Characterization, Robustness and Aggregation of Signed Choquet Integrals

Ruodu Wang[*] Yunran Wei[†] Gordon E. Willmot[‡]

June 12, 2019

Abstract

This article contains various results on a class of non-monotone law-invariant risk functionals, called the signed Choquet integrals. A functional characterization via comonotonic additivity is established, along with some theoretical properties including six equivalent conditions for a signed Choquet integral to be convex. We proceed to address two practical issues currently popular in risk management, namely, robustness (continuity) issues and risk aggregation with dependence uncertainty, for signed Choquet integrals. Our results generalize in several directions those in the literature of risk functionals. From the results obtained in this paper, we see that many profound and elegant mathematical results in the theory of risk measures hold for the general class of signed Choquet integrals; thus they do not rely on the assumption of monotonicity.

Keywords: comonotonicity; Choquet integrals; risk functionals; risk aggregation; robustness

1 Introduction

Over the past few decades, measures of risk and variability are introduced to quantify various characteristics of random financial losses of a financial institution. These measures are mappings from the set of random variables to real numbers (thus, *risk functionals*). Typical examples of risk measures include the Value-at-Risk, the Expected Shortfall and various coherent or convex risk measures as introduced by Artzner et al. (1999) and Föllmer and Schied (2002), and typical examples of variability measures include the variance, the standard deviation, the mean absolute

[*]Department of Statistics and Actuarial Science, University of Waterloo (email: `wang@uwaterloo.ca`).

[†]Department of Statistics and Actuarial Science, University of Waterloo (email: `y27wei@uwaterloo.ca`)

[‡]Department of Statistics and Actuarial Science, University of Waterloo (email: `gewillmot@uwaterloo.ca`)

deviation, and various deviation measures as introduced by Rockafellar et al. (2006). We refer to McNeil et al. (2015) for a comprehensive treatment of the use of risk measures in modern risk management.

In the practice of risk measurement, one very often assesses a risk through its distribution, which is obtained via statistical and simulation analysis. In academic terms, this means commonly used measures of risk and variability are *law-invariant*. From the work of Kusuoka (2001) and Grechuk et al. (2009), a class of risk functionals becomes the building block of law-invariant risk measures: the (law-invariant) *signed Choquet integrals*.

In this paper, a signed Choquet integral $I_h : L^\infty \to \mathbb{R}$ is defined as

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

where L^∞ is the set of bounded random variables in a probability space, and $h : [0,1] \to \mathbb{R}$ is a function of bounded variation with $h(0) = 0$. The notion of signed Choquet integrals without law-invariance originates from Choquet (1954) in the framework of *capacities*, and is further characterized and studied in decision theory by Schmeidler (1986, 1989) and extended by Cerreia-Vioglio et al. (2012, 2015) to general spaces.

From a risk management perspective, we focus on law-invariant functionals in this paper. In what follows we shall omit the term "law-invariant", as all risk functionals we discuss are law-invariant. There has been an extensive literature on a subclass of signed Choquet integrals, in which h is increasing and $h(1) = 1$; we simply call this class of functionals *increasing Choquet integrals*. In different contexts, such functionals I_h are referred to as L-functionals (Huber and Ronchetti (2009)) in statistics, Yaari's dual utilities (Yaari (1987)) in decision theory, distorted premium principles (Denneberg (1994) and Wang et al. (1997)) in insurance, and distortion risk measures (Kusuoka (2001) and Acerbi (2002)) in finance. In particular, the two most important risk measures used in current banking and insurance regulation, the Value-at-Risk and the Expected Shortfall, are increasing Choquet integrals. For properties and recent advances on various issues related to increasing Choquet integrals, we refer to Dhaene et al. (2012), Wang et al. (2015), Kou and Peng (2016), Delbaen et al. (2016) and Ziegel (2016).

On the other hand, there has been relatively limited research on signed Choquet integrals compared to that on increasing Choquet integrals. The major difference between an increasing Choquet integral and a signed one is that the latter, being more general, is not necessarily monotone. We are particularly interested in signed Choquet integrals for various practical and theoretical reasons. First, although a suitable risk measure should be monotone as argued by Artzner et al.

(1999), this issue is irrelevant for a measure of variability. Indeed, all practical measures of variability are not monotone (for instance, variance, standard deviation, or deviation measures in Rockafellar et al. (2006) and Grechuk et al. (2009)). Therefore, instead of obtaining an increasing-Choquet-integral-based representation for law-invariant coherent risk measures as in Kusuoka (2001), one naturally arrives at a signed-Choquet-integral-based representation of deviation measures as in Grechuk et al. (2009). In other words, signed Choquet integrals are relevant as long as a measure of variability is concerned. Second, there are many preferences or risk measures used in practice which are not monotone. A prominent example is the mean-variance and the mean-standard-deviation preferences as already studied by Markowitz (1952); see also Filipović and Svindland (2008) for a study of risk sharing with non-monotone risk measures, and Furman et al. (2017) for the class of Gini Shortfall risk measures, which are not necessarily monotone. Third, in economic decision theory, signed Choquet integrals appear naturally in many rank-based decision making; see Quiggin (1982), Gilboa and Schmeidler (1989) and De Waegenaere and Wakker (2001). Fourth, from a mathematical perspective, we aim to generalize some elegant results, which are known to hold true for increasing Choquet integrals, to the broader class of signed Choquet integrals.

The main contributions that we offer are summarized below. In Section 2, we establish a characterization of signed Choquet integrals via comonotonic additivity based on the seminal work of Schmeidler (1986). Furthermore, various theoretical properties of signed Choquet integrals are studied, such as monotonicity, additivity, quantile representations, convexity, quasi-convexity, convex order consistency, and mixture-concavity. The characterization and properties are partially known in the literature; yet we are unaware of a good summarizing article (hopefully this paper serves as one). In particular, few results were found for an atomless probability space.

We proceed to discuss in Sections 3 and 4 two practically relevant and currently popular problems concerning signed Choquet integrals: robustness issues and risk aggregation with dependence uncertainty. As pointed out by the recent Basel accords (see BCBS (2016)), model uncertainty and robustness become a focal point in both academic research and industry practice of risk assessment over the past few years. We refer to Embrechts et al. (2014) and Emmer et al. (2015) for a summary on these issues and their relation to the recent development in banking and insurance regulation. For more on robustness of risk measures, see Cont et al. (2010), Kou et al. (2013), Krätschmer et al. (2014) and Embrechts et al. (2015), and for more on risk aggregation with dependence uncertainty, see Embrechts et al. (2013), Bernard et al. (2014) and Cai et al. (2018). Our results generalize those of Cont et al. (2010), Embrechts et al. (2015) and Pesenti et al. (2016) on robustness of distortion risk measures and L-statistics, and those of Wang et al. (2015) and Cai et al. (2018) on extreme risk aggregation for distortion risk measures. In particular, the detailed analysis in Wang et al. (2015) used to characterize an extreme-aggregation measure cannot be applied to signed Choquet integrals, and in this paper, we develop a completely different and more systemic approach based on some recent results on asymptotics of the set of risk aggregation.

From the results obtained in this paper, we clearly see that many profound and elegant mathematical results in the theory of risk functionals remain valid for the general class of signed Choquet integrals; they do not rely on the common assumption of monotonicity. Hopefully, our results serve as a building block for future theoretical developments and applications of signed Choquet integrals.

In this paper, our discussions are confined to the space of bounded random variables L^∞, in order for signed Choquet integrals to be properly defined, and for all results to be concisely stated. Some results involve norm-continuity on the space or an operation (addition or subtraction) on several signed Choquet integrals, and hence we need to fix a suitable domain upfront. Certainly, many results can be naturally generalized to functional-specific spaces such as Λ-spaces (Lorentz (1951)) and Orlicz spaces (e.g. Rao and Ren (1991)). We leave this direction of research for future work.

2 Characterization and properties

2.1 Notation and definition

We first list some notation which will be used throughout. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be an atomless probability space. Consistently with the literature on risk measures, we work with the space L^∞ of essentially bounded random variables in $(\Omega, \mathcal{A}, \mathbb{P})$ equipped with L^∞-norm $||\cdot||_\infty$; this choice of common domain ensures all functionals we encounter are well-defined. A functional $\rho: L^\infty \to \mathbb{R}$ is *law-invariant* if $\rho(X) = \rho(Y)$ for any $X, Y \in L^\infty$ that have the same distribution under $\mathbb{P}$ (denoted as $X \overset{\mathrm{d}}{=} Y$). For all functionals discussed in this paper, we assume law-invariance. Moreover, we denote by $\mathcal{M}$ the set of distribution functions of $X \in L^\infty$. Terms of "increasing" and "decreasing" are in the non-strict sense.

For $F \in \mathcal{M}$, we write $X \sim F$ for $X \in L^\infty$ and X has distribution F. The left-continuous generalized inverse of F (left-quantile) is denoted by

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

whereas its right-continuous generalized inverse (right-quantile) is

$$F^{-1+}(t) = \sup\{x \in \mathbb{R} : F(x) \le t\},\ t \in [0,1),\ \text{and } F^{-1+}(1) = F^{-1}(1).$$

For any random variable X, we use F_X to denote its distribution function. Further, write

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

Next we present the definition of a signed Choquet integral, which originates from the seminal work of Choquet (1954) in the theory of capacities without assuming law-invariance.

Definition 1. A *signed Choquet integral* $I_h : L^\infty \to \mathbb{R}$ is defined as

$$I_h(X) = \int_{-\infty}^{0} \left(h(\mathbb{P}(X \ge x)) - h(1)\right) \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X \ge x))\, \mathrm{d}x, \tag{2}$$

where $h \in \mathcal{H}$. The function h is called the *distortion function* of I_h.

We first note that I_h is always finite on L^∞. For $X \in L^\infty$, since $h \in \mathcal{H}$ is of bounded variation, it is measurable and bounded. We can take $M > 0$ such that $|X| \le M$ and hence

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

As h is bounded, we have $|I_h(X)| < \infty$.

Remark 1. I_h has an alternative formulation by replacing $\mathbb{P}(X \ge x)$ with $\mathbb{P}(X > x)$ in (2). For $X \in L^\infty$, the functions $h(\mathbb{P}(X \ge x))$ and $h(\mathbb{P}(X > x))$ are equal almost everywhere for $x \in \mathbb{R}$, and therefore

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

In different places we shall use either of (2) and (3), whichever is more convenient.

From (2), it is clear that $I_{ah_1+bh_2} = aI_{h_1} + bI_{h_2}$ for $h_1, h_2 \in \mathcal{H}$ and $a, b \in \mathbb{R}$. In particular, for any $h \in \mathcal{H}$, we have $I_h = I_{h_+} - I_{h_-}$, where $h_+ \in \mathcal{H}$ and $h_- \in \mathcal{H}$ are increasing functions such that $h = h_+ - h_-$ via the Jordan decomposition. This decomposition will be used repeatedly in this paper, as often results are available in the literature for Choquet integrals with an increasing distortion function.

Before we proceed with characterizing signed Choquet integrals, we present some more terminology used throughout the paper. A most relevant concept to signed Choquet integrals is comonotonicity. Random variables X and Y are said to be *comonotonic* if there exists $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that $\omega, \omega' \in \Omega_0$,

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

For a functional $\rho : L^\infty \to \mathbb{R}$, we say that ρ is *comonotonic-additive*, if for any comonotonic random variables $X, Y \in L^\infty$, $\rho(X + Y) = \rho(X) + \rho(Y)$; ρ is *positively homogeneous*, if for $X \in L^\infty$ and constant $\lambda > 0$, $\rho(\lambda X) = \lambda \rho(X)$; ρ is *(uniformly) norm-continuous*, if it is (uniformly) continuous with respect to L^∞-norm; ρ is *quasi-convex* if $\rho(\lambda X + (1 - \lambda)Y) \leq \max\{\rho(X), \rho(Y)\}$ for all $X, Y \in L^\infty$ and $\lambda \in [0, 1]$.

A random variable X is said to be smaller than a random variable Y in *convex order*, denoted by $X \leq_{\mathrm{cx}} Y$, if $\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)]$ for all convex $\phi : \mathbb{R} \to \mathbb{R}$, provided that both expectations exist. The following fact about comonotonicity and convex order is well-known (see e.g. Theorem 3.5 of Rüschendorf (2013)): for any integrable random variables X, Y, X^c and Y^c such that $X \stackrel{\mathrm{d}}{=} X^c$, $Y \stackrel{\mathrm{d}}{=} Y^c$, and X^c and Y^c are comonotonic, one has $X + Y \leq_{\mathrm{cx}} X^c + Y^c$. We say that a functional $\rho : L^\infty \to \mathbb{R}$ is *convex order consistent* if $\rho(X) \leq \rho(Y)$ for all random variables $X, Y \in L^\infty$ satisfying $X \leq_{\mathrm{cx}} Y$.

2.2 Characterization

In the following, we establish a functional characterization for signed Choquet integrals. As far as we are aware of, this characterization is not known to the literature without assuming monotonicity. We shall first show that a law-invariant, comonotonic-additive and uniformly norm-continuous functional from L^∞ to $\mathbb{R}$ is necessarily a signed Choquet integral, based on a remarkable result of Schmeidler (1986), which we list as Theorem 7 in the appendix for completeness. The converse is also true, but it will be verified later as we establish some further properties of the signed Choquet integrals.

