(\mu)$ with the metric topology of convergence in probability.[1] For this large space we establish that continuity and convexity cannot be simultaneously satisfied for non-trivial and complete preferences (Theorem 1). If we want to model non-trivial preferences over the random variables in $L^0(\mu)$, we must thus give up at least one of the three fundamental principles of continuity, convexity, or completeness, respectively. Under the additional assumption of transitivity, Theorem 2 establishes that continuity is neither compatible with quasi-concave nor with quasi-convex preferences. It is also not compatible with preference for diversification (Theorem 3).

For standard utility representations our incompatibility results imply the following restrictions:[2]

- Any non-trivial expected utility representation with

  $$\int_\Omega u(X) \, d\mu \tag{2}$$

  for all $X \in L^0(\mu)$ is neither compatible with a concave nor with a convex utility function u.

- Any non-trivial Choquet expected utility representation with

  $$\int_\Omega^C u(X) \, d\nu \tag{3}$$

[1] Our analytical findings can be analogously derived for the 'smaller' $L^p(\mu)$ spaces with $0 < p < 1$ (cf. Remark 3).

[2] The first restriction is a special case of the second one. For the formal definition of Choquet integration with respect to the non-additive probability measure (=capacity) ν see Section 4.

for all $X \in L^0(\mu)$ is neither compatible with (i) a concave utility function u combined with a convex capacity ν nor with (ii) a convex utility function u combined with a concave capacity ν.

Consequently, there cannot exist any expected utility representation over all random variables that expresses global risk-aversion (or global risk-seeking, for that matter). Neither can there exist a Choquet expected utility representation over all random variables which expresses global risk-aversion combined with global ambiguity/uncertainty aversion (or global risk-seeking combined with global ambiguity/uncertainty seeking).[3]

One possible way of dealing with this incompatibility result is to give up on continuous utility representations altogether and consider complete preferences over all random variables that are convex but not continuous (cf. Examples 4 and 5 in Section 5). From an applicational point of view, however, the lack of a continuous utility representation is not very attractive.

In case one wants to keep continuous utility representations, there are two alternative approaches for getting around the incompatibility between convexity and continuity. The first approach is to give up complete preferences on $L^0(\mu)$. By restricting attention to suitable subsets of random variables, continuity and convexity may become compatible for preferences that are complete for these subsets only (cf. Examples 2 and 3 in Section 5). A straightforward example for this approach would be preferences on $L^1(\mu)$ that are represented by the random variables' expected values. These preferences are (weakly) convex and continuous as well as complete on $L^1(\mu)$ (but not on $L^0(\mu)$!) because, by definition, every random variable in $L^1(\mu)$ comes with an expected value. For an example of convex and continuous preferences that are complete for the non-negative random variables in $L^1(\mu)$, let us quote from Nielsen (1984):

> "The conclusion of an exchange between Ryan (1974) and Arrow (1974) was that if u is a concave and increasing function on the non-negative real line, and if Z is a random variable on the non-negative real line with finite expected value, then the expected value of $u(Z)$ is finite." (p.202)

The second approach is to give up on convex preferences by restricting attention to utility functions u that are bounded and therefore neither concave nor convex (cf. Examples 6 and 7 in Section 5). Peter Wakker (1993) already reflects on these two alternative approaches while exploring the role of bounded utility in Savage's (1954) subjective expected utility theory:

[3] We adopt here Schmeidler's (1989) interpretation of ambiguity/uncertainty aversion (resp. seeking) of a CEU decision maker in terms of a convex (resp. concave) capacity.

> "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)

The main insight from our analysis is that the conflict between the three fundamental principles of (i) continuity, (ii) convexity, and (iii) completeness is not specific to any given utility representation such as, e.g., expected or, more generally, Choquet expected utility. Instead, this conflict is a mathematical necessity that affects any model of preferences over random variables.[4]

The remainder of this paper is organized as follows. Section 2 introduces our formal framework. Section 3 derives our main incompatibility results whose implications for utility representations are discussed in Section 4. Section 5 presents several examples which illustrate our analytical findings. Finally, in Section 6 we argue in favor of our topological choice compared to alternative topologies whose definitions of continuity would be compatible with convexity. All formal proofs are relegated to the Appendix.

2 Our topological space of all random variables

We endow the set of all random variables $L^0(\mu)$ with the *topology of convergence in probability* (cf. Chapters 13.10 and 13.11 in Aliprantis and Border 2006). This topology is generated by the translation-invariant metric $d_0 : L^0(\mu) \times L^0(\mu) \to [0,1)$ such that[5]

$$d_0(X,Y) = \int_\Omega \frac{|X-Y|}{1+|X-Y|} d\mu. \tag{4}$$

That is, for any sequence of random variables $\{X_n\}_{n=1}^{\infty}$ we have that

$$d_0(X_n, X) \to 0 \text{ iff } \forall \epsilon > 0,\ \mu(|X_n - X| > \epsilon) \to 0. \tag{5}$$

In Section 6 we will come back to our topology of choice and compare it to alternative topologies in the light of our analytical results.

[4] Another example for this conflict are *risk measures* from the mathematical finance literature. By our incompatibility result, there cannot exist risk measures defined over all random variables that are simultaneously convex and continuous (cf. Example 7 in Section 5).

[5] The distance between any two random variables is obviously zero under this (essential) metric whenever both random variables coincide *μ-almost everywhere*; that is, we distinguish between equivalence classes of μ-measurable functions rather than between functions themselves.

Note that $L^0(\mu)$ is a vector space because the operations of addition and scalar multiplication for all its members are well defined. To state the obvious, $Z = \lambda X + (1-\lambda) Y$ means

$$Z(\omega) = \lambda X(\omega) + (1-\lambda) Y(\omega), \ \mu\text{-a.e.} \tag{6}$$

so that the 'mixture operation' on $L^0(\mu)$ is an 'averaging' of real-valued outcomes in a given state.[6]

Recall that a subset of random variables $L \subseteq L^0(\mu)$ is *convex* if, and only if,

$$Y_1, ..., Y_n \in L \text{ implies } \lambda_1 Y_1 + ... + \lambda_n Y_n \in L \text{ for all } \lambda_i \geq 0 \text{ s.t. } \sum_{i=1}^{n} \lambda_i = 1. \tag{7}$$

Next recall that the *interior* of a given subset of a topological space is the largest (in the sense of set-inclusion) open set included in this subset. The following proposition will be crucial for deriving our subsequent incompatibility result.

Proposition 1. *The only convex subset of $L^0(\mu)$ with non-empty interior is the set $L^0(\mu)$ itself.*

Sketch of the proof (for details see the Appendix): If there exists some Y in the interior of a convex subset of $L^0(\mu)$, denoted L, we can (for an atomless μ) find for any given $Z \in L^0(\mu)$ a collection of Y_i, $i = 1, .., n$, such that (i) all Y_i are sufficiently close to Y with respect to the d_0-metric, implying $Y_i \in L$ for all i, and (ii) $Y + Z = \frac{1}{n} Y_1 + ... + \frac{1}{n} Y_n$. By convexity of L, we have thus that $Y + Z$ for an arbitrary $Z \in L^0(\mu)$ must also belong to L, implying that the $Y + Z \in L$ generate for any Y the whole set $L^0(\mu)$ so that $L^0(\mu) = L$.

