Asymptotic Equivalence of Risk Measures under Dependence Uncertainty

Jun Cai[*], Haiyan Liu[†] and Ruodu Wang[‡]

March 23, 2016[§]

Abstract

In this paper we study the aggregate risk of inhomogeneous risks with dependence uncertainty, evaluated by a generic risk measure. We say that a pair of risk measures are asymptotically equivalent if the ratio of the worst-case values of the two risk measures is almost one for the sum of a large number of risks with unknown dependence structure. The study of asymptotic equivalence is particularly important for a pair of a non-coherent risk measure and a coherent risk measure, since the worst-case value of a non-coherent risk measure under dependence uncertainty is typically very difficult to obtain. The main contribution of this paper is that we establish general asymptotic equivalence results for the classes of distortion risk measures and convex risk measures under different mild conditions. The results implicitly suggest that it is only reasonable to implement a coherent risk measure for the aggregation of a large number of risks with uncertainty in the dependence structure, a relevant situation for risk management practice.

Key-words: risk aggregation; distortion risk measures; convex risk measures; dependence uncertainty; diversification.

AMS Classification: 91B30; 91G10; 62P05.

[*]Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L3G1, Canada (email: jcai@uwaterloo.ca).

[†]Corresponding author. Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L3G1, Canada (email: h262liu@uwaterloo.ca).

[‡]Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L3G1, Canada (email: wang@uwaterloo.ca).

[§]The authors thank two anonymous referees, an Associate Editor, and an Editor for comments and suggestions that improved the presentation of the paper. Jun Cai and Ruodu Wang acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-250031-2011 and RGPIN-435844-2013, respectively). Haiyan Liu acknowledges financial support from the University of Waterloo and the China Scholarship Council.

1 Introduction

In the past two decades, risk measures have been the standard tool for financial institutions in both calculating regulatory capital requirement and internal risk management. In particular, the two most popular risk measures in practice are the Value-at-Risk (VaR) and the Expected Shortfall (ES). There have been extensive discussions recently around the comparative advantages of VaR and ES in regulation; the reader is referred to the survey papers Embrechts et al. (2014), Emmer et al. (2015), and Föllmer and Weber (2015). Related debates in regulatory documents by the Basel Committee on Banking Supervision and the International Association of Insurance Supervisors can be found in BCBS (2013) and IAIS (2014).

A risk measure is a mapping from a set of risks to numbers, and it has to be implemented with certain models, either internal models of a financial institution or external models designed by the regulator. By specifying a model, uncertainty always arises as an important issue in practice. One particular type of uncertainty that we focus on in this paper is the dependence uncertainty in risk aggregation. In the framework of dependence uncertainty, we assume that in a joint model $(X_1, \ldots, X_n)$, the marginal distribution of each of $X_1, \ldots, X_n$ is known, but the joint distribution is unknown. This is due to statistical and modeling challenges in obtaining precise information on the dependence structure of a joint model; see Embrechts et al. (2014) for more illustrations. Denote by $\mathcal{F}$ the set of univariate distribution functions. For $F_1, \ldots, F_n \in \mathcal{F}$, let

$$\mathcal{S}_n = \mathcal{S}_n(F_1, \ldots, F_n) = \{X_1 + \cdots + X_n : X_i \in L^0,\ X_i \sim F_i,\ i = 1, \ldots, n\}.$$

That is, $\mathcal{S}_n$ is the set of aggregate risks with given marginal distributions, but an arbitrary dependence structure. Some properties of the set $\mathcal{S}_n$ are given in Bernard et al. (2014).

For a given risk measure $\rho : \mathcal{X} \to (-\infty, +\infty]$, where the set $\mathcal{X}$ is a convex cone of risks, we are interested in the value of the risk aggregation $\rho(X_1 + \cdots + X_n)$ for some joint model $(X_1, \ldots, X_n)$ with unknown dependence structure. Obviously, $\rho(X_1 + \cdots + X_n)$ lies in a range, and often the worst-case value and the best-case value are of particular interest. The value $\bar{\rho}(\mathcal{S}_n) := \sup_{S \in \mathcal{S}_n} \rho(S)$ represents the worst-case measurement of the aggregate risk in the presence of dependence uncertainty. If ρ is not convex, the value of $\bar{\rho}(\mathcal{S}_n)$ is in general very difficult to calculate. For the case of VaR, some analytical results are given in Wang et al. (2013) and Jakobsons et al. (2016). It is common to calculate $\overline{\mathrm{VaR}}_p(\mathcal{S}_n)$ by numerical calculation and a popular algorithm is the Rearrangement Algorithm in Embrechts et al. (2013). If partial dependence information is available, one can study the values of risk measures in constrained subsets of $\mathcal{S}_n$; see Bernard et al. (2016), Bernard and Vanduffel (2015), and Bignozzi et al. (2015) for research along this direction. In this paper, we focus on the full set $\mathcal{S}_n$, that is, no dependence information.

We are particularly interested in the case when n goes to infinity, that is, a very large number of risks. On one hand, this setting provides mathematical tractability for the behaviour of risk aggregation; on the other hand, dependence uncertainty among a very large number of risks is a practical setting due to the statistical challenges arising in high-dimensional models.

Under this setting, a particularly elegant result is that the VaR and the ES at the same confidence level are asymptotically equivalent. That is, for a given sequence of distributions $F_1, F_2 \ldots$ and $p \in (0, 1)$,

$$\lim_{n\to\infty} \frac{\sup_{S\in\mathcal{S}_n} \mathrm{VaR}_p(S)}{\sup_{S\in\mathcal{S}_n} \mathrm{ES}_p(S)} = 1 \tag{1.1}$$

holds under some conditions; some references on (1.1) are mentioned below. First let us define the risk measures VaR and ES as used in this paper. The VaR at confidence level $p \in (0, 1)$ is defined as

$$\mathrm{VaR}_p(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leqslant x) \geqslant p\}, \quad X \in L^0,$$

and the ES at confidence level $p \in (0, 1)$ is defined as

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q, \;\; X \in L^0,$$

where L^0 is the set of all random variables in a probability space which we formally introduce in Section 2. Note that in general ES_p can be infinite for non-integrable random variables. For properties of the two regulatory risk measures, see for instance McNeil et al. (2015).

The equivalence (1.1) is known in a series of papers under some particular conditions. (1.1) is first shown under a homogeneous setting (that is, $F_1 = F_2 = \cdots$) in Puccetti and Rüschendorf (2014) under an assumption of *complete mixability* (Wang and Wang (2011)). It is then generalized by for instance Puccetti et al. (2013) and Wang and Wang (2015) under different conditions. The case of inhomogeneous setting is finally obtained in Embrechts et al. (2015) under some general moment conditions on the marginal distributions $F_1, F_2, \ldots$.

An immediate question is whether the asymptotic equivalence in (1.1) is not only true for the pair (VaR_p, ES_p), but it also holds for much larger classes of risk measures. We say that a risk measure ρ^* dominates ρ if they are defined on the same set $\mathcal{X}$ and $\rho \leqslant \rho^*$ on $\mathcal{X}$. It is well known that ES_p

is the smallest *law-invariant coherent risk measure* (see Section 2 below for a definition) dominating VaR_p; see Kusuoka (2001). For a law-invariant risk measure ρ, denote by ρ^* the smallest law-invariant coherent risk measure dominating ρ, if such a risk measure exists. It is natural to ask whether the following equivalence

$$\lim_{n\to\infty} \frac{\sup_{S\in\mathcal{S}_n} \rho(S)}{\sup_{S\in\mathcal{S}_n} \rho^*(S)} = 1, \tag{1.2}$$

holds and under what conditions. A result of type (1.2) is called an *asymptotic equivalence* for risk measures ρ and ρ^*.

In this paper, we focus on two popular classes of risk measures. The class of *distortion risk measures*, including VaR and ES above, is extensively studied as tools for capital calculation (see e.g. Acerbi (2002) and Cont et al. (2010)), insurance premium calculation (see e.g. Wang et al. (1997)), and decision making (see e.g. Yaari (1987)). The class of *convex risk measures*, introduced by Föllmer and Schied (2002) and Frittelli and Rossaza Gianin (2002) as an extension of coherent risk measures, is able to reflect non-linearity in the increase of the size of risks, such as risky positions in a financial market with limited liquidity. See Section 2 below for precise definitions, and for more discussions on the use of these two classes of risk measures, see Föllmer and Schied (2011, Chapter 4).

The main results in Wang et al. (2015) imply that (1.2) holds in the homogeneous model ($F_1 = F_2 = \cdots$) if ρ is a distortion risk measure or a convex risk measure. The assumption of homogeneity is nice for mathematical analysis; however, it is not a realistic assumption for practical applications. In this paper, our aim is to show (1.2) in inhomogeneous models for general risk measures. This requires some regularity conditions on the marginal distributions, which we will specify later.