Theorem 1. *A functional $I : L^\infty \to \mathbb{R}$ is law-invariant, comonotonic-additive and uniformly norm-continuous if and only if I is a signed Choquet integral.*

Proof. (i) "$\Rightarrow$": By Theorem 7 (Proposition 2 of Schmeidler (1986)), a comonotonic-additive and norm-continuous functional I has a representation

$$I(X) = \int_{-\infty}^{0} (v(X \geq x) - v(\Omega))\, \mathrm{d}x + \int_{0}^{\infty} v(X \geq x)\, \mathrm{d}x, \quad X \in L^\infty, \tag{4}$$

where the set function $v : \mathcal{A} \to \mathbb{R}$ is given by $v(E) = I(\mathbb{1}_E)$, $E \in \mathcal{A}$. Note that I is law-invariant, which means $I(\mathbb{1}_E) = h(\mathbb{P}(E))$ for some function $h : [0,1] \to \mathbb{R}$. Hence $v(E) = h(\mathbb{P}(E))$ for $E \in \mathcal{A}$, and (4) can be rewritten as

$$I(X) = \int_{-\infty}^{0} (h(\mathbb{P}(X \geq x)) - h(1))\, \mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X \geq x))\, \mathrm{d}x. \tag{5}$$

Next we verify $h \in \mathcal{H}$, so that I is indeed a signed Choquet integral. Noting that comonotonic additivity gives $I(0) + I(0) = I(0) = h(0)$, we have $h(0) = 0$. It remains to verify that h is of bounded variation. Let U be a uniform random variable on $[0,1]$. First, notice that $h(t) = I(\mathbb{1}_{\{U<t\}}) < \infty$ for $t \in [0,1]$. Thus h is finite. As I is uniformly norm-continuous, for a fixed $\epsilon > 0$, there exists a $\delta > 0$ such that $|I(X) - I(Y)| < \epsilon$, whenever $\|X - Y\|_\infty < \delta$. Let $\mathcal{P} = \{t_0, \dots, t_n\}$ be an arbitrary partition of $[0,1]$, where $0 = t_0 < \cdots < t_n = 1$. In the summation $\sum_{i=1}^{n} |h(t_i) - h(t_{i-1})|$, there are exactly n terms of $h(x)$, $x \in \mathcal{P}$ with a positive sign, and n terms of $h(y)$, $y \in \mathcal{P}$ with a negative sign. Therefore, we can write two increasing sequences $\{x_1, \dots, x_n\} \subset \mathcal{P}$ and $\{y_1, \dots, y_n\} \subset \mathcal{P}$ such that

$$\sum_{i=1}^{n} |h(t_i) - h(t_{i-1})| = \sum_{i=1}^{n} h(x_i) - \sum_{i=1}^{n} h(y_i).$$

Since positive and negative terms in the summation $\sum_{i=1}^{n} |h(t_i) - h(t_{i-1})|$ appear in pairs, we have $x_i, y_i \in [t_{i-1}, t_i]$, $i = 1, \dots, n$.

Next, let $f : [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}$ be given by $f(t) = \sum_{i=1}^{n} \mathbb{1}_{\{t > 1 - x_i\}}$ and $g(t) = \sum_{i=1}^{n} \mathbb{1}_{\{t > 1 - y_i\}}$. Let $X = \delta f(U)$ and $Y = \delta g(U)$. Clearly, $\|X - Y\|_\infty \leq \delta$ because $x_i, y_i \in [t_{i-1}, t_i]$ for each $i = 1, \dots, n$. It is straightforward to calculate

$$I(X) = \int_{0}^{\infty} h(\mathbb{P}(\delta f(U) > x))\, \mathrm{d}x = \delta \int_{0}^{\infty} h(\mathbb{P}(f(U) > y))\, \mathrm{d}y = \delta \sum_{i=1}^{n} h(x_i),$$

and similarly $I(Y) = \delta \sum_{i=1}^{n} h(y_i)$. Noting that $\|X - Y\|_\infty < \delta$, we have

$$|I(X) - I(Y)| = \delta \left| \sum_{i=1}^{n} h(x_i) - \sum_{i=1}^{n} h(y_i) \right| < \epsilon.$$

It follows that

$$\sum_{i=1}^{n} |h(t_i) - h(t_{i-1})| < \frac{\epsilon}{\delta} < \infty,$$

and this holds for an arbitrary partition $\mathcal{P} = \{t_0, \dots, t_n\}$. Thus h has bounded variation.

(ii) " $\Leftarrow$": Law-invariance is obvious. The uniform norm-continuity of a signed Choquet integral is verified by Lemma 4 in Section 3. Comonotonic additivity is implied by Lemma 3 below; see Remark 4. □

Remark 2. Theorem 4.2 of Murofushi et al. (1994) characterizes signed Choquet integrals that are not necessarily law-invariant. Comparing the above result with our Theorem 1, our result suggests that this extra law-invariance condition implies the existence of a function h such that $I = I_h$, and h has bounded variation on $[0,1]$. The corresponding condition in Murofushi et al. (1994) is that the set function μ, which is $h \circ \mathbb{P}$ in our paper, has bounded variation on $(\Omega, \mathcal{A})$. In fact, we can verify from the definition of total variation in Murofushi et al. (1994) that the total variation of h on $[0,1]$ is equal to the total variation of $h \circ \mathbb{P}$ on $(\Omega, \mathcal{A})$.

We can also compare Theorem 1 with Theorem 22 of Cerreia-Vioglio et al. (2015) which characterizes signed Choquet integrals on general spaces without law-invariance. In the latter result, a property of functional bounded variation is imposed, instead of the uniform norm-continuity in Theorem 1. Generally, uniform norm-continuity is not sufficient for functional bounded variation used in Cerreia-Vioglio et al. (2012, 2015). The assumption of law-invariance provides extra regularity and continuity for the underlying functional, due to a huge dimension reduction resulting from mapping random variables to their distributions. This phenomenon is well documented in the risk management literature, see e.g. Jouini et al. (2006) for the case of convex risk measures on L^∞ and more recently, Gao et al. (2018) and Gao and Xanthos (2018) for the case of convex risk measures on Orlicz hearts. Generally speaking, assuming the same set of other properties, law-invariant functionals have better regularity conditions than non-law-invariant ones.

Remark 3. From the proof of Theorem 1, we see that, if uniform norm-continuity of I is weakened to norm-continuity, a representation of the form (5) holds with a function h not necessarily of bounded variation (thus, not a signed Choquet integral according to our definition). Indeed, for a positively homogeneous functional, uniform continuity is equivalent to Lipschitz continuity; see also Lemma 4.

2.3 Basic properties

In this and the next few sections, we give several basic properties of signed Choquet integrals which will be useful in Sections 3 and 4. These properties are partially known in the literature (see e.g. De Waegenaere and Wakker (2001) and Acerbi (2002) for special cases), and can be derived from classic properties of increasing Choquet integrals; for the sake of completeness we provide short self-contained proofs in the appendix.

Lemma 1. *For $h_1, h_2 \in \mathcal{H}$, if $h_1(1) = h_2(1)$, then*

$$h_1 \leq h_2 \text{ on } [0,1] \quad \Leftrightarrow \quad I_{h_1} \leq I_{h_2} \text{ on } L^\infty.$$

In particular, $h_1 = h_2$ *holds if and only if* $I_{h_1} = I_{h_2}$ *on* L^∞.

For $h \in \mathcal{H}$, I_h is said to be *increasing* (or *decreasing*) if, for all random variables $X, Y \in L^\infty$, $X \le Y$ implies $I_h(X) \le I_h(Y)$ (or $I_h(X) \ge I_h(Y)$, respectively).

Lemma 2. *For* $h \in \mathcal{H}$,

(i) I_h *is increasing (respectively decreasing) if and only if* h *is increasing (respectively decreasing);*

(ii) for $X \in L^\infty$ *and* $c \in \mathbb{R}$, $I_h(X + c) = I_h(X) + ch(1)$;

(iii) for $X \in L^\infty$ *and* $\lambda > 0$, $I_h(\lambda X) = \lambda I_h(X)$;

(iv) for $X \in L^\infty$, $I_h(-X) = I_{\hat{h}}(X)$, *where* $\hat{h} : [0, 1] \to \mathbb{R}$ *is given by* $\hat{h}(x) = h(1 - x) - h(1)$.

In the context of non-law-invariant comonotonic-additive functionals, similar results to Lemma 2 (i)-(iii) can be found in Proposition 4.11 of Marinacci and Montrucchio (2004).

2.4 Quantile representation

In this section, we present an important property of signed Choquet integrals, namely, the quantile representation. This result will be referred to repeatedly in this paper. In particular, it is used to show the following properties of a signed Choquet integral: comonotonic additivity (Theorem 1 above), convex order consistency (Theorem 2 below), continuity with respect to weak convergence (Theorem 4 below), and extreme-aggregation for heterogeneous portfolios (Theorem 6 below). In the following Lemma, the first two conditions (i) and (ii) for a quantile representation are known for increasing Choquet integrals (see e.g. Denneberg (1994) and Theorems 4 and 6 of Dhaene et al. (2012)). Although (i) and (ii) can be obtained from corresponding results on increasing Choquet integrals via a Jordan decomposition, we give an independent short proof here.

Lemma 3. *For* $h \in \mathcal{H}$ *and* $X \in L^\infty$,

(i) if h *is right-continuous, then* $I_h(X) = \int_0^1 F_X^{-1+}(1 - p)\,\mathrm{d}h(p)$;

(ii) if h *is left-continuous, then* $I_h(X) = \int_0^1 F_X^{-1}(1 - p)\,\mathrm{d}h(p)$;

(iii) if F_X^{-1} *is continuous, then* $I_h(X) = \int_0^1 F_X^{-1}(1 - p)\,\mathrm{d}h(p)$.

Proof. (i) Without loss of generality, we may assume $X \ge 0$, and the general case can be easily obtained via Lemma 2. Noting that h is right-continuous, $h(\mathbb{P}(X > x)) = \int_0^{\mathbb{P}(X>x)} \mathrm{d}h(p)$.

Since h is of bounded variation, one can apply Fubini's theorem to the following Lebesgue-Stieltjes integral,

$$I_h(X) = \int_0^\infty \int_0^{\mathbb{P}(X>x)} \mathrm{d}h(p)\,\mathrm{d}x = \int_0^1 \int_0^{F_X^{-1+}(1-p)} \mathrm{d}x\,\mathrm{d}h(p) = \int_0^1 F_X^{-1+}(1-p)\,\mathrm{d}h(p),$$

where the second equality is due to $p \le \mathbb{P}(X > x) \Leftrightarrow x \le F_X^{-1+}(1-p)$.

(ii) Note that $\hat{h}$ in part (iii) of Lemma 2 is right-continuous, and $F_X^{-1}(p) = -F_{-X}^{-1+}(1-p)$. Applying part (iii) of Lemma 2, we obtain

$$I_h(X) = I_{\hat{h}}(-X) = \int_0^1 F_{-X}^{-1+}(1-p)\,\mathrm{d}\hat{h}(p) = -\int_0^1 F_X^{-1}(p)\,\mathrm{d}h(1-p) = \int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p).$$

(iii) As h can be replaced by its Jordan decomposition $h = h_+ - h_-$, it suffices to show the representation for h increasing. First note that $\int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p)$ is finite, and through integration-by-parts,

$$\int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p) = F_X^{-1}(0)h(1) - \int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p).$$

For $p \in [0,1]$, we have

$$p \in [\mathbb{P}(X > F_X^{-1}(1-p)), \mathbb{P}(X \ge F_X^{-1}(1-p))].$$

Define $g_1^* : \mathbb{R} \to \mathbb{R}$ by

$$g_1^*(x) = \sup\{h(y) \in \mathbb{R} : y \in [\mathbb{P}(X > x), \mathbb{P}(X \ge x)]\}.$$

For $p \in [0,1]$,

$$h(p) \le g_1^*(F_X^{-1}(1-p)) = \sup\{h(y) \in \mathbb{R} : y \in [\mathbb{P}(X > F_X^{-1}(1-p)), \mathbb{P}(X \ge F_X^{-1}(1-p))]\},$$

and therefore,

$$\begin{aligned}\int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p) \le \int_0^1 g_1^*(F_X^{-1}(1-p))\,\mathrm{d}F_X^{-1}(1-p) &= \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} g_1^*(t)\,\mathrm{d}t \\ &= \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X \ge t))\,\mathrm{d}t.\end{aligned}$$

Via a symmetric argument through replacing g_1^* by $g_2^* : \mathbb{R} \to \mathbb{R}$,

$$g_2^*(x) = \inf\{h(y) \in \mathbb{R} : y \in [\mathbb{P}(X > x), \mathbb{P}(X \ge x)]\},$$

we obtain

$$\int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X>x))\,\mathrm{d}x \le \int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p).$$

Note that

$$\int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X>x))\,\mathrm{d}x = \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X\ge x))\,\mathrm{d}x,$$

and therefore we have

$$\int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p) = \int_{F_X^{-1}(1)}^{F_X^{-1}(0)} h(\mathbb{P}(X\ge x))\,\mathrm{d}x.$$

Finally, for $X \in L^\infty$,

$$\begin{aligned} I_h(X) &= \int_{F_X^{-1}(0)}^{F_X^{-1}(1)} h(\mathbb{P}(X\ge x))\,\mathrm{d}x - \int_{F_X^{-1}(0)}^{0} h(1)\,\mathrm{d}x \\ &= -\int_0^1 h(p)\,\mathrm{d}F_X^{-1}(1-p) + F_X^{-1}(0)h(1) \\ &= \int_0^1 F_X^{-1}(1-p)\,\mathrm{d}h(p). \end{aligned}$$

This completes the proof. □

Remark 4. Part (i) of Lemma 3 implies comonotonic additivity of a signed Choquet integral I_h. First, we decompose $h = h_l + h_r$, where h_l and h_r are left-continuous and right-continuous, respectively. This is always possible as h has countably many points of discontinuity. Then, it follows from Lemma 3 that I_{h_l} and I_{h_r} are both comonotonic-additive, as the left- and right-quantiles are comonotonic-additive (a well-known fact; see e.g. Proposition 7.20 of McNeil et al. (2015) for the case of left-quantiles).

