Distortion Riskmetrics on General Spaces

Qiuqi Wang* Ruodu Wang† Yunran Wei‡

May 2020§

Abstract

The class of distortion riskmetrics is defined through signed Choquet integrals, and it includes many classic risk measures, deviation measures, and other functionals in the literature of finance and actuarial science. We obtain characterization, finiteness, convexity, and continuity results on general model spaces, extending various results in the existing literature on distortion risk measures and signed Choquet integrals. This paper offers a comprehensive toolkit of theoretical results on distortion riskmetrics which are ready for use in applications.

Keywords: comonotonicity; Choquet integrals; convexity; convex order; continuity

JEL classification: C6, D8, G00

1 Introduction

In this paper we study *distortion riskmetrics* on general model spaces. A distortion riskmetric is a real-valued functional ρ with the following form

$$\rho(X) = \int_{-\infty}^{0} \left(h(\mathbb{P}(X \geq x)) - h(1)\right) \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X \geq x))\, \mathrm{d}x, \tag{1}$$

where h is a function of bounded variation on $[0,1]$ with $h(0) = 0$ and X is a random variable in the domain of ρ; a precise definition is given in Definition 1 below.

Let us first explain our somewhat unusual choice of terminology, "distortion riskmetrics". Clearly, the term "*distortion*" addresses the dominating role played by the (not necessarily monotone) distortion function h in (1), whereas the term "*riskmetrics*" is chosen to distinguish ρ from

*Department of Statistics and Actuarial Science, University of Waterloo (email: q428wang@uwaterloo.ca)

†Department of Statistics and Actuarial Science, University of Waterloo (email: wang@uwaterloo.ca)

‡Department of Statistics and Actuarial Science, Northern Illinois University (email: ywei1@niu.edu)

§A minor correction is made in May 2023.

the classic notions of risk measures and deviation measures. For instance, risk measures are often required to be monotone and translation-invariant in the sense of Artzner et al. (1999), and deviation measures are required to be convex in the sense of Rockafellar et al. (2006). Insurance risk measures and premium principles are typically assumed to be monotone with some other properties as in e.g., Gerber (1974) or Wang et al. (1997). Our notion of distortion riskmetrics does not require monotonicity, translation-invariance or convexity, and it unifies risk measures, deviation measures, and many other functionals in the literature of finance and insurance.

This paper is not the first to study functionals in (1) in risk management. Historically, such functionals, assuming monotonicity, were studied by Yaari (1987) in the economic literature and by Denneberg (1994) and Wang et al. (1997) in the actuarial literature. More recently, for non-monotone h, Wang et al. (2020) called the functional in (1) a *signed Choquet integral* on the space L^∞ of bounded random variables. To be precise, a signed Choquet integral refers to the right-hand side of (1). We note that a signed Choquet integral should be interpreted as an "integral", thus a mathematical operation, and not a functional. Mathematically, signed Choquet integrals can be formulated for any random variable, leading to a finite, infinite or undefined value in (1), whereas a distortion riskmetric is defined on a domain of financial relevance. The difference is negligible if the study is confined to L^∞, but it becomes delicate in the case of a larger space such as an L^p-space; see Section 2. Moreover, the term "distortion riskmetric" better describes the practical purpose of these functionals in risk management. For the above reasons, we decided to invent the term "distortion riskmetrics", which will hopefully be the standard term for the object in (1) in the future.

As hinted above, monotone (increasing) distortion riskmetrics have been studied for decades under different names: the L-functionals (Huber and Ronchetti (2009)) in statistics, Yaari's dual utilities (Yaari (1987)) in decision theory, distorted premium principles (Denneberg (1994), Wang et al. (1997) and Denuit et al. (2005)) in insurance, and distortion risk measures (Kusuoka (2001) and Acerbi (2002)) in finance. Some specific examples of distortion risk measures include the Value-at-Risk (VaR), the Expected Shortfall (ES, or TVaR/CVaR), the performance measures in Cherny and Madan (2009), the GlueVaR in Belles-Sampera et al. (2014), and the economic risk measures in Kou and Peng (2016). Non-monotone examples of signed Choquet integrals include the mean-median deviation, the Gini deviation, the inter-quantile range, some deviation measures of Rockafellar et al. (2006), and the Gini Shortfall of Furman et al. (2017). We collect some examples of one-dimensional distortion riskmetrics in Table 1.

name (notation)	formula for $X \in \mathcal{X}$ and parameters	distortion function for $t \in [0,1]$	domain $\mathcal{X}$	convex?	monotone?
mean ($\mathbb{E}$)	$\mathbb{E}[X]$	t	L^1	yes	yes
Value-at-Risk (VaR_α)	$F_X^{-1}(\alpha), \quad \alpha \in (0,1)$	$\mathbb{1}_{\{t>1-\alpha\}}$	L^0	no	yes
ES/TVaR/CVaR (ES_α)	$\frac{1}{1-\alpha}\int_\alpha^1 F_X^{-1}(t)\,\mathrm{d}t, \quad \alpha \in (0,1)$	$\frac{t}{1-\alpha} \wedge 1$	$L^{0,1}$	yes	yes
Gini deviation	$\frac{1}{2}\mathbb{E}[\lvert X^* - X^{**}\rvert]$	$t - t^2$	L^1	yes	no
mean-median deviation	$\min_{x\in\mathbb{R}} \mathbb{E}[\lvert X - x\rvert]$	$t \wedge (1-t)$	L^1	yes	no
essential supremum (ess sup)	$F_X^{-1}(1)$	$\mathbb{1}_{\{0<t\le 1\}}$	$L^{0,\infty}$	yes	yes
essential infimum (ess inf)	$F_X^{-1+}(0)$	$\mathbb{1}_{\{t=1\}}$	$L^{\infty,0}$	no	yes
range	$F_X^{-1}(1) - F_X^{-1+}(0)$	$\mathbb{1}_{\{0<t<1\}}$	L^∞	yes	no
inter-quantile range (IQR_α)	$F_X^{-1+}(\alpha) - F_X^{-1}(1-\alpha), \quad \alpha \in [1/2,1)$	$\mathbb{1}_{\{1-\alpha\le t\le\alpha\}}$	L^0	no	no
inter-ES range (IER_α)	$\mathrm{ES}_\alpha(X) + \mathrm{ES}_\alpha(-X), \quad \alpha \in (0,1)$	$\frac{t}{1-\alpha} \wedge 1 + \frac{\alpha-t}{1-\alpha} \wedge 0$	L^1	yes	no
Range Value-at-Risk ($\mathrm{RVaR}_{\alpha,\beta}$)	$\frac{1}{\beta-\alpha}\int_\alpha^\beta F_X^{-1}(t)\,\mathrm{d}t, \quad 0<\alpha<\beta<1$	$\frac{(t-1+\beta)_+}{\beta-\alpha} \wedge 1$	L^0	no	yes
Gini Shortfall ($\mathrm{GS}_\alpha^\lambda$)	$\mathrm{ES}_\alpha(X) + \lambda\mathbb{E}[\lvert X_\alpha^* - X_\alpha^{**}\rvert]$ $\alpha \in (0,1)$, $\lambda \ge 0$	$\frac{t}{1-\alpha} \wedge 1 + \frac{2\lambda t(1-t-\alpha)_+}{(1-\alpha)^2}$	$L^{0,1}$	$\lambda \le 1/2$	$\lambda \le 1/2$
proportional hazard principle/MAXVAR	$\frac{1}{\alpha}\int_0^1 (1-t)^{(1-\alpha)/\alpha} F_X^{-1}(t)\,\mathrm{d}t, \quad \alpha \ge 1$	$t^{1/\alpha}$	$\cup_{p>\alpha} L^{1,p} \subset \mathcal{X}$	yes	yes
dual power principle/MINVAR	$\alpha\int_0^1 t^{\alpha-1} F_X^{-1}(t)\,\mathrm{d}t, \quad \alpha \ge 1$	$1-(1-t)^\alpha$	$\cup_{q>1/\alpha} L^{q,1} \subset \mathcal{X}$	yes	yes
GlueVaR ($\mathrm{GlueVaR}_{\beta,\alpha}^{\omega_1,\omega_2}$)	$\omega_1\mathrm{ES}_\alpha(X) + \omega_2\mathrm{ES}_\beta(X) + \omega_3\mathrm{VaR}_\alpha(X)$ $0<\alpha\le\beta<1$, $(\omega_1,\omega_2,\omega_3) \in \Delta_3$	$\omega_1(\frac{t}{1-\alpha} \wedge 1) + \omega_2(\frac{t}{1-\beta} \wedge 1) + \omega_3\mathbb{1}_{\{t>1-\alpha\}}$	$L^{0,1}$	no	yes

Table 1: Some examples of one-dimensional distortion riskmetrics

Notation. $F_X^{-1}(\alpha) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \le x) \ge \alpha\}$ for $\alpha \in (0,1]$ and $F_X^{-1+}(\alpha) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \le x) > \alpha\}$ for $\alpha \in [0,1)$. $L^{p,q} = \{X \in L^0 : X_- \in L^p,\ X_+ \in L^q\}$ for $p,q \ge 0$. $\Delta_n = \{(x_1,\dots,x_n) \in (0,1)^n : x_1 + \cdots + x_n = 1\}$ is the interior of the standard n-simplex. X^*, X^{**} are iid copies of X and $X_\alpha^*, X_\alpha^{**}$ are iid copies of $F_X^{-1}(U_\alpha)$ where $U_\alpha \sim \mathrm{U}[\alpha,1]$.

Moreover, distortion riskmetrics serve as the building block of law-invariant convex risk functionals in the sense that any law-invariant convex risk functional can be written as a supremum of signed Choquet integrals plus constants (Liu et al. (2020)) and this includes all law-invariant convex risk measures in Föllmer and Schied (2016) and all law-invariant deviation measures in Grechuk et al. (2009), as well as the classic mean-variance and mean-standard deviation principles in insurance.