The result of asymptotic equivalence in (1.2) has two practical merits. First, it suggests that using a non-coherent risk measure would lead to roughly the same worst-case value as its coherent partner if the dependence structure is unknown for a joint model of high dimension; therefore a regulator may want to directly implement a coherent risk measure instead. This point is very much relevant to the general question of searching for risk measures in the recent regulatory documents BCBS (2013) and IAIS (2014). Second, the value $\bar{\rho}^*(\mathcal{S}_n)$ can be analytically calculated without specifying a dependence structure, since the worst-case value for ρ^* is often simply the sum of the values of $\rho^*(X_1), \ldots, \rho^*(X_n)$ with corresponding marginal distributions $X_i \sim F_i$, $i = 1, \ldots, n$. As a consequence, (1.2) can be used to approximate $\bar{\rho}(\mathcal{S}_n)$ if needed. These merits provide a powerful tool for evaluating model uncertainty for risk aggregation with non-coherent risk measures.

Mathematically, the main result in this paper generalizes not only the results in Embrechts et al. (2015) for VaR and ES, but also those in Wang et al. (2015) for general risk measures in the homogeneous setting. More importantly, our methods unify the two streams of research in this field. A significant mathematical challenge arises as the method in Wang et al. (2015) relies on the study of the quantity

$$\Gamma_\rho(X) = \lim_{n\to\infty} \frac{1}{n}\sup\{\rho(S) : S \in \mathcal{S}_n(F, \dots, F)\}, \quad X \sim F,$$

which cannot be naturally generalized to an inhomogeneous setting. In this paper, we use an alternative method by constructing a specific $S_n \in \mathcal{S}_n$ such that $\rho(S_n)$ and $\rho^*(S_n)$ are close. It should be noted that the case for distortion risk measures is technically much more involved than the case for convex risk measures, since we know the worst-case dependence structure for convex risk measures is comonotonicity, but not for non-coherent distortion risk measures in general. The main theorem and its proof thereby reveal the worst-case dependence structure for general distortion risk measures (Choquet integrals). This dependence structure is valuable to many other fields where probability distortion is involved, for instance in decision theory (see for instance Yaari (1987) and Quiggin (1993)), behavioral finance (see for instance He and Zhou (2016)), reinsurance (see for instance Bernard et al. (2015)) and insurance pricing (see for instance Wang et al. (1997)).

The structure of the paper is as follows. In Section 2, we give some definitions and preliminaries on risk measures, and present two examples showing that without some regularity conditions the asymptotic equivalence may fail to hold. In Section 3, we study the asymptotic equivalence for distortion risk measures under some regularity conditions. In Section 4, we study the asymptotic equivalence for convex risk measures under general conditions. Brief conclusions are stated in Section 5. We put a complete proof of the main theorem of Section 3 in the Appendix.

2 Preliminaries

2.1 Preliminaries on risk measures

We work with an atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let L^p be the set of all random variables in $(\Omega, \mathcal{A}, \mathbb{P})$ with finite p-th moment, $p \in [0, \infty)$, L^∞ be the set of essentially bounded random variables, and L^+ be the set of non-negative random variables. A positive (negative) value of $X \in L^0$ represents a financial loss (profit) in this paper.

A *risk measure* is a function $\rho : \mathcal{X} \to (-\infty, \infty]$, where the set $\mathcal{X}$ is a convex cone such that

$L^\infty \subset \mathcal{X} \subset L^0$ ($\subset$ is the non-strict set inclusion). Below we list some standard properties studied in the literature of risk measures. For any $X, Y \in \mathcal{X}$:

(a) *Monotonicity*: if $X \leqslant Y$ $\mathbb{P}$-a.s, then $\rho(X) \leqslant \rho(Y)$;

(b) *Cash-invariance*: for any $m \in \mathbb{R}$, $\rho(X - m) = \rho(X) - m$;

(c) *Convexity*: for any $\lambda \in [0, 1]$, $\rho(\lambda X + (1-\lambda)Y) \leqslant \lambda\rho(X) + (1-\lambda)\rho(Y)$;

(d) *Subadditivity*: $\rho(X + Y) \leqslant \rho(X) + \rho(Y)$;

(e) *Positive homogeneity*: for any $\alpha > 0$, $\rho(\alpha X) = \alpha\rho(X)$;

(f) *Law-invariance*: if X and Y have the same distribution under $\mathbb{P}$, denoted as $X \overset{d}{=} Y$, then $\rho(X) = \rho(Y)$.

We refer to Föllmer and Schied (2011, Chapter 4) and Delbaen (2012) for the interpretations of these standard properties of risk measures.

Definition 2.1. A *monetary risk measure* is a risk measure satisfying (a) and (b), a *convex risk measure* is a risk measure satisfying (a)-(c), and a *coherent risk measure* is a risk measure satisfying (a)-(e).

For a distribution function F, write

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

In the next we introduce the class of *distortion risk measures*, including VaR and ES defined in Section 1 as special cases. Let $\mathcal{H}$ be the set of increasing (in the non-strict sense) function h with $h(0) = h(0^+) = 0$ and $h(1^-) = h(1) = 1$. A distortion risk measure $\rho_h : \mathcal{X} \to (-\infty, \infty]$ with a *distortion function* $h \in \mathcal{H}$ is defined as

$$\rho_h(X) = \int_{\mathbb{R}} x \mathrm{d}h(F(x)), \quad X \in \mathcal{X},\ X \sim F, \tag{2.1}$$

provided that (2.1) is well-posed for all $X \in \mathcal{X}$. Note that for a given set $\mathcal{X}$, h may need to satisfy some conditions to avoid some ill-posed cases. If $\mathcal{X}$ is either L^∞ or L^+, (2.1) is well-posed for all $h \in \mathcal{H}$ and $X \in \mathcal{X}$.

When h is continuous, through a change of variable, ρ_h can be written as

$$\rho_h(X) = \int_0^1 F^{-1}(t)\mathrm{d}h(t) = \int_0^1 \mathrm{VaR}_t(X)\mathrm{d}h(t), \quad X \in \mathcal{X}. \tag{2.2}$$

Any distortion risk measure ρ_h is monotone, cash-invariant, positively homogeneous, and law-invariant. ρ_h is subadditive if and only if h is convex; this dates back to Yaari (1987, Theorem 2). The key feature which characterizes ρ_h is *comonotonic additivity*. Let us first recall the definition of comonotonic random variables.

Definition 2.2. Two random variables X and Y are *comonotonic* if

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geqslant 0 \text{ for } (\omega, \omega') \in \Omega \times \Omega \ (\mathbb{P} \times \mathbb{P})\text{-a.s.}$$

Comonotonicity of X and Y is equivalent to the existence of a random variable $Z \in L^0$ and two non-decreasing functions f and g, such that $X = f(Z)$ and $Y = g(Z)$ almost surely. See Dhaene et al. (2002) for an overview on comonotonicity.

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

Comonotonic additive law-invariant monetary risk measures are equivalent to distortion risk measures. This result essentially dates back to the property of Choquet integrals; see Yaari (1987) and Theorem 4.88 of Föllmer and Schied (2011). For a subadditive risk measure ρ interpreted as a tool for capital calculation, comonotonic additivity is particularly important: For comonotonic risks X and Y, the lack of comonotonic additivity (that is, $\rho(X + Y) < \rho(X) + \rho(Y)$) means a diversification benefit (reduction in capital) for non-diversified risks, an undesirable property for risk management.

Finally, we give the Fatou property, an important property related to convex risk measures, which will be used in the proof of our results in Section 4.

(h) *(L^1-)Fatou property*: $\liminf_{n\to\infty} \rho(X_n) \geqslant \rho(X)$ if $X, X_1, X_2, \dots \in \mathcal{X} = L^1$ and $X_n \xrightarrow{L^1} X$ as $n \to \infty$.

2.2 Vanishing risks and exploding risks

Before we move on to the main result of this paper, we present two counter-examples of asymptotic equivalence to help the reader to understand the nature of the problem. Let Γ be the set of all pairs (ρ_1, ρ_2) where ρ_1 is a non-coherent monetary risk measure on $\mathcal{X}$ and ρ_2 is a coherent risk measure on $\mathcal{X}$ dominating ρ_1. For $(\rho, \rho^*) \in \Gamma$, in order to have the general asymptotic equivalence

$$\lim_{n\to\infty} \frac{\sup_{S\in\mathcal{S}_n} \rho(S)}{\sup_{S\in\mathcal{S}_n} \rho^*(S)} = 1, \tag{2.3}$$

some regularity conditions have to be imposed to avoid the following cases of vanishing risks and exploding risks. Note that both the two cases are typically irrelevant in practice.