2.5 Convexity, convex order consistency, and mixture-concavity

Next we show that convex order consistency of a signed Choquet integral is equivalent to its distortion function being concave. For increasing Choquet integrals, this result is established by Yaari (1987).

Theorem 2. *For random variables $X, Y \in L^\infty$, $X \le_{\mathrm{cx}} Y$ if and only if $I_h(X) \le I_h(Y)$ for all concave functions $h \in \mathcal{H}$.*

Proof. (i) "$\Rightarrow$": Given $X, Y \in L^\infty$ with distributions F and G respectively, let

$$a = \operatorname{ess\,inf}\{X\} \wedge \operatorname{ess\,inf}\{Y\},\ \ b = \operatorname{ess\,sup}\{X\} \vee \operatorname{ess\,sup}\{Y\}$$

and $f = 1 - F,\ g = 1 - G$. If $X \leq_{\rm cx} Y$, by Equation (3.A.7) of Shaked and Shanthikumar (2007),

$$\mathbb{E}[X] = \mathbb{E}[Y] \text{ and } \int_x^\infty \overline{F}(u)\,\mathrm{d}u \leq \int_x^\infty \overline{G}(u)\,\mathrm{d}u \text{ for all } x,$$

which is

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t \quad \text{and} \quad \int_a^x f(t)\,\mathrm{d}t \geq \int_a^x g(t)\,\mathrm{d}t \quad \text{for} \quad a \leq x \leq b.$$

A concave function h defined on $[0,1]$ is necessarily continuous on $(0,1)$. Define $\phi : [0,1] \to \mathbb{R}$ by

$$\phi(x) = \begin{cases} h(x) & \text{for } 0 < x < 1; \\ \lim\limits_{x \downarrow 0} h(x) & \text{for } x = 0; \\ \lim\limits_{x \uparrow 1} h(x) & \text{for } x = 1. \end{cases}$$

Since h has bounded variation, it can be written as the difference of two increasing functions. As the bounded monotone functions have finite limits, $\lim_{x\downarrow 0} h(x)$ and $\lim_{x\uparrow 1} h(x)$ are well defined. Note that ϕ is a continuous concave function, $\phi = h$ on $(0,1)$ and $\phi \geq h$ on $[0,1]$. By the classic Hardy-Littlewood-Pólya inequality (listed as Theorem 8 for the sake of completeness),

$$\int_a^b \phi(f(x))\,\mathrm{d}x \leq \int_a^b \phi(g(x))\,\mathrm{d}x.$$

By Equation (3.A.12) of Shaked and Shanthikumar (2007), $a = \operatorname{ess\,inf}\{Y\}$ and $b = \operatorname{ess\,sup}\{Y\}$, and therefore $h(g(x)) = \phi(g(x))$ for $x \in (a,b)$. Moreover, $h(f(x)) = h(g(x))$ for $x > b$ or $x < a$. Utilizing the above observations, we have

$$\begin{aligned} I_h(X) - I_h(Y) &= \int_a^b \left(h(f(x)) - h(g(x))\right)\,\mathrm{d}x \\ &= \int_a^b \left(h(f(x)) - \phi(g(x))\right)\,\mathrm{d}x \\ &\leq \int_a^b \left(\phi(f(x)) - \phi(g(x))\right)\,\mathrm{d}x \leq 0. \end{aligned}$$

(ii) " $\Leftarrow$": For all $p \in [0,1],\ t \in [0,1]$, let $h(t) = -\mathbb{1}_{\{t \geq 1-p\}}(t - 1 + p)$, and then h is concave and in $\mathcal{H}$. For fixed p, by Lemma 3,

$$I_h(X) = -\int_{1-p}^1 F_X^{-1}(1-t)\,\mathrm{d}t = -\int_0^p F_X^{-1}(u)\,\mathrm{d}u.$$

Thus for all $p \in [0,1]$, $I_h(X) \leq I_h(Y)$ results in

$$\int_0^p F^{-1}(t)\,\mathrm{d}t \geq \int_0^p G^{-1}(t)\,\mathrm{d}t,$$

which implies $X \leq_{\rm cx} Y$ by Theorem 3.A.5 of Shaked and Shanthikumar (2007). □

Remark 5. The forward implication of Theorem 2 can also be deduced by noticing that $\nu = h \circ \mathbb{P}$ defines a submodular game (see Marinacci and Montrucchio (2004)) whenever h is concave. Then an application of Corollary 4.2 of Marinacci and Montrucchio (2004) and Theorem 4.1 of Dana (2005) establishes the claim.

At this point, we are ready to establish six equivalent conditions characterizing the convexity of a signed Choquet integral. For a law-invariant functional ρ on L^∞, define $\tilde{\rho} : \mathcal{M} \to \mathbb{R}$ by $\tilde{\rho}(F) = \rho(X)$ where $X \sim F$, and we say that ρ is *concave on mixtures* if $\tilde{\rho}$ is concave.

Theorem 3. *For $h \in \mathcal{H}$, the following are equivalent: (i) h is concave; (ii) I_h is convex order consistent; (iii) I_h is subadditive; (iv) I_h is convex; (v) I_h is quasi-convex; (vi) I_h is concave on mixtures.*

Proof. We complete the proof in the order $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$: Guaranteed by Theorem 2.

$(ii) \Rightarrow (iii)$: By Theorem 1, I_h is law-invariant and comonotonic-additive. We take random variables X, Y and comonotonic random variables X^c, Y^c whose distribution functions are identical to X, Y, respectively. Then $X + Y \leq_{\rm cx} X^c + Y^c$ as mentioned in Section 2.1. Thus

$$I_h(X+Y) \leq I_h(X^c + Y^c) = I_h(X^c) + I_h(Y^c) = I_h(X) + I_h(Y),$$

and I_h is subadditive.

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

$(iv) \Rightarrow (v)$: Convexity is stronger than quasi-convexity by definition.

$(v) \Rightarrow (vi)$: Take any $x, y \in [0,1]$, $x \leq y$. Define random variables X, Y, Z by

$$\mathbb{P}(X=0) = 1-y,\ \mathbb{P}(X=1/2) = y-x,\ \mathbb{P}(X=1) = x,$$

and the joint distribution function of Y and Z is given by

$$\mathbb{P}(Y=0, Z=0) = 1-y,\ \mathbb{P}(Y=1, Z=0) = \mathbb{P}(Y=0, Z=1) = \frac{y-x}{2},\ \mathbb{P}(Y=1, Z=1) = x.$$

Clearly $X \overset{\mathrm{d}}{=} \frac{1}{2}Y + \frac{1}{2}Z$ and $Y \overset{\mathrm{d}}{=} Z$. Since I_h is quasi-convex and law-invariant, we have

$$I_h(X) = I_h\left(\frac{1}{2}Y + \frac{1}{2}Z\right) \leq \max\{I_h(Y), I_h(Z)\} = I_h(Y).$$

Note that

$$\mathbb{P}(X \geq t) = \begin{cases} 1 & t \leq 0; \\ y & 0 < t \leq \frac{1}{2}; \\ x & \frac{1}{2} < t \leq 1; \\ 0 & t > 1, \end{cases} \quad \text{and} \quad \mathbb{P}(Y \geq t) = \begin{cases} 1 & t \leq 0; \\ \frac{1}{2}x + \frac{1}{2}y & 0 < t \leq 1; \\ 0 & t > 1. \end{cases}$$

As

$$I_h(X) = \int_0^{\frac{1}{2}} h(y)\,\mathrm{d}t + \int_{\frac{1}{2}}^1 h(x)\,\mathrm{d}t = \frac{1}{2}h(x) + \frac{1}{2}h(y),$$

and

$$I_h(Y) = \int_{-\infty}^0 (h(1) - h(1))\,\mathrm{d}t + \int_0^1 h\left(\frac{1}{2}x + \frac{1}{2}y\right)\,\mathrm{d}t + \int_1^\infty h(0)\,\mathrm{d}t = h\left(\frac{1}{2}x + \frac{1}{2}y\right),$$

$I_h(X) \le I_h(Y)$ leads to $\frac{1}{2}h(x)+\frac{1}{2}h(y) \le h\left(\frac{1}{2}x + \frac{1}{2}y\right)$; thus h is mid-point concave. By the Sierpinski Theorem (see page 12 of Donoghue (1969)), a mid-point concave and Lebesgue measurable function is a concave function. Therefore h is concave; (i) holds. With concavity of h, (vi) is straightforward from the definition of Choquet integral in (2).

$(vi) \Rightarrow (i)$: For $p, q, \lambda \in [0, 1]$, let F be a Bernoulli distribution with mean p and G be a Bernoulli distribution with mean q. Then $\lambda F + (1-\lambda)G$ is the Bernoulli distribution with mean $\lambda p + (1-\lambda)q$. It follows from simple calculation that

$$\lambda h(p) + (1-\lambda)h(q) = \lambda \tilde{I}_h(F) + (1-\lambda)\tilde{I}_h(G) \le \tilde{I}_h(\lambda F + (1-\lambda)G) = h(\lambda p + (1-\lambda)q),$$

and thus h is concave. □

Remark 6. Concavity on mixtures (mixture-concavity) is a natural property for risk functionals, especially measures of variability, as it assigns a higher risk value to a mixture of two distributions with equal risk value; see e.g. Acciaio and Svindland (2013). This property is satisfied by classic variability measures, such as the variance, the standard deviation and the Gini deviation; see Section 2.6 below. Although being equivalent for signed Choquet integrals, mixture-concavity is essentially different from convexity for general functionals, in terms of both mathematical and economic interpretations. For instance, taking a supremum over convex signed Choquet integrals preserves convexity and may lose mixture-concavity, whereas taking an infimum over convex signed Choquet integrals preserves mixture-concavity and may lose convexity.

2.6 Some examples

Example 1. We first present some examples of signed Choquet integrals used as measures of distributional variability. Note that all distortion functions below are concave but not monotone.

(i) The range:

$$\mathrm{Range}(X) = \operatorname{ess\,sup}(X) - \operatorname{ess\,inf}(X), \quad X \in L^\infty.$$

The range is a signed Choquet integral with a concave distortion function h given by $h(t) = \mathbb{1}_{\{0<t<1\}},\ t \in [0,1]$. This fact can be checked by straightforward calculation:

$$I_h(X) = \int_{\operatorname{ess\,inf}(X)}^{0} \mathrm{d}x + \int_{0}^{\operatorname{ess\,sup}(X)} \mathrm{d}x = \operatorname{ess\,sup}(X) - \operatorname{ess\,inf}(X) = \mathrm{Range}(X).$$

(ii) The mean median difference:

$$\mathrm{MD}(X) = \min_{x\in\mathbb{R}} \mathbb{E}[|X-x|] = \mathbb{E}\left[\left|X - F_X^{-1}\left(\frac{1}{2}\right)\right|\right], \quad X \in L^\infty.$$

The mean median difference is a signed Choquet integral with a concave distortion function h given by $h(t) = \min\{t, 1-t\},\ t \in [0,1]$. This can be checked using Lemma 3:

$$\begin{aligned} I_h(X) &= \int_{\frac{1}{2}}^{1} F_X^{-1}(u)\,\mathrm{d}u - \int_0^{\frac{1}{2}} F_X^{-1}(u)\,\mathrm{d}u \\ &= \frac{1}{2}F_X^{-1}\left(\frac{1}{2}\right) - \int_0^{\frac{1}{2}} F_X^{-1}(u)\,\mathrm{d}u + \int_{\frac{1}{2}}^{1} F_X^{-1}(u)\,\mathrm{d}u - \frac{1}{2}F_X^{-1}\left(\frac{1}{2}\right) \\ &= \int_0^1 \left|F_X^{-1}(u) - F_X^{-1}\left(\frac{1}{2}\right)\right| \mathrm{d}u = \mathrm{MD}(X). \end{aligned}$$

(iii) The Gini deviation:

$$\mathrm{Gini}(X) = \frac{1}{2}\mathbb{E}[|X_1 - X_2|], \quad X \in L^\infty,\ X_1, X_2, X \text{ are iid.}$$

The Gini deviation is a signed Choquet integral with a concave distortion function h given by $h(t) = t - t^2,\ t \in [0,1]$. This is due to its alternative form (see e.g. Denneberg (1990))

$$\mathrm{Gini}(X) = \int_0^1 F_X^{-1}(t)(2t-1)\,\mathrm{d}t.$$

Example 2. Next we present some examples of signed Choquet integrals used as measures of risk. The first two popular risk measures used in regulation are increasing signed Choquet integrals. The last one does not necessarily have an increasing distortion function.