We already mentioned that characterization and various properties of distortion riskmetrics are studied on L^∞ by Wang et al. (2020). As a follow-up of the previous work, the main purpose of this paper is to extend the domain of distortion riskmetrics to more general spaces, including L^p-spaces for $p \in [1, \infty)$. In many applications, risk measures such as the industry standard VaR and ES are defined on spaces beyond L^∞ to include unbounded loss distributions, e.g., normal, Pareto or t-distributions. Furthermore, for many convex risk measures, their natural domains on which key properties are preserved are Banach spaces much larger than L^∞; see e.g., Filipović and Svindland (2012), Pichler (2013) and Liebrich and Svindland (2017). Indeed, there is an extensive literature on risk measures defined on general spaces (e.g., Delbaen (2002), Föllmer and Schied (2002) and Ruszczyński and Shapiro (2006)) and in particular on L^p-spaces (Frittelli and Rosazza Gianin (2002)) or Orlicz spaces (Cheridito and Li (2009)). Different from the previous literature, we consider many functionals that are not necessarily monotone or convex. Notably, as a special example, the inter-quantile range (see Table 1) is not monotone, convex, or L^p-continuous, but it is a popular measure of dispersion in statistics, and it belongs to the class of distortion riskmetrics. Finally, we extend distortion riskmetrics to a multi-dimensional setting, where the concepts of elicitability and convex level sets has been popular recently; see Fissler and Ziegel (2016), Frongillo and Kash (2018) and Wang and Wei (2020).

Most results in this paper are similar to those in the literature in terms of both statements and proofs, and our findings that these results hold on general spaces are not surprising. However, most of the results in previous literature on L^∞, especially those in Wang et al. (2020), may not be convenient to directly use in practice where most applications require results on more general spaces of random variables. As such, more general results are in need, and this paper can be viewed as a convenient toolkit for future studies and applications of distortion riskmetrics. Nevertheless, there are several additions to the existing literature. The similarity of this paper with Wang et al. (2020) and the new results of this paper are summarized in Table 2.

Below we briefly explain the new results. First, an ES-based representation of convex distortion riskmetric ρ in Theorem 5 is new to the literature. Four other new results, all requiring the

corresponding results			new results
this paper (on general spaces)		Wang et al. (2020) (on L^∞)	this paper (on general spaces)
Theorem 1	$\longleftrightarrow$	Theorem 1	Proposition 1
Theorem 2	$\longleftrightarrow$	Theorem 2	Proposition 3
Proposition 2	$\longleftrightarrow$	Lemmas 2 and 3	Theorem 4
Theorem 3	$\longleftrightarrow$	Theorem 3	Theorem 5
Theorem 6	$\longleftrightarrow$	Theorem 4	Proposition 4

Table 2: Comparison with results in Wang et al. (2020)

considered domain to be larger than L^∞, are the finiteness condition in Proposition 1, the domain of convex distortion riskmetrics in Proposition 3, the existence of dominating convex functionals in Theorem 4, and the L^p-continuity in Proposition 4. Moreover, the condition in Theorem 6 is slightly weakened compared to a similar result on L^∞ in Wang et al. (2020).

The paper is organized as follows. In Section 2, we collect basic definitions needed for our paper, and present a functional characterization of distortion riskmetrics. In Section 3, results related to convexity, convex order consistency, and mixture concavity are presented. Section 4 contains results on continuity properties of distortion riskmetrics and and Section 5 extends the discussions to the multi-dimensional setting. To facilitate the main purpose of the paper as a toolkit, most proofs are self-contained and are relegated to the appendix.

2 Distortion riskmetrics and their characterization

2.1 Notation and definition

Throughout the paper, let $(\Omega, \mathcal{A}, \mathbb{P})$ be an atomless probability space. Two random variables X and Y have the same distribution under $\mathbb{P}$ is denoted by $X \overset{\mathrm{d}}{=} Y$. For $x, y \in \mathbb{R}$, we write $x \vee y = \max\{x, y\}$, $x \wedge y = \min\{x, y\}$, $x_+ = x \vee 0$ and $x_- = (-x) \vee 0$. For $p \in [1, \infty)$, L^p is the space of random variables with finite p-th moment, and L^∞ is that of essentially bounded random variables. Throughout, the set $\mathcal{X} \supset L^\infty$ is a law-invariant convex cone, that is, for all random variables X and Y,

(i) if $X \in \mathcal{X}$ and $X \overset{\mathrm{d}}{=} Y$, then $Y \in \mathcal{X}$;

(ii) if $X \in \mathcal{X}$, then $\lambda X \in \mathcal{X}$ for all $\lambda > 0$;

(iii) if $X, Y \in \mathcal{X}$, then $X + Y \in \mathcal{X}$.

Let $\mathcal{M}$ be the set of distribution functions of random variables in $\mathcal{X}$. For $F \in \mathcal{M}$, $X \sim F$ means that $X \in \mathcal{X}$ has distribution F. Denote by F_X the distribution function of the random variable X. We define the left-continuous generalized inverse of F (left-quantile) as

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad t \in (0, 1],$$

while the right-continuous generalized inverse of F (right-quantile) is defined as

$$F^{-1+}(t) = \inf\{x \in \mathbb{R} : F(x) > t\}, \quad t \in [0, 1).$$

For simplicity, we also let $F^{-1}(0) = F^{-1+}(0)$ and $F^{-1+}(1) = F^{-1}(1)$.

Next, we define the *distortion riskmetric* using the signed Choquet integral (Choquet (1954)) on a general space. Denote by

$$\mathcal{H} = \{h : h \text{ maps } [0, 1] \text{ to } \mathbb{R},\ h \text{ is of bounded variation, } h(0) = 0\}.$$

Definition 1. A functional $\rho_h : \mathcal{X} \to \mathbb{R}$, whose domain $\mathcal{X} \supset L^\infty$ is a law-invariant convex cone, is a *distortion riskmetric* if there exists $h \in \mathcal{H}$ such that $\rho_h(X) = \int X \,\mathrm{d}h \circ \mathbb{P}$, where $\int X \,\mathrm{d}h \circ \mathbb{P}$ is a signed Choquet integral defined by

$$\int X \,\mathrm{d}h \circ \mathbb{P} = \int_{-\infty}^{0} (h(\mathbb{P}(X \geq x)) - h(1)) \,\mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X \geq x)) \,\mathrm{d}x. \tag{2}$$

The function h is called the *distortion function* of ρ_h.

Generally, the two integrals in (2) may not be finite, and hence $\int X \,\mathrm{d}h \circ \mathbb{P}$ may be infinite or even not well-defined (i.e., $\infty - \infty$). We emphasize that according to our definition, a distortion riskmetric $\rho_h : \mathcal{X} \to \mathbb{R}$ is only defined when $\int X \,\mathrm{d}h \circ \mathbb{P}$ is finite (i.e., both integrals are finite), and hence the two terms "distortion riskmetrics" and "signed Choquet integrals" are no longer interchangeable, in contrast to the case of L^∞ studied by Wang et al. (2020). In other words, $\mathcal{X}$ and h have to be compatible, making (2) finite. In Section 2.2 below we will give a sufficient condition for (2) to be finite. The notion of a distortion function h we use in this paper is broader than the classical sense in which h is assumed increasing with $h(1) = 1$.

For a given distortion riskmetric $\rho_h : \mathcal{X} \to \mathbb{R}$, the distortion function $h \in \mathcal{H}$ is unique. To see this, suppose that $\rho_{h_1}(X) = \rho_{h_2}(X)$ for all $X \in \mathcal{X}$. Choose a random variable $X \sim \text{Bernoulli}(p)$ with a fixed $p \in [0, 1]$. It follows that

$$\rho_{h_i}(X) = h_i(p) + \int_1^\infty h_i(0) \,\mathrm{d}x = h_i(p), \quad i = 1, 2.$$

Since p is arbitrary, we get $h_1 = h_2$ on $[0, 1]$.

Remark 1. A distortion riskmetric ρ_h can be equivalently expressed as

$$\rho_h(X) = \int_{-\infty}^{0} (h(\mathbb{P}(X > x)) - h(1))\,\mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X > x))\,\mathrm{d}x. \tag{3}$$

Indeed, since $\mathbb{P}(X > x) = \mathbb{P}(X \geq x)$ almost everywhere on $\mathbb{R}$, we know $h(\mathbb{P}(X > x)) = h(\mathbb{P}(X \geq x))$ almost everywhere on $\mathbb{R}$.

2.2 Quantile representation and finiteness of signed Choquet integrals

The quantile representation of signed Choquet integrals is obtained in Lemma 3 of Wang et al. (2020) on L^∞ and Theorems 4 and 6 of Dhaene et al. (2012) for increasing h. Combining the above results, we have the following quantile representation of signed Choquet integrals on a general space with distortion functions not necessarily increasing.

Lemma 1. *For $h \in \mathcal{H}$ and $X \in L^0$ such that $\int X\,\mathrm{d}h \circ \mathbb{P}$ is well-defined (it may take values $\pm\infty$),*

(i) if h is right-continuous, then $\int X\,\mathrm{d}h \circ \mathbb{P} = \int_0^1 F_X^{-1+}(1-t)\,\mathrm{d}h(t)$;

(ii) if h is left-continuous, then $\int X\,\mathrm{d}h \circ \mathbb{P} = \int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h(t)$;

(iii) if F_X^{-1} is continuous on $(0,1)$, then $\int X\,\mathrm{d}h \circ \mathbb{P} = \int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h(t) = \int_0^1 F_X^{-1+}(1-t)\,\mathrm{d}h(t)$.

Now we focus on L^p-spaces for $p \in [1,\infty]$. Define a set of distortion functions $\mathcal{H}_1$ as

$$\mathcal{H}_1 = \{h \in \mathcal{H} : h \text{ is absolutely continuous on } [0,\epsilon) \cup (1-\epsilon, 1] \text{ for some } \epsilon \in (0,1)\}.$$

Note that $\mathcal{H}_1$ excludes only special examples such as the essential supremum, the essential infimum, and the range in Table 1. Moreover, noticing that h is differentiable almost everywhere on $[0,1]$ due to bounded variation, we let

$$\mathcal{H}_q = \left\{h \in \mathcal{H}_1 : h' \in L^q((0,\epsilon) \cup (1-\epsilon, 1)) \text{ for some } \epsilon \in (0,1)\right\},$$

where h' is (in a.e. sense) the derivative of h and q is the conjugate of $p \in [1,\infty]$ (i.e., $1/p + 1/q = 1$). Next, we give a sufficient condition for ρ_h to be well defined, which is almost necessary in case that h is concave.

Proposition 1. *For $p \in [1,\infty)$, q being its conjugate,*

(i) $\int X\,\mathrm{d}h \circ \mathbb{P}$ is finite for all $X \in L^p$ if $h \in \mathcal{H}_q$;

(ii) if $h \in \mathcal{H}$ is concave and $\int X\,\mathrm{d}h \circ \mathbb{P}$ is finite for all $X \in L^p$, then $h \in \mathcal{H}_r$ for all $r < q$.