Example 2.1 (Vanishing risks)**.** For a pair $(\rho,\rho^*) \in \Gamma$, take $X \in \mathcal{X}$ such that $0 < \rho(X) < \rho^*(X)$; such X always exists since ρ is not coherent and hence $\rho \neq \rho^*$ for some subset of $\mathcal{X}$. Write $a = \rho(X)$ and $b = \rho^*(X)$. Let F_1 be the distribution of X. For $i = 2, 3, \dots$, let F_i be a distribution supported in $[0, k_i]$, where $\{k_i,\ i = 2, 3, \dots\}$ is a sequence of positive numbers such that $\sum_{i=2}^{\infty} k_i < (b-a)/2$. From the monotonicity and cash-invariance of ρ and ρ^*, we have

$$\sup_{S\in\mathcal{S}_n} \rho(S) \leqslant \rho(X_1) + \sum_{i=2}^{n} k_i \leqslant a + \frac{1}{2}(b-a) = \frac{1}{2}(a+b)$$

and

$$\sup_{S\in\mathcal{S}_n} \rho^*(S) \geqslant \rho^*(X_1) = b.$$

Then for $n \in \mathbb{N}$,

$$\frac{\sup_{S\in\mathcal{S}_n} \rho(S)}{\sup_{S\in\mathcal{S}_n} \rho^*(S)} \leqslant \frac{a+b}{2b} < 1.$$

That is, (2.3) does not hold. This example suggests that for (2.3) to hold, a regularity condition has to be imposed to avoid vanishing risks, that is, the scale of individual risks shrinks too fast as $n \to \infty$.

Example 2.2 (Exploding risks)**.** For the purpose of illustration we take $(\rho,\rho^*) \in \Gamma$ where ρ is positive homogeneous. This example includes, for instance, a distortion risk measure and its dominating coherent distortion risk measure; see Section 3 below. Take a random variable $X \in \mathcal{X}$ supported in a compact interval $[0,1]$ such that $\rho(X) < \rho^*(X)$; such X always exists since both ρ and ρ^* are positive homogeneous and $\rho \neq \rho^*$ for some subset of $\mathcal{X}$. Write $a = \rho(X)$ and $b = \rho^*(X)$. Now, let $\{k_i,\ i \in \mathbb{N}\}$ be a sequence of positive numbers such that $k_1 = 1$ and $2\sum_{i=1}^{n} k_i < (b-a)k_{n+1}$ for all $n \in \mathbb{N}$. Let F_i be the distribution of $k_i X$ for $i \in \mathbb{N}$.

From the monotonicity and the cash-invariance of ρ and ρ^*, we have

$$\sup_{S\in\mathcal{S}_n} \rho(S) \leqslant k_n\rho(X) + \sum_{i=1}^{n-1} k_i = k_n a + \sum_{i=1}^{n-1} k_i < k_n a + \frac{1}{2}k_n(b-a) = \frac{1}{2}k_n(a+b)$$

and

$$\sup_{S\in\mathcal{S}_n} \rho^*(S) \geqslant \rho^*(X_n) = k_n\rho^*(X) = k_n b.$$

Therefore,

$$\frac{\sup_{S\in\mathcal{S}_n} \rho(S)}{\sup_{S\in\mathcal{S}_n} \rho^*(S)} \leqslant \frac{k_n(a+b)}{2k_n b} = \frac{a+b}{2b} < 1.$$

That is, (2.3) does not hold. This example suggests that for (2.3) to hold, a regularity condition has to be imposed to avoid exploding risks, that is, the scale of individual risks grows too fast as $n \to \infty$.

3 Asymptotic equivalence for distortion risk measures

Throughout this section, we take $\mathcal{X} = L^+$. Since monetary risk measures are cash-invariant, this assumption is technically equivalent to assuming that each risk is uniformly bounded from below (bounded gain). Gains are typically not relevant when regulatory risk measures such as VaR and ES are applied, and hence this is a common assumption in risk management.

3.1 Some lemmas

Before approaching the main result of this section, we first provide some necessary lemmas on distortion risk measures and on the set $\mathcal{S}_n$. A key object for our analysis is the largest convex distortion function dominated by h, defined as

$$h^*(t) = \sup\{g(t):\ g:[0,1]\to[0,1],\ g\leqslant h,\ g \text{ is increasing and convex on } [0,1]\},\quad t\in[0,1]. \tag{3.1}$$

We will use the notation h^* throughout Section 3.

The first lemma formulates an order in two distortion risk measures from the order in their respective distortion functions.

Lemma 3.1. *For two distortion functions $h_1, h_2 \in \mathcal{H}$, if $h_1(t) \leqslant h_2(t)$ for all $t \in [0,1]$, then*

$$\rho_{h_1}(X) \geqslant \rho_{h_2}(X),\quad X\in\mathcal{X}.$$

Proof. Let F be the distribution of $X \in \mathcal{X}$. For $x \in \mathbb{R}$ and $i = 1, 2$, let $g_i(x) = (h_i \circ F)(x+) = \lim_{y\to x^+} h_i(F(y))$, that is, g_i is the right-continuous correction of $h_i \circ F$. Since $h_1 \leqslant h_2$ on $[0,1]$, we have $g_1 \leqslant g_2$ on $\mathbb{R}$. Let Y_i be a random variable with distribution function g_i, $i = 1, 2$. Then we have $\mathbb{E}[Y_1] \geqslant \mathbb{E}[Y_2]$ from $g_1 \leqslant g_2$. Finally, we obtain

$$\rho_{h_1}(X) = \int_{\mathbb{R}} x\mathrm{d}(h_1\circ F)(x) = \int_{\mathbb{R}} x\mathrm{d}(h_1\circ F)(x+) = \int_{\mathbb{R}} x\mathrm{d}g_1(x) = \mathbb{E}[Y_1] \geqslant \mathbb{E}[Y_2] = \rho_{h_2}(X),$$

as desired, where the second equality is due to the facts that the integrand $x \to x \in \mathbb{R}$ is continuous, $X \in L^+$, and $h_1 \circ F$ is increasing. □

The next lemma gives ρ_{h^*} as the smallest coherent distortion risk measure dominating ρ_h. It was given in Wang et al. (2015) for right-continuous $h \in \mathcal{H}$; however from there it is a simple exercise to see that the lemma holds for all $h \in \mathcal{H}$. In the latter paper it is also shown that ρ_{h^*} is the smallest law-invariant coherent risk measure dominating ρ_h.

Lemma 3.2 (Lemma 3.1 of Wang et al. (2015))**.** *For any $h \in \mathcal{H}$, h^* as in (3.1) is a continuous distortion function. Moreover, the smallest coherent distortion risk measure dominating ρ_h exists and has distortion function h^*, that is,*

$$\rho_{h^*}(X) = \int_0^1 \mathrm{VaR}_t(X)\mathrm{d}h^*(t), \quad X \in \mathcal{X}. \tag{3.2}$$

The following lemma provides a building block for the dependence structure that we need for the asymptotic equivalence.

Lemma 3.3 (Corollary A.3 of Embrechts et al. (2015))**.** *Suppose that $\{F_i,\ i \in \mathbb{N}\}$ is a sequence of distributions with bounded support, then there exist random variables $X_i \sim F_i$, $i \in \mathbb{N}$ such that for each $n \in \mathbb{N}$,*

$$|S_n - \mathbb{E}[S_n]| \leqslant L_n, \tag{3.3}$$

where $S_n = X_1 + \cdots + X_n$ and L_n is the largest length of the support of F_i, $i = 1, \ldots, n$.

Finally, the following lemma from convex analysis provides an important geometric property of the pair (h, h^*).

Lemma 3.4 (Lemma 5.1 of Brighi and Chipot (1994))**.** *Suppose $h \in \mathcal{H}$ is continuous and h^* is defined in (3.1). The set $\{t \in [0, 1] : h(t) \neq h^*(t)\}$ is the union of some disjoint open intervals, and h^* is linear on each of the intervals.*

3.2 Asymptotic equivalence for distortion risk measures

For a given $h \in \mathcal{H}$ and h^* defined in (3.1), we list the two conditions for a sequence of distribution functions $\{F_i,\ i \in \mathbb{N}\}$ that we work with. In the following, $X_i \sim F_i$, $i \in \mathbb{N}$.

Condition A1. $\liminf\limits_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \rho_{h^*}(X_i) > 0$.

Condition A2. $\lim_{q\to 1} \sup_{i\in\mathbb{N}} \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t) = 0$.

Condition A1 requires that ρ_{h^*} of the marginal risks does not vanish, and thereby it eliminates the case of vanishing risks as in Example 2.1. Condition A2 requires that the marginal risks are uniformly integrable with respect to h^*, and thereby it eliminates the case of exploding risks as in Example 2.2. A1 automatically holds for marginal risks uniformly bounded below away from zero and A2 automatically holds for marginal risks uniformly bounded above. The following theorem contains the main result of this paper.

Theorem 3.5. *For $h \in \mathcal{H}$ and a sequence of distribution functions $\{F_i,\ i \in \mathbb{N}\}$ supported in $\mathbb{R}_+$ and satisfying Conditions A1-A2, we have*

$$\lim_{n\to\infty} \frac{\sup\{\rho_h(S) : S \in \mathcal{S}_n\}}{\sup\{\rho_{h^*}(S) : S \in \mathcal{S}_n\}} = 1, \tag{3.4}$$

where h^* is defined in (3.1).