(i) The Value-at-Risk (VaR) for $p \in (0,1)$:

$$\mathrm{VaR}_p(X) = \inf\{x : \mathbb{P}(X \le x) \ge p\}, \quad X \in L^\infty.$$

VaR_p for $p \in (0,1)$ is a signed Choquet integral with distortion function h given by $h(t) = \mathbb{1}_{\{t>1-p\}},\ t \in [0,1]$. This can be directly checked via Lemma 3:

$$I_h(X) = F_X^{-1}(p) = \mathrm{VaR}_p(X).$$

(ii) The Expected Shortfall (ES) for $p \in (0,1)$:

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_t(X)\,\mathrm{d}t, \quad X \in L^\infty.$$

ES_p for $p \in (0,1)$ is a signed Choquet integral with distortion function h given by $h(t) = \min\{\frac{t}{1-p}, 1\},\ t \in [0,1]$. This can be directly checked via Lemma 3:

$$I_h(X) = \frac{1}{1-p}\int_p^1 F_X^{-1}(t)\,\mathrm{d}t = \mathrm{ES}_p(X).$$

(iii) The Gini Shortfall (GS) for $p \in [0,1)$ and $\lambda \geq 0$:

$$\mathrm{GS}_p^\lambda(X) = \mathrm{ES}_p(X) + \lambda \mathrm{TGini}_p(X), \quad X \in L^\infty,$$

where TGini_p is the tail-Gini functional,

$$\mathrm{TGini}_p(X) = \frac{2}{(1-p)^2}\int_p^1 F_X^{-1}(t)(2t-(1+p))\,\mathrm{d}t, \quad X \in L^\infty.$$

By Theorem 4.1 of Furman et al. (2017), GS_p^λ for $p \in [0,1)$ and $\lambda \geq 0$ is a signed Choquet integral with distortion function h given by

$$h(t) = \frac{1}{(1-p)^2}\left((1-p)t + 4\lambda t\left(1 - \frac{t}{2} - \frac{1+p}{2}\right)\right)\mathbb{1}_{\{t \leq 1-p\}} + \mathbb{1}_{\{t > 1-p\}},\ t \in [0,1].$$

A Gini Shortfall is an increasing Choquet integral if and only if $\lambda \in [0, \frac{1}{2}]$.

Example 3. Below we look at the standard deviation, which is not a signed Choquet integral, but a supremum over some signed Choquet integrals. Let

$$\tilde{\mathcal{H}} = \left\{h \in \mathcal{H} : h(1) = 0,\ \int_0^1 (h'(t))^2\,\mathrm{d}t \leq 1,\ h \text{ is concave}\right\}.$$

The standard deviation, defined as

$$\sigma(X) = \sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}, \quad X \in L^\infty,$$

has the following representation

$$\sigma(X) = \sup_{h \in \tilde{\mathcal{H}}} I_h(X), \quad X \in L^\infty, \tag{6}$$

and hence it is the supremum over a class of signed Choquet integrals.

To show this, let $\mathcal{Z} = \{Z \in L^\infty : \mathbb{E}[Z] = 0,\ \mathbb{E}[Z^2] \leq 1\}$. It is clear that, for $X \in L^\infty$,

$$\sigma(X) = \frac{\mathbb{E}[X(X - \mathbb{E}[X])]}{\sigma(X)} = \mathbb{E}\left[X\frac{X - \mathbb{E}[X]}{\sigma(X)}\right] \leq \sup_{Z \in \mathcal{Z}} \mathbb{E}[XZ],$$

and for any $Z \in \mathcal{Z}$, we have $\mathbb{E}[XZ] = \mathrm{cov}(X, Z) \le \sigma(Z)\sigma(X) \le \sigma(X)$. Therefore, by the Fréchet-Hoeffding inequality (see e.g. Lemma 4.60 of Föllmer and Schied (2016)),

$$\sigma(X) = \sup_{Z \in \mathcal{Z}} \mathbb{E}[XZ] = \sup_{Z \in \mathcal{Z}} \int_0^1 F_Z^{-1}(u) F_X^{-1}(u)\,\mathrm{d}u.$$

Via the relation $h(t) = \int_0^t F_Z^{-1}(1-u)\,\mathrm{d}u,\ t \in [0,1]$, we establish a one-to-one mapping from the distributions of $Z \in \mathcal{Z}$ to functions in $\tilde{\mathcal{H}}$. Therefore, using Lemma 3,

$$\sigma(X) = \sup_{Z \in \mathcal{Z}} \int_0^1 F_Z^{-1}(u) F_X^{-1}(u)\,\mathrm{d}u = \sup_{h \in \tilde{\mathcal{H}}} \int_0^1 F_X^{-1}(1-u)\,\mathrm{d}h(u) = \sup_{h \in \tilde{\mathcal{H}}} I_h(X), \quad X \in L^\infty.$$

Note that $h \in \tilde{\mathcal{H}}$ is concave. As a consequence, each $I_h,\ h \in \tilde{\mathcal{H}}$ is subadditive, convex and consistent with the convex order (see Theorem 3), and so is the standard deviation σ by noting that these properties are preserved when taking a supremum.

Indeed, all law-invariant deviation measures in the sense of Rockafellar et al. (2006) admit a signed Choquet integral representation similar to (6); this result is established in Grechuk et al. (2009).

Example 4. We look at two further examples of measures of distributional variability based on risk measures in Example 2. They will be revisited in Sections 3 and 4.

(i) The inter-quantile range (IQR) for $p \in (1/2, 1)$:

$$\mathrm{IQR}_p(X) = \mathrm{VaR}_p(X) - \mathrm{VaR}_{1-p}(X), \quad X \in L^\infty.$$

The inter-quantile range is a commonly used measure of dispersion in statistics, and the typical choice of p is 0.75, yielding the difference between the first quarter and the third quarter quantiles. IQR_p for $p \in (1/2, 1)$ is a signed Choquet integral with distortion function h given by $h(t) = \mathbb{1}_{\{1-p<t\le p\}}$, $t \in [0,1]$; see Figure 1. Unlike the other measures of variability in Example 1, the distortion function h of IQR_p is not concave, and hence IQR_p is not convex or convex-order consistent by Theorem 3. For $X \in L^\infty$ with a continuous quantile at $1-p$, noting that $F_X^{+}(1-p) = -F_{-X}(p)$ by Lemma 2 (iv), we can alternatively write

$$\mathrm{IQR}_p(X) = \mathrm{VaR}_p(X) + \mathrm{VaR}_p(-X). \tag{7}$$

(ii) The inter-ES range (IER) for $p \in (1/2, 1)$:

$$\mathrm{IER}_p(X) = \frac{1}{1-p}\left(\int_p^1 \mathrm{VaR}_t(X)\,\mathrm{d}t - \int_0^{1-p} \mathrm{VaR}_t(X)\,\mathrm{d}t\right), \quad X \in L^\infty.$$

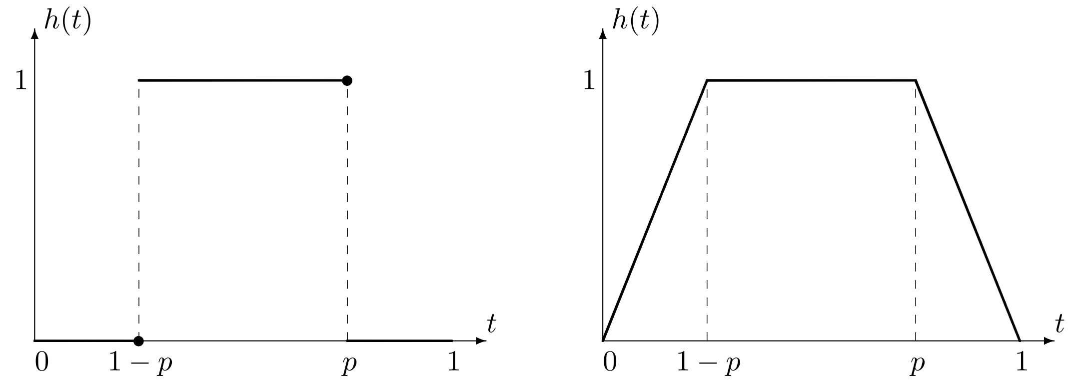

Figure 1: Distortion functions of IQR_p (left) and IER_p (right)

Similarly to (7), we can write, without assuming a continuous quantile,

$$\mathrm{IER}_p(X) = \mathrm{ES}_p(X) + \mathrm{ES}_p(-X), \quad X \in L^\infty.$$

IER_p for $p \in (1/2, 1)$ is a signed Choquet integral with distortion function h given by $h(t) = \min\{\frac{t}{1-p}, 1\} + \min\{\frac{p-t}{1-p}, 0\}$, $t \in [0, 1]$; see Figure 1. In sharp contrast to IQR_p, IER_p has a concave distortion function and hence it is convex and convex-order consistent.

3 Continuity

In this section, we discuss some issues related to continuity of signed Choquet integrals. We first demonstrate the simple fact that a signed Choquet integral is Lipschitz-continuous with respect to L^∞-norm. This result completes the proof of Theorem 1 above, and will be used later in Section 4 to study risk aggregation. This result can be derived (with a small effort) from Proposition 4.11 of Marinacci and Montrucchio (2004) on the continuity of signed Choquet integrals without law-invariance. A simple self-contained proof is put in the appendix.

Lemma 4. *For $h \in \mathcal{H}$ and $X, Y \in L^\infty$,*

$$|I_h(X) - I_h(Y)| \le \mathrm{TV}_h \|X - Y\|_\infty, \tag{8}$$

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

Next we study continuity with respect to convergence in distribution (equivalently, weak convergence in the set of distributions $\mathcal{M}$). In general, a signed Choquet integral is not necessarily continuous with respect to convergence in distribution in L^∞, a well-known property of L-statistics; see Cont et al. (2010) for a discussion on increasing Choquet integrals (termed distortion risk measures) in risk management.

In risk management practice, convergence in distribution is the most common type of convergence, due to the statistical nature of data analysis and simulation studies. This issue is closely related to the notion of *qualitative robustness* of statistical functionals as pioneered by Hampel (1971). It would then be of interest to study under what extra conditions a risk functional can be *robust*, thus continuous with respect to convergence in distribution. This direction of research is explored by Embrechts et al. (2015), Pesenti et al. (2016) and Krätschmer et al. (2017).

The following uniform integrability condition turns out to be relevant. A set $\mathcal{D} \subset L^\infty$ is *h-uniformly integrable* for $h \in \mathcal{H}$, if

$$\lim_{k \downarrow 0} \sup_{X \in \mathcal{D}} \int_0^k \left| F_X^{-1}(1-t) \right| \mathrm{d}h(t) = 0, \tag{9}$$

and

$$\lim_{k \uparrow 1} \sup_{X \in \mathcal{D}} \int_k^1 \left| F_X^{-1}(1-t) \right| \mathrm{d}h(t) = 0. \tag{10}$$

Note that if $h \in \mathcal{H}$ is linear and non-constant in some neighborhoods of 0 and 1, then h-uniform integrability reduces to the usual uniform integrability.

Theorem 4. *For $h \in \mathcal{H}$ and $X, X_1, X_2, \dots \in L^\infty$, assume that $X_n \to X$ in distribution as $n \to \infty$ and $\{X, X_1, X_2, \dots\}$ is h-uniformly integrable. If (i) h is continuous, or (ii) X has a continuous inverse distribution function, then $I_h(X_n) \to I_h(X)$ as $n \to \infty$.*

Proof. For $n \in \mathbb{N}$, let F_n and F be the distribution functions of X_n and X, respectively.

We first assume (i). By Lemma 3, we have

$$I_h(X_n) = \int_0^1 F_n^{-1}(1-p)\,\mathrm{d}h(p) \quad \text{and} \quad I_h(X) = \int_0^1 F^{-1}(1-p)\,\mathrm{d}h(p). \tag{11}$$

As h can be replaced by its Jordan decomposition $h = h_+ - h_-$, it suffices to show the statement for an increasing and continuous h. The increasing function h induces a finite Borel measure μ on $[0,1]$ via $\mu([0,x]) = h(x)$, $x \in [0,1]$. Since $F_n^{-1} \to F^{-1}$ as $n \to \infty$ almost everywhere on $\mathbb{R}$ and h is continuous, the convergence is also μ-almost surely. Moveover, the h-uniform integrability of $\{X_i\}_{i \in \mathbb{N}}$ implies that $\{F_n^{-1}\}_{n \in \mathbb{N}}$ is uniformly integrable with respect to the measure μ. Therefore, using Vitali's Convergence Theorem (Rudin (1987, p. 133)), we have $I_h(X_n) \to I_h(X)$ as $n \to \infty$.