As a consequence of Proposition 1, if $h \in \mathcal{H}$ is absolutely continuous and $\int_0^1 |h'(t)|^q\,\mathrm{d}t < \infty$, then $\int X\,\mathrm{d}h \circ \mathbb{P}$ is finite for all $X \in L^p$. In particular, the case $p = q = 2$ gives a sufficient condition for the finiteness of $\int X\,\mathrm{d}h \circ \mathbb{P}$ for $X \in L^2$.

2.3 Characterization and basic properties

Before we further characterize distortion riskmetrics, we list some terminology and properties for random variables and functionals. Recall that random variables X and Y are *comonotonic* if there exists $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for each $\omega, \omega' \in \Omega_0$,

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0.$$

A functional $\rho : \mathcal{X} \to \mathbb{R}$ may satisfy the following properties, where the statements hold for all random variables $X, Y \in \mathcal{X}$.

(a) *Law-invariance:* $\rho(X) = \rho(Y)$ for $X \stackrel{\mathrm{d}}{=} Y$.

(b) *Comonotonic-additivity*: $\rho(X + Y) = \rho(X) + \rho(Y)$ if X and Y are comonotonic.

(c) *Continuity at infinity*: $\lim_{M\to\infty} \rho((X \wedge M) \vee (-M)) = \rho(X)$.

(d) *Uniform sup-continuity:* For any $\epsilon > 0$, there exists $\delta > 0$, such that $|\rho(X)-\rho(Y)| < \epsilon$ whenever $\operatorname{ess\,sup} |X - Y| < \delta$, where $\operatorname{ess\,sup}(\cdot)$ is the essential supremum in Table 1.

The above four properties are satisfied by distortion riskmetrics, and moreover, they indeed characterize distortion riskmetrics, similarly to the case of bounded random variables studied by Wang et al. (2020) and the case of increasing Choquet integrals in Wang et al. (1997) and Kou and Peng (2016), all based on a classic result of Schmeidler (1986).

Theorem 1. *A functional $\rho : \mathcal{X} \to \mathbb{R}$ is law-invariant, comonotonic-additive, continuous at infinity and uniformly sup-continuous if and only if ρ is a distortion riskmetric.*

Remark 2. From the proof of necessity part of Theorem 1 in Appendix A, we can see a distortion riskmetric $\rho_h : \mathcal{X} \to \mathbb{R}$ is, in fact, Lipschitz-continuous with respect to L^∞-norm with Lipschitz constant TV_h, the total variation of h on $[0, 1]$. This continuity is stronger than uniform sup-continuity. This point will be further developed in Section 4.

Below we present some basic properties of distortion riskmetrics which are useful in later sections. They are well-established for random variables in L^∞ and $h \in \mathcal{H}$. In what follows, a functional ρ is said to be *increasing* (or *decreasing*) if $X \leq Y$ almost surely implies $\rho(X) \leq \rho(Y)$ (or $\rho(X) \geq \rho(Y)$, respectively).[1]

[1] The terms "increasing" and "decreasing" in this paper are always in the non-strict sense.

Proposition 2. *For $h, h_1, h_2 \in \mathcal{H}$,*

(i) if $h_1(1) = h_2(1)$, then $h_1 \leq h_2$ on $[0, 1] \Leftrightarrow \rho_{h_1} \leq \rho_{h_2}$ on $\mathcal{X}$. In particular, $h_1 = h_2$ on $[0, 1] \Leftrightarrow \rho_{h_1} = \rho_{h_2}$ on $\mathcal{X}$;

(ii) ρ_h is increasing (resp. decreasing) if and only if h is increasing (resp. decreasing);

(iii) for all $c \in \mathbb{R}$ and $X \in \mathcal{X}$, $\rho_h(X + c) = \rho_h(X) + ch(1)$;

(iv) for all $\lambda > 0$ and $X \in \mathcal{X}$, $\rho_h(\lambda X) = \lambda \rho_h(X)$;

(v) for all $X \in \mathcal{X}$, $\rho_h(-X) = \rho_{\hat{h}}(X)$, where $\hat{h} : [0, 1] \to \mathbb{R}$ is given by $\hat{h}(x) = h(1 - x) - h(1)$ for all $x \in [0, 1]$.

3 Convexity, convex order consistency and mixture concavity

In this section, we study the important class of convex distortion riskmetrics and their related properties. A functional $\rho : \mathcal{X} \to \mathbb{R}$ is *convex* if $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$ for all $X, Y \in \mathcal{X}$ and $\lambda \in [0, 1]$. As shown in Theorem 3 of Wang et al. (2020), the following properties: convexity, convex order consistency, and mixture concavity, on L^∞, are equivalent to concavity of the distortion function. We establish a similar result on general spaces, as well as a few new results on convex distortion riskmetrics.

We first justify that for a convex distortion riskmetric, if its domain $\mathcal{X}$ is a linear space, then it is contained in L^1; hence, it makes sense to confine our study to subsets of L^1. Note also that L^1 is the canonical space for law-invariant convex risk measures (e.g., Filipović and Svindland (2012)).

Proposition 3. *Suppose that $\mathcal{X}$ is a linear space and $\rho_h : \mathcal{X} \to \mathbb{R}$ is a convex distortion riskmetric. Then $\mathcal{X} \subset L^1$ unless $\rho_h = 0$ on $\mathcal{X}$.*

The assumption that $\mathcal{X}$ is a linear space in Proposition 3 is not dispensable. An important example is the Expected Shortfall (ES) in Table 1 at level $\alpha \in (0, 1)$, defined as

$$\mathrm{ES}_\alpha(X) = \frac{1}{1 - \alpha} \int_\alpha^1 F_X^{-1}(t) \, \mathrm{d}t, \quad X \in \mathcal{X}, \tag{4}$$

where its domain $\mathcal{X}$ can be chosen as $\{X \in L^0 : X_+ \in L^1\}$, which is larger than L^1. In addition, we let $\mathrm{ES}_0 = \mathbb{E}$ which is finite on L^1 and ES_1 be the essential supremum which is finite on the set of random variables bounded from above. For $\alpha \in [0, 1]$, ES_α is a convex distortion riskmetric with distortion function h given by

$$h(t) = \frac{t}{1 - \alpha} \wedge 1, \quad t \in [0, 1], \ \alpha \in [0, 1)$$

and $h(t) = \mathbb{1}_{\{t>0\}}$ if $\alpha = 1$. These facts will be useful later.

Next, we fix some terminology. A random variable X is said to be smaller than a random variable Y in *convex order*, denoted by $X \le_{\mathrm{cx}} Y$, if $\mathbb{E}[\phi(X)] \le \mathbb{E}[\phi(Y)]$ for all convex $\phi : \mathbb{R} \to \mathbb{R}$, provided that both expectations exist. For a functional $\rho : \mathcal{X} \to \mathbb{R}$ and all random variables $X, Y \in \mathcal{X}$, ρ is *quasi-convex* if $\rho(\lambda X + (1-\lambda)Y) \le \rho(X) \vee \rho(Y)$ for all $\lambda \in [0,1]$; ρ is *convex order consistent* if $\rho(X) \le \rho(Y)$ for $X \le_{\mathrm{cx}} Y$. For a law-invariant functional ρ, define $\tilde{\rho} : \mathcal{M} \to \mathbb{R}$ such that $\tilde{\rho}(F) = \rho(X)$ where $X \sim F$, and ρ is *concave on mixtures* if $\tilde{\rho}$ is concave. The following result characterizes convex order using distortion riskmetrics. For a version of this result for increasing h, see Theorem 5.2.1 of Dhaene et al. (2006).

Theorem 2. *For all random variables $X, Y \in L^1$, $X \le_{\mathrm{cx}} Y$ if and only if $\rho_h(X) \le \rho_h(Y)$ for all concave functions $h \in \mathcal{H}$ such that X and Y are in the domain of ρ_h.*

In the following theorem, we present six equivalent conditions about convexity of a distortion riskmetric on a general space, similar to Theorem 3 of Wang et al. (2020). Recall that $\mathcal{X}$ is a law-invariant convex cone containing L^∞. In the following result, we further assume $\mathcal{X} \subset L^1$ as discussed above.

Theorem 3. *For a distortion riskmetric $\rho_h : \mathcal{X} \to \mathbb{R}$ where $\mathcal{X} \subset L^1$, the following are equivalent: (i) h is concave; (ii) ρ_h is convex order consistent; (iii) ρ_h is subadditive; (iv) ρ_h is convex; (v) ρ_h is quasi-convex; (vi) ρ_h is concave on mixtures.*

A few well-known characterization results in risk management can be directly obtained from Theorem 1 and 3. For a history of these results, see Föllmer and Schied (2016). Following the terminology of Föllmer and Schied (2016), we say a functional $\rho : \mathcal{X} \to \mathbb{R}$ is *cash-invariant* if $\rho(X + c) = \rho(X) + c$ for all $X \in \mathcal{X}$ and $c \in \mathbb{R}$. A *coherent risk measure* is a functional that is increasing, cash-invariant, positively homogeneous, and convex.

Corollary 1. *Suppose that $\mathcal{X} \subset L^1$. A functional $\rho : \mathcal{X} \to \mathbb{R}$ is law-invariant, increasing, cash-invariant, continuous at infinity, and comonotonic-additive if and only if ρ is a distortion riskmetric ρ_h for an increasing h with $h(1) = 1$. In addition, ρ satisfies any of the properties (ii)-(vi) in Theorem 3 if and only if h is concave, and in that case ρ is a coherent risk measure.*

Note that in Corollary 1 we do not assume uniform sup-continuity as it is implied by monotonicity and cash-invariance. In case $\mathcal{X} = L^\infty$, continuity at infinity can also be removed from the statement. In Corollary 1, $\rho = \rho_h$ is a distortion risk measure or a dual utility (Yaari (1987)). If h

is concave, then $\rho = \rho_h$ is commonly known as a spectral risk measure; see Acerbi (2002) where h is additionally assumed to be continuous at 0.

In the next result, we consider the relationship between a distortion riskmetric ρ_h and a convex one dominating ρ_h. For this purpose, we introduce the *concave envelope* $h^* : [0,1] \to \mathbb{R}$ of $h \in \mathcal{H}$, defined as

$$h^*(t) = \inf \{g(t) : \ g \in \mathcal{H}, \ g \geq h, \ g \text{ is concave on } [0,1]\}.$$

One can check that h^* is concave, $h^*(0) = 0$ and $h^*(1) = h(1)$; see Wang et al. (2020) for a simple justification. Theorem 3 yields that $\rho_{h^*} : \mathcal{X} \to \mathbb{R}$ is a convex distortion riskmetric if $\mathcal{X} \subset L^1$. We also know that $\rho_{h^*} \geq \rho_h$ on their common domain due to Proposition 2. The next theorem shows that ρ_{h^*} is actually the smallest law-invariant, convex and continuous-at-infinity functional dominating ρ_h; note that it is not obvious whether such a functional exists and whether it is a distortion riskmetric. Below, we say that ρ_{h^*} is finite on $\mathcal{X}$, if the signed Choquet integral $\int X \, \mathrm{d}h^* \circ \mathbb{P}$ is finite for all $X \in \mathcal{X}$.