Proof. The proof of this theorem is very technical and depends on the geometrical relationship between h and h^*. Here we give the proof for the following nice case, from which the reader should be able to grasp the main ideas. A full proof is put in the Appendix.

Case 1. Assume that h is continuous and there exists $p \in (0,1)$ such that $h(t) = h^*(t)$ for all $t \in [p,1]$.

Proof of the Theorem for Case 1. Since h is continuous, we directly work with (2.2). From Lemma 3.4, there exist disjoint open intervals $(a_k, b_k),\ k \in K \subset \mathbb{N}$ on which $h \neq h^*$, and furthermore, p can be taken as $p = \sup_{k\in K} b_k < 1$. Note that $h(t) = h^*(t)$ for $t \in [p,1]$ and h^* is linear on each of $[a_k, b_k]$, $k \in K$. Define $\mathrm{I}_k = (a_k, b_k)$, $k \in K$. For some $U \sim \mathrm{U}[0,1]$, let

$$S_n^c = F_1^{-1}(U) + \cdots + F_n^{-1}(U), \tag{3.5}$$

and

$$R_n = \begin{cases} F_1^{-1}(U) + \cdots + F_n^{-1}(U), & \text{if } U \notin \cup_{k\in K}\mathrm{I}_k, \\ \mathbb{E}\left[F_1^{-1}(U) + \cdots + F_n^{-1}(U) \mid U \in \mathrm{I}_k\right], & \text{if } U \in \mathrm{I}_k,\ k \in K. \end{cases} \tag{3.6}$$

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

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

we have

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

It follows that

$$\begin{aligned} \rho_{h^*}(S_n^c) - \rho_{h^*}(R_n) &= \int_0^p F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) - \int_0^p F_{R_n}^{-1}(t)\mathrm{d}h^*(t) \\ &= \sum_{k\in K} \left[\int_{a_k}^{b_k} F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) - \int_{a_k}^{b_k} F_{R_n}^{-1}(t)\mathrm{d}h^*(t)\right] = 0, \end{aligned} \tag{3.7}$$

that is, $\rho_{h^*}(S_n^c) = \rho_{h^*}(R_n)$. As $F_i^{-1}(U)$ is bounded for $U \in \mathrm{I}_k$, $k \in K$, by Lemma 3.3, for each k, we can find random variables $Y_{1k}, \dots, Y_{nk}$, independent of U, such that Y_{ik} is identically distributed as $F_i^{-1}(U)|U \in \mathrm{I}_k$, $i = 1, \dots, n$, and

$$\left| Y_{1k} + \cdots + Y_{nk} - \mathbb{E}\left[F_1^{-1}(U) + \cdots + F_n^{-1}(U) \mid U \in \mathrm{I}_k \right] \right| \leqslant \max_{i=1,\dots,n} \{F_i^{-1}(b_k) - F_i^{-1}(a_k)\}. \tag{3.8}$$

Let $X_i^* = F_i^{-1}(U)\mathrm{I}_{\{U \notin \cup_{k\in K}\mathrm{I}_k\}} + \sum_{k\in K} Y_{ik}\mathrm{I}_{\{U\in \mathrm{I}_k\}}$, $i = 1, \dots, n$. It is easy to check that $X_i^* \sim F_i$, $i = 1, \dots, n$. Denote by

$$S_n^* = X_1^* + \cdots + X_n^*. \tag{3.9}$$

Clearly, $S_n^* \in \mathcal{S}_n$ and

$$\left| R_n - S_n^* \right| \leqslant \max_{i=1,\dots,n} \{F_i^{-1}(b_k) - F_i^{-1}(a_k)\} \leqslant \max_{i=1,\dots,n} \left\{ F_i^{-1}(p) \right\}. \tag{3.10}$$

Since $h^* \leqslant h$ and ρ_{h^*} is coherent and hence subadditive, by Lemma 3.1, we have $\rho_h(S_n^*) \leqslant \rho_{h^*}(S_n^*) \leqslant \rho_{h^*}(S_n^c)$. Integration by parts yields

$$\begin{aligned} \int_0^p F_{R_n}^{-1}(t)\mathrm{d}h^*(t) - \int_0^p F_{R_n}^{-1}(t)\mathrm{d}h(t) &= \int_0^p (h(t) - h^*(t))\, \mathrm{d}F_{R_n}^{-1}(t) \\ &= \sum_{k\in K} \int_{(a_k,b_k)} (h(t) - h^*(t))\, \mathrm{d}F_{R_n}^{-1}(t) = 0 \end{aligned} \tag{3.11}$$

where the last equality follows since $F_{R_n}^{-1}(t)$ is a constant for t in each (a_k, b_k).

$$\begin{aligned} \rho_{h^*}(S_n^c) - \rho_h(S_n^*) &= \int_0^p F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) - \int_0^p F_{S_n^*}^{-1}(t)\mathrm{d}h(t) \quad \text{since } h(t) = h^*(t) \text{ on } [p, 1] \\ &= \left(\int_0^p F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) - \int_0^p F_{R_n}^{-1}(t)\mathrm{d}h^*(t) \right) + \left(\int_0^p F_{R_n}^{-1}(t)\mathrm{d}h^*(t) - \int_0^p F_{R_n}^{-1}(t)\mathrm{d}h(t) \right) \\ &\quad + \left(\int_0^p F_{R_n}^{-1}(t)\mathrm{d}h(t) - \int_0^p F_{S_n^*}^{-1}(t)\mathrm{d}h(t) \right) \\ &\leqslant \max_{i=1,\dots,n} \left\{ F_i^{-1}(p) \right\}, \end{aligned} \tag{3.12}$$

where the last inequality follows from (3.7), (3.10), and (3.11). Condition A2 implies that for any $\varepsilon > 0$, there exists $q > p$ such that

$$\sup_{i\in\mathbb{N}} \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t) < \varepsilon.$$

Hence, by noting that $h^*(q) < 1$,

$$\max_{i=1,\dots,n} \left\{ F_i^{-1}(p) \right\} \leqslant \max_{i=1,\dots,n} \left\{ F_i^{-1}(q) \right\} < \frac{\varepsilon}{1 - h^*(q)}. \tag{3.13}$$

By Condition A1, $\lim_{n\to\infty}\sum_{i=1}^n \rho_{h^*}(X_i)=\infty$. Therefore, as $n\to\infty$,

$$\left|\frac{\sup\{\rho_h(S): S\in\mathcal{S}_n\}}{\sup\{\rho_{h^*}(S): S\in\mathcal{S}_n\}}-1\right| \leqslant \frac{\max_{i=1,\ldots,n}\{F_i^{-1}(p)\}}{\sum_{i=1}^n \rho_{h^*}(X_i)} \to 0. \tag{3.14}$$

The desired result follows. □

From the above proof, we can see that for this nice case, Condition A1 can be weakened to $\lim_{n\to\infty}\sum_{i=1}^n \rho_{h^*}(X_i)=\infty$ and Condition A2 can be weakened to $\max_{i=1,\ldots,n}\left\{F_i^{-1}(p)\right\}<\infty$. Conditions A1 and A2 in full power will be used in the proof for other cases discussed in Appendix. For Case 1, indeed we can give a more intuitive condition which is also easy to verify.

Condition A3. For a pre-assigned $p\in(0,1)$,

$$\lim_{n\to\infty}\frac{\max_{i=1,\ldots,n}\left\{\mathrm{VaR}_p(X_i)\right\}}{\sum_{i=1}^n \mathrm{VaR}_p(X_i)}=0. \tag{3.15}$$

Condition A3 simply says that there is no single risk which dominates the sum of all other risks in terms of VaR_p, a reasonable assumption for a joint model of high dimension. A3 is not strictly comparable to A1 and A2, but it has an important merit: it does not depend on h or h^* except for a point $p\in(0,1)$ given beforehand, which may be based on h and h^*. For a practical choice of $\{F_i,\ i\in\mathbb{N}\}$, it is often that (3.15) holds for all $p\in(0,1)$.

Theorem 3.6. *Suppose that $h\in\mathcal{H}$ is continuous and there exists $p\in(0,1)$ such that $h(t)=h^*(t)$ for all $t\in[p,1]$. For a sequence of distribution functions $\{F_i,\ i\in\mathbb{N}\}$ supported in $\mathbb{R}_+$ satisfying Condition A3, we have*

$$\lim_{n\to\infty}\frac{\sup\{\rho_h(S): S\in\mathcal{S}_n\}}{\sup\{\rho_{h^*}(S): S\in\mathcal{S}_n\}}=1, \tag{3.16}$$

where h^* is defined in (3.1).