Next we assume (ii). In this case the convergence $F_n^{-1} \to F^{-1}$ is point-wise on $(0,1)$. Suppose for the moment that h is left-continuous. By Lemma 3, (11) holds. Similarly to the case above, we assume that h is increasing, and it induces a finite Borel measure μ on $[0,1]$ via $\mu([0,x)) = h(x)$, $x \in [0,1]$. Note that the h-uniform integrability of $\{X, X_1, X_2, \dots\}$ implies that, if $\mu(\{0\}) > 0$, then $F_n^{-1}(1) \to 0$ and $F^{-1}(1) = 0$. Analogously, if $\mu(\{1\}) > 0$, then $F_n^{-1}(0) \to 0$ and $F^{-1}(0) = 0$. Combining the above facts, $F_n^{-1} \to F^{-1}$ μ-almost surely as $n \to \infty$. Using Vitali's Convergence Theorem, we have $I_h(X_n) \to I_h(X)$ as $n \to \infty$.

If h is right-continuous, define the Borel measure μ on $[0,1]$ via $\mu([0,x]) = h(x)$, $x \in [0,1]$ and use the representation in Lemma 3 (i). The conclusion follows analogously.

Finally, for a general h, we decompose $h = h_l + h_r$, where h_l and h_r are left-continuous and right-continuous, respectively. Then we have

$$|I_h(X_n) - I_h(X)| \le |I_{h_l}(X_n) - I_{h_l}(X)| + |I_{h_r}(X_n) - I_{h_r}(X)| \to 0, \quad \text{as } n \to \infty.$$

The proof is complete. □

Next we present a condition on h which implies the h-uniform integrability for all random variables. We say that $h \in \mathcal{H}$ is *flat in neighborhoods of* 0 *and* 1 if it satisfies the following condition: if there exists some $\delta > 0$ such that for all $0 < \epsilon < \delta$, $h(\epsilon) = h(0)$ and $h(1-\epsilon) = h(1)$. In this case, clearly, any set of random variables is h-uniformly integrable. This condition is satisfied if, for instance, I_h is a finite linear combination of some quantile functionals.

Corollary 1. *For $h \in \mathcal{H}$ and $X, X_1, X_2, \dots \in L^\infty$, assume that $X_n \to X$ in distribution as $n \to \infty$ and h is flat in neighborhoods of* 0 *and* 1*. If (i) h is continuous, or (ii) X has a continuous inverse distribution function, then $I_h(X_n) \to I_h(X)$ as $n \to \infty$.*

Robustness properties of the two popular classes of risk measures VaR and ES (defined in Section 2.6) are well-studied in the literature. With respect to convergence in distribution, it is known that VaR_p is continuous at random variables with a continuous quantile function, and ES_p is continuous at random variables among a uniformly integrable set. These are special cases of Theorem 4 and Corollary 1.

Theorem 4 and Corollary 1 generalize Theorem 1 of Cont et al. (2010) for distortion risk measures, Theorem 2.5 of Embrechts et al. (2015) on robustness in the set of risk aggregation, and Theorem 3.5 of Pesenti et al. (2016) for finite-valued convex risk measures. Moreover, different from the settings of Krätschmer et al. (2014, 2017) and Pesenti et al. (2016), our results do not rely on any convexity assumptions.

Example 5 (Continuity of measures of distributional variability)**.** As mentioned above, continuity of risk measures with respect to convergence in distribution are well studied in the recent risk management literature. Below, we apply Theorem 4 and Corollary 1 to the measures of variability in Examples 1, 3 and 4.

(i) The range is generally not continuous with respect to convergence in distribution. Note that for the distortion function h of the range, given by $h(t) = \mathbb{1}_{\{0<t<1\}}$, $t \in [0,1]$, the h-uniform integrality condition in (9) and (10) implies $X = 0$ a.s. for $X \in \mathcal{D}$, which is very restrictive.

(ii) The mean median difference has a continuous distortion function h given by $h(t) = \min\{t, 1-t\}$, $t \in [0,1]$. Since h is linear in neighbourhoods of 0 and 1, the h-uniform integrability is equivalent to the usual uniform integrability. Hence, by Theorem 4, the mean median difference is continuous with respect to convergence in distribution over any uniformly integrable set.

(iii) The Gini deviation has a continuous distortion function h given by $h(t) = t - t^2$, $t \in [0,1]$. For this distortion function h, as it has non-zero (one-sided) derivatives at 0 and 1, the h-uniform integrability is equivalent to the usual uniform integrability. Therefore, by Theorem 4, the Gini deviation is also continuous with respect to convergence in distribution over any uniformly integrable set.

(iv) The inter-quantile range for $p \in (1/2, 1)$ has a distortion function h given by $h(t) = \mathbb{1}_{\{1-p<t\le p\}}$, $t \in [0,1]$. Note that h is flat in neighborhoods of 0 and 1, but it is not continuous. Hence, by Corollary 1, the inter-quantile range is continuous with respect to convergence in distribution over any set of random variables with continuous quantile functions.

(v) The inter-ES range for $p \in (1/2, 1)$ has a distortion function h given by $h(t) = \min\{\frac{t}{1-p}, 1\} + \min\{\frac{p-t}{1-p}, 0\}$, $t \in [0,1]$. Since h is linear in neighbourhoods of 0 and 1, the h-uniform integrability is equivalent to the usual uniform integrability. Hence, by Theorem 4, the inter-ES range is continuous with respect to convergence in distribution over any uniformly integrable set.

(vi) The standard deviation is continuous with respect to convergence in distribution over any uniformly square-integrable set. One can show this statement by applying Vitali's Convergence Theorem to the first and second moments. Theorem 4 does not directly lead to this statement. Nevertheless, in Example 3 we have seen $\sigma(X) = \sup_{h\in\tilde{\mathcal{H}}} I_h(X)$, $X \in L^\infty$. By Hölder's inequality, uniform square-integrability implies h-uniform integrability for each $h \in \tilde{\mathcal{H}}$. Hence, Theorem 4 implies that I_h is continuous with respect to convergence in distribution over any uniformly square-integrable set.

Remark 7. The continuity results of signed Choquet integrals in Theorem 4 also hold on a set larger than L^∞, as long as I_h is well-defined on the corresponding set. In this paper, due to the limitation of space, we focus on random variables in L^∞. For robustness properties of risk functionals defined on Orlicz hearts, we refer to Krätschmer et al. (2014, 2017).

4 Risk aggregation under uncertainty

In the literature of risk management, risk aggregation concerns quantities related to the sum $S = X_1 + \cdots + X_n$ (e.g. the distribution or a risk measure of S) of a risk vector $(X_1, \ldots, X_n)$ representing random losses from a certain portfolio. A currently popular direction of research is *risk aggregation with dependence uncertainty*, where for each $i = 1, \ldots, n$, the marginal distribution F_i of X_i, is known while the joint distribution of $(X_1, ..., X_n)$ remains unspecified. We refer to Embrechts et al. (2013, 2014) and Wang et al. (2013) for the case of the risk measure VaR (defined in Section 2.6), Bernard et al. (2017a,b) for some recent development, and Section 8.4 of McNeil et al. (2015) for a general discussion. As the precise distribution of S is unknown, one typically studies the worst-case value of the aggregate risk evaluated by a risk measure ρ, that is,

$$\sup\{\rho(X_1 + \cdots + X_n) : X_i \sim F_i,\ i = 1, ..., n\}. \tag{12}$$

Another important quantity related to portfolio diversification is the worst-case diversification ratio, defined as

$$\sup\left\{\frac{\rho(X_1 + \cdots + X_n)}{\rho(X_1) + \cdots + \rho(X_n)} : X_i \sim F_i,\ i = 1, ..., n\right\}. \tag{13}$$

If the functional ρ is not convex, the quantities in (12)-(13) are generally difficult to analytically compute. In this section we investigate them for signed Choquet integrals.

4.1 Homogeneous portfolios and the extreme-aggregation measure

To investigate the asymptotic behaviour of the values in (12)-(13) for homogeneous portfolios, Wang et al. (2015) introduced the extreme-aggregation measure as follows. Denote the set of possible sums of n F-distributed random variables by $\mathcal{S}_n(F) = \{X_1 + \cdots + X_n : X_i \sim F,\ i = 1, ..., n\},\ n \in \mathbb{N}$.

Definition 2. The *extreme-aggregation measure* Γ_ρ *induced by a law-invariant functional* $\rho : L^\infty \to \mathbb{R}$ is defined as

$$\Gamma_\rho : L^\infty \to (-\infty, \infty],\ \Gamma_\rho(X) = \limsup_{n\to\infty} \left\{ \frac{1}{n} \sup\{\rho(S) : S \in \mathcal{S}_n(F_X)\} \right\}.$$

Γ_ρ provides a limit of (12)-(13) for homogenous portfolios. The VaR-ES relation $\Gamma_{\mathrm{VaR}_p} = \mathrm{ES}_p$ for $p \in (0,1)$ is shown in Wang and Wang (2015) via direct construction; some first special cases of this relation are established by Puccetti and Rüschendorf (2014). Generalizations to inhomogeneous portfolios are given in Embrechts et al. (2015) (VaR and ES) and Cai et al. (2018) (distortion risk measures and convex risk measures). For a distortion risk measure (equivalently, an increasing Choquet integral) ρ, Wang et al. (2015) obtained an explicit expression for Γ_ρ, which is the smallest subadditive distortion risk measure dominating ρ. The proof used in Wang et al. (2015) is based on analyzing the precise form of h, which requires a lot of delicate analysis and random variable construction. Below we give a much more concise proof, generalizing the characterization of Γ_ρ to signed Choquet integrals.

To present our main result, for $h \in \mathcal{H}$, define its *concave envelope*

$$h^*(t) = \inf\{g(t) : g \text{ is a concave function on } [0,1] \text{ and } g \geq h\}, \quad t \in [0,1]. \tag{14}$$

Note that calculating h^* for a given $h \in \mathcal{H}$ is equivalent to finding the convex hull of the set $\{(x,y) \in [0,1] \times \mathbb{R} : h(x) \geq y\}$.

It is clear that h^* is concave as it is an infimum of concave functions. Further, $h^*(0) = h(0) = 0$ and $h^*(1) = h(1)$; to see this, as $h \in \mathcal{H}$ is bounded, we can define a concave function $g : [0,1] \to \mathbb{R}$ as

$$g(t) = \begin{cases} 0 & t = 0 \\ \sup_{t\in[0,1]} h(t) & 0 < t < 1 \\ h(1) & t = 1 \end{cases}$$

Then one has $0 = h(0) \leq h^*(0) \leq g(0) = 0$ and $h(1) \leq h^*(1) \leq g(1) = h(1)$.

Our main result is the following theorem, which generalizes Theorem 3.2 of Wang et al. (2015) for increasing Choquet integrals.

Theorem 5. *For $h \in \mathcal{H}$, the extreme-aggregation measure induced by I_h is I_{h^*}, and it is the smallest law-invariant convex functional on L^∞ dominating I_h.*

The key to our proof of Theorem 5 is to show that the law-invariant functional Γ_{I_h} is comonotonic-additive and uniformly norm-continuous, and from there we can rely on Theorem 1 to justify that it is a signed Choquet integral. We first demonstrate some useful facts built on several results in Mao and Wang (2015). Denote for a distribution function $F \in \mathcal{M}$,

$$\mathcal{B}_n(F) = \left\{ \frac{1}{n}(X_1 + \cdots + X_n) : X_i \sim F,\ i = 1, ..., n \right\},$$

and

$$\mathcal{C}(F) = \{X : X \leq_{\text{cx}} Y,\ \text{where } Y \sim F\}.$$

Lemma 5. *For $h \in \mathcal{H}$, the following statements hold.*

(i) For $F \in \mathcal{M}$,

$$\limsup_{n\to\infty} \{\sup\{I_h(X) : X \in \mathcal{B}_n(F)\}\} = \sup\{I_h(X) : X \in \mathcal{C}(F)\}. \tag{15}$$

(ii) The functional on L^∞, $X \mapsto \sup_{Y\in\mathcal{C}(F_X)} I_h(Y)$ is comonotonic-additive and convex order consistent.

Proof. (i) Lemma 3.4 of Mao and Wang (2015) states $\mathcal{B}_n(F) \subset \mathcal{C}(F)$ for $n \in \mathbb{N}$. Therefore

$$\limsup_{n\to\infty} \{\sup\{I_h(T) : T \in \mathcal{B}_n(F)\}\} \leq \sup\{I_h(T) : T \in \mathcal{C}(F)\}. \tag{16}$$

On the other hand, by Proposition 3.6 of Mao and Wang (2015), $\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)}^* = \mathcal{C}(F)$, where $\overline{B}^*$ is the L^∞-closure of a set B. It follows that, for each $Y \in \mathcal{C}(F)$, $\epsilon > 0$ and $m \in \mathbb{N}$, there exists $k \in \mathbb{N}$ and $X_k \in \mathcal{B}_k(F)$ such that $||X_k - Y||_\infty < \epsilon$. Hence, by Lemma 4,

$$\sup\{I_h(X) : X \in \mathcal{B}_k(F)\} \geq \sup\{I_h(X) : X \in \mathcal{C}(F)\} - \epsilon \text{TV}_h.$$

Therefore,

$$\limsup_{n\to\infty} \{\sup\{I_h(X) : X \in \mathcal{B}_n(F)\}\} \geq \sup\{I_h(X) : X \in \mathcal{C}(F)\} - \epsilon \text{TV}_h.$$

As ϵ is arbitrary, we obtain

$$\limsup_{n\to\infty} \{\sup\{I_h(X) : X \in \mathcal{B}_n(F)\}\} \geq \sup\{I_h(X) : X \in \mathcal{C}(F)\}. \tag{17}$$

Combining (16)-(17), we obtain (15).