Theorem 4. *For a distortion riskmetric $\rho_h : \mathcal{X} \to \mathbb{R}$ where $\mathcal{X} \subset L^1$, if ρ_{h^*} is finite on $\mathcal{X}$, then ρ_{h^*} is the smallest law-invariant, convex and continuous-at-infinity functional dominating ρ_h. If ρ_{h^*} is not finite on $\mathcal{X}$, then there is no real-valued law-invariant, convex and continuous-at-infinity functional dominating ρ_h.*

Theorem 4 implies in particular that ES_α in (4) is the smallest law-invariant and continuous-at-infinity convex functional dominating VaR_α (Table 1); see Theorem 9 of Kusuoka (2001) and Theorem 4.67 of Föllmer and Schied (2016) for this statement on the set of bounded random variables.

In the next result, we establish a new ES-based representation of convex distortion riskmetrics, which covers the classic ES-based representation of coherent distortion risk measures in Theorem 4.93 of Föllmer and Schied (2016) on L^∞. As far as we are aware of, the representation (5) is new to the literature.

Theorem 5. *A functional $\rho : \mathcal{X} \to \mathbb{R}$ where $\mathcal{X} \subset L^1$ is a convex distortion riskmetric if and only if there exist finite Borel measures μ, ν on $[0,1]$ such that*

$$\rho(X) = \int_0^1 \mathrm{ES}_\alpha(X) \, \mathrm{d}\mu(\alpha) + \int_0^1 \mathrm{ES}_\alpha(-X) \, \mathrm{d}\nu(\alpha). \tag{5}$$

Moreover, if ρ is increasing, then we can take $\nu = 0$.

Remark 3. In case ν in (5) satisfies $\beta := \int_0^1 \frac{1}{1-\alpha}\,\mathrm{d}\nu(\alpha) < \infty$, using the equality

$$\mathrm{ES}_\alpha(-X) = \frac{1}{1-\alpha}(\alpha \mathrm{ES}_{1-\alpha}(X) - \mathbb{E}[X]), \quad X \in L^1,$$

we can rewrite (5) as

$$\rho(X) = \int_0^1 \mathrm{ES}_\alpha(X)\,\mathrm{d}\hat{\mu}(\alpha) - \beta\mathbb{E}[X], \qquad X \in \mathcal{X}, \tag{6}$$

where $\hat{\mu}$ is another finite Borel measure on $[0,1]$. Note that the condition $\beta < \infty$ is not automatically satisfied for a general convex distortion riskmetric ρ. An example of a convex distortion riskmetric that does not admit the form in (6) is $\rho : L^\infty \to \mathbb{R},\ X \mapsto -F_X^{-1}(0)$. Note that ρ admits the form in (5) with $\mu = 0$ and $\nu = \delta_1$, where δ_1 is the point-mass at 1; of course, $\beta = \infty$ in this case.

Finally, we mention the related concept of the *convex level sets* (CxLS) property. A functional ρ has CxLS if the level set $\{F \in \mathcal{M} : \tilde{\rho}(F) = x\}$ of $\tilde{\rho}$ is convex for each $x \in \mathbb{R}$. The CxLS property is a necessary condition for the notions of elicitability, identifiability and backtestability; see Wang and Wei (2020, Section 6) for an explanation. The above three concepts, referring to the quality and validity of risk forecasts, are notably popular in current banking regulation and model risk management. We refer to Gneiting (2011), Fissler and Ziegel (2016) and Acerbi and Szekely (2017) for more discussions on these concepts. Theorem 1 of Wang and Wei (2020) characterizes a signed Choquet integral with CxLS on a convex set $\mathcal{M}$ that contains all three-point distributions, which naturally applies to our distortion riskmetrics on general spaces. In short, up to a constant multiplier, distortion riskmetrics with CxLS only have three forms: the mean, a mixture of left and right α-quantiles, $\alpha \in (0,1)$, and a mixture of the essential supremum and the essential infimum.

4 Continuity of distortion riskmetrics

In this section, we examine continuity of distortion riskmetrics. It is already shown in Remark 2 that a distortion riskmetric is Lipschitz-continuous with respect to L^∞-norm. Namely, for $h \in \mathcal{H}$ and $X, Y \in \mathcal{X}$,

$$|\rho_h(X) - \rho_h(Y)| \leq \operatorname{ess\,sup}|X - Y| \cdot \mathrm{TV}_h,$$

where TV_h is the total variation of h on $[0,1]$.

We are then interested in continuity of a distortion riskmetric with respect to convergence in distribution, or equivalently, weak convergence in the set of distributions $\mathcal{M}$. This is closely related to robustness of a risk functional in risk management; see Krätschmer et al. (2014). Before stating the result of such continuity, we write the following relevant definition of h-uniform integrability. Given a convex cone $\mathcal{X}$ and $h \in \mathcal{H}$, a set $\mathcal{D} \subset \mathcal{X}$ is called *h-uniformly integrable* if

$$\lim_{k \downarrow 0} \sup_{X \in \mathcal{D}} \int_0^k |F_X^{-1}(1-t)| \, \mathrm{d}h(t) = 0$$

and

$$\lim_{k \uparrow 1} \sup_{X \in \mathcal{D}} \int_k^1 |F_X^{-1}(1-t)| \, \mathrm{d}h(t) = 0.$$

Note that h-uniform integrability reduces to the usual uniform integrability when $h \in \mathcal{H}$ is linear and nonconstant in some neighborhoods of 0 and 1. We give the following result for continuity of distortion riskmetrics with respect to convergence in distribution.

Theorem 6. *For $h \in \mathcal{H}$ and $X, X_1, X_2, \dots \in \mathcal{X}$, suppose that $X_n \to X$ in distribution as $n \to \infty$ and the set $\{X, X_1, X_2, \dots\}$ is h-uniformly integrable. If for all $t \in (0,1)$, either $s \mapsto h(s)$ or $s \mapsto F_X^{-1}(1-s)$ is continuous at t, then $\rho_h(X_n) \to \rho_h(X)$ as $n \to \infty$.*

Next, we consider the L^p-continuity of distortion riskmetrics (i.e., continuity with respect to the L^p-norm, defined as $||X||_p = (\mathbb{E}[|X|^p])^{1/p}$, $X \in L^p$). We give a sufficient condition for a distortion riskmetric to be L^p-continuous without assuming convexity of the functional, as is typically done in the literature.

Proposition 4. *For $p \in [1, \infty)$ and continuous $h \in \mathcal{H}$, a distortion riskmetric $\rho_h : L^p \to \mathbb{R}$ is L^p-continuous if $h \in \mathcal{H}_q$ where q is the conjugate of p.*

We remark that all convex distortion riskmetrics (i.e., the ones with concave h by Theorem 3) on L^p are L^p-continuous; see Rüschendorf (2013, Corollary 7.10) for the L^p-continuity of the finite-valued convex risk measures on L^p.

5 Multi-dimensional distortion riskmetrics

In this section, we discuss distortion riskmetrics in a multi-dimensional setting. The importance of multi-dimensional riskmetrics arises in a statistical context, where multi-dimensional forecasting and elicitation of statistical quantities (jointly) has become a popular topic; see Lambert et al. (2008), Fissler and Ziegel (2016) and Frongillo and Kash (2018). Here, multi-dimensionality refers to the range, rather than the domain, of the riskmetrics; in other words, our riskmetrics map $\mathcal{X}$ to $\mathbb{R}^d$ for some $d \geq 2$. This formulation is motivated by the statistical applications mentioned above, and in particular, estimating, forecasting, and testing multiple quantities depending on a random object.

In this section, we simply extend the results in Section 2 to multi-dimensional distortion riskmetrics. There is essentially nothing new; nevertheless, in view of the importance of multi-dimensional riskmetrics and their applications, we collect some basic results. The distortion riskmetrics of dimension $d \geq 2$ are defined as follows.

Definition 2. A d-dimensional distortion riskmetric $\rho_{\mathbf{h}} : \mathcal{X} \to \mathbb{R}^d$ is defined as

$$\rho_{\mathbf{h}}(X) = (\rho_{h_1}(X), \dots, \rho_{h_d}(X)),$$

where $\mathbf{h} = (h_1, \dots, h_d) \in \mathcal{H}^d$. Obviously, each ρ_{h_i} for $i = 1, \dots, d$ is a one-dimensional distortion riskmetric on $\mathcal{X}$.

Properties (a)-(d) in Section 2.3 can be equivalently formulated for d-dimensional distortion riskmetrics. More precisely, $\rho_{\mathbf{h}} : \mathcal{X} \to \mathbb{R}^d$ with $\mathbf{h} = (h_1, \dots, h_d)$ satisfies some of the properties (a)-(d) in Section 2.3 if and only if each one-dimensional distortion riskmetric ρ_{h_i}, $i = 1, \dots, d$, satisfies the respective properties. We can now provide the characterization result for multi-dimensional distortion riskmetrics. The same representation on L^∞ is given by Proposition 5 of Wang and Wei (2020).

Proposition 5. *A functional $\rho : \mathcal{X} \to \mathbb{R}^d$ is law-invariant, comonotonic-additive, continuous at infinity and uniformly sup-continuous if and only if ρ is a d-dimensional distortion riskmetric.*

Similarly to Theorem 6, the continuity of multi-dimensional distortion riskmetrics with respect to weak convergence is summarized below.