Remark 1. Our proof of Proposition 1 uses the whole space $L^0(\mu)$ and it does not necessarily go through for large subsets of $L^0(\mu)$ (but see Remark 3). Consider for example the subset of $L^0(\mu)$ that only contains non-negative random variables

$$L^0_+(\mu) \equiv \left\{X \in L^0(\mu) \mid X(\omega) \geq 0 \ \mu\text{-a.e.}\right\} \tag{8}$$

endowed with (4). For an open, non-empty, convex subset $L \subseteq L^0_+(\mu)$ we only have that $Y + Z \in L$ for all $Z \in L^0_+(\mu)$ but not for non-positive $Z \in L^0(\mu)$. That is, for any $Y \neq \mathbf{0}$ the $Y + Z \in L$ do no longer generate the whole set $L^0_+(\mu)$ so that we cannot conclude that $L^0_+(\mu) = L$. Consequently, one cannot prove the statement of Proposition 1 if

[6] This mixture operation is different from the Anscombe-Aumann (1963) mixture operation which 'averages' in any given state over probability distributions defined as outcomes of Anscombe-Aumann acts thereby resulting in a new distribution instead of a new real number.

one substitutes $L_+^0(\mu)$ for $L^0(\mu)$. Our subsequent incompatibility results will therefore not apply to preferences restricted to the non-negative random variables in $L_+^0(\mu)$ (cf. Example 3 in Section 5).

Remark 2. Observe that Proposition 1 implies that $L^0(\mu)$ is not *locally convex* (Example 8.47 (3) in Aliprantis and Border 2006). This implies in turn that, except for $\mathbf{0}$, there does not exist any continuous functional $f : L^0(\mu) \to \mathbb{R}$ which is linear, i.e., that satisfies for all $X, Y \in L^0(\mu)$,

$$f(\alpha X + \beta Y) = \alpha f(X) + \beta f(Y) \text{ for all } \alpha, \beta \in \mathbb{R}. \tag{9}$$

On the other hand, there exist non-zero continuous linear functionals which separate points from closed convex subsets for the locally convex spaces $L^p(\mu)$ with $1 \leq p \leq \infty$ (Corollary 5.80 in Aliprantis and Border 2006). Proposition 1 can thus not be extended to $L^p(\mu)$ spaces with $1 \leq p \leq \infty$. As a consequence, our incompatibility results will not apply to these spaces.

Remark 3. Note that $L^p(\mu)$ spaces with $0 < p < 1$ endowed with the metric

$$d_p(X, Y) = \int_\Omega |X - Y|^p \, d\mu \tag{10}$$

are also locally non-convex spaces on which only $\mathbf{0}$ exists as continuous linear functional (Theorem 1 in Day 1940). The statement of Proposition 1 can be shown to hold also for $L^p(\mu)$ spaces with $0 < p < 1$ (for a proof see paragraph 1.47 in Rudin 1991). As a consequence, our subsequent incompatibility results obtained for $L^0(\mu)$ can be analogously derived for $L^p(\mu)$ spaces with $0 < p < 1$.

3 Main results

3.1 Incompatibility of convexity and continuity

Consider a binary preference relation $\preceq$ on $L^0(\mu)$ whereby we treat random variables as identical objects if they coincide μ-almost everywhere. The standard interpretations and notational conventions apply: $X \preceq Y$ means that Y is at least as desirable as X; an agent is indifferent between X and Y, denoted $X \sim Y$, iff $X \preceq Y$ and $Y \preceq X$; in addition, we have strict preference, i.e., $X \prec Y$, whenever $X \preceq Y$ holds whereas $Y \preceq X$ does not. We assume that $\prec$ is *asymmetric* (i.e., for all $X, Y \in L^0(\mu)$, $X \prec Y$ implies not $Y \prec X$) and that $\preceq$ is *reflexive* (i.e., for all $X \in L^0(\mu)$, $X \preceq X$). At this point, we neither assume *completeness* nor *transitivity* of $\preceq$ on $L^0(\mu)$ (see below).

Let us introduce the *super-level* (*=weakly better*) set of $X \in L^0(\mu)$ which contains all random variables that are at least as desirable as X:

$$S(X) \equiv \left\{ Z \in L^0(\mu) \,|X \preceq Z \right\}. \tag{11}$$

Similarly, the *sub-level* (*=weakly worse*) set of $X \in L^0(\mu)$ contains all random variables that are weakly less desirable than X:

$$s(X) \equiv \left\{ Z \in L^0(\mu) \,|Z \preceq X \right\}. \tag{12}$$

Note that, by reflexivity of $\preceq$, both sets $s(X)$ and $S(X)$ are non-empty for all $X \in L^0(\mu)$.

Next consider the following definitions of possible properties that a preference relation $\preceq$ on $L^0(\mu)$ may or may not satisfy.

- *Non-triviality*: $\exists X, Y, Z \in L^0(\mu)$ such that $Y \prec X$ and $X \prec Z$.
- *Completeness*: $\forall X, Y \in L^0(\mu)$, $X \preceq Y$ or $Y \preceq X$.
- *d_0-continuity*: $\forall X \in L^0(\mu)$, the super-level set $S(X)$ and the sub-level set $s(X)$ are *closed* sets with respect to the topology of convergence in probability.
- *S-convexity*: $\forall X \in L^0(\mu)$, the super-level set $S(X)$ is convex.
- *s-convexity*: $\forall X \in L^0(\mu)$, the sub-level set $s(X)$ is convex.

Without non-triviality the preference relation $\preceq$ is not very interesting. By completeness, the decision maker is capable of making decisions in any situation. Although completeness might not always be plausible in empirical situations[7], the whole point of this paper is to assume that a decision maker may have preferences over all random variables in $L^0(\mu)$ and study the consequences of this assumption.

In behavioral terms continuity ensures that small changes, with respect to our chosen metric d_0, will not lead to abrupt changes in a decision maker's choice. More precisely, d_0-continuity ensures that whenever a sequence of random variables $\{Y_k\}_{k \in \mathbb{N}}$ with $X \prec Y_k$ for all k converges in probability to a random variable Y, then also $X \preceq Y$, i.e., preferences will not be reversed in the limit. From an applicational perspective, d_0-continuity is necessary for any representation of complete preferences on $L^0(\mu)$ by some continuous utility function (see Section 4).

[7] In support of the empirical relevance of incomplete preferences see, e.g., Danan et al. (2015) and references therein.

S-convexity means that the decision maker likes to mix over the outcomes of random variables; a feature that is closely associated with behavioral concepts like risk- or/and uncertainty aversion as well as preference for diversification. s-convexity means the opposite and is associated with a risk- or/and uncertainty seeking and aversion against diversification.

To sum up: None of these five properties is behaviorally implausible (whereby S-convexity is empirically far more relevant than s-convexity). Nevertheless, convexity and continuity turn out to be incompatible with one another whenever preferences are non-trivial and complete.

Theorem 1. *Consider a binary preference relation $\preceq$ on $L^0(\mu)$ that is non-trivial and complete.*

(a) *The preference relation $\preceq$ cannot simultaneously satisfy d_0-continuity and S-convexity.*

(b) *Neither can $\preceq$ simultaneously satisfy d_0-continuity and s-convexity.*

(c) *If non-triviality or completeness are dropped, then $\preceq$ might simultaneously satisfy d_0-continuity and S-convexity (resp. s-convexity).*

Sketch of the proof (for details see the Appendix): Define the *strictly better* and *strictly worse* sets of $X \in L^0(\mu)$ as follows

$$S^*(X) \equiv \left\{Z \in L^0(\mu) \,|\, X \prec Z\right\}, \tag{13}$$

$$s^*(X) \equiv \left\{Z \in L^0(\mu) \,|\, Z \prec X\right\}. \tag{14}$$

Completeness ensures that the topological structure of $L^0(\mu)$ determines open, resp. closed, sets with respect to the preference relation $\preceq$ so that Proposition 1 becomes applicable (cf. Remark 1). In particular, by completeness, d_0-continuity implies that the sets $S^*(X)$ and $s^*(X)$ must be open in the topology of convergence in probability. But by Proposition 1, $S^*(X)$ and $s^*(X)$ cannot be open if they are non-empty, convex, strict subsets of $L^0(\mu)$. Non-triviality ensures non-emptiness of $S^*(X)$ and $s^*(X)$ as well as $S^*(X), s^*(X) \subsetneq L^0(\mu)$.