(ii) Corollary 4.3 of Mao and Wang (2015) states that the functional $X \mapsto \sup_{Y\in\mathcal{C}(F_X)} \rho(Y)$ is comonotonic-additive and convex order consistent if ρ is comonotonic-additive, which is the case if $\rho = I_h$. □

Proof of Theorem 5. For any $X \in L^\infty$, by positive homogeneity of I_h and the definition of $\mathcal{B}_n(F_X)$, we have

$$\begin{aligned}\Gamma_{I_h}(X) &= \limsup_{n\to\infty}\left\{\sup\left\{I_h\left(\frac{1}{n}S\right) : S \in \mathcal{S}_n(F_X)\right\}\right\}\\ &= \limsup_{n\to\infty}\left\{\sup\{I_h(T) : T \in \mathcal{B}_n(F_X)\}\right\}.\end{aligned}$$

Applying Lemma 5 (i), it is

$$\Gamma_{I_h}(X) = \sup\{I_h(T) : T \in \mathcal{C}(F_X)\}, \quad X \in L^\infty.$$

By Lemma 5 (ii), Γ_{I_h} is comonotonic-additive and convex order consistent.

Next we verify that Γ_{I_h} is uniformly norm-continuous. Fix $n \in \mathbb{N}$. For any $S \in \mathcal{S}_n(F_X)$, write $S = X_1 + \cdots + X_n$, where $X_i \sim F_X,\ i = 1, \dots, n$. Let $U_1, \dots, U_n$ be uniform random variables on $[0,1]$ such that $F_X^{-1}(U_i) = X_i$ almost surely, $i = 1, \dots, n$. The existence of such $U_1, \dots, U_n$ is given by, for instance, Lemma A.32 of Föllmer and Schied (2016). Let $Z = F_Y^{-1}(U_1) + \cdots + F_Y^{-1}(U_n)$. Clearly, $Z \in \mathcal{S}_n(F_Y)$. By Lemma 4,

$$\begin{aligned}&I_h(S) - \sup\{I_h(T) : T \in \mathcal{S}_n(F_Y)\}\\ &\le I_h(S) - I_h(Z)\\ &\le \mathrm{TV}_h\|S - Z\|_\infty\\ &= \mathrm{TV}_h\left\|\left(F_X^{-1}(U_1) + \cdots + F_X^{-1}(U_n)\right) - \left(F_Y^{-1}(U_1) + \cdots + F_Y^{-1}(U_n)\right)\right\|_\infty\\ &\le n\mathrm{TV}_h\|X - Y\|_\infty,\end{aligned}$$

where the last inequality is due to the well-known fact that $\|F_X^{-1}(U_1) - F_Y^{-1}(U_1)\|_\infty \le \|X - Y\|_\infty$ (see e.g. Lemma 8.2 of Bickel and Freedman (1981)). It follows from taking a supremum over $S \in \mathcal{S}_n(F_X)$ that

$$\frac{1}{n}\sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\} \le \frac{1}{n}\sup\{I_h(T) : T \in \mathcal{S}_n(F_Y)\} + \mathrm{TV}_h\|X - Y\|_\infty.$$

Therefore,

$$\Gamma_{I_h}(X) \le \Gamma_{I_h}(Y) + \mathrm{TV}_h\|X - Y\|_\infty.$$

By symmetry, we have $|\Gamma_{I_h}(X) - \Gamma_{I_h}(Y)| \le \mathrm{TV}_h\|X - Y\|_\infty$; thus Γ_{I_h} is uniformly norm-continuous.

At this point, we know that the law-invariant functional Γ_{I_h} is norm-continuous, comonotonic-additive and convex order consistent. By Theorem 1, there exists $g \in \mathcal{H}$ such that Γ_{I_h} is identified with a signed Choquet integral $I_g = \Gamma_{I_h}$. Note that

$$I_g(X) = \sup_{T \in \mathcal{C}(F_X)} I_h(T) \ge I_h(X),$$

and therefore $g \geq h$ by Lemma 1. Since Γ_{I_h} is convex order consistent, g is concave by Theorem 3. From the definition of h^*, $h^* \leq g$, and this implies $I_{h^*} \leq I_g$ by Lemma 1 again.

On the other hand, $h^* \geq h$, and hence $I_{h^*} \geq I_h$. Noting that h^* is concave and thus I_{h^*} is also convex order consistent, we have

$$I_{h^*}(X) = \sup_{T \in \mathcal{C}(F_X)} I_{h^*}(T) \geq \sup_{T \in \mathcal{C}(F_X)} I_h(T) = I_g(X)$$

Therefore, we conclude that $I_{h^*} = I_g = \Gamma_{I_h}$.

Finally we show that I_{h^*} is the smallest law-invariant convex functional on L^∞ dominating I_h. Suppose that $I : L^\infty \to \mathbb{R}$ is a law-invariant convex functional and $I \geq I_h$. For any $n \in \mathbb{N}$ and $X \in L^\infty$,

$$\begin{aligned}
\sup\{I_h(T) : T \in \mathcal{B}_n(F_X)\} &\leq \sup\{I(T) : T \in \mathcal{B}_n(F_X)\} \\
&\leq \sup\left\{\frac{1}{n}I(X_1) + \cdots + \frac{1}{n}I(X_n) : X_i \sim F_X, 1 \leq i \leq n\right\} \\
&= \frac{1}{n}nI(X) = I(X).
\end{aligned}$$

By taking a limit on both sides of the above equation, we conclude that $I_{h^*} \leq I$. Thus, I_{h^*} is the smallest law-invariant convex functional dominating I_h. □

Example 6. Theorem 5 implies two well-known facts in the literature of risk measures on the relation between VaR_p and ES_p for $p \in (0, 1)$: First, the worst-case aggregation of VaR_p is asymptotically equivalent to that of ES_p (Corollary 3.7 of Wang and Wang (2015)). Second, ES_p is the smallest law-invariant convex risk measure dominating VaR_p (Theorem 9 of Kusuoka (2001); see also Theorem 4.67 of Föllmer and Schied (2016)). Theorem 5 generalizes these results to all signed Choquet integrals, and our approach is different from those in the literature.

Example 7 (The inter-quantile range and inter-ES range)**.** In Examples 4 and 5 we have already seen many differences between the two measures of variability IQR_p and IER_p in terms of convexity, convex-order consistency, and continuity. Next, we will see, by applying Theorem 5, an interesting connection between the two signed Choquet integrals. Recall that for $p \in (1/2, 1)$, IQR_p has a (non-concave) distortion function h given by $h(t) = \mathbb{1}_{\{1-p<t\leq p\}}$, $t \in [0, 1]$. It is straightforward that the smallest concave function h^* dominating h is given by $h^*(t) = \min\{\frac{t}{1-p}, 1\} + \min\{\frac{p-t}{1-p}, 0\}$, $t \in [0, 1]$, which is the distortion function of IER_p; see Figure 1 for these distortion functions. Therefore, for $I_h = \mathrm{IQR}_p$, we have $I_{h^*} = \mathrm{IER}_p$. By Theorem 5, IER_p is the extreme-aggregation measure of IQR_p; in other words, the worst-case value of the aggregate risk evaluated by IQR_p is asymptotically

equivalent to that evaluated by IER_p. This relationship will be further illustrated by the numerical example in Section 4.3.

Remark 8. Wang et al. (2015) gives a few conditions for the upper limit in the definition of Γ_ρ to be replaced by a supremum or a limit. If ρ is a positively homogeneous functional, then the upper limit can be replaced by a supremum. Furthermore, for $h \in \mathcal{H}$ and $X \in L^\infty$, the upper limit can be replaced by either a limit or a supremum, namely

$$\Gamma_{I_h}(X) = \lim_{n\to\infty} \left\{ \frac{1}{n} \sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\} \right\} = \sup_{n\in\mathbb{N}} \left\{ \frac{1}{n} \sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\} \right\}.$$

We conclude this section by showing that the result in Theorem 5 can be generalized to suprema over a set of signed Choquet integrals. Indeed, suprema over a set of signed Choquet integrals represent all continuous law-invariant coherent risk measures and deviation measures, as established in Kusuoka (2001) and Grechuk et al. (2009), respectively.

Corollary 2. *Define a functional $\rho = \sup_{h\in\mathcal{H}_0} I_h$, where $\mathcal{H}_0 \subset \mathcal{H}$. Then $\Gamma_\rho = \sup_{h\in\mathcal{H}_0} I_{h^*}$ and Γ_ρ is the smallest law-invariant convex functional on L^∞ dominating ρ.*

Proof. For $X \in L^\infty$, as ρ is positively homogeneous, we can write

$$\Gamma_\rho(X) = \sup_{n\in\mathbb{N}}\{\sup\{\rho(T) : T \in \mathcal{B}_n(F_X)\}\};$$

see Remark 8. By exchanging the order of suprema, we have

$$\Gamma_\rho(X) = \sup_{n\in\mathbb{N}} \sup_{T\in\mathcal{B}_n(F_X)} \sup_{h\in\mathcal{H}_0} I_h(T) = \sup_{h\in\mathcal{H}_0} \sup_{n\in\mathbb{N}} \sup_{T\in\mathcal{B}_n(F_X)} I_h(T) = \sup_{h\in\mathcal{H}_0} \Gamma_{I_h}(X) = \sup_{h\in\mathcal{H}_0} I_{h^*}(X),$$

where the last equality is due to Theorem 5. The last statement of the corollary can be obtained via an argument analogous to the last part of the proof of Theorem 5. □

4.2 Heterogeneous portfolios

For a sequence of distribution functions $\{F_i\}_{i\in\mathbb{N}} \subset \mathcal{M}$, denote the set of possible sums of n random variables with respective distributions by

$$\mathcal{S}_n(F_1, \dots, F_n) = \{X_1 + \cdots + X_n : X_i \sim F_i,\ i = 1, ..., n\}, \quad n \in \mathbb{N}.$$

To investigate risk aggregation for heterogeneous portfolios, we study an asymptotic equivalence of the following type,

$$\lim_{n\to\infty} \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}}{\sup\{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \dots, F_n)\}} = 1. \tag{18}$$

The asymptotic equivalence (18) is established in Theorem 3.5 of Cai et al. (2018) for increasing Choquet integrals under some regularity conditions. Interpreting (18), to evaluate large portfolios with dependence uncertainty via a non-convex functional I_h, one can replace I_h by a convex functional I_{h^*}, the extreme-aggregation measure induced by I_h, which is much easier to calculate due to its comonotonic-additivity and subadditivity. It is clear that if $F_1 = F_2 = \cdots = F_X$, then (18) reads as

$$\lim_{n\to\infty} \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_X)\}}{nI_{h^*}(X)} = 1,$$

which is precisely Theorem 5 if $I_{h^*}(X) \neq 0$. For the same relation to hold for heterogeneous portfolios, one needs some regularity conditions on $\{F_i\}_{i\in\mathbb{N}}$ and $h \in \mathcal{H}$.

Condition C1 (non-vanishing). $\lim_{n\to\infty} |\sum_{i=1}^n I_{h^*}(X_i)| = \infty$, where $X_i \sim F_i$, $i \in \mathbb{N}$.

Condition C2 (bounded ranges). $\sup_{i\in\mathbb{N}}\{F_i^{-1}(1) - F_i^{-1}(0)\} < \infty$.

Sections 2.2 and 3.3 of Cai et al. (2018) contain counter-examples where (18) fails to hold without some regularity conditions. Next we present the asymptotic equivalence for signed Choquet integrals with a continuous distortion function h under Conditions C1-C2.