Proposition 6. *Let $\mathbf{h} = (h_1, \dots, h_d)$ with $h_i \in \mathcal{H}$, $i = 1, \dots, d$. For $X, X_1, X_2, \dots \in \mathcal{X}$, suppose that $X_n \to X$ in distribution as $n \to \infty$ and the set $\{X, X_1, X_2, \dots\}$ is h_i-uniformly integrable for all $i = 1, \dots, d$. If for any given $i = 1, \dots, d$ and for all $t \in (0, 1)$, either $s \mapsto h_i(s)$ or $s \mapsto F_X^{-1}(1 - s)$ is continuous at t, then $\rho_{\mathbf{h}}(X_n) \to \rho_{\mathbf{h}}(X)$ as $n \to \infty$.*

Convexity and concavity cannot be naturally formulated for multi-dimensional functionals due to the lack of complete order in $\mathbb{R}^d$. On the other hand, the CxLS property can be naturally defined for multi-dimensional functionals. Similarly to Section 3, a multi-dimensional functional ρ has CxLS if the level set $\{F \in \mathcal{M} : \tilde{\rho}(F) = x\}$ is convex for each $x \in \mathbb{R}^d$. As in the case of dimension one, multi-dimensional CxLS serves as a necessary condition for multi-dimensional elicitability, and hence it is important in the recent study of statistical elicitation.

Unlike the other properties in this section, which do not need new mathematical treatment for multi-dimensional distortion riskmetrics, the multi-dimensional CxLS is highly non-trivial to study or characterize. For instance, one-dimensional distortion riskmetrics with CxLS are characterized by Theorem 1 of Wang and Wei (2020), whereas a full characterization of multi-dimensional distortion riskmetrics with CxLS is a well-known difficult open question; see Fissler and Ziegel (2016) and Kou and Peng (2016). As far as we are aware of, the only existing characterization result on multi-dimensional distortion riskmetrics is given in Theorem 2 of Wang and Wei (2020), which identifies the form of ρ_h such that $(\rho_h, \mathrm{VaR}_\alpha)$ has CxLS; note that $(\rho_h, \mathrm{VaR}_\alpha)$ is a two-dimensional distortion riskmetric.

Remark 4. Another direction of multi-dimensional generalization of riskmetrics is to consider mappings from $\mathcal{X}^d$ to $\mathbb{R}^m$ where m is a positive integer, usually equal to d or 1. This relates to the study of measures of multivariate risks; see e.g., Embrechts and Puccetti (2006). Our formulation in this section should not be confused with the above one. We stick to the domain $\mathcal{X}$ for the main reason that probability distortion is usually defined and well-understood in dimension one; see the recent work Liu et al. (2020) for a characterization of probability distortion in dimension one.

Acknowledgements

The authors thank Andreas Tsanakas, Steven Vanduffel, the Editor Mario Wüthrich, and the anonymous referees for helpful comments on an earlier version of the paper. Ruodu Wang acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03823, RGPAS-2018-522590) and the University of Waterloo CAE Research Grant from the Society of Actuaries.

A Proofs of all results

Proof of Lemma 1. (i) and (ii) can be obtained by combining the results of Lemma 3 in Wang et al. (2020) and Theorems 4 and 6 of Dhaene et al. (2012). We only prove (iii). We first suppose that h is right-continuous. Since F_X^{-1} is continuous on $(0,1)$, we have

$$F_X^{-1}(1-t) = F_X^{-1+}(1-t), \quad \text{for all } t \in [0,1].$$

It then follows from (i) that

$$\int X \,\mathrm{d}h \circ \mathbb{P} = \int_0^1 F_X^{-1+}(1-t)\,\mathrm{d}h(t) = \int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h(t).$$

Then suppose that h is left-continuous. According to (ii), it is straightforward that

$$\int X \, \mathrm{d}h \circ \mathbb{P} = \int_0^1 F_X^{-1}(1-t) \, \mathrm{d}h(t).$$

Then consider a general h. Since h is of bounded variation, it has countably many points of discontinuity. Then we can always decompose $h = h_r + h_l$, where h_r and h_l are right-continuous and left-continuous parts of h, respectively. From (2), it is obvious that

$$\int X \, \mathrm{d}(ah_1 + bh_2) \circ \mathbb{P} = a \int X \, \mathrm{d}h_1 \circ \mathbb{P} + b \int X \, \mathrm{d}h_2 \circ \mathbb{P}$$

for all $h_1, h_2 \in \mathcal{H}$ and $a, b \in \mathbb{R}$. According to the above discussion,

$$\begin{aligned}\int X \, \mathrm{d}h \circ \mathbb{P} &= \int X \, \mathrm{d}h_r \circ \mathbb{P} + \int X \, \mathrm{d}h_l \circ \mathbb{P} \\ &= \int_0^1 F_X^{-1}(1-t) \, \mathrm{d}h_r(t) + \int_0^1 F_X^{-1}(1-t) \, \mathrm{d}h_l(t) = \int_0^1 F_X^{-1}(1-t) \, \mathrm{d}h(t).\end{aligned}$$

The other equality is similar. $\square$

Proof of Proposition 1. (i) Recall the quantile representation of the integral $\int X \, \mathrm{d}h \circ \mathbb{P}$,

$$\int X \, \mathrm{d}h \circ \mathbb{P} = \int_0^1 F_X^{-1+}(1-t) \, \mathrm{d}h_r(t) + \int_0^1 F_X^{-1}(1-t) \, \mathrm{d}h_l(t). \tag{7}$$

We show finiteness of the first term in (7) and finiteness of the second term follows similarly. For any $\epsilon \in (0,1)$ such that h is absolutely continuous in $[0,\epsilon) \cup (1-\epsilon, 1]$ and

$$h' \in L^q((0,\epsilon) \cup (1-\epsilon, 1)),$$

we have $|F_X^{-1+}(1-t)| < \infty$ for all $t \in [\epsilon, 1-\epsilon]$. It follows that

$$\left| \int_\epsilon^{1-\epsilon} F_X^{-1+}(1-t) \, \mathrm{d}h_r(t) \right| < \infty$$

since h is of bounded variation. It then suffices to show that

$$\left| \int_{[0,\epsilon)\cup(1-\epsilon,1]} F_X^{-1+}(1-t) \, \mathrm{d}h_r(t) \right| = \left| \int_{[0,\epsilon)\cup(1-\epsilon,1]} F_X^{-1+}(1-t) h_r'(t) \, \mathrm{d}t \right| < \infty.$$

Since $X \in L^p$, the right-quantile $F_X^{-1+} \in L^p([0,1])$. Note that $h_r' \in L^q((0,\epsilon) \cup (1-\epsilon, 1))$ and $1/p + 1/q = 1$. By Hölder's inequality,

$$\begin{aligned}&\left| \int_{[0,\epsilon)\cup(1-\epsilon,1]} F_X^{-1+}(1-t) h_r'(t) \, \mathrm{d}t \right| \\ &\leq \int_{[0,\epsilon)\cup(1-\epsilon,1]} |F_X^{-1+}(1-t)| \cdot |h_r'(t)| \, \mathrm{d}t \\ &\leq \left(\int_{[0,\epsilon)\cup(1-\epsilon,1]} |F^{-1+}(1-t)|^p \, \mathrm{d}t \right)^{\frac{1}{p}} \left(\int_{[0,\epsilon)\cup(1-\epsilon,1]} |h_r'(t)|^q \, \mathrm{d}t \right)^{\frac{1}{q}} < \infty.\end{aligned}$$

We then conclude that

$$\left|\int_0^1 F_X^{-1+}(1-t)\,\mathrm{d}h_r(t)\right| < \infty.$$

By similar arguments, $|\int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h_l(t)| < \infty$ holds naturally. Therefore, $\int X\,\mathrm{d}h\circ\mathbb{P}$ is finite.

(ii) Concavity of h implies that h is absolutely continuous on $(0,1)$. Suppose that h is not continuous at 0. Take $X_0 \sim \mathrm{N}(0,1)$ and $X = X_0^{1/p}$. It follows that $F_X^{-1}(1) = \infty$. By Lemma 1 (iii),

$$\left|\int X\,\mathrm{d}h\circ\mathbb{P}\right| = \left|\int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h(t)\right| = \infty,$$

which leads to a contradiction. Therefore, h is continuous at 0. Continuity of h at 1 holds analogously. h is thus absolutely continuous on $[0,1]$. Since h is of bounded variation, we can always use Jordan decomposition $h = h_+ - h_-$, where h_+ and h_- are increasing functions. Moreover, h can always be decomposed into $h = h_r + h_l$. It then suffices to prove the property for all increasing and right-continuous h.

Since h is concave, we have $h' \in L^1([0,1])$. Let

$$q' = \sup\{r \geq 1 : h' \in L^r((0,\epsilon)\cup(1-\epsilon,1)) \text{ for some } \epsilon \in (0,1)\}$$

and suppose for the purpose of contradiction that $q' < q$. Note that we have $q'/(q'-1) > p$. Hence, there exists $\delta > 0$ such that

$$q' + \delta < q \quad \text{and} \quad \frac{q'}{q'+\delta-1} > p.$$

Let $q^* = q' + \delta$ and $p^* = q^*/(q^*-1) > p$. Note that $q^*p/p^* = (q'+\delta-1)p < q'$. Construct a random variable X such that

$$\left|F_X^{-1}(1-t)\right| = |h'(t)|^{\frac{q^*}{p^*}},$$

for almost everywhere $t \in [0,1]$. This is always possible due to concavity of h, which implies that h' is decreasing and h' has countably many discontinuity points. Since $q^*p/p^* < q'$, we have $h' \in L^{(q^*p/p^*)}((0,\epsilon)\cup(1-\epsilon,1))$ for some $\epsilon > 0$, and hence $X \in L^p$. Noting that $h' \notin L^{q^*}((0,\epsilon)\cup(1-\epsilon,1))$, we have

$$\left|\int_{[0,\epsilon)\cup(1-\epsilon,1]} F_X^{-1}(1-t)h'(t)\,\mathrm{d}t\right| = \int_{[0,\epsilon)\cup(1-\epsilon,1]} |h'(t)|^{\frac{q^*}{p^*}+1}\,\mathrm{d}t = \int_{[0,\epsilon)\cup(1-\epsilon,1]} |h'(t)|^{q^*}\,\mathrm{d}t = \infty,$$

which leads to a contradiction. Therefore, $q' \geq q$. $\square$

Proof of Theorem 1. (i) "⇒": For all $X \in \mathcal{X}$, we define a random variable

$$X_M = X\mathbb{1}_{\{|X|\le M\}} + M\mathbb{1}_{\{X>M\}} - M\mathbb{1}_{\{X<-M\}}, \quad M \ge 0.$$