3.2 Quasi-concave and quasi-convex preferences

Note that the incompatibility result of Theorem 1 does not require *transitivity* of $\preceq$ which is defined as follows:

- *Transitivity*: $\forall X, Y, Z \in L^0(\mu)$ if $X \preceq Y$ and $Y \preceq Z$, then $X \preceq Z$.

Transitivity is a standard rationality requirement for economic agents that precludes the possibility of simple money pumps (cf. Cubit and Sugden 2001). Next consider the following possible properties of preferences.

- *Quasi-concavity:* $\forall X, Y \in L^0(\mu)$ if $X \preceq Y$, then $X \preceq \alpha X + (1-\alpha) Y$ for all $\alpha \in [0,1]$.
- *Quasi-convexity:* $\forall X, Y \in L^0(\mu)$ if $X \preceq Y$, then $\alpha X + (1-\alpha) Y \preceq Y$ for all $\alpha \in [0,1]$.

The concept of quasi-concavity–formally defined as "uncertainty aversion" over acts in the Anscombe-Aumann (1963) framework–goes back to Gilboa and Schmeidler (1989, Axiom A.5) and Schmeidler (1989).[8] Because our formal definition of quasi-concavity applies to the outcomes of random variables, the meaning of our definition is different from the original one formulated for the Anscombe-Aumann framework.

Note that S-convexity implies quasi-concavity. Similarly, s-convexity implies quasi-convexity. In what follows we establish that these relationships also hold in the other direction whenever transitivity is satisfied.

Proposition 2. *Assume that $\preceq$ on $L^0(\mu)$ is complete and transitive.*

(a) *Then quasi-concavity implies S-convexity.*

(b) *Then quasi-convexity implies s-convexity.*

Combining Theorem 1 and Proposition 2 gives the following results.

Theorem 2. *Assume that $\preceq$ is non-trivial, complete, transitive, and d_0-continuous.*

(a) *Then $\preceq$ must violate quasi-concavity.*

(b) *Then $\preceq$ must violate quasi-convexity.*

[8] As motivation for his definition, Schmeidler (1989) writes: "Intuitively, uncertainty aversion means that "smoothing" or averaging utility distributions makes the decision maker better off. Another way is to say that substituting objective mixing for subjective mixing makes the decision maker better off." (p.582) For an alternative approach to uncertainty aversion defined over random variables (i.e., Savage acts) rather than over Anscombe-Aumann acts see Epstein (1999).

3.3 Preference for diversification

Eddie Dekel (1989) has introduced the following definition in the context of portfolio choices:[9]

- *Preference for diversification:* $\forall X, Y \in L^0(\mu)$ if $X \sim Y$, then $X \preceq \alpha X + (1-\alpha) Y$ for all $\alpha \in [0,1]$.

Quasi-concavity implies preference for diversification. The proof of the following result establishes that preference for diversification implies quasi-concavity under transitivity and d_0-continuity.

Theorem 3. *Assume that $\preceq$ on $L^0(\mu)$ is complete, transitive, and d_0-continuous. Then $\preceq$ must violate preference for diversification.*

Remark 4. One implication of the above analysis is that so-called *convex/coherent/subadditive* risk measures defined over all $L^0(\mu)$ cannot be continuous (see, e.g., Föllmer and Schied 2002; Delbaen 2002, 2009; Assa 2016). The next section discusses implications of our analytical findings for utility representations.

4 Implications for utility representations

This section assumes that preferences on $L^0(\mu)$ are represented by some utility functional.

Assumption 1. *Fix some non-trivial and complete preference relation $\preceq$ on $L^0(\mu)$ and suppose that there exists some functional $U : L^0(\mu) \to \mathbb{R}$ such that, for all $X, Y \in L^0(\mu)$,*

$$X \preceq Y \text{ iff } U(X) \leq U(Y). \tag{15}$$

We say that U is continuous (in probability) if $d_0(Y_k, Y) \to 0$ implies $\lim_k U(Y_k) = U(Y)$. For quasi-concave U it must hold that

$$\forall X, Y \in L^0(\mu), \alpha \in (0,1), U(\alpha X + (1-\alpha) Y) \geq \min\{U(X), U(Y)\} \tag{16}$$

whereas we have for quasi-convex U that

$$\forall X, Y \in L^0(\mu), \alpha \in (0,1), \max\{U(X), U(Y)\} \geq U(\alpha X + (1-\alpha) Y). \tag{17}$$

[9]For extensions of Dekel's (1989) approach see Chateauneuf and Lakhnati (2007) and Chateauneuf and Tallon (2002). For an excellent survey on this literature see De Giorgi and Mahmoud (2016).

Proposition 4. *Suppose that Assumption 1 holds. If U is continuous, then U can neither be quasi-concave nor quasi-convex.*

In the remainder of this section we discuss implications for expected utility and Choquet expected utility, respectively.

4.1 Expected utility

Suppose that the utility representation (15) is of the expected utility (EU) form, i.e., for all $X \in L^0(\mu)$,

$$U(X) \equiv E(u(X)) \tag{18}$$

$$= \int_\Omega u(X(\omega))\, d\mu \tag{19}$$

for some utility function $u : \mathbb{R} \to \mathbb{R}$. Since the EU representation (18) is continuous, the represented preferences $\preceq$ satisfy d_0-continuity. Next observe that quasi-concavity of EU preferences holds if u is concave thereby formally expressing risk-aversion of the EU decision maker. Conversely, quasi-convexity of EU preferences holds if u is convex thereby expressing risk-seeking. By Proposition 4, we thus obtain the following result.

Corollary 1. *Suppose that Assumption 1 holds such that U is of the EU form (18). Then the utility function u can neither be concave nor convex.*

By Corollary 1, an EU representation over all random variables can thus neither express global risk-aversion nor global risk-seeking. We will come back to this point in our Examples 3 and 6 in Section 5.

Remark 5. The quintessence of Corollary 1 already appears in the EU literature in the form of existence conditions for the integral (19) (cf. Nielsen 1984; Wakker 1993; Delbaen, Drapeau and Kupper 2011 and references therein). A main insight from this literature is that boundedness of u is required for any EU representation defined over all random variables: for unbounded u we can always find random variables for which the integral (19) does not exist.[10] To see the connection between this literature and our Corollary 1, observe that any (non-constant) concave u is unbounded from below whereas any (non-constant) convex u is unbounded from above.

[10] Cf. Wakker (1993, p.448): "The underlying problem was already observed by Menger (1934). As soon as utility is unbounded, there exist acts with unbounded expected utility[...]."

Remark 6. Beyond this purely decision-theoretic literature, macro-economists have observed that 'model uncertainty' may easily lead to exploding moments of expected utility functions (or of the stochastic discount factor) for unbounded utility functions that are standard in the literature (cf. Geweke 2001; Weitzman 2007). This insight culminated in Weitzman's (2009) *Dismal Theorem* about modeling preferences over random consumption streams: "Seemingly thin-tailed probability distributions (like the normal), which are actually only thin-tailed conditional on known structural parameters of the model (like the standard deviation), become tail-fattened (like the Student-t) after integrating out the structural-parameter uncertainty. This core issue is generic and cannot be eliminated in any clean way." (p.9)

4.2 Choquet expected utility

Consider now a utility representation (41) which is of the Choquet expected utility (CEU) form, i.e., for all $X \in L^0(\mu)$,

$$U(X) \equiv E^C(u(X)) \tag{20}$$

$$= \int_\Omega^C u(X(\omega))\, dv \tag{21}$$

where the integral in (20) is the Choquet integral with respect to a nonatomic capacity ν on (Ω, Σ) that is equivalent to μ. The Choquet integral is formally defined as

$$\int_\Omega^C u(X(\omega))\, dv \equiv \int_0^\infty \nu(u(X(\omega)) \geq x)\, dx - \int_{-\infty}^0 (1 - \nu(u(X(\omega)) \geq x))\, dx \tag{22}$$

(for details on Choquet integration and properties of the Choquet integral see Schmeidler 1986 for bounded u and, more generally, Wakker 1993).