Proof. Following the same proof in Case 1 of the above theorem, we obtain

$$0\leqslant \rho_{h^*}(S_n^c)-\rho_h(S_n^*)\leqslant \max_{i=1,\ldots,n}\left\{\mathrm{VaR}_p(X_i)\right\}.$$

Since

$$\rho_{h^*}(X_i)=\int_0^1 \mathrm{VaR}_t(X_i)\mathrm{d}h^*(t)\geqslant \int_p^1 \mathrm{VaR}_t(X_i)\mathrm{d}h^*(t)\geqslant \mathrm{VaR}_p(X_i)(1-h^*(p)),$$

we have

$$\left|\frac{\sup\{\rho_h(S): S\in\mathcal{S}_n\}}{\sup\{\rho_{h^*}(S): S\in\mathcal{S}_n\}}-1\right|\leqslant \frac{\max_{i=1,\ldots,n}\{\mathrm{VaR}_p(X_i)\}}{\sum_{i=1}^n \rho_{h^*}(X_i)}\leqslant \frac{\max_{i=1,\ldots,n}\{\mathrm{VaR}_p(X_i)\}}{(1-h^*(p))\sum_{i=1}^n \mathrm{VaR}_p(X_i)}\to 0 \text{ as } n\to\infty$$

by (3.15). □

Remark 3.1. The worst-case dependence structure for general distortion risk measures is revealed via the construction of S_n^*. For $n \to \infty$, to obtain a sum of S_n^* one needs comonotonicity on the set $(\bigcup_{k\in K} \mathrm{I}_k)^c$ and an *extreme negative dependence* conditional on each of the intervals I_k, $k \in K$. For a fixed n, the worst-case dependence structure for a general distortion risk measure is still not clear, since an extreme negative dependence may not be properly defined for fixed n unless the marginal distributions satisfies a notion of *joint mixability*; see Puccetti and Wang (2015) for related discussions on the above two notions of negative dependence.

3.3 Remarks on the conditions

In addition to Examples 2.1 and 2.2, we give a more subtle example to show that the uniform integrability condition A2 is essential. We compare our conditions with the ones in Embrechts et al. (2015) for VaR and ES. Theorem 3.3 of Embrechts et al. (2015) shows that

$$\lim_{n\to\infty} \frac{\sup\left\{\mathrm{VaR}_p(S) : S \in \mathcal{S}_n\right\}}{\sup\left\{\mathrm{ES}_p(S) : S \in \mathcal{S}_n\right\}} = 1, \tag{3.17}$$

if for $X_i \sim F_i$, $i \in \mathbb{N}$, the following two conditions are satisfied:

(a*) $\sup_{i\in\mathbb{N}} \mathbb{E}[|X_i|^k] < \infty$ for some $k > 1$,

(b*) $\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathrm{ES}_p(X_i) > 0$.

A natural question is whether k in (a*) can be taken as 1, that is,

(a') $\sup_{i\in\mathbb{N}} \mathbb{E}[|X_i|] < \infty$.

In comparison with the conditions in Embrechts et al. (2015), another question is whether A2 in this paper can be weakened to

(A2') $\sup_{i\in\mathbb{N}} \rho_{h^*}(X_i) < \infty$.

For the pair $(\rho_h, \rho_{h^*}) = (\mathrm{VaR}_p, \mathrm{ES}_p)$, (b*) is equivalent to our condition A1, and (a') is equivalent to A2' if we only consider $\mathcal{X} = L^+$.

The answer to both questions turns out to be negative. In the following example, Conditions A1and A2' are satisfied; in other words, conditions (a') and (b*) are satisfied. We will see that (3.17) fails to hold for all $p \in (0, 1)$.

Example 3.1. Suppose that the probability space is the Lebesgue unit interval $([0,1],\mathcal{B}([0,1]),\mathbb{P})$, where $\mathbb{P}$ is the Lebesgue measure. For $i\in\mathbb{N}$, let

$$F_i(x)=\begin{cases}0 & \text{if } x<0,\\ 1-\frac{1}{i^2} & \text{if } 0\leqslant x<i^2,\\ 1 & \text{if } i^2\leqslant x.\end{cases}$$

Clearly the support of F_i is nonnegative, $i\in\mathbb{N}$. One can calculate

$$\mathrm{VaR}_\alpha(X_i)=i^2\mathrm{I}_{\{\alpha\in(1-1/i^2,1)\}},\quad i\in\mathbb{N}.$$

For $X_i\sim F_i, i\in\mathbb{N}$, $\sup_{i\in\mathbb{N}}\mathbb{E}[X_i]=1<\infty$. One can also check that for $i\geqslant 1/\sqrt{1-p}$, $\mathrm{ES}_p(X_i)=\frac{1}{1-p}$. As a consequence,

$$\lim_{n\to\infty}\sup\left\{\mathrm{ES}_p\left(S/n\right):\ S\in\mathcal{S}_n\right\}=\lim_{n\to\infty}\frac{\sum_{i=1}^n\mathrm{ES}_p(X_i)}{n}=\frac{1}{1-p}.$$

Thus, (a'), (b*), A1 and A2' are all satisfied.

Next we will show that

$$\lim_{n\to\infty}\sup\left\{\mathrm{VaR}_p(S/n):S\in\mathcal{S}_n\right\}=0.$$

Note that $\sum_{i=1}^\infty\frac{1}{i^2}<\infty$. For any $\varepsilon>0$, which we choose as $\varepsilon=1-p$, there exists an N such that for $n\geqslant N$, we have

$$\sum_{i=n}^{\infty}\frac{1}{i^2}<\varepsilon. \tag{3.18}$$

Take a fixed number $k>N$ such that $\sum_{i=1}^N i^2<k^2$, we have for any $n>N$,

$$\begin{aligned}\mathbb{P}(S_n>k^2)&=\mathbb{P}(X_1+\cdots+X_N+X_{N+1}+\cdots+X_n>k^2)\\ &\leqslant\mathbb{P}(\text{at least one } X_i>0,\ i=N+1,\dots,n)\\ &\leqslant\sum_{i=N+1}^{n}\mathbb{P}(X_i>0)=\sum_{i=N+1}^{n}\frac{1}{i^2}<\varepsilon.\end{aligned}$$

Thus, $\mathrm{VaR}_{1-\varepsilon}(S_n)\leqslant k^2$. Therefore,

$$0\leqslant\lim_{n\to\infty}\sup\left\{\mathrm{VaR}_p(S/n):S\in\mathcal{S}_n\right\}=\lim_{n\to\infty}\frac{\sup\left\{\mathrm{VaR}_p(S):S\in\mathcal{S}_n\right\}}{n}\leqslant\lim_{n\to\infty}\frac{k^2}{n}=0.$$

In summary,

$$\lim_{n\to\infty}\frac{\sup\left\{\mathrm{VaR}_p(S):S\in\mathcal{S}_n\right\}}{\sup\left\{\mathrm{ES}_p(S):S\in\mathcal{S}_n\right\}}=0.$$

4 Asymptotic equivalence for convex risk measures

In this section we study asymptotic equivalence for convex risk measures. Compared to the previous section, the result in this section is much less technically involved since the worst-case dependence structure for convex risk measures is explicitly known as comonotonicity. We assume $\mathcal{X} = L^1$, since the canonical space for law-invariant convex risk measures is L^1; see Filipović and Svindland (2012).

4.1 Some lemmas

First, we recall the Kusuoka representation of law-invariant convex risk measures as established in Frittelli and Rosazza Gianin (2005) for $\mathcal{X} = L^\infty$. The extension of the representation to L^p, $p \in [1, \infty)$ is established in Svindland (2009). The Fatou property (FP) has to be assumed throughout for the representation to hold.

Lemma 4.1 (Lemma 2.14 of Svindland (2009))**.** *A law-invariant convex risk measure ρ mapping L^1 to $\mathbb{R}$ with the Fatou property has a representation*

$$\rho(X) = \sup_{\mu \in \mathcal{P}} \left\{ \int_0^1 \mathrm{ES}_p(X) \mathrm{d}\mu(p) - v(\mu) \right\}, \quad X \in L^1, \tag{4.1}$$

where $\mathcal{P}$ is the set of all probability measures on $[0,1]$ and v is a function from $\mathcal{P}$ to $\mathbb{R} \cup \{+\infty\}$, called a penalty function of ρ.

From now on, we denote by ρ^v a convex risk measure with penalty function v which maps L^1 to $\mathbb{R}$. For a law-invariant convex risk measure, without loss of generality we can assume $\rho^v(0) = 0$, or equivalently, in (4.1), $\inf\{v(\mu) : \mu \in \mathcal{P}\} = 0$. If one is interested in a law-invariant convex risk measure ρ with $\rho(0) = c \neq 0$, one can define $\tilde{\rho}(\cdot) = \rho(\cdot) - c$ so that $\tilde{\rho}$ is a law-invariant convex risk measure and $\tilde{\rho}(0) = 0$. A result on $\tilde{\rho}$ would simply lead to a result on ρ.