Theorem 6. *For a continuous $h \in \mathcal{H}$ and $\{F_i\}_{i\in\mathbb{N}} \subset \mathcal{M}$ satisfying Conditions C1-C2, we have*

$$\lim_{n\to\infty} \frac{\sup\{I_h(S) : S \in \mathcal{S}_n(F_1,\dots,F_n)\}}{\sup\{I_{h^*}(S) : S \in \mathcal{S}_n(F_1,\dots,F_n)\}} = 1. \tag{19}$$

Proof. Our proof is similar to that of Theorem 3.5 of Cai et al. (2018), although the conditions in the latter result are different from Conditions C1-C2. Since h is continuous, we directly work with the quantile representation in Lemma 3 (ii). By Lemma 5.1 of Brighi and Chipot (1994), there exist disjoint open intervals (a_k, b_k), $k \in K \subset \mathbb{N}$ on which $h \neq h^*$, and h^* is linear on each of $[a_k, b_k]$, $k \in K$. Define $A_k = (1 - b_k, 1 - a_k)$, $k \in K$. Let U, V be independent U[0, 1] random variables, and

$$S_n^c = F_1^{-1}(U) + \cdots + F_n^{-1}(U),$$

and

$$R_n = \begin{cases} F_1^{-1}(U) + \cdots + F_n^{-1}(U), & \text{if } U \notin \cup_{k\in K} A_k, \\ \mathbb{E}\left[F_1^{-1}(U) + \cdots + F_n^{-1}(U) \mid U \in A_k\right], & \text{if } U \in A_k,\ k \in K. \end{cases}$$

Clearly, $F_i^{-1}(U) \sim F_i$, $i = 1, \dots, n$, and hence $S_n^c \in \mathcal{S}_n(F_1, \dots, F_n)$. Since

$$\mathbb{E}\left[F_i^{-1}(U) \mid U \in A_k\right] = \frac{\int_{(a_k,b_k)} F_i^{-1}(1-t)\mathrm{d}t}{b_k - a_k} \quad \text{and} \quad F_{S_n^c}^{-1}(t) = \sum_{i=1}^n F_i^{-1}(t) \quad \text{for } t \in (0,1),$$

we have

$$\begin{aligned}&\int_{(a_k,b_k)} F^{-1}_{S^c_n}(1-t)\mathrm{d}h^*(t) - \int_{(a_k,b_k)} F^{-1}_{R_n}(1-t)\mathrm{d}h^*(t)\\&= \frac{h^*(b_k)-h^*(a_k)}{b_k-a_k}\sum_{i=1}^n \int_{(a_k,b_k)} F_i^{-1}(1-t)\mathrm{d}t - \sum_{i=1}^n \frac{\int_{(a_k,b_k)} F_i^{-1}(1-t)\mathrm{d}t}{b_k-a_k}\int_{(a_k,b_k)} \mathrm{d}h^*(t) = 0.\end{aligned}$$

It follows that

$$\begin{aligned}I_{h^*}(S^c_n) - I_{h^*}(R_n) &= \int_0^1 F^{-1}_{S^c_n}(1-t)\mathrm{d}h^*(t) - \int_0^1 F^{-1}_{R_n}(1-t)\mathrm{d}h^*(t)\\&= \sum_{k\in K}\left[\int_{a_k}^{b_k} F^{-1}_{S^c_n}(1-t)\mathrm{d}h^*(t) - \int_{a_k}^{b_k} F^{-1}_{R_n}(1-t)\mathrm{d}h^*(t)\right] = 0. \qquad (20)\end{aligned}$$

By Corollary A.3 of Embrechts et al. (2015), for each k, we can find random variables $Y_{1k},\dots,Y_{nk}$, independent of U (guaranteed by the existence of V), such that Y_{ik} is identically distributed as $F_i^{-1}(U)|U\in A_k$, $i=1,\dots,n$, and

$$\left|Y_{1k}+\cdots+Y_{nk} - \mathbb{E}\left[F_1^{-1}(U)+\cdots+F_n^{-1}(U) \mid U\in A_k\right]\right| \le \max_{i=1,\dots,n}\{F_i^{-1}(1-a_k)-F_i^{-1}(1-b_k)\}.$$

Let $X_i^* = F_i^{-1}(U)1_{\{U\notin\cup_{k\in K}A_k\}} + \sum_{k\in K} Y_{ik}1_{\{U\in A_k\}}$, $i=1,\dots,n$. It is easy to check that $X_i^*\sim F_i$, $i=1,\dots,n$. Denote by $S_n^* = X_1^*+\cdots+X_n^*$. Clearly, $S_n^*\in\mathcal{S}_n(F_1,\dots,F_n)$ and by definition

$$|R_n - S_n^*| \le \sup_{k\in K}\max_{i=1,\dots,n}\{F_i^{-1}(1-a_k)-F_i^{-1}(1-b_k)\} \le M,$$

where $M=\sup_{i\in\mathbb{N}}\{F_i^{-1}(1)-F_i^{-1}(0)\}$ and $M<\infty$ by Condition C2. Therefore, by Lemma 4, we have

$$|I_h(R_n) - I_h(S_n^*)| \le \mathrm{TV}_h \times M. \qquad (21)$$

Integration by parts yields

$$\begin{aligned}I_{h^*}(R_n) - I_h(R_n) &= \int_0^1 F^{-1}_{R_n}(1-t)\mathrm{d}h^*(t) - \int_0^1 F^{-1}_{R_n}(1-t)\mathrm{d}h(t)\\&= \int_0^1 (h(t)-h^*(t))\,\mathrm{d}F^{-1}_{R_n}(1-t)\\&= \sum_{k\in K}\int_{(a_k,b_k)} (h(t)-h^*(t))\,\mathrm{d}F^{-1}_{R_n}(1-t) = 0, \qquad (22)\end{aligned}$$

where the last equality follows as $F^{-1}_{R_n}(1-t)$ is constant for t in each (a_k,b_k). Combining (20)-(22), we have

$$\begin{aligned}|I_{h^*}(S^c_n) - I_h(S^*_n)| &= |(I_{h^*}(S^c_n) - I_{h^*}(R_n)) + (I_{h^*}(R_n) - I_h(R_n)) + (I_h(R_n) - I_h(S^*_n))|\\&= |0+0+(I_h(R_n)-I_h(S^*_n))| = \mathrm{TV}_h\times M. \qquad (23)\end{aligned}$$

Since $h \le h^*$, we have

$$\sup \{I_h(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\} \le \sup \{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\} = I_{h^*}(S_n^c)$$

and hence (23) implies $|I_{h^*}(S_n^c) - \sup \{I_h(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\}| \le \mathrm{TV}_h \times M$. By Condition C1, $\lim_{n\to\infty} |I_{h^*}(S_n^c)| = \infty$. Therefore, as $n \to \infty$,

$$\left| \frac{\sup \{I_h(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\}}{\sup \{I_{h^*}(S) : S \in \mathcal{S}_n(F_1, \ldots, F_n)\}} - 1 \right| \le \frac{\mathrm{TV}_h \times M}{|I_{h^*}(S_n^c)|} \to 0.$$

The desired result follows. □

Remark 9. We can compare Theorem 6 with Theorem 3.5 of Cai et al. (2018), and there are several major differences on the assumptions. First, the latter result is about increasing Choquet integrals. Second, random variables are non-negative in Cai et al. (2018) because they focus on random losses and risk measures. For signed Choquet integrals, non-negativity seems irrelevant, and we assume instead that the random variables have a bounded sequence of ranges. Third, our Condition C1 is weaker than their Condition A1, and our Condition C2 is stronger than their Condition A2. Fourth, we assume h to be continuous for technical convenience.

4.3 Numerical illustration

In this section, we present numerical examples of risk aggregation under dependence uncertainty for the inter-quantile range and the inter-ES range. As explained in Example 7, IQR_p and IER_p are asymptotically equivalent in terms of the worst-case risk aggregation under dependence uncertainty; this also holds for inhomogeneous portfolios as implied by Theorem 6. Although we work with bounded random variables throughout the paper to establish the theoretical results, the numerical examples in this section are built for unbounded risks to be more realistic for risk management practice. As we shall see below, the results of Theorem 6 are numerically valid for unbounded risks although they do not satisfy Condition C2.

We consider the following three representative models studied in Embrechts et al. (2015). The portfolios in Models (A) and (B) are inhomogeneous whereas the portfolio in Model (C) is homogeneous and very heavy-tailed.

(A) (Mixed portfolio) $F_i = \text{Pareto}(2 + 0.1i)$, $i = 1, \ldots, 5$; $F_i = \text{Exp}(i - 5)$, $i = 6, \ldots, 10$; $F_i = \text{Log-Normal}(0, (0.1(i - 10))^2)$, $i = 11, \ldots, 20$.

(B) (Light-tailed portfolio) $F_i = \text{Exp}(i)$, $i = 1, \ldots, 5$; $F_i = \text{Weibull}(i - 5, 1/2)$, $i = 6, \ldots, 10$; $F_i = F_{i-10}$, $i = 11, \ldots, 20$.

(C) (Very heavy-tailed portfolio) $F_i = \text{Pareto}(1.5)$, $i = 1, \cdots, 50$.

As the common choice of p for the inter-quantile range is 0.75, we compare the values of $\mathrm{IQR}_{0.75}$ and $\mathrm{IER}_{0.75}$ in each of the above models. We look at the influence on the number of risks in the portfolio ($n = 5, 10, 20$ for Models (A) and (B) and $n = 5, 10, 20, 50$ for Model (C)), and on different dependence structures. We report the following quantities for the sum of random variables $X_i \sim F_i$, $i = 1, \ldots, n$.

(i) $\mathrm{IQR}_{0.75}(S_n^{\perp})$: $S_n^{\perp} = \sum_{i=1}^n X_i$ and we assume $X_1, \ldots, X_n$ are independent.

(ii) $\mathrm{IER}_{0.75}(S_n^{\perp})$: $S_n^{\perp}$ is same as in (i).

(iii) $\mathrm{IQR}_{0.75}(S_n^c)$: $S_n^c = \sum_{i=1}^n X_i$ and we assume $X_1, \ldots, X_n$ are comonotonic. By comonotonic-additivity, $\mathrm{IQR}_{0.75}(S_n^c) = \sum_{i=1}^n \mathrm{IQR}_{0.75}(X_i)$.

(iv) $\overline{\mathrm{IQR}}_{0.75}(S_n)$: the worst-case value of $\mathrm{IQR}_{0.75}(S)$ over $S \in \mathcal{S}_n(F_1, \ldots, F_n)$.

(v) $\overline{\mathrm{IER}}_{0.75}(S_n)$: the worst-case value of $\mathrm{IER}_{0.75}(S)$ over $S \in \mathcal{S}_n(F_1, \ldots, F_n)$. By comonotonic-additivity and subadditivity, $\overline{\mathrm{IER}}_{0.75}(S_n) = \mathrm{IER}_{0.75}(S_n^c) = \sum_{i=1}^n \mathrm{IER}_{0.75}(X_i)$.

(vi) $\frac{\overline{\mathrm{IER}}_{0.75}(S_n)}{\overline{\mathrm{IQR}}_{0.75}(S_n)}$: the ratio of the worst-case value of $\mathrm{IER}_{0.75}(S)$ to that of $\mathrm{IQR}_{0.75}(S)$.

The calculation for the independence model in (i) and (ii) is carried out through a Monte-Carlo simulation with sample size $N = 10^6$, and the marginal values in (iii) and (v) are carried out by analytical formulas. The numerical calculation of the worst-case value of $\mathrm{IQR}_{0.75}$ in (iv) is carried out through the Rearrangement Algorithm (RA) of Embrechts et al. (2013) with tail discretization parameter $N = 10^6$ (`R` package: `QRM`)[1]. The numerical results are reported in Tables 1-2.

From Tables 1-2, we make the following observations.

(i) In all models, $\overline{\mathrm{IQR}}_{0.75}(S_n)$ is much larger than $\mathrm{IQR}_{0.75}(S_n^c)$ and $\mathrm{IER}_{0.75}(S_n^{\perp})$. This suggests that neither independence or comonotonicity serves as a conservative benchmark when studying risk aggregation with dependence uncertainty for I_h with a non-concave h such as $I_h = \mathrm{IQR}_p$.

(ii) The ratio of $\overline{\mathrm{IER}}_{0.75}(S_n)$ to $\overline{\mathrm{IQR}}_{0.75}(S_n)$ goes to 1 as n grows for all models (for bounded risks this is shown in Theorem 6). The convergence is very fast for the light-tailed model (B) and relatively slow for the heavy-tailed model (C).

[1]Although the RA is designed for the worst-case risk aggregation of VaR_p, it also works for IQR_p since $\mathrm{VaR}_p(S)$ and $-\mathrm{VaR}_{1-p}(S)$ can be simultaneously maximized over $S \in \mathcal{S}_n(F_1, \ldots, F_n)$; this is because the worst-case scenario for quantiles concerns only tail events; see e.g. Theorem 4.6 of Bernard et al. (2014).

Table 1: Numerical values for two inhomogeneous portfolio models

	Model (A)			Model (B)		
	$n = 5$	$n = 10$	$n = 20$	$n = 5$	$n = 10$	$n = 20$
$\mathrm{IQR}_{0.75}(S_n^{\perp})$	2.7108	3.2664	5.3565	1.4649	1.6836	2.4712
$\mathrm{IER}_{0.75}(S_n^{\perp})$	6.7298	7.5964	11.4492	2.9024	3.2967	4.7444
$\mathrm{IQR}_{0.75}(S_n^{c})$	3.4939	6.0024	13.7364	2.5085	3.9262	7.8523
$\overline{\mathrm{IQR}}_{0.75}(S_n)$	9.7960	15.3198	33.3608	4.9976	7.8054	15.7044
$\overline{\mathrm{IER}}_{0.75}(S_n)$	11.0144	16.1504	33.8089	5.1360	7.8541	15.7082
$\frac{\overline{\mathrm{IER}}_{0.75}(S_n)}{\overline{\mathrm{IQR}}_{0.75}(S_n)}$	1.1243	1.0542	1.0134	1.0277	1.0062	1.0002

Table 2: Numerical values for a very heavy-tailed portfolio model

	Model (C)			
	$n = 5$	$n = 10$	$n = 20$	$n = 50$
$\mathrm{IQR}_{0.75}(S_n^{\perp})$	6.4091	11.6276	20.4428	41.3704
$\mathrm{IER}_{0.75}(S_n^{\perp})$	21.9388	37.3421	62.6127	121.5008
$\mathrm{IQR}_{0.75}(S_n^{c})$	6.5421	13.0843	26.1686	65.4214
$\overline{\mathrm{IQR}}_{0.75}(S_n)$	22.6459	51.7782	111.7977	296.4094
$\overline{\mathrm{IER}}_{0.75}(S_n)$	32.3112	64.6225	129.2450	323.1125
$\frac{\overline{\mathrm{IER}}_{0.75}(S_n)}{\overline{\mathrm{IQR}}_{0.75}(S_n)}$	1.4268	1.2481	1.1561	1.0901

(iii) The values of $\overline{\mathrm{IER}}_{0.75}(S_n)$ and $\overline{\mathrm{IQR}}_{0.75}(S_n)$ are very close for the light-tailed model (B) even for small n such as $n = 5$.