We follow the literature and call ν *convex* iff, for all $A, B \in \Sigma$,

$$\nu(A \cup B) + \nu(A \cap B) \geq \nu(A) + \nu(B). \tag{23}$$

ν is called *concave* iff the inequality in (23) is reversed. CEU has been axiomatized in Schmeidler (1989) and in Gilboa (1987) whereby Schmeidler (1989) associates convex (resp. concave) ν with ambiguity/uncertainty aversion (resp. seeking) (cf. Footnote 8).

Since the CEU representation (20) is continuous, Proposition 4 implies that (20) cannot represent preferences that satisfy either quasi-concavity or quasi-convexity. For the EU representation (18) quasi-concavity (resp. quasi-convexity) of U is simply implied by concavity (resp. convexity) of u. The case is more complicated for the CEU representation (20) for which we must additionally consider properties of ν. In what follows we derive conditions that imply quasi-concavity of any CEU representation.

By Proposition 3 in Schmeidler (1986), convexity of ν implies, for any $\lambda \in [0,1]$,

$$E^C(\lambda u(X) + (1-\lambda) u(Y)) \geq E^C(\lambda u(X)) + E^C((1-\lambda) u(Y)) \text{ for all } X, Y \in L^0(\mu), \tag{24}$$

which is equivalent to

$$E^C(\lambda u(X) + (1-\lambda) u(Y)) \geq \lambda E^C(u(X)) + (1-\lambda) E^C(u(Y)) \text{ for all } X, Y \in L^0(\mu) \tag{25}$$

since the Choquet integral is homogeneous of degree one. For concave u we have that

$$u(\lambda X + (1-\lambda) Y) \geq \lambda u(X) + (1-\lambda) u(Y) \tag{26}$$

so that monotonicity of the Choquet integral implies, by (25),

$$E^C(u(\lambda X + (1-\lambda) Y)) \geq \lambda E^C(u(X)) + (1-\lambda) E^C(u(Y)) \text{ for all } X, Y \in L^0(\mu). \tag{27}$$

For all $X, Y \in L^0(\mu)$ such that $E^C(u(X)) \geq E^C(u(Y))$, we obtain from (27) that

$$E^C(u(\lambda X + (1-\lambda) Y)) \geq E^C(u(Y)), \tag{28}$$

which is the definition of a quasi-concave $E^C(u(\cdot))$. Collecting the above arguments gives the following result (the argument for quasi-convexity proceeds analogously).

Corollary 2. *Suppose that Assumption 1 holds such that U is of the CEU form (20). Then we can neither have (i) concavity of the utility function u combined with convexity of the capacity ν nor (ii) convexity of the utility function u combined with concavity of the capacity ν.*

Loosely speaking, Corollary 2 establishes that a CEU representation over all random variables can neither express (i) global risk aversion combined with global ambiguity/uncertainty aversion nor (ii) global risk seeking combined with global ambiguity/uncertainty seeking.

Remark 7. Corollary 1 is obviously a special case of Corollary 2. Whenever the capacity ν is simultaneously convex and concave it becomes the additive measure μ so that the CEU representation (20) becomes the EU representation (18). Similar to our Remark 5 about unbounded utility functions in an EU representation, Schmeidler (1989,p.580) ensures existence of (20) by restricting attention to a bounded utility function u.

5 Examples

This section illustrates our analytical results through examples that relax different assumptions of Theorem 1 in order to ensure existence of preferences and/or their utility representations.

Example 1. [Relaxing non-triviality]. Just consider a degenerate preference relation such that $\forall X, Y \in L^0(\mu), X \sim Y$. This preference relation is (trivially) complete, d_0-continuous as well as S-convex (resp. s-convex). Moreover, it can be represented by any constant functional $U : L^0(\mu) \to \mathbb{R}$.$\square$

Example 2. [Relaxing completeness: Monotonicity]. We say that Y *dominates* (μ-a.e.) X, denoted $X \leq Y$, iff

$$X(\omega) \leq Y(\omega) \;\; \mu\text{-a.e.} \tag{29}$$

Let $\forall X, Y \in L^0(\mu), X \preceq Y$ iff $X \leq Y$ and observe that continuity and convexity hold for these incomplete preferences.$\square$

Example 3. [Relaxing completeness: Risk-averse expected utility for non-negative random variables]. Suppose that $\preceq$ is only defined on the set of non-negative random variables

$$L^0_+(\mu) \equiv \left\{ X \in L^0(\mu) \mid X(\omega) \geq 0 \;\mu\text{-a.e.} \right\} \tag{30}$$

and consider an expected utility decision maker with the following utility function

$$u(x) = \frac{x}{1+x}, \text{ for } x \geq 0. \tag{31}$$

The expected utility of any $X \in L^0_+(\mu)$ is given as the distance (4) of X from the constantly zero random variable:

$$\int_\Omega u(X(\omega))\, d\mu = \int_\Omega \frac{|X(\omega) - 0|}{1 + |X(\omega) - 0|} d\mu \tag{32}$$

$$= d_0(X, \mathbf{0}) \in [0, 1). \tag{33}$$

This decision maker's preferences on $L^0_+(\mu)$ are continuous and, by strict concavity of u on $\mathbb{R}_+$, they are also S-convex on $L^0_+(\mu)$.

On the one hand, this example shows that continuity and convexity can be easily reconciled if we restrict attention to preferences that are only complete on a suitable subset of $L^0(\mu)$ like $L^0_+(\mu)$ (cf. Remark 1). On the other hand, however, this example also demonstrates that incompleteness can be a very unnatural assumption: Why should the decision maker not have preferences over random variables with losses (negative x) in their support? We come back to this situation in Example 6 where we consider complete preferences.□

Example 4. [Relaxing continuity: Lexicographic preferences]. Define *(strict) dominance on an event* $E \in \Sigma$ as follows: $\forall X, Y \in L^0(\mu)$

$$X \leq_E Y \text{ iff } X(\omega) \leq Y(\omega) \;\; \mu\text{-a.e. on } E;$$
$$X <_E Y \text{ iff } X \leq_E Y \text{ and } X(\omega) < Y(\omega) \text{ on some } E' \subseteq E \text{ with } \mu(E) > 0.$$

Fix a collection $\Omega_1, \Omega_2, ...$ of nested events in Σ such that $\Omega_{i+1} \subset \Omega_i$, $\mu(\Omega_1) = 1$ and $\mu(\Omega_i) > \mu(\Omega_{i+1}) > 0$ for all i. Define the following lexicographic preferences:

if $X <_{\Omega_1} Y$ then $X \prec Y$,
if neither $Y <_{\Omega_i} X$ nor $X <_{\Omega_i} Y$ for any $i < j$ but $X <_{\Omega_j} Y$, then $X \prec Y$,
$X \sim Y$, else.

First, let us show that the (complete and non-trivial) preference relation $\preceq$ is S-convex. If not, then $X \preceq Y$ but $\lambda Y + (1-\lambda) X \prec X$ for some λ. Focus on the strict case $X \prec Y$. Then there exists some $i \geq 1$ and X, Y such that $X <_{\Omega_i} Y$ but neither $Y <_{\Omega_j} X$ nor $X <_{\Omega_j} Y$ for $j < i$. Note that $X <_{\Omega_i} Y$ implies $X <_{\Omega_i} \lambda Y + (1-\lambda) X$. Similarly, neither $Y <_{\Omega_j} X$ nor $X <_{\Omega_j} Y$ implies neither $Y <_{\Omega_j} \lambda Y + (1-\lambda) X$ nor $X <_{\Omega_j} \lambda Y + (1-\lambda) X$ for $j < i$. Consequently, $X \prec \lambda Y + (1-\lambda) X$, a contradiction. Now focus on $X \sim Y$ so that, by the same argument, neither $Y <_{\Omega_j} \lambda Y + (1-\lambda) X$ nor $X <_{\Omega_j} \lambda Y + (1-\lambda) X$ for any j, implying $\lambda Y + (1-\lambda) X \sim X$.