Since ρ is continuous at infinity, we have $\rho(X_M) \to \rho(X)$. Note that $X_M \in L^\infty$ for any $M \ge 0$. It follows from Theorem 1 of Wang et al. (2020) that on L^∞, the law-invariant, comonotonic-additive and uniformly sup-continuous functional ρ can be represented by a signed Choquet integral

$$\begin{aligned}\rho(X_M) &= \int_{-\infty}^{0} \left(h(\mathbb{P}(X_M \ge x)) - h(1)\right) \mathrm{d}x + \int_0^\infty h(\mathbb{P}(X_M \ge x))\,\mathrm{d}x \\ &= \int_{-M}^{0} \left(h(\mathbb{P}(X \ge x)) - h(1)\right) \mathrm{d}x + \int_0^M h(\mathbb{P}(X \ge x))\,\mathrm{d}x,\end{aligned} \tag{8}$$

where $h \in \mathcal{H}$. Specifically, $h(t) = \rho(\mathbb{1}_{\{U<t\}}) < \infty$ for $t \in [0,1]$, where U is a uniform random variable on $[0,1]$. Letting $M \to \infty$, we have

$$\rho(X) = \int_{-\infty}^{0} \left(h(\mathbb{P}(X \ge x)) - h(1)\right) \mathrm{d}x + \int_0^\infty h(\mathbb{P}(X \ge x))\,\mathrm{d}x.$$

(ii) "⇐": Law-invariance is straightforward. Comonotonic-additivity follows from (7), since the left- and right-quantiles are well-known to be comonotonic-additive (see Proposition 7.20 of McNeil et al. (2015) for the case of left-quantile). Continuity at infinity holds simply by

$$\begin{aligned}\rho_h(X_M) &= \int_{-\infty}^{0} \left(h(\mathbb{P}(X_M \ge x)) - h(1)\right) \mathrm{d}x + \int_0^\infty h(\mathbb{P}(X_M \ge x))\,\mathrm{d}x \\ &= \int_{-M}^{0} \left(h(\mathbb{P}(X \ge x)) - h(1)\right) \mathrm{d}x + \int_0^M h(\mathbb{P}(X \ge x))\,\mathrm{d}x \xrightarrow{M\to\infty} \rho_h(X).\end{aligned}$$

To see the uniform sup-continuity, we take any two random variables $X, Y \in \mathcal{X}$. By representation (7), we have

$$\begin{aligned}&|\rho_h(X) - \rho_h(Y)| \\ &\le \left|\int_0^1 \left(F_X^{-1+}(1-t) - F_Y^{-1+}(1-t)\right) \mathrm{d}h_r(t)\right| + \left|\int_0^1 \left(F_X^{-1}(1-t) - F_Y^{-1}(1-t)\right) \mathrm{d}h_l(t)\right| \\ &\le \operatorname{ess\,sup}|X-Y| \cdot \mathrm{TV}_h,\end{aligned}$$

where TV_h is the total variation of the function h on $[0,1]$. □

Proof of Proposition 2. (i) Sufficiency is straightforward from the definition of distortion risk-metrics. Necessity can be checked by Bernoulli random variables.

(ii) "$\Rightarrow$": We take $X = \mathbb{1}_{\{U \le t_1\}}$ and $Y = \mathbb{1}_{\{U \le t_2\}}$ for all $t_1, t_2 \in [0,1]$ such that $t_1 \le t_2$, where $U \sim \mathrm{U}[0,1]$. Then we have $X \le Y$. Suppose that ρ_h is increasing (resp. decreasing). We have $h(t_1) = \rho_h(X) \le \rho_h(Y) = h(t_2)$ (resp. $h(t_1) = \rho_h(X) \ge \rho_h(Y) = h(t_2)$). Thus h is increasing (resp. decreasing).

"$\Leftarrow$": For any random variables $X, Y \in \mathcal{X}$ such that $X \le Y$, we have $\mathbb{P}(X \ge x) \le \mathbb{P}(Y \ge x)$ for all $x \in \mathbb{R}$. If h is increasing (resp. decreasing), then $h(\mathbb{P}(X \ge x)) \le h(\mathbb{P}(Y \ge x))$ (resp. $h(\mathbb{P}(X \ge x)) \ge h(\mathbb{P}(Y \ge x))$) for all $x \in \mathbb{R}$. It implies that $\rho_h(X) \le \rho_h(Y)$ (resp. $\rho_h(X) \ge \rho_h(Y)$).

(iii) For all $c \in \mathbb{R}$, we first calculate

$$\begin{aligned}\rho_h(c) &= \int_{-\infty}^{0} (h(\mathbb{P}(c \ge x)) - h(1))\,\mathrm{d}x + \int_0^\infty h(\mathbb{P}(c \ge x))\,\mathrm{d}x \\ &= \int_{0\wedge c}^{0} (-h(1))\,\mathrm{d}x + \int_0^{0\vee c} h(1)\,\mathrm{d}x = ch(1).\end{aligned}$$

Note that any random variable $X \in \mathcal{X}$ and c are comonotonic. By comonotonic-additivity of ρ_h, we have $\rho_h(X + c) = \rho_h(X) + \rho_h(c) = \rho_h(X) + ch(1)$.

(iv) For all $\lambda > 0$ and all $X \in \mathcal{X}$,

$$\begin{aligned}\rho_h(\lambda X) &= \int_{-\infty}^{0} (h(\mathbb{P}(\lambda X \ge x)) - h(1))\,\mathrm{d}x + \int_0^\infty h(\mathbb{P}(\lambda X \ge x))\,\mathrm{d}x \\ &= \int_{-\infty}^{0} (h(\mathbb{P}(X \ge \frac{1}{\lambda}x)) - h(1))\,\mathrm{d}x + \int_0^\infty h(\mathbb{P}(X \ge \frac{1}{\lambda}x))\,\mathrm{d}x \\ &= \lambda \int_{-\infty}^{0} (h(\mathbb{P}(X \ge u)) - h(1))\,\mathrm{d}u + \lambda \int_0^\infty h(\mathbb{P}(X \ge u))\,\mathrm{d}u = \lambda \rho_h(X).\end{aligned}$$

(v) This property is already shown in the proof of Lemma 1 (ii). $\square$

Proof of Proposition 3. Since ρ_h is convex on $\mathcal{X}$, we know that it is convex on L^∞, which implies that h is concave by Theorem 3 of Wang et al. (2020).

Suppose that there exists $X \in \mathcal{X}$ such that $\mathbb{E}[|X|] = \infty$. Note that $\mathbb{E}[|X|] = \infty$ implies either $\mathbb{E}[X_+] = \infty$ or $\mathbb{E}[X_-] = \infty$. If $\mathbb{E}[X_+] = \infty$, then $Y = -X \in \mathcal{X}$ since $\mathcal{X}$ is a linear space, and $\mathbb{E}[Y_-] = \infty$. Similarly, if $\mathbb{E}[X_-] = \infty$, then $\mathbb{E}[Y_+] = \infty$. Therefore, we know that there exist $X, Y \in \mathcal{X}$ such that $\mathbb{E}[X_+] = \mathbb{E}[Y_-] = \infty$.

Take $X \in \mathcal{X}$ with $\mathbb{E}[X_+] = \infty$. Since

$$\rho_h(X) = \int_{-\infty}^{0} (h(\mathbb{P}(X \ge x)) - h(1))\,\mathrm{d}x + \int_0^\infty h(\mathbb{P}(X \ge x))\,\mathrm{d}x \in \mathbb{R},$$

both $\int_{-\infty}^{0}(h(\mathbb{P}(X \geq x))-h(1))\,\mathrm{d}x$ and $\int_0^\infty h(\mathbb{P}(X \geq x))\,\mathrm{d}x$ have to be finite. Since X is unbounded from above, this implies that h is continuous at 0. Similarly, take $Y \in \mathcal{X}$ with $\mathbb{E}[Y_-]=\infty$, and we obtain h is continuous at 1. Further by concavity, h is continuous on $[0,1]$. Using Lemma 1, we get

$$\rho_h(X)=\int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h(t).$$

There exists $\delta>0$ such that $F_X^{-1}(1-\epsilon)>0$ for all $\epsilon \in (0,\delta)$. Moreover,

$$\epsilon \int_0^\epsilon F_X^{-1}(1-t)\,\mathrm{d}t=\infty$$

for all $\epsilon \in (0,\delta)$. Let $h'(t)$ be the right-derivative of h at $t \in [0,1)$. Assume that $h'(0)>0$. Since h is concave and continuous, there exists $\epsilon>0$ such that $h'(t)>\epsilon$ for $t\in[0,\epsilon]$. It follows that

$$\int_0^\epsilon F_X^{-1}(1-t)\,\mathrm{d}h(t) \geq \epsilon \int_0^\epsilon F_X^{-1}(1-t)\,\mathrm{d}t=\infty,$$

contradicting the fact that $\rho_h(X)$ is finite. Therefore, $h'(0)\leq 0$. Using similar arguments as above for Y, we obtain $h'(1)\geq 0$ where $h'(1)$ is the left derivative of h at 1. Since h is concave, these two conditions imply that $h=0$ on $[0,1]$, and hence $\rho_h=0$ on $\mathcal{X}$. $\square$

Proof of Theorem 2. (i) "$\Rightarrow$": Suppose that $X \leq_{\mathrm{cx}} Y$. We first consider the case where $h \in \mathcal{H}$ is increasing. For an increasing concave function $h \in \mathcal{H}$, it is well-known (e.g., Theorem 1 of Williamson (1956)) that there exists some finite Borel measure μ on $[0,1]$, such that

$$h(t)=\int_0^1 \frac{1}{u} h_u(t)\,\mathrm{d}\mu(u), \quad t\in[0,1], \tag{9}$$

where $h_u(t)=t\wedge u$ for $t,u\in[0,1]$ and we use the convention $h_u(t)/u=\mathbb{1}_{\{t>0\}}$ if $u=0$. By the quantile representation of a distortion riskmetric,

$$\rho_{h_u}(X)=\int_0^u F_X^{-1}(1-t)\,\mathrm{d}t=\int_{1-u}^1 F_X^{-1}(u)\,\mathrm{d}u \leq \int_{1-u}^1 F_Y^{-1}(u)\,\mathrm{d}u=\rho_{h_u}(Y),$$

where the third inequality holds by Theorem 3.A.5 of Shaked and Shanthikumar (2007). It follows that

$$\rho_h(X)=\int_0^1 \frac{1}{u}\rho_{h_u}(X)\,\mathrm{d}\mu(u) \leq \int_0^1 \frac{1}{u}\rho_{h_u}(Y)\,\mathrm{d}\mu(u)=\rho_h(Y).$$