Similarly to the case of distortion risk measures, a convex risk measure is dominated by a coherent risk measure. The following simple lemma is a combination of Theorem 4.1 and Corollary 4.2 of Wang et al. (2015).

Lemma 4.2 (Wang et al. (2015))**.** *The smallest law-invariant coherent risk measure dominating ρ^v exists, and it is given by*

$$\rho^{v*}(X) = \sup_{\mu \in \mathcal{P}_v} \left\{ \int_0^1 \mathrm{ES}_p(X) \mathrm{d}\mu(p) \right\}, \quad X \in L^1, \tag{4.2}$$

where $\mathcal{P}_v = \{\mu \in \mathcal{P} : v(\mu) < +\infty\}$.

Remark 4.1. A popular subclass of law-invariant convex risk measures is the class of *convex shortfall risk measures* in Föllmer and Schied (2011). It is shown that for all convex shortfall risk measures ρ^v, the smallest dominating coherent risk measure ρ^{v*} is always a coherent *expectile*; see Proposition 4.3 of Wang et al. (2015).

Unlike the case of general distortion risk measures, the dependence structure of $(X_1, \ldots, X_n)$ which gives the maximum value of $\rho^v(X_1+\cdots+X_n)$ for given marginal distributions is always comonotonicity. Hence, an explicit expression of $\sup\{\rho^v(S_n) : S_n \in \mathcal{S}_n\}$ can be obtained. This creates a huge technical convenience to study asymptotic equivalence for convex risk measures.

Lemma 4.3. *For a sequence of distribution functions* $\{F_i,\ i \in \mathbb{N}\}$*,*

$$\sup\{\rho^v(S) : S \in \mathcal{S}_n\} = \sup_{\mu\in\mathcal{P}}\left\{\sum_{i=1}^n \int_0^1 \mathrm{ES}_p(X_i)\mathrm{d}\mu(p) - v(\mu)\right\}, \tag{4.3}$$

where $X_i \sim F_i$*,* $i = 1, \ldots, n$*.*

Proof. Let $Y_1, \ldots, Y_n \in L^1$ be comonotonic random variables such that $Y_i \sim F_i$, $i = 1, \ldots, n$. We have $\rho^v(X_1 + \cdots + X_n) \leqslant \rho^v(Y_1 + \cdots + Y_n)$; see Lemma 5.2 of Bäuerle and Müller (2006). It follows from Lemma 4.1 that

$$\begin{aligned}\sup\{\rho^v(S) : S \in \mathcal{S}_n\} &= \rho^v(Y_1 + \cdots + Y_n)\\ &= \sup_{\mu\in\mathcal{P}}\left\{\int_0^1 \mathrm{ES}_p\left(\sum_{i=1}^n Y_i\right)\mathrm{d}\mu(p) - v(\mu)\right\}\\ &= \sup_{\mu\in\mathcal{P}}\left\{\sum_{i=1}^n \int_0^1 \mathrm{ES}_p(Y_i)\,\mathrm{d}\mu(p) - v(\mu)\right\}.\end{aligned}$$

We obtain (4.3) since $\mathrm{ES}_p(X_i) = \mathrm{ES}_p(Y_i)$, $p \in (0, 1)$, $i = 1, \ldots, n$. □

Lemma 4.4. *For given* $\varepsilon > 0$*,* $n \in \mathbb{N}$*, and a sequence of distribution functions* $\{F_i,\ i \in \mathbb{N}\}$ *such that* $\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\} < \infty$*, there exists* $\mu_n \in \mathcal{P}_v$ *such that*

$$\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\} - \sum_{i=1}^n \int_0^1 \mathrm{ES}_p(X_i)\mathrm{d}\mu_n(p) < \varepsilon, \tag{4.4}$$

where $X_i \sim F_i$*,* $i \in \mathbb{N}$*.*

Proof. By applying Lemma 4.3 to the coherent risk measure ρ^{v*}, we obtain

$$\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\} = \sup_{\mu\in\mathcal{P}_v}\left\{\sum_{i=1}^n \int_0^1 \mathrm{ES}_p(X_i)\mathrm{d}\mu(p)\right\}.$$

By definition, there exists $\mu_n \in \mathcal{P}_v$ such that (4.4) holds. □

4.2 Asymptotic equivalence for convex risk measures

Similarly to Section 3, we need to assume some conditions on a sequence of distribution functions $\{F_i,\ i \in \mathbb{N}\}$ for the result of asymptotic equivalence to hold. In the following, $X_i \sim F_i$, $i \in \mathbb{N}$.

Condition B1. $\sum_{i=1}^n \mathbb{E}[X_i] \to \infty$ as $n \to \infty$.

Condition B2. $\rho^{v*}(\sum_{i=1}^n F_i^{-1}(U)) < \infty$ for some $U \sim \mathrm{U}[0,1]$ and all $n \in \mathbb{N}$.

Condition B3. There exist $\varepsilon > 0$ and a sequence $\mu_n \in \mathcal{P}_v,\ n \in \mathbb{N}$ satisfying (4.4), such that

$$\lim_{n\to\infty} \frac{v(\mu_n)}{\sum_{i=1}^n \int_0^1 \mathrm{ES}_p(X_i)\mathrm{d}\mu_n(p)} = 0.$$

Condition B1 is assumed to avoid the vanishing risks in Example 2.1. Condition B2 is trivial since we need the denominator in the asymptotic equivalence (1.2) to be finite for any given n. Condition B3 is somewhat an artificial technical condition to guarantee the convergence in our proof. Note that if $v(\mu)$ is bounded for $\mu \in \mathcal{P}_v$, then B3 is automatically satisfied whenever B1 holds.

Theorem 4.5. *Given a sequence of distribution functions $\{F_i,\ i \in \mathbb{N}\}$ satisfying Conditions B1-B3, we have*

$$\lim_{n\to\infty} \frac{\sup\{\rho^v(S) : S \in \mathcal{S}_n\}}{\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\}} = 1. \tag{4.5}$$

Proof. First note that for any $S_n \in \mathcal{S}_n$, due to Lemma 4.1 and B2, we have

$$\infty > \rho^{v*}(S_n) \geqslant \rho^v(S_n) \geqslant \sum_{i=1}^n \mathbb{E}[X_i],$$

and hence both $\sup\{\rho^v(S) : S \in \mathcal{S}_n\}$ and $\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\}$ are positive for large n, and

$$\lim_{n\to\infty} \frac{\sup\{\rho^v(S) : S \in \mathcal{S}_n\}}{\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\}} \leqslant 1. \tag{4.6}$$

Write $\lambda_n = \sum_{i=1}^n \int_0^1 \mathrm{ES}_p(X_i)\mathrm{d}\mu_n(p) \geqslant \sum_{i=1}^n \mathbb{E}[X_i]$. We have $\lambda_n \to \infty$ as $n \to \infty$ from Condition B1. From Lemmas 4.3 and 4.4, we have

$$\lim_{n\to\infty} \frac{\sup\{\rho^v(S) : S \in \mathcal{S}_n\}}{\sup\{\rho^{v*}(S) : S \in \mathcal{S}_n\}} \geqslant \lim_{n\to\infty} \frac{\lambda_n - v(\mu_n)}{\lambda_n + \varepsilon} = \lim_{n\to\infty} \frac{\lambda_n}{\lambda_n + \varepsilon} = 1. \tag{4.7}$$

Combining (4.6) and (4.7) we obtain (4.5). □

5 Conclusion

In this paper, we show that the asymptotic equivalence of VaR and ES in Embrechts et al. (2015) and preceding papers can be generalized to general risk measures for inhomogeneous models under some regularity conditions. The risk measures that we study include the class of distortion risk measures and the class of convex risk measures. The main result in this paper is that under dependence uncertainty in the aggregation of a large number of risks, the worst-case value of a non-coherent risk measure is asymptotically equivalent to that of a corresponding coherent risk measure. This result helps to analyze risk aggregation under dependence uncertainty in financial regulation and internal risk management.

A Full proof of Theorem 3.5

Proof. We show the theorem in two steps. First we assume that h is continuous, and then we approximate the general case by the result for continuous h.

For some intervals $\{\mathrm{I}_k,\ k \in K\}$ which will be specified later, let S_n^c, R_n, and S_n^* be as defined in (3.5), (3.6) and (3.9).

The proof in the case of continuous $h \in \mathcal{H}$.

Depending on the set $\{t \in [0,1] : h(t) \neq h^*(t)\}$, we have the following three cases:

Case 1: For some $p \in (0,1)$, $h(t) = h^*(t)$ for all $t \in [p,1]$. This case is dealt with in Section 3.

Case 2: $h \neq h^*$ in the intervals (a_k, b_k), $k \in K \subset \mathbb{N}$, where $\sup_{k\in K} b_k = 1$. Moerover, for all $p \in (0,1)$, there exist $t_0, t_1 \in (p,1)$ such that $h^*(t_0) = h(t_0)$ and $h^*(t_1) \neq h(t_1)$.