Next observe that $\preceq$ is not d_0-continuous. To see this, let $\Omega_1 = E_1 \cup E_2$, $\Omega_2 = E_1$ and consider the following random variables:

	E_1	E_2
X	1	0
Y_k	$1-\frac{1}{k}$	1
Y	1	1

Note that $Y_k \prec X$ for all k but $X \prec Y$ whereby $d_0(Y_k, Y) \to 0$.□

Example 5. [Relaxing continuity: Preferences generated by a linear functional]. Suppose that there exists a non-zero linear functional f on $L^0(\mu)$. Then we can use f to construct a non-trivial, complete, and convex preference relation as follows:

$$X \preceq_f Y \text{ iff } f(X) \leq f(Y). \tag{34}$$

This preference relation is non-trivial since f is non-zero (and by linearity thus non-constant). It is complete since for all $X, Y \in L^0(\mu)$ we have either $f(X) \leq f(Y)$ or $f(Y) \leq f(X)$. It is convex since, for all $X, Y, Z \in L^0(\mu)$, if $f(Z) \leq f(X)$ and $f(Z) \leq f(Y)$ then

$$
\begin{aligned}
f(Z) &= \lambda f(Z) + (1-\lambda) f(Z) & (35)\\
&\leq \lambda f(X) + (1-\lambda) f(Y) = f(\lambda X + (1-\lambda) Y). & (36)
\end{aligned}
$$

Recall from our Remark 2 that there does not exist any (non-zero) continuous linear functional on $L^0(\mu)$. However, that does not mean that there does not exist any linear functional on this space at all. In what follows, we prove the existence of a linear function on $L^0(\mu)$ whereby we use Zorn's lemma (cf. pp.65-66 in Komjàth and Totik 2006):

Zorn's Lemma. *Suppose that a non-empty partially ordered set $(\mathcal{Z}, R)$ has the property that every chain has an upper bound, i.e., for any totally ordered set $\mathcal{C} \subseteq \mathcal{Z}$ there exists $\mathcal{M}_{\mathcal{C}}$ such that $\mathcal{X} R \mathcal{M}_{\mathcal{C}}$ for all $\mathcal{X} \in \mathcal{C}$. Then the set $\mathcal{Z}$ contains at least one maximal element $\mathcal{M}$, i.e., there is no $\mathcal{X} \in \mathcal{Z}$ with $\mathcal{M} R \mathcal{X}$ and $\neg \mathcal{X} R \mathcal{M}$.*

Let $\mathcal{O}$ be the set of all linearly independent subsets of $L^0(\mu)$ that contains the constant random variable $\mathbf{1}$. Because of $\{\mathbf{1}\} \in \mathcal{O}$, $\mathcal{O}$ is non-empty. In Zorn's lemma let $\mathcal{Z} = \mathcal{O}$ and $R = \subseteq$. Since $\mathcal{O}$ is a set of subsets of $L^0(\mu)$, $(\mathcal{O}, \subseteq)$ is partially ordered. On the other hand, for any chain $\mathcal{C}$ one can see that $\mathcal{M}_{\mathcal{C}} = \cup_{A \in \mathcal{C}} A$ is an upper bound. By Zorn's lemma, there must thus exist a maximal set $\mathcal{M}$ of linearly independent members in $L^0(\mu)$ that also contains $\mathbf{1}$. We claim that $\mathcal{M}$ is a basis for $L^0(\mu)$. If not, there exists some $X \in L^0(\mu)$ such that X cannot be written as linear combination of members in $\mathcal{M}$. That means X is linearly independent from members of $\mathcal{M}$. But if we introduce $\mathcal{X}' = \mathcal{M} \cup \{X\}$, then $\mathcal{M} \subsetneq \mathcal{X}'$, which contradicts the maximality of $\mathcal{M}$.

Now let us construct a linear functional f_1 as follows: for every $X \in L^0(\mu)$, there are real numbers $\{x_m\}_{m \in \mathcal{M}}$ such that $X = \sum_{m \in \mathcal{M}} x_m m$. Let $f_1(X) := x_1$. Since $\mathcal{M}$ is a basis, the representation $X = \sum_{m \in \mathcal{M}} x_m m$ is unique, and as a result f_1 is well defined and linear. $\square$

Example 6. [Relaxing convexity: Expected utility with a reference point at zero]. Recall the situation of Example 3 but assume now a complete preference ordering on $L^0(\mu)$. Define the following (once-differentiable) utility function:

$$
u(x) = \begin{cases} \frac{x}{1+x} & \text{if } x \geq 0 \\ \frac{x}{1-x} & \text{if } x \leq 0 \end{cases} \tag{37}
$$

resulting in an EU representation of continuous preferences $\preceq$ on $L^0(\mu)$. As under Example 3, the expected utility of any μ-a.e. positive X is its distance $d_0(X, \mathbf{0})$ from the constant zero random variable. For an μ-a.e. negative Y we have

$$\int_\Omega u(Y(\omega))\, d\mu = \int_\Omega \frac{(-)\,|Y(\omega) - 0|}{1 + |Y(\omega) - 0|} d\mu = -d_0(Y, \mathbf{0})\,; \tag{38}$$

that is, the expected utility of the negative Y is the negative of its distance from this zero random variable. Consequently, $U(X) \in (-1, 1)$ for any $X \in L^0(\mu)$.

Observe that u is strictly concave for all $x > 0$ and strictly convex for all $x < 0$ so that the EU decision maker is risk-averse for positive and risk-seeking for negative outcomes. From Corollary 1 we know that an EU representation of a preference relation $\preceq$ on $L^0(\mu)$ is impossible for an utility function that is concave (or convex) on the whole domain $\mathbb{R}$. This example shows that we can have an EU representation of preferences on $L^0(\mu)$ when we are prepared to give up S-convexity (corresponding to a concave u, i.e., risk-aversion) as well as s-convexity (corresponding to a convex u, i.e., risk-seeking) as global properties.

Finally, let us interpret u as a value function from *prospect theory* (cf. Wakker 2010) such that positive x correspond to gains with respect to the reference point zero whereas negative x stand for losses. Under this interpretation giving up on S- and s-convexity for the above preferences is nothing else than the standard assumption of prospect theory according to which the bounded value function for gains is (strictly) concave whereas it is (strictly) convex for losses (cf. Vendrik and Woltjer 2007 and references therein).□

Example 7. [Relaxing convexity: Value-at-Risk]. Recall the definition of *Value at Risk* (VaR) as a popular risk measure in financial applications which is not sub-additive:

$$\mathrm{VaR}_\alpha(X) = -\sup\{x \in \mathbb{R} | P(X \geq x) \geq \alpha\} \tag{39}$$

for a fixed confidence level $1 - \alpha \in (0, 1)$. Let $\forall X, Y \in L, X \preceq Y$ iff $\mathrm{VaR}_\alpha(X) \geq \mathrm{VaR}_\alpha(Y)$. It is easy to see that (the complete and non-trivial) $\preceq$ is d_0-continuous because d_0-continuity implies convergence in distribution. The following example taken from Embrechts et al. (2002) shows that S-convexity is violated. Let X, Y be two independent Pareto distributed random variables with $F_X(x) = F_Y(x) = 1 - x^{-1/2}, x \geq 1$ and 0, otherwise. Then it is easy to see that $P(X + Y \leq z) = 1 - \frac{2\sqrt{z-1}}{z} < P(2X \leq z)$, for $z \geq 2$. Consequently, $\mathrm{VaR}_\alpha(\frac{X+Y}{2}) > \mathrm{VaR}_\alpha(X) = \frac{\mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y)}{2}$. That is, we have $X, Y \in S(X)$ but not $\frac{X+Y}{2} \in S(X)$ so that S-convexity fails.