(iv) The difference between $\overline{\mathrm{IQR}}_{0.75}(S_n)$ and $\mathrm{IQR}_{0.75}(S_n^c)$ is more pronounced for the heavy-tailed model (C), compared to the light-tailed model (B).

Acknowledgements

The authors are grateful to various comments from two anonymous referees and the editors, which greatly improved the quality of the paper. Ruodu Wang acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) grants RGPIN-2018-03823 and RGPAS-2018-522590, and the Center of Actuarial Excellence Research Grant from the Society of Actuaries. Gordon E. Willmot acknowledges financial support from NSERC grant RGPIN-2016-03654 and from the Munich Reinsurance Company of Canada. Yunran Wei is supported by the above NSERC grants and the Hickman Scholar Program of the Society of Actuaries.

A Appendix

A.1 Two classic results

Here we list two classic results used in this paper for the sake of completeness. The first result is used in the proof of Theorem 1. The original choice of notation in Schmeidler (1986) is kept here.

Theorem 7 (Proposition 2 of Schmeidler (1986))**.** *Let Σ denote a nonempty algebra of subsets of a set S, let B denote the set of all bounded, real-valued, Σ-measurable functions on S. Suppose that $I : B \to \mathbb{R}$ is comonotonic additive and continuous with respect to supremum norm in B. Then, for any E in Σ, defining $v(E) = I(\mathbb{1}_E)$ on Σ, we have for all $a \in B$,*

$$I(a) = \int_{-\infty}^{0} (v(a \geq \alpha) - v(S)) \,\mathrm{d}\alpha + \int_{0}^{\infty} v(a \geq \alpha) \,\mathrm{d}\alpha.$$

The second result is the classic Hardy-Littlewood-Pólya inequality (see e.g. page 22 of Olkin and Marshall (2016)), which is used in the proof of Theorem 2.

Theorem 8 (Hardy-Littlewood-Pólya)**.** *Let f and g be two decreasing integrable functions on $[a, b]$, taking values in $[0, 1]$. Then*

$$\int_a^b \phi(f(x)) \,\mathrm{d}x \leq \int_a^b \phi(g(x)) \,\mathrm{d}x$$

holds for all continuous concave function ϕ (for which both functions $\phi \circ f$ and $\phi \circ g$ are integrable) if and only if

$$\int_a^x f(t)\,\mathrm{d}t \geq \int_a^x g(t)\,\mathrm{d}t \qquad \text{for } a \leq x \leq b$$

and

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t.$$

A.2 Proofs of some lemmas

Proof of Lemma 1. (i) " $\Rightarrow$": This is trivial from the definition of signed Choquet integrals.

(ii) " $\Leftarrow$": Fix $p \in [0,1]$, we take a Bernoulli random variable X such that

$$\mathbb{P}(X=0)=p,\ \mathbb{P}(X=1)=1-p.$$

It follows that

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

As $I_{h_1} \leq I_{h_2}$ and p is arbitrary, we conclude $h_1 \leq h_2$. □

Proof of Lemma 2. (i) " $\Rightarrow$": We only show the case when I_h is increasing. Take $U \sim \mathrm{U}[0,1]$, for any $t_1, t_2 \in [0,1]$, $t_1 \leq t_2$, we let $X = \mathbb{1}_{\{U \leq t_1\}}$ and $Y = \mathbb{1}_{\{U \leq t_2\}}$. $X \leq Y$ implies $h(t_1) = I_h(X) \leq I_h(Y) = h(t_2)$. This shows that h is increasing.

" $\Leftarrow$": We only show the case when h is increasing. For random variables $X, Y \in L^\infty$, $x \in \mathbb{R}$, if $X \leq Y$, then $\mathbb{P}(X \geq x) \leq \mathbb{P}(Y \geq x)$. If h is increasing, then

$$h(\mathbb{P}(X \geq x)) \leq h(\mathbb{P}(Y \geq x)),$$

which implies $I_h(X) \leq I_h(Y)$. Hence I_h is increasing.

(ii) By straightforward calculation,

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

Since I_h is comonotonic-additive, and X and c are comonotonic, $I_h(X+c) = I_h(X) + I_h(c) = I_h(X) + ch(1)$.

(iii) By (2),

$$I_h(\lambda X) = \int_{-\infty}^{0} (h(\mathbb{P}(\lambda X \geq x)) - h(1))\,\mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(\lambda X \geq x))\,\mathrm{d}x$$
$$= \lambda \int_{-\infty}^{0} (h(\mathbb{P}(X \geq y)) - h(1))\,\mathrm{d}y + \lambda \int_{0}^{\infty} h(\mathbb{P}(X \geq y))\,\mathrm{d}y = \lambda I_h(X).$$

(iv) By (2) and (3),

$$I_h(-X) = \int_{-\infty}^{0} (h(\mathbb{P}(X < -x)) - h(1))\,\mathrm{d}x + \int_{0}^{\infty} h(\mathbb{P}(X < -x))\,\mathrm{d}x$$
$$= \int_{-\infty}^{0} \hat{h}(\mathbb{P}(X \geq -x))\,\mathrm{d}x + \int_{0}^{\infty} \left(\hat{h}(\mathbb{P}(X \geq -x)) - \hat{h}(1)\right)\,\mathrm{d}x$$
$$= \int_{0}^{\infty} \hat{h}(\mathbb{P}(X \geq x))\,\mathrm{d}x + \int_{-\infty}^{0} \left(\hat{h}(\mathbb{P}(X \geq x)) - \hat{h}(1)\right)\,\mathrm{d}x = I_{\hat{h}}(X). \qquad \square$$

Proof of Lemma 4. Replace h by its Jordan decomposition $h = h_+ - h_-$, where h_+, h_- are increasing. We have $\mathrm{TV}_{h_+} + \mathrm{TV}_{h_-} = \mathrm{TV}_h$, and

$$|I_h(X) - I_h(Y)| = |I_{h_+}(X) - I_{h_-}(X) - I_{h_+}(Y) + I_{h_-}(Y)| \leq |I_{h_+}(X) - I_{h_+}(Y)| + |I_{h_-}(X) - I_{h_-}(Y)|.$$

Therefore, it suffices to show (8) for h increasing. From Lemma 2, we have

$$I_h(Y) \leq I_h(X + ||X - Y||_\infty) = I_h(X) + h(1)||X - Y||_\infty = I_h(X) + \mathrm{TV}_h ||X - Y||_\infty.$$

Therefore, by symmetry, (8) holds. $\square$

References

Acciaio, B. and Svindland, G. (2013). Are law-invariant risk functions concave on distributions? *Dependence Modeling*, **1**, 54–64.

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

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

BCBS (2016). *Standards. Minimum Capital Requirements Market Risk. January 2016.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Bernard, C., Jiang, X. and Wang, R. (2014). *Insurance: Mathematics and Economics*, **54**, 93–108.

Bernard, C., Rüschendorf, L. and Vanduffel, S. (2017a). VaR bounds with variance constraint. *Journal of Risk and Insurance*, **84**(3), 923–959.

Bernard, C., Rüschendorf, L., Vanduffel, S. and Wang, R. (2017b). Risk bounds for factor models. *Finance and Stochastics*, **21**(3), 631–659.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, 1196–1217.

Brighi, B. and Chipot, M. (1994): Approximated convex envelope of a function. *SIAM Journal on Numerical Analysis*, **31**, 128–148.

Cai, J., Liu, H. and Wang, R. (2018). Asymptotic equivalence of risk measures under dependence uncertainty. *Mathematical Finance*, **28**(1), 29–49.

Cerreia-Vioglio, S. and Maccheroni, F. and Marinacci, M. and Montrucchio, L. (2012). Signed integral representations of comonotonic additive functionals. *Journal of Mathematical Analysis and Applications*, **385**(2), 895–912.

Cerreia-Vioglio, S. and Maccheroni, F. and Marinacci, M. and Montrucchio, L. (2015). Choquet integration on Riesz spaces and dual comonotonicity. *Transactions of the American Mathematical Society*, **367**(12), 8521–8542.

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

Cont, R., Deguest, R. and Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, **10**(6), 593–606.

Dana, R.A. (2005). A representation result for concave Schur concave functions. *Mathematical Finance*, **15**(4), 613–634.

Delbaen, F., Bellini, F., Bignozzi, V. and Ziegel, J. (2016). Risk measures with the CxLS property. *Finance and Stochastics*, **20**(2), 433–453.

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.

De Waegenaere, A. and Wakker, P. P. (2001). Nonmonotonic Choquet integrals. *Journal of Mathematical Economics*, **36**, 45–60.

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

Donoghue, W. (1969). *Distributions and Fourier Transforms.* Academic press, New York.

Embrechts, P., Puccetti, G. and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, **37**(8), 2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A. (2014). An academic response to Basel 3.5. *Risks*, **2**(1), 25–48.

Embrechts, P., Wang, B. and Wang, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, **19**(4), 763–790.

Emmer, S., Kratz, M. and Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, **18**(2), 31–60.

Filipović, D. and Svindland, G. (2008). Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance and Stochastics*, **12**, 423–439.

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, Berlin, Fourth Edition.

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.

Gao, N. and Leung, D. H. and Munari, C. and Xanthos, F. (2018). Fatou Property, representations, and extensions of law-invariant risk measures on general Orlicz spaces. *Finance and Stochastics*, **22**(2), 395–415.

Gao, N. and Xanthos, F. (2018). On the C-property and w^*-representations of risk measures. *Mathematical Finance*, **28**(2), 748–754.

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

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

Hampel, F. (1971). A general qualitative definition of robustness. *Annals of Mathematical Statistics*, **42**(6), 1887–1896.

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

Jouini, E., Schachermayer, W. and Touzi, N. (2006). Law invariant risk measures have the Fatou property. *Advances in Mathematical Economics*, **9**, 49–71.

Kou, S., Peng, X. and Heyde, C. C. (2013). External risk measures and Basel accords. *Mathematics of Operations Research*, **38**(3), 393–417.

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 quantitative robustness for law-invariant risk measures. *Finance and Stochastics*, **18**(2), 271–295.

Krätschmer, V., Schied, A. and Zähle, H. (2017). Domains of weak continuity of statistical functionals with a view toward robust statistics. *Journal of Multivariate Analysis*, **158**, 1–19.

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

Lorentz, G. G. (1951). On the theory of spaces Λ. *Pacific Journal of Mathematics*, **1**(3), 411–429.

Mao, T. and Wang, R. (2015). On aggregation sets and lower-convex sets. *Journal of Multivariate Analysis*, **136**, 12–25.

Marinacci, M. and Montrucchio L. (2004). Introduction to the mathematics of ambiguity. In *Uncertainty in Economic Theory*, I. Gilboa, Ed. Routledge, New York, NY, USA. 46–107.

Markowitz, H. (1951). Portfolio selection, *Journal of Finance*, **7**(1), 77–91.

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

Murofushi, T., Sugeno, M. and Machida, M. (1994). Non-monotonic fuzzy measures and the Choquet integral. *Fuzzy sets and Systems*, **64**(1), 73–86.

Olkin, I. and Marshall, A. W. (2016). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

Pesenti, S. M., Millossovich, P. and Tsanakas A. (2016). Robustness regions for measures of risk aggregation. *Dependence Modeling*, **4**, 348–367.

Puccetti, G. and Rüschendorf, L. (2014). Asymptotic equivalence of conservative value-at-risk- and expected shortfall-based capital charges. *Journal of Risk*, **16**(3), 1–19.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, **3**(4), 323–343.

Rao, M. M. and Ren, Z. D. (1991). *Theory of Orlicz Spaces*. New York: M. Dekker.

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.

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

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

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

Wang, B. and Wang, R. (2015). Extreme negative dependence and risk aggregation. *Journal of Multivariate Analysis*. **136**, 12–25.

Wang, R., Bignozzi, V. and Tsakanas, A. (2015). How superadditive can a risk measure be? *SIAM Jounral on Financial Mathematics*, **6**, 776–803.

Wang, R., Peng, L. and Yang, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, **17**(2), 395–417.

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

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

Ziegel, J. (2016). Coherence and elicitability. *Mathematical Finance*, **26**, 901–918.