Condition A2 and the above property of h and h^* impliy that for any $\varepsilon > 0$, there exists q such that

$$\text{(A.1)} \qquad \sup_{i\in\mathbb{N}} \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t) < \varepsilon \quad \text{and} \quad h(q) = h^*(q).$$

Let I_k in (3.6) be $(a_k, b_k) \cap [0,q]$, $k \in K$. Then $\rho_{h^*}(S_n^c) = \rho_{h^*}(R_n)$ and

$$\left|S_n^* - R_n\right| \leqslant \max_{i=1,\ldots,n} \{F_i^{-1}(q)\},$$

which implies

$$\left|\int_0^q F_{S_n^*}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{R_n}^{-1}(t)\mathrm{d}h(t)\right| \leqslant \max_{i=1,\ldots,n} \{F_i^{-1}(q)\}.$$

$$\left|\int_0^q F_{R_n}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{R_n}^{-1}(t)\mathrm{d}h^*(t)\right| = \left|F_{R_n}^{-1}(t)[h(t)-h^*(t)]\,\Big|_0^q - \int_0^q [h(t)-h^*(t)]\mathrm{d}F_{R_n}^{-1}(t)\right|$$
$$= \left|F_{R_n}^{-1}(q)[h(q)-h^*(q)] - \sum_{k\in K}\int_{\mathrm{I}_k}[h(t)-h^*(t)]\mathrm{d}F_{R_n}^{-1}(t)\right| = 0.$$

By (3.7),

$$\left|\int_0^q F_{R_n}^{-1}(t)\mathrm{d}h^*(t) - \int_0^q F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right| = \left|\sum_{k\in K}\left[\int_{\mathrm{I}_k} F_{R_n}^{-1}(t)\mathrm{d}h^*(t) - \int_{\mathrm{I}_k} F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right]\right| = 0.$$

Thus,

$$\begin{aligned}
&\left|\int_0^q F_{S_n^*}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right| \\
&\leqslant \left|\int_0^q F_{S_n^*}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{R_n}^{-1}(t)\mathrm{d}h(t)\right| + \left|\int_0^q F_{R_n}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{R_n}^{-1}(t)\mathrm{d}h^*(t)\right| \\
&+ \left|\int_0^q F_{R_n}^{-1}(t)\mathrm{d}h^*(t) - \int_0^q F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right| \\
&\leqslant \max_{i=1,\ldots,n}\{F_i^{-1}(q)\}.
\end{aligned} \tag{A.2}$$

On the other hand,

$$\left|\int_q^1 F_{S_n^*}^{-1}(t)\mathrm{d}h(t) - \int_q^1 F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right| \leqslant \int_q^1 F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) = \sum_{i=1}^n \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t). \tag{A.3}$$

By Condition A1, $s := \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \rho_{h^*}(X_i) > 0$. Then for the above $\varepsilon > 0$, there exists $N > 0$ such that for $n > N$,

$$\frac{\sum_{i=1}^n \rho_{h^*}(X_i)}{n} > s - \varepsilon. \tag{A.4}$$

Hence, for any $\varepsilon > 0$ and $n > \max\{N, 1/(1-h^*(q))\}$, from (A.1)–(A.4), we have

$$\begin{aligned}
\left|\frac{\sup\{\rho_h(S) : S \in \mathcal{S}_n\}}{\sup\{\rho_{h^*}(S) : S \in \mathcal{S}_n\}} - 1\right| &\leqslant \frac{\left|\rho_h(S_n^*) - \rho_{h^*}(S_n^c)\right|}{\sum_{i=1}^n \rho_{h^*}(X_i)} \\
&\leqslant \frac{\max_{i=1,\ldots,n}\{F_i^{-1}(q)\}}{\sum_{i=1}^n \rho_{h^*}(X_i)} + \frac{\sum_{i=1}^n \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t)}{\sum_{i=1}^n \rho_{h^*}(X_i)} \\
&\leqslant \frac{\varepsilon}{n(1-h^*(q))(s-\varepsilon)} + \frac{\varepsilon}{(s-\varepsilon)} \\
&\leqslant \frac{2\varepsilon}{(s-\varepsilon)}.
\end{aligned}$$

Since ε is arbitrary, (3.4) follows.

Case 3: $h \neq h^*$ in the intervals (a_k, b_k), $k \in K \subset \mathbb{N}$, where $\sup_{k\in K} b_k = 1$. Moreover, there exists a $p \in (0,1)$ such that $h(t) \neq h^*(t)$ for all $t \in [p,1)$ and h^* is linear on $[p,1]$ with slope $c > 0$.

Recall that $h(1^-) = h(1) = 1$ and $h^*(1^-) = h^*(1) = 1$. For any $\varepsilon > 0$, take $q \in [p,1]$ such that

$$|h(q) - 1| < \frac{\varepsilon}{2}, \qquad \left|h^*(q) - 1\right| < \frac{\varepsilon}{2}, \tag{A.5}$$

$$\sup_{i\in\mathbb{N}} \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t) < c\varepsilon. \tag{A.6}$$

(A.6) implies that

$$(1-q)\sup_{i\in\mathbb{N}} F_i^{-1}(q) < \varepsilon.$$

Let I_k in (3.6) be $(a_k, b_k) \cap [0,q]$. Then $\rho_{h^*}(S_n^c) = \rho_{h^*}(R_n)$. Similarly to Case 2, we have

$$\left|\int_0^q F_{S_n^*}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{R_n}^{-1}(t)\mathrm{d}h(t)\right| \leqslant \max_{i=1,\ldots,n}\{F_i^{-1}(q)\},$$

$$\left|\int_0^q F_{R_n}^{-1}(t)\mathrm{d}h^*(t) - \int_0^q F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right| = 0,$$

$$\left|\int_q^1 F_{S_n^*}^{-1}(t)\mathrm{d}h(t) - \int_q^1 F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t)\right| \leqslant \sum_{i=1}^n \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t).$$

Moreover,

$$\begin{aligned}\left|\int_0^q F_{R_n}^{-1}(t)\mathrm{d}h(t) - \int_0^q F_{R_n}^{-1}(t)\mathrm{d}h^*(t)\right| &= F_{R_n}^{-1}(q)\left|h(q) - h^*(q)\right| \\ &\leqslant F_{R_n}^{-1}(q)\varepsilon = \varepsilon\sum_{i=1}^n \frac{\int_p^1 F_i^{-1}(t)\mathrm{d}t}{1-p},\end{aligned}$$

where the last inequality follows by (A.5). Thus, for any $\varepsilon > 0$, $n > \max\{N, 1/(1-q)\}$,

$$\begin{aligned}&\left|\frac{\sup\{\rho_h(S) : S \in \mathcal{S}_n\}}{\sup\{\rho_{h^*}(S) : S \in \mathcal{S}_n\}} - 1\right| \\ &\leqslant \frac{\max_{i=1,\ldots,n}\{F_i^{-1}(q)\}}{\sum_{i=1}^n \rho_{h^*}(X_i)} + \frac{\frac{\sum_{i=1}^n \int_p^1 F_i^{-1}(t)\mathrm{d}t}{1-p}\varepsilon}{\sum_{i=1}^n \rho_{h^*}(X_i)} + \frac{\sum_{i=1}^n \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t)}{\sum_{i=1}^n \rho_{h^*}(X_i)} \\ &\leqslant \frac{\varepsilon}{n(s-\varepsilon)(1-q)} + \frac{\frac{\varepsilon}{1-p}\sum_{i=1}^n \int_p^q F_i^{-1}(t)\mathrm{d}t + (\frac{\varepsilon}{1-p} + c)\sum_{i=1}^n \int_q^1 F_i^{-1}(t)\mathrm{d}t}{n(s-\varepsilon)} \\ &\leqslant \frac{\varepsilon}{n(s-\varepsilon)(1-q)} + \frac{\frac{\varepsilon n(q-p)}{1-p}\sup_{i\in\mathbb{N}} F_i^{-1}(q) + (\frac{\varepsilon}{1-p} + c)n\varepsilon}{n(s-\varepsilon)} \\ &\leqslant \frac{\varepsilon}{s-\varepsilon} + \frac{\varepsilon^2\frac{q-p}{(1-p)(1-q)} + (\frac{\varepsilon}{1-p} + c)\varepsilon}{s-\varepsilon}.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the result follows.

The proof in the case of general $h \in \mathcal{H}$.

Denote $\bar{\rho}_h(n) = \sup\{\rho_h(S) : S \in \mathcal{S}_n\}$ for any $h \in \mathcal{H}$. S_n^c is defined as in (3.5). Clearly $\rho_{h^*}(S_n^c) = \bar{\rho}_{h^*}(n)$. For any $h \in \mathcal{H}$, let $h_\delta \in \mathcal{H}$ be continuous such that $h_\delta \geqslant h$ on $[0,1]$ and $h_\delta \to h$ weakly as $\delta \to 0^+$. The existence of such h_δ is an exercise for mathematical analysis.