When $h\in\mathcal{H}$ is decreasing, similar to (9), we have

$$h(t)=\int_0^1 \frac{1}{1-u}(h_u(t)-t)\,\mathrm{d}\nu(u), \quad t\in[0,1]$$

for some finite Borel measure ν on $[0,1]$ where the convention is $(h_u(t)-t)/(1-u)=-\mathbb{1}_{\{t=1\}}$ if $u=1$. By definition of $X\leq_{\mathrm{cx}} Y$, it implies that $\mathbb{E}[X]=\mathbb{E}[Y]$. It then follows that

$$\rho_h(X)=\int_0^1 \frac{1}{1-u}(\rho_{h_u}(X)-\mathbb{E}[X])\,\mathrm{d}\nu(u) \leq \int_0^1 \frac{1}{1-u}(\rho_{h_u}(Y)-\mathbb{E}[Y])\,\mathrm{d}\nu(u)=\rho_h(Y).$$

For any concave function h on $[0,1]$, there always exists $\hat{x} \in [0,1]$, such that $h(\hat{x}) \geq h(x)$ for all $x \in [0,1]$. Then h can always be decomposed by $h = h_\uparrow + h_\downarrow$, where

$$h_\uparrow(x) = h(x)\mathbb{1}_{\{0 \leq x < \hat{x}\}} + h(\hat{x})\mathbb{1}_{\{\hat{x} \leq x \leq 1\}} \text{ and } h_\downarrow(x) = [h(x) - h(\hat{x})]\mathbb{1}_{\{\hat{x} \leq x \leq 1\}}.$$

Notice that $h_\uparrow$ and $h_\downarrow$ are increasing and decreasing concave functions, respectively, with

$$h_\uparrow(0) = h_\downarrow(0) = 0.$$

According to the above arguments, we have

$$\rho_h(X) = \rho_{h_\uparrow}(X) + \rho_{h_\downarrow}(X) \leq \rho_{h_\uparrow}(Y) + \rho_{h_\downarrow}(Y) = \rho_h(Y).$$

(ii) "$\Leftarrow$": Suppose that $\rho_h(X) \leq \rho_h(Y)$ for all concave functions $h \in \mathcal{H}$. For all $t, u \in [0,1]$, choose a concave $h \in \mathcal{H}$ such that $h(t) = h_u(t) = t \wedge u$. Then for all $u \in [0,1]$,

$$\rho_h(X) = \int_{1-u}^{1} F_X^{-1}(u)\,\mathrm{d}u \text{ and } \rho_h(Y) = \int_{1-u}^{1} F_Y^{-1}(u)\,\mathrm{d}u.$$

It follows that

$$\int_{1-u}^{1} F_X^{-1}(u)\,\mathrm{d}u \leq \int_{1-u}^{1} F_Y^{-1}(u)\,\mathrm{d}u \text{ for all } u \in [0,1],$$

which is equivalent to $X \leq_{\mathrm{cx}} Y$ by Theorem 3.A.5 of Shaked and Shanthikumar (2007). □

Proof of Theorem 3. $(i) \Rightarrow (ii)$ is shown by Theorem 2. We proceed in the order $(ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (i)$, and the arguments are based on Theorem 3 of Wang et al. (2020).

$(ii) \Rightarrow (iii)$: Take random variables $X, Y, X^c, Y^c \in \mathcal{X}$, such that $X \stackrel{\mathrm{d}}{=} X^c$, $Y \stackrel{\mathrm{d}}{=} Y^c$ and X^c and Y^c are comonotonic. By Theorem 3.5 of Rüschendorf (2013), we have $X + Y \leq_{\mathrm{cx}} X^c + Y^c$. It then follows from law-invariance, comonotonic-additivity and convex order consistency of ρ_h that

$$\rho_h(X+Y) \leq \rho_h(X^c + Y^c) = \rho_h(X^c) + \rho_h(Y^c) = \rho_h(X) + \rho_h(Y).$$

$(iii) \Rightarrow (iv)$: As ρ_h is positively homogeneous, subadditivity is equivalent to convexity.

$(iv) \Rightarrow (v)$: Directly from the definition of convexity and quasi-convexity.

$(v) \Rightarrow (vi)$: Theorem 3 of Wang et al. (2020) gives that quasi-convexity of I_h on L^∞ implies that h is concave. Concavity on mixtures follows directly from the concavity of h by the definition of a distortion riskmetric.

$(vi) \Rightarrow (i)$: Theorem 3 of Wang et al. (2020) gives that mixture-concavity of I_h on L^∞ implies that h is concave. □

Proof of Theorem 4. Suppose that $\rho : \mathcal{X} \to (-\infty, \infty]$ is a law-invariant, convex and continuous-at-infinity functional dominating ρ_h. Using Theorem 5 of Wang et al. (2020), we know that, on L^∞, ρ_{h^*} is the smallest law-invariant convex functional dominating ρ_h. Therefore, $\rho \geq \rho_{h^*}$ on L^∞. If ρ_{h^*} is finite on $\mathcal{X}$, then both ρ and ρ_{h^*} are continuous at infinity on $\mathcal{X}$, and hence $\rho \geq \rho_{h^*}$ on $\mathcal{X}$. If ρ_{h^*} is not finite on $\mathcal{X}$, then we know that $\int X \,\mathrm{d}h^* \circ \mathbb{P} = \infty$ (but not $-\infty$ since $\rho_{h^*} \geq \rho_h$) for some $X \in \mathcal{X}$. Let

$$X_M = X\mathbb{1}_{\{|X| \leq M\}} + M\mathbb{1}_{\{X > M\}} - M\mathbb{1}_{\{X < -M\}}, \quad M \geq 0.$$

Using (8), $\rho = \rho_{h^*}$ on L^∞ and $\int X \,\mathrm{d}h^* \circ \mathbb{P} = \infty$, we have, as $M \to \infty$,

$$\rho(X_M) = \rho_{h^*}(X_M) = \int_{-M}^{0} \left(h^*(\mathbb{P}(X \geq x)) - h(1)\right) \mathrm{d}x + \int_0^M h^*(\mathbb{P}(X \geq x)) \,\mathrm{d}x \to \infty.$$

The continuity at infinity of ρ implies $\rho(X) = \infty$, and hence ρ cannot be real-valued on $\mathcal{X}$. $\square$

Proof of Theorem 5. Note that $X \mapsto \mathrm{ES}_\alpha(X)$ and $X \mapsto \mathrm{ES}_\alpha(-X)$ are convex distortion riskmetrics for all $\alpha \in [0,1]$. As a mixture of $X \mapsto \mathrm{ES}_\alpha(X)$ and $X \mapsto \mathrm{ES}_\alpha(-X)$, ρ defined by (5) satisfies convexity, comonotonic-additivity, law-invariance, continuity at infinity, and uniform sup-continuity. Hence, ρ is a convex distortion riskmetric. Next we show the "only-if" statement. Denote by h the distortion function of ρ, which by Theorem 3 is a concave function. Following the same argument in the proof of Theorem 2, we can write for some finite Borel measures γ, ν on $[0,1]$,

$$h(t) = \int_0^1 \frac{1}{\alpha} h_\alpha(t) \,\mathrm{d}\gamma(\alpha) + \int_0^1 \frac{1}{1-\alpha}(h_\alpha(t) - t) \,\mathrm{d}\nu(\alpha), \quad t \in [0,1], \tag{10}$$

where $h_\alpha(t) = t \wedge \alpha$. Note that $\frac{1}{\alpha} h_\alpha$ is the distortion function of $\mathrm{ES}_{1-\alpha}$. By Proposition 2, the distortion function of $X \mapsto \mathrm{ES}_\alpha(-X)$ is given by

$$g_\alpha(t) = \frac{1-t}{1-\alpha} \wedge 1 - 1 = \frac{(\alpha - t) \wedge 0}{1-\alpha} = \frac{1}{1-\alpha}(h_\alpha(t) - t), \quad t \in [0,1].$$

Therefore, (10) gives

$$\rho(X) = \int_0^1 \mathrm{ES}_{1-\alpha}(X) \,\mathrm{d}\gamma(\alpha) + \int_0^1 \mathrm{ES}_\alpha(-X) \,\mathrm{d}\nu(\alpha), \qquad X \in \mathcal{X}.$$

Thus (5) holds with $\mathrm{d}\mu(\alpha) = \mathrm{d}\gamma(1-\alpha)$. $\square$

Proof of Theorem 6. Since $h \in \mathcal{H}$ is of bounded variation, it can be decomposed into $h = h_+ - h_-$ where h_+ and h_- are increasing functions. It then suffices to prove the result for all increasing function h. We denote the distribution function of X_n by F_n for $n \in \mathbb{N}$.

(i) If h is left-continuous and increasing, it induces a Borel measure μ on $[0,1]$ such that $h(t) = \mu([0,t))$, $t \in [0,1]$. By quantile representation of a distortion riskmetric,

$$\rho_h(X_n) = \int_0^1 F_n^{-1}(1-t)\,\mathrm{d}h(t) \text{ and } \rho_h(X) = \int_0^1 F_X^{-1}(1-t)\,\mathrm{d}h(t).$$

Since $X_n \to X$ in distribution, $F_n^{-1} \to F_X^{-1}$ almost everywhere on $[0,1]$, where F_X^{-1} is continuous. Let

$$A = \{t \in (0,1) : s \mapsto F_X^{-1}(1-s) \text{ is not continuous at } t\}.$$

According to the assumption, h must be continuous on the set A, which implies μ has no point mass on A and $\mu(A) = 0$. It remains to consider the points 0 and 1. Notice that h-uniform integrability implies that when $\mu(\{0\}) > 0$, $F_n^{-1}(1) \to F_X^{-1}(1)$ as $n \to \infty$ since $F_n^{-1}(1) = F_X^{-1}(1) = 0$ for all $n \in \mathbb{N}$. Similarly, when $\mu(\{1\}) > 0$, $F_n^{-1}(0) \to F_X^{-1}(0) = 0$ as $n \to \infty$. Therefore, $F_n^{-1} \to F_X^{-1}$ μ-almost surely. In addition, h-uniform integrability of $\{X_1, X_2, \dots\}$ is equivalent to uniform integrability of $\{F_1^{-1}, F_2^{-1}, \dots\}$ with respect to the measure μ. It then follows from Vitali's Convergence Theorem (Rudin (1987, p. 133)) that $\rho_h(X_n) \to \rho_h(X)$ as $n \to \infty$.