As the basis for the Basel II and III capital requirement formula, the VaR criterion has been heavily criticized in the mathematical finance literature because it does not satisfy preference for diversification (cf. Artzner et al. 1997, 1999). On the other hand, VaR has the nice feature to ensure continuity of preferences on $L^0(\mu)$, which is impossible for convex/coherent/subadditive risk measures (see Remark 4).□

6 Discussion: Our topology of choice

Mathematical continuity is a relative concept that is determined by the topology we impose on $L^0(\mu)$. We will show in a moment that it is easy to come up with topologies on $L^0(\mu)$ that can reconcile convexity with mathematical continuity with respect to these topologies. This raises the question why we have chosen the topology of convergence in probability.

The remainder of this section presents three arguments in favor of the d_0-metric as our topology of choice. These arguments can be summarized as follows:

1. A utility representation over the distributions of random variables is continuous if, and only if, d_0-continuity holds.

2. The d_0-metric is behaviorally plausible and it translates the standard convergence behavior of random variables from familiar $L^p(\mu)$ spaces into the larger $L^0(\mu)$ space.

3. Any alternative topologies we can think of that reconcile convexity with mathematical continuity require behaviorally implausible notions of convergence.

6.1 Continuous utility representation over distributions

Let us assume that a non-trivial and complete preference relation on $L^0(\mu)$ can be represented by some utility function defined over the distributions of all random variables in $L^0(\mu)$.[11] Recall that the *distribution* F_Z of any $Z \in L^0(\mu)$ is a probability measure on the Borel subsets of the real line satisfying

$$F_Z(A) \equiv \mu(\{\omega \in \Omega \mid Z(\omega) \in A\}). \tag{40}$$

Assumption 2. *Fix some non-trivial and complete preference relation $\preceq$ on $L^0(\mu)$ and suppose that there exists some real-valued U such that, for all $X, Y \in L^0(\mu)$,*

$$X \preceq Y \text{ iff } U(F_X) \leq U(F_Y). \tag{41}$$

[11] The majority of utility representations reduces preferences over random variables to preferences over distributions. Notable exceptions are *state-dependent utility models*. For a good textbook treament of state-dependent expected utility see Chapter 6.E in Mas-Collel et al. (1995). For a recent overview on objective and subjective models with state-dependent utility see Karni and Schmeidler (2016) and references therein.

For a sequence of random variables $\{Y_k\}_{k\in\mathbb{N}}$ we write $F_{Y_k} \Rightarrow F_Y$ whenever the Y_k *converge in distribution* to Y, i.e., whenever the cumulative distribution functions (=cdf) of the Y_k converge weakly to the cdf of Y.[12] We say that U is continuous in distribution if $F_{Y_k} \Rightarrow F_Y$ implies $\lim_k U(F_{Y_k}) = U(F_Y)$.

Proposition 5. *Suppose that Assumption 2 holds. U is continuous in distribution if, and only if, $\preceq$ is d_0-continuous.*

Most decision-theoretic applications are concerned with the maximization of utility functions over distributions whereby–mainly out of analytical convenience–these utility functions are supposed to be continuous. By Proposition 5, such analytical convenience would not be at hand without d_0-continuity.

6.2 $L^p(\mu)$ spaces and the d_0-metric

Beyond the mere mathematical definition of continuity there is also a behavioral interpretation of what it means that a decision maker has ‘continuous preferences’. According to this behavioral interpretation of continuity, preferences should not abruptly switch in the limit of converging random variables. A good behavioral concept of continuity should therefore be based on a behaviorally plausible concept of convergence that closely captures what real-life decision makers may perceive as convergence of random variables.

Let us consider the familiar $L^p(\mu)$ spaces with $1 \leq p \leq \infty$ which only contain random variables that come with an expected value.[13] The standard topology imposed on these spaces is generated by the L_p-*norm*

$$\|X\|_p = \begin{cases} \left[\int_\Omega |X|^p \, d\mu\right]^{\frac{1}{p}} & \text{for } 1 \leq p < \infty \\ \inf\{\alpha \in [0,\infty) \mid \mu(\{\omega \in \Omega \mid |X(\omega)| > \alpha\}) = 0\} & \text{for } p = \infty \end{cases} \tag{42}$$

with corresponding metric

$$d_p(X,Y) = \|X - Y\|_p \text{ for all } X, Y \in L^p(\mu). \tag{43}$$

Arguably, most decision-theorists would agree that convergence in the d_p-metric is a behaviorally plausible notion for the convergence behavior of random variables in $L^p(\mu)$.

[12] Denote by CDF_Z the cdf of Z, formally defined as

$$CDF_Z(x) \equiv F_Z(-\infty, x] \text{ for all } x \in \mathbb{R}.$$

The CDF_{Y_k} converge weakly to the CDF_Y iff $CDF_{Y_k}(x) \to CDF_Y(x)$ for all x such that $\mu(Y = x) = 0$; (for more details see Chapter 14 in Billingsley 1995).

[13] For properties of $L^p(\mu)$ spaces with $1 \leq p \leq \infty$ see Section 19 in Billingsley (1995).

When we move from an $L^p(\mu)$ space to the large $L^0(\mu)$ space, where the metric d_p is no longer available in general, it would be desirable to have a metric for $L^0(\mu)$ that guarantees for any sequence $\{Y_k\}_{k\in\mathbb{N}} \subset L^p(\mu)$ the same convergence behavior in $L^0(\mu)$ as under the d_p-metric. The following proposition shows that the d_0-metric is accomplishing this task.

Proposition 6. *Fix some $L^p(\mu)$ space with $1 \leq p \leq \infty$. Convergence in the d_p-metric implies convergence in the d_0-metric, i.e.,*

$$d_p(Y_k, Y) \to 0 \text{ implies } d_0(Y_k, Y) \to 0. \tag{44}$$

6.3 Alternative topologies that establish compatibility between convexity and continuity

To see that it is actually trivial to ensure compatibility of convexity with some notion of mathematical continuity, let us first consider the *discrete topology* on $L^0(\mu)$ generated by the discrete metric metric $d^D : L^0(\mu) \times L^0(\mu) \to [0, 1)$ such that

$$d^D(X, Y) = \begin{cases} 0 & X = Y,\ \mu\text{-}a.e. \\ 1 & \text{else.} \end{cases} \tag{45}$$

In this topology all subsets of $L^0(\mu)$ are closed to the effect that convexity and d^D-continuity are compatible. To impose d^D-continuity, however, comes with the drawback that the corresponding notion of convergence is behaviorally not very plausible.

Example 8. [Convergence under d^D-continuity]. Let us revisit the lexicographic preferences of Example 4 which satisfy convexity but violate d_0-continuity. These (like any other) preferences trivially satisfy d^D-continuity because only (eventually) constant sequences of random variables converge under the discrete topology. For the random variables

	E_1	E_2
Y_k	$1 - \frac{1}{k}$	1
Y	1	1

we can thus have under d^D-continuity that $Y_k \prec X$ for all k as well as $X \prec Y$ because the Y_k no longer converge to Y, i.e.,

$$\lim_{k\to\infty} d^D(Y_k, Y) = 1. \tag{46}$$

Arguably, most real-life decision makers would judge that (or: behave as if) the Y_k were increasingly resembling Y for larger k whereby the difference between the Y_k and Y becomes negligible in the limit. But then any behaviorally relevant concept of continuity should be based on the notion that the Y_k are indeed converging to Y, which is not the case under d^D-continuity.□

The discrete topology stands for the largest topology under which any given convex preference relation over equivalence classes of random variables becomes continuous. Alternatively, we might consider the smallest topology under which a given convex preference relation becomes continuous. More precisely, fix some convex preference relation $\preceq$ and introduce the smallest topology whose closed sets consist of a basis given by super- and sub-level sets $s(X), S(X), \forall X \in L^0(\mu)$. Indeed, this topology is the smallest topology under which $\preceq$ is continuous and it is also included in any such topology. However, the same criticism as under Example 8 applies: Making the (convex) lexicographic preferences of Example 4 continuous is incompatible with any topology in which Y belongs to some closed set containing all Y_k. As in the case of the discrete topology, the notion of convergence required to make the preferences of Example 4 continuous is therefore not plausible from a behavioral perspective.