(ii) If h is right-continuous, we define the Borel measure ν on $[0,1]$ by $\nu([0,t]) = h(t)$, $t \in [0,1]$. We write the distortion riskmetrics as

$$\rho_h(X_n) = \int_0^1 F_n^{-1+}(1-t)\,\mathrm{d}h(t) \text{ and } \rho_h(X) = \int_0^1 F_X^{-1+}(1-t)\,\mathrm{d}h(t).$$

Note that the set

$$\begin{aligned} B &= \{t \in (0,1) : s \mapsto F_X^{-1+}(1-s) \text{ is not continuous at } t\} \\ &= \{t \in (0,1) : s \mapsto F_X^{-1}(1-s) \text{ is not continuous at } t\}. \end{aligned}$$

This implies $\nu(B) = 0$. By similar argument as (i), we get $F_n^{-1+} \to F_X^{-1+}$ ν-almost surely and $\rho_h(X_n) \to \rho_h(X)$ as $n \to \infty$.

(iii) For a general h, we can write ρ_h by (7), where h_r and h_l are taken such that the collection of discontinuity points of h_r and h_l coincides with that of h. To see that it is always possible, we define countable sets

$$C = \{t \in [0,1] : s \mapsto h(s) \text{ is not continuous at } t\},$$

$$C^+ = \{t \in C : s \mapsto h(s) \text{ is right-continuous at } t\} \text{ and } C^- = C \setminus C^+.$$

Take

$$h_r(x) = \sum_{t \in C^+} [h(t^+) - h(t^-)] \mathbb{1}_{\{x > t\}} + h(x) \mathbb{1}_{\{x \notin C\}} \text{ and } h_l(x) = \sum_{t \in C^-} [h(t^+) - h(t^-)] \mathbb{1}_{\{x \geq t\}}$$

for $x \in [0, 1]$. Thus h_r and h_l are as desired. It follows that

$$|\rho_h(X_n) - \rho_h(X)| \leq |\rho_{h_r}(X_n) - \rho_{h_r}(X)| + |\rho_{h_l}(X_n) - \rho_{h_l}(X)| \to 0$$

as $n \to \infty$. This implies $\rho_h(X_n) \to \rho_h(X)$ as $n \to \infty$ in general. □

Proof of Proposition 4. Suppose that we have random variables $X_1, X_2, \dots \in L^p$ such that $X_n \to X$ in L^p as $n \to \infty$. Let F_n be the distribution function of X_n for $n \in \mathbb{N}$. Since $h \in \mathcal{H}_q$, there exists $\epsilon \in (0, 1)$ such that $h' \in L^q((0, \epsilon) \cup (1 - \epsilon, 1))$. Then we have

$$\begin{aligned}|\rho_h(X_n) - \rho_h(X)| \leq & \left| \int_{[0,\epsilon) \cup (1-\epsilon,1]} (F_n^{-1}(1-t) - F_X^{-1}(1-t)) \, \mathrm{d}h(t) \right| \\ & + \left| \int_{[\epsilon,1-\epsilon]} (F_n^{-1}(1-t) - F_X^{-1}(1-t)) \, \mathrm{d}h(t) \right|. \end{aligned} \tag{11}$$

By Hölder's inequality, the first term of (11) satisfies

$$\begin{aligned} & \left| \int_{[0,\epsilon) \cup (1-\epsilon,1]} (F_n^{-1}(1-t) - F_X^{-1}(1-t)) \, \mathrm{d}h(t) \right| \\ & \leq \int_{[0,\epsilon) \cup (1-\epsilon,1]} \left| F_n^{-1}(1-t) - F_X^{-1}(1-t) \right| \cdot |h'(t)| \, \mathrm{d}t \\ & \leq \left(\int_{[0,\epsilon) \cup (1-\epsilon,1]} \left| F_n^{-1}(1-t) - F_X^{-1}(1-t) \right|^p \mathrm{d}t \right)^{\frac{1}{p}} \left(\int_{[0,\epsilon) \cup (1-\epsilon,1]} |h'(t)|^q \, \mathrm{d}t \right)^{\frac{1}{q}} \xrightarrow{n \to \infty} 0. \end{aligned}$$

It remains to show the second term of (11) converges to zero. Note that

$$\begin{aligned} \left| \int_{[\epsilon,1-\epsilon]} (F_n^{-1}(1-t) - F_X^{-1}(1-t)) \, \mathrm{d}h(t) \right| &= \left| \int_0^1 (F_n^{-1}(1-t) - F_X^{-1}(1-t)) \, \mathrm{d}\tilde{h}(t) \right| \\ &= |\rho_{\tilde{h}}(X_n) - \rho_{\tilde{h}}(X)|, \end{aligned}$$

where

$$\tilde{h}(t) = \begin{cases} 0 & t \in [0, \epsilon), \\ h(t) - h(\epsilon) & t \in [\epsilon, 1 - \epsilon], \\ h(1 - \epsilon) - h(\epsilon) & t \in (1 - \epsilon, 1]. \end{cases}$$

Clearly, $\{X, X_1, X_2, \dots\}$ is uniformly $\tilde{h}$-integrable since $\tilde{h}$ stays constant in some neighborhood of 0 and 1. Also, $X_n \to X$ in L^p implies $X_n \to X$ in distribution and $\tilde{h}$ is continuous due to h being continuous. It then follows from Theorem 6 that

$$|\rho_{\tilde{h}}(X_n) - \rho_{\tilde{h}}(X)| \to 0 \quad \text{as } n \to \infty.$$

Therefore, the second term of (11) also converges to zero. We conclude that $\rho_h(X_n) \to \rho_h(X)$ as $n \to \infty$, which proves the proposition. $\square$

Proof of Proposition 5. The proposition follows by applying Theorem 2 to each dimension of ρ. $\square$

Proof of Proposition 6. The proposition follows by applying Theorem 6 to each dimension of ρ. $\square$

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, **26**(7), 1505–1518.

Acerbi, C. and Szekely, B. (2017). General properties of backtestable statistics. *SSRN*: 2905109.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.

Belles-Sampera, J., Guillén, M. and Santolino, M. (2014). Beyond value-at-risk: GlueVaR distortion risk measures. *Risk Analysis*, **34**(1), 121-134.

Cheridito, P. and Li, T. (2009). Risk measures on Orlicz hearts. *Mathematical Finance*, **19**(2), 189–214.

Cherny, A. S. and Madan, D. (2009). New measures for performance evaluation. *Review of Financial Studies*, **22**(7), 2571–2606.

Choquet, G. (1954). Theory of capacities. *Annales de l'institut Fourier*, **5**, 131–295.

Delbaen, F. (2002). Coherent risk measures on general probability spaces. In *Advances in Finance and Stochastics* (pp. 1–37). Springer, Berlin.

Denneberg, D. (1990). Premium calculation: why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin*, **20**, 181–190.

Denneberg, D. (1994). *Non-additive Measure and Integral.* Springer Science & Business Media.

Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). *Actuarial Theory for Dependent Risks.* Wiley: Chichester, UK.

Dhaene, J., Kukush, A., Linders, D. and Tang, Q. (2012). Remarks on quantiles and distortion risk measures. *European Actuarial Journal*, **2**(2), 319–328.

Dhaene, J., Vanduffel, S., Goovaerts, M.J., Kaas, R., Tang, Q. and Vynche, D. (2006). Risk measures and comonotonicity: a review. *Stochastic Models*, **22**, 573–606.

Embrechts, P. and Puccetti, G. (2006). Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, **97**(2), 526–547.

Filipović, D. and Svindland, G. (2012). The canonical model space for law-invariant convex risk measures is L^1. *Mathematical Finance*, **22**(3), 585–589.

Fissler, T. and Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. *Annals of Statistics*, **44**(4), 1680–1707.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**(4), 429–447.

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

Frittelli, M. and Rosazza Gianin, E. (2002). Putting order in risk measures. *Journal of Banking and Finance*, **26**,1473–1486.

Frongillo, R. and Kash, I. A. (2018). Elicitation complexity of statistical properties. *arXiv:* 1506.07212v2.

Furman, E., Wang, R. and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocation and heavy-tailed risks. *Journal of Banking and Finance*, **83**, 70–84.

Gerber, H. U. (1974). On additive premium calculation principles. *ASTIN Bulletin*, **7**(3), 215–222.

Gneiting T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, **106**(494),746–762.

Grechuk, B., Molyboha, A. and Zabarankin, M. (2009). Maximum entropy principle with general deviation measures. *Mathematics of Operations Research*, **34**(2), 445–467.

Huber, P. J. and Ronchetti, E. M. (2009). *Robust Statistics.* Second Edition, Wiley Series in Probability and Statistics. Wiley, New Jersey.

Kou, S. and Peng, X. (2016). On the measurement of economic tail risk. *Operations Research*, **64**(5): 1056-1072.

Krätschmer, V., Schied, A. and Zähle, H. (2014). Comparative and quantitiative robustness for law-invariant risk measures. *Finance and Stochastics*, **18**(2), 271–295.

Kusuoka, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83–95.

Lambert, N., Pennock, D. M. and Shoham, Y. (2008). Eliciting properties of probability distributions. *Proceedings of the 9th ACM Conference on Electronic Commerce*, 129–138.

Liebrich, F-B. and Svindland, G. (2017). Model spaces for risk measures. *Insurance: Mathematics and Economics*, **77**, 150-165.

Liu, F., Cai, J., Lemieux, C. and Wang, R. (2020). Convex risk functionals: representation and applications. *Insurance: Mathematics and Economics*, **90**, 66–79.

Liu, P., Schied, A. and Wang, R. (2020). Distributional transforms, probability distortions, and their applications. *Mathematic of Operations Research*, forthcoming. *SSRN*: 3419388.

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

Pichler, A. (2013). The natural Banach space for version independent risk measures. *Insurance: Mathematics and Economics*, **53**(2), 405-415.

Rockafellar, R. T., Uryasev, S. and Zabarankin, M. (2006). Generalized deviation in risk analysis. *Finance and Stochastics*, **10**, 51–74.

Rudin, W. (1987). *Real and Complex Analysis.* Tata McGraw-Hill Education.

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

Portfolios. Springer, Heidelberg.

Ruszczyński, A. and Shapiro, A. (2006). Optimization of convex risk functions. *Mathematics of Operations Research*, **31**(3), 433–452.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society*, **97**(2), 255–261.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer Series in Statistics.

Wang, R. and Wei, Y. (2020). Risk functionals with convex level sets. *Mathematical Finance*, forthcoming. *SSRN*: 3292661.

Wang, R., Wei, Y. and Willmot, G. (2020). Characterization, robustness and aggregation of signed Choquet integrals. *Mathematic of Operations Research*, published online. DOI: 10.1287/moor.2019.1020.

Wang, S., Young, V. R. and Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, **21**(2), 173–183.

Williamson, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Mathematical Journal*, **23**(2), 189-257.

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