So far we have considered topologies that treat random variables which coincide μ-almost everywhere as identical objects. If we are prepared to give up this notion of equivalence classes of random variables, preference relations on $L^0(\mu)$ become possible that can combine convexity with mathematical continuity.

Example 9. [Abandoning equivalence classes of μ-a.e. random variables]. Let $\Omega = [0, 1)$ and consider the topology for which convergence means (i) convergence in d_0 and (ii) for all points in $[0, 1)$, i.e.,

$$\text{for any net } X_\lambda \to X \text{ iff } d_0(X_\lambda, X) \to 0 \text{ and } X_\lambda(\omega) \to X(\omega), \forall\omega \in [0, 1). \tag{47}$$

The analysis in Aliprantis and Burkinshaw (1978, p.114) implies that, for any $\omega \in [0, 1)$, $f_\omega(X) = X(\omega)$ is a continuous functional. Consequently, for any fixed $\omega \in [0, 1)$, the complete preference relation $\preceq$ defined by

$$X \prec Y \text{ iff } X(\omega) < Y(\omega) \tag{48}$$

is continuous. Obviously, this preference relation is also convex.□

Preferences described under Example 9 suffer from the interpretational drawback that the decision maker must care about probability zero events under our assumption of an atomless μ. In our opinion, it is behaviorally more plausible for decision makers to treat random variables as identical objects in case they are identical almost everywhere.

Remark 8. Example 9 also demonstrates why the assumption of a nonatomic measure space is crucial to our analysis. Suppose, for example, that $\mu(\omega) > 0$ with ω given by (48). Then the preferences of Example 9 are (i) convex as well as continuous whereby (ii) the decision maker's preferences no longer depend on a probability zero event.

Appendix: Formal proofs

Proof of Proposition 1. Let L be a convex subset of $L^0(\mu)$ with non-empty interior and suppose that $Y \in L^0(\mu)$ belongs to the interior of L. Fix some $\epsilon > 0$ such that $X \in L$ whenever $d_0(Y, X) \leq \epsilon$. Pick some partition $\{\Omega_1, ..., \Omega_n\}$ of Ω such that $\mu(\Omega_i) \leq \epsilon, i = 1, ..., n$, which always exists for nonatomic μ. Choose $Z \in L^0(\mu)$ arbitrarily and introduce $Y_i = Y + nZ1_{\Omega_i}$ where 1_{Ω_i} denotes the indicator function on Ω_i. For any $i = 1, ..., n$ we have

$$d_0(Y, Y_i) = \int_\Omega \frac{|Y - Y_i|}{1 + |Y - Y_i|} d\mu = \int_\Omega \frac{|nZ1_{\Omega_i}|}{1 + |nZ1_{\Omega_i}|} d\mu = \int_\Omega \frac{|nZ|}{1 + |nZ|} 1_{\Omega_i} d\mu \quad (49)$$

$$< \int_\Omega 1_{\Omega_i} d\mu = \mu(\Omega_i) \leq \epsilon. \quad (50)$$

Consequently, we have $Y_i \in L$ for all $i = 1, ..., n$.

Next note that

$$\frac{1}{n}\sum_{i=1}^n Y_i = Y + \sum_{i=1}^n Z1_{\Omega_i} \quad (51)$$

$$= Y + Z. \quad (52)$$

By convexity of L, we thus have

$$Y + Z = \frac{1}{n}Y_1 + ... + \frac{1}{n}Y_n \in L. \quad (53)$$

Since $Z \in L^0(\mu)$ was chosen arbitrarily, we obtain $L = L^0(\mu)$, which proves the lemma.□□

Proof of Theorem 1. Ad part (a).

Step 1. Suppose that $S(X)$ is convex. By non-triviality, we have $Y \notin S(X)$ which implies $S(X) \neq L^0(\mu)$. By Proposition 1, $S(X)$ must thus have an empty interior with respect to the topology of convergence in probability on $L^0(\mu)$.

Step 2. By non-triviality, we also have that the set

$$S^*(X) \equiv \left\{X' \in L^0(\mu) \,|\, X \prec X'\right\} \subset S(X) \quad (54)$$

is non-empty because of $Z \in S^*(X)$.

Step 3. Combining Step 1 and Step 2 establishes that $S^*(X)$ cannot be an open set in the topology of convergence in probability. However, by completeness,

$$s(X) = L^0(\mu) \backslash S^*(X) \quad (55)$$

so that $s(X)$ cannot be a closed set, which contradicts d_0-continuity.□

Ad part (b). Just observe that non-triviality implies (i), by $s(X) \neq L^0(\mu)$ and Proposition 1, that $s(X)$ has an empty interior as well as (ii) non-emptiness of

$$s^*(X) \equiv \left\{X' \in L^0(\mu) \,|X' \prec X\right\} \subset s(X). \tag{56}$$

By an analogue argument as under Step 3, the set

$$S(X) = L^0(\mu) \setminus s^*(X) \tag{57}$$

is thus not closed.□

Ad part (c). The validness of this statement is demonstrated through the examples in Section 3. □□

Proof of Proposition 2. We prove part (a). If $S(X) = \{X\}$ there is nothing to prove so let us assume that $Y, Z \in S(X)$ with $Y \neq Z$. Without loss of generality, suppose that, by completeness, $Y \preceq Z$. If quasi-concavity holds, we have $Y \preceq \alpha Z + (1-\alpha) Y$. Finally, since $X \preceq Y$, transitivity implies $X \preceq \alpha Z + (1-\alpha) Y$.□□

Proof of Theorem 3. By Theorem 2 it is sufficient to show that quasi-concave preferences follow from preference for diversification under the assumptions of Theorem 3.

Step 1. Without loss of generality, suppose that $X \preceq Y$ with $X \neq Y$ (again: if $S(X) = \{X\}$, we don't have anything to prove). We have to show that preference for diversification implies

$$X \preceq \alpha X + (1-\alpha) Y \tag{58}$$

for $\alpha \in [0,1]$. If $X \sim Y$, we immediately obtain (58). So, let us assume $X \prec Y$.

Step 2. Because any metric is continuous (3.16 Theorem in Aliprantis and Border 2006), we obtain:

Lemma 1. *Fix some $\epsilon \geq 0$. For any $X, Y \in L^0(\mu)$, there exists some $\delta > 0$ such that*

$$d_0(q'X + (1-q')Y, qX + (1-q)Y) \leq \epsilon \tag{59}$$

for all $|q' - q| \leq \delta$.

Step 3. Introduce

$$q^* = \max\{q \in [0,1] \,|X \preceq \alpha X + (1-\alpha) Y, \forall \alpha \in [0,q]\}. \tag{60}$$

By transitivity, we have

$$X \preceq \alpha X + (1-\alpha) Y \tag{61}$$

iff $\alpha \in [0, q^*]$. If $q^* = 1$, we have the desired result (58). Suppose now $0 \leq q^* < 1$. By $X \prec Y$, d_0–continuity and completeness implies that the set

$$S^*(X) \equiv \left\{Z \in L^0(\mu) \,|X \prec Z\right\} \tag{62}$$

is open. Consequently, there exists some number $\epsilon > 0$ such that $d_0(Y, Z) \leq \epsilon$ implies $X \prec Z$, i.e., $Z \in S^*(X)$. By Lemma 1, there exists some $\delta > 0$ such that $d_0(Y, \alpha X + (1-\alpha) Y) \leq \epsilon$ for all $\alpha \leq \delta$. Consequently, for all $\alpha \leq \delta$, $X \prec \alpha X + (1-\alpha) Y$ implying $q^* \geq \delta > 0$. That is, we can henceforth assume that $0 < q^* < 1$.

Step 4. We claim that $q^* < 1$ implies $X \sim q^* X + (1-q^*) Y$. We prove this claim by way of contradiction. First, suppose that $X \prec q^* X + (1-q^*) Y$. By Lemma 1 and openness of the set $S^*(X)$, there exists some $\delta > 0$ such that

$$d_0\left(q^{'} X + \left(1 - q^{'}\right) Y, q^* X + (1-q^*) Y\right) \leq \epsilon \tag{63}$$

for all $|q' - q^*| \leq \delta$. Let $q' = \min\left\{1, q^* + \frac{1}{2}\delta\right\}$ and observe that $q' > q^*$ as well as $X \prec q^{'} X + \left(1 - q^{'}\right) Y$. But this contradicts the definition of q^*.

Next, suppose that $X \succ q^* X + (1-q^*) Y^*$. An analogous argument as above results in some q' such that $q' < q^*$ as well as $q^{'} X + \left(1 - q^{'}\right) Y \prec X$. Again, a contradiction to the definition of q^*.

Step 5. In Step 4 we have proven that $X \sim q^* X + (1-q^*) Y$ whenever $q^* < 1$ By preference for diversification, we thus obtain

$$X \preceq \beta X + (1-\beta)(q^* X + (1-q^*) Y) \tag{64}$$

$$\Leftrightarrow$$

$$X \preceq (\beta + (1-\beta) q^*) X + (1-\beta)(1-q^*) Y \tag{65}$$

for all $\beta \in [0, 1]$. By definition of q^*,

$$\beta + (1-\beta) q^* \leq q^* \tag{66}$$

for all $\beta \in [0, 1]$, which only holds for $q^* = 1$. But this contradicts $q^* < 1$ and gives us the desired result (58).□□

Proof of Proposition 4. d_0-continuity is violated if, and only if, there exists some sequence of random variables $\{Y_k\}_{k \in \mathbb{N}}$ with $d_0(Y_k, Y) \to 0$ such that $X \preceq Y_k$ for all k but $Y \prec X$. By Assumption 1, we then have that $U(X) \leq U(Y_k)$ for all k and $U(Y) < U(X)$, which violates continuity of U. Consequently, continuity of U requires d_0-continuity. Moreover, by Assumption 1, quasi-concave (resp. quasi-convex) preferences require a quasi-concave (resp. quasi-convex) U. The proposition then follows from Theorem 2.□□

Proof of Proposition 5. The 'if'-part is easy since convergence in the d_0-metric implies convergence in distribution; that is, $d_0(Y_k, Y) \to 0$ implies $F_{Y_k} \Rightarrow F_Y$ (cf., e.g., Theorem 25.2 in Billingsley 1995).

The 'only if' part is less obvious as convergence in distribution on the same probability space does not necessarily imply convergence in the d_0-metric. Suppose that $F_{Y_k} \Rightarrow F_Y$. Then $F_{Y_k}^{-1}$ converges point-wise to F_Y^{-1} where, for any Z, F_Z^{-1} denotes the left inverse of CDF_Z. Let us fix a uniform random variable V on $(0,1)$ which exists because the probability space is nonatomic. By construction, the random variable $F_Z^{-1}(V)$ has the same distribution as the random variable Z, implying, by (41), $F_{Y_k}^{-1}(V) \sim Y_k$ and $F_Y^{-1}(V) \sim Y$. Since the $F_{Y_k}^{-1}(V)$ converge point-wise to $F_Y^{-1}(V)$, they also converge in probability (i.e., in d_0). By law-invariance of U, we thus have

$$\lim_k U(Y_k) = \lim_k U(F_{Y_k}^{-1}(V)) = U(F_Y^{-1}(V)) = U(Y). \tag{67}$$

□□

Proof of Proposition 6. Suppose that $f \in L^p(\mu)$ and $g \in L^q(\mu)$ with either $\frac{1}{p} + \frac{1}{q} = 1$ or $p = \infty$, $q = 1$. By Hölder's inequality, we have that

$$\int_\Omega |f \cdot g| \, d\mu \leq \|f\|_p \cdot \|g\|_q . \tag{68}$$

For any $X, Y \in L^p(\mu)$, let

$$f \equiv |X - Y| , \tag{69}$$

$$g \equiv \frac{1}{1 + |X - Y|} \tag{70}$$

so that (68) becomes

$$d_0(X, Y) \leq d_p(X, Y) \cdot \|g\|_q . \tag{71}$$

Since $\|g\|_q \leq 1$, convergence in d_p implies convergence in d_0 on $L^p(\mu)$.□□

References

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

Aliprantis, C.D., Burkinshaw, O. (1978) *Locally Solid Riesz Spaces.* New York: Academic Press, Inc.

Anscombe, F.J., Aumann, R.J. (1963) "A Definition of Subjective Probability" *Annals of American Statistics* **34**, 199-205.

Arrow K.J. (1974) "The Use of Unbounded Utility Functions in Expected-Utility Maximization: Response" *The Quarterly Journal of Economics* **88**, 136-138.

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. (2016) "Natural Risk Measures" *Mathematics and Financial Economics* **10**, 441–456.

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

Chateauneuf, A., Lakhnati, G. (2007) "From Sure to Strong Diversification" *Economic Theory* **32**, 511-522.

Chateauneuf, A., Tallon, J.-M. (2002) "Diversification, Convex Preferences and Nonempty Core in the Choquet Expected Utility Model" *Economic Theory* **19**, 509-523.

Cubit, R., Sugden, R. (2001) "On Money Pumps" *Games and Economic Behavior* **37**, 121-160.

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.

De Giorgi, E.G., Mahmoud, O. (2016) "Diversification Preferences in the Theory of Choice" *Decisions in Economics and Finance* **39**, 143-174.

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.

Delbaen, F., Drapeau, S. Kupper, M. (2011) "A von Neumann-Morgenstern Representation Result without Weak Continuity Assumption" *Journal of Mathematical Economics* **47**, 401–408.

Embrechts, P., McNeil, A., Straumann, D. (2002) "Correlation and Dependence in Risk Management: Properties and Pitfalls" In *RISK Management: Value at Risk and Beyond*, pp. 176–223. Cambridge University Press.

Epstein, L.G. (1999) "A Definition of Uncertainty Aversion" *The Review of Economic Studies* **66**, 579-608.

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

Geweke, J. (2001) "A Note on Some Limitations of CRRA Utility" *Economics Letters* **71**, 341–345.

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

Karni, E., Schmeidler, D. (2016) "An Expected Utility Theory for State-dependent Preferences" *Theory and Decision* **81**, 467–478.

Komjàth, P., Totik, V. (2006) *Problems and Theorems in Classical Set Theory.* Berlin: Springer.

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

Menger, K. (1934) "Das Unsicherheitsmoment in der Wertlehre. Betrachtungen im Anschluss an das sogenannte Petersburger Spiel" *Zeitschrift für Nationalökonomie* **5**, 459–485.

Nielsen, L., (1984) "Unbounded Expected Utility and Continuity" *Mathematical Social Sciences* **8(3)**, 201–216.

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

Ryan K.J. (1974) "The Use of Unbounded Utility Functions in Expected-Utility Maximization: Comment" *The Quarterly Journal of Economics* **88**, 133-135.

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

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.

Vendrik, M., Woltjer, G. (2007) "Happiness and Loss Aversion: Is Utility Concave or Convex in Relative Income?" *Journal of Public Economics* **91**, 1423–1448.

Wakker, 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.

Weitzman, M.L. (2007) "Subjective Expectations and Asset-Return Puzzles" *American Economic Review* **97**, 1102–1130.

Weitzman, M.L. (2009) "On Modeling and Interpreting the Economics of Catastrophic Climate Change" *The Review of Economics and Statistics* **91**, 1–19.