By Lemma A.5 of Wang et al. (2015), for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sup_{t\in[0,1]} \left| h_\delta^*(t) - h^*(t) \right| \leqslant \varepsilon. \tag{A.7}$$

Condition A2 implies that for any $\varepsilon > 0$, there exists $q \in (0,1)$ such that

$$\sup_{i\in\mathbb{N}} \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t) < \varepsilon.$$

Note that $\sup_{i\in\mathbb{N}} F_i^{-1}(q) < \frac{\varepsilon}{1-h^*(q)} < \infty$. Take $M = \sup_{i\in\mathbb{N}} F_i^{-1}(q)$. Then

$$\rho_{h^*}(S_n^c I_{\{S_n^c > Mn\}}) = \int_{\{S_n^c > Mn\}} F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) \leqslant \int_q^1 F_{S_n^c}^{-1}(t)\mathrm{d}h^*(t) = \sum_{i=1}^n \int_q^1 F_i^{-1}(t)\mathrm{d}h^*(t) \leqslant n\varepsilon.$$

Condition A1 implies that for $\varepsilon > 0$, there exists $N_1 \in \mathbb{N}$ and $s > 0$ such that for $n \geqslant N_1$, $\rho_{h^*}(S_n^c) > ns$. By comonotonic additivity and monotonicity of distortion risk measures,

$$\rho_{h^*}(S_n^c) = \rho_{h^*}(S_n^c \wedge (Mn)) + \rho_{h^*}((S_n^c - Mn)I_{\{S_n^c > Mn\}}) \leqslant \rho_{h^*}(S_n^c \wedge (Mn)) + n\varepsilon.$$

Thus,

$$\frac{\rho_{h^*}(S_n^c \wedge (Mn))}{\rho_{h^*}(S_n^c)} \geqslant 1 - \frac{\varepsilon}{s}, \qquad \text{for all } n \geqslant N_1. \tag{A.8}$$

Let $Y = S_n^c \wedge (Mn)$.

$$\begin{aligned}
\rho_{h^*}(S_n^c \wedge (Mn)) - \rho_{h_\delta^*}(S_n^c \wedge (Mn)) &= \int_0^1 F_Y^{-1}(t)\mathrm{d}h^*(t) - \int_0^1 F_Y^{-1}(t)\mathrm{d}h_\delta^*(t) \\
&= \int_0^1 F_Y^{-1}(t)\mathrm{d}(-(1-h^*(t))) - \int_0^1 F_Y^{-1}(t)\mathrm{d}(-(1-h_\delta^*(t))) \\
&= \int_0^1 [h_\delta^*(t) - h^*(t)]\mathrm{d}F_Y^{-1}(t) \leqslant \varepsilon M n,
\end{aligned}$$

where the last inequality follows from (A.7). Thus,

$$\frac{\rho_{h^*}(S_n^c \wedge (Mn)) - \rho_{h_\delta^*}(S_n^c \wedge (Mn))}{\rho_{h^*}(S_n^c \wedge (Mn))} \leqslant \frac{\varepsilon M n}{(1-\varepsilon/s)ns} = \frac{\varepsilon M}{s-\varepsilon},$$

which implies

$$\frac{\rho_{h_\delta^*}(S_n^c \wedge (Mn))}{\rho_{h^*}(S_n^c \wedge (Mn))} \geqslant 1 - \frac{\varepsilon M}{s-\varepsilon} \qquad \text{for all } n \geqslant N_1.$$

Since $\rho_{h^*_\delta}(S^c_n \wedge (Mn)) \leqslant \rho_{h^*_\delta}(S^c_n)$ and by the above inequality, we have

$$\frac{\rho_{h^*_\delta}(S^c_n)}{\rho_{h^*}(S^c_n \wedge (Mn))} \geqslant 1 - \frac{\varepsilon M}{s-\varepsilon} \quad \text{for all } n \geqslant N_1. \tag{A.9}$$

From the first half of the proof, for any $\varepsilon > 0$, there exists $N_2 \in \mathbb{N}$ such that for $n \geqslant N_2$,

$$\frac{\bar\rho_{h_\delta}(n)}{\bar\rho_{h^*_\delta}(n)} \geqslant 1 - \varepsilon. \tag{A.10}$$

Thus for any $\varepsilon > 0$, there exist $\delta > 0$ and $N = N_1 \vee N_2$ such that for $n \geqslant N$,

$$\begin{aligned}
\frac{\bar\rho_{h_\delta}(n)}{\rho_{h^*}(S^c_n)} &= \frac{\bar\rho_{h_\delta}(n)}{\bar\rho_{h^*_\delta}(n)} \times \frac{\bar\rho_{h^*_\delta}(n)}{\rho_{h^*}(S^c_n \wedge (Mn))} \times \frac{\rho_{h^*}(S^c_n \wedge (Mn))}{\rho_{h^*}(S^c_n)} \\
&= \frac{\bar\rho_{h_\delta}(n)}{\bar\rho_{h^*_\delta}(n)} \times \frac{\rho_{h^*_\delta}(S^c_n)}{\rho_{h^*}(S^c_n \wedge (Mn))} \times \frac{\rho_{h^*}(S^c_n \wedge (Mn))}{\rho_{h^*}(S^c_n)} \\
&\geqslant (1-\varepsilon)\left(1 - \frac{\varepsilon M}{s-\varepsilon}\right)\left(1 - \frac{\varepsilon}{s}\right) \\
&\geqslant 1 - \left(1 + \frac{M}{s-\varepsilon} + \frac{1}{s}\right)\varepsilon,
\end{aligned}$$

where the inequality follows from (A.8-A.10). Note that $\bar\rho_h(n) \geqslant \bar\rho_{h_\delta}(n)$. For any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for $n \geqslant N$,

$$\frac{\bar\rho_h(n)}{\bar\rho_{h^*}(n)} \geqslant 1 - \left(1 + \frac{M}{s-\varepsilon} + \frac{1}{s}\right)\varepsilon,$$

that is,

$$\lim_{n\to\infty} \frac{\sup\{\rho_h(S) : S \in \mathcal{S}_n\}}{\sup\{\rho_{h^*}(S) : S \in \mathcal{S}_n\}} = 1.$$

□

References

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.

Bäuerle, N. and Müller, A. (2006): Stochastic orders and risk measures: Consistency and bounds. *Insurance: Mathematics and Economics*, **38**(1), 132–148.

BCBS (2013): *Consultative Document October 2013. Fundamental review of the trading book: A revised market risk framework.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Bernard, C., He, X., Yan, J. A. and Zhou, X. Y. (2015): Optimal insurance design under rank-dependent expected utility. *Mathematical Finance*, **25**, 154–186.

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

Bernard, C., Rüschendorf, L. and Vanduffel, S. (2016): VaR bounds with variance constraint. *Journal of Risk and Insurance*, forthcoming.

Bernard, C. and Vanduffel, S. (2015): A new approach to assessing model risk in high dimensions, *Journal of Banking and Finance*, **58**, 166–178.

Bignozzi, V., Puccetti, G. and Rüschendorf L. (2015): Reducing model risk via positive and negative dependence assumptions. *Insurance: Mathematics and Economics*, **61**(1), 17–26.

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

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

Delbaen, F. (2012): *Monetary Utility Functions*. Osaka University Press, Osaka.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R. and Vyncke, D. (2002): The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, **31**(1), 3–33.

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

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. (2011): *Stochastic Finance: An Introduction in Discrete Time.* Third Edition. Walter de Gruyter, Berlin.

Föllmer, H. and Weber, S. (2015): The axiomatic approach to risk measures for capital determination. *The Annual Review of Financial Economics*, **7**, 301–337.

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

Frittelli M. and Rosazza Gianin E. (2005): Law invariant convex risk measures. *Advances in Mathematical Economics*, **7**, 33–46.

He, X. and Zhou, X. Y. (2016): Hope, fear and aspirations. *Mathematical Finance*, forthcoming.

IAIS (2014): *Consultation Document December 2014. Risk-based global insurance capital standard*. International Association of Insurance Supervisors.

Jakobsons, E., Han, X. and Wang, R. (2016): General convex order on risk aggregation. *Scandinavian Actuarial Jounral*, forthcoming.

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

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

Puccetti, G. and Rüschendorf, L. (2014): Asymptotic equivalence of conservative VaR- and ES-based capital charges. *Journal of Risk*, **16**(3), 3–22.

Puccetti, G., Wang, B. and Wang, R. (2013): Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics*, **53**(3), 821–828.

Puccetti, G. and Wang R. (2015): Extremal dependence concepts. *Statistical Science*, **30**(4), 485–517.

Quiggin, J. (1993): *Generalized Expected Utility Theory: The Rank-dependent Model.* Kluwer, the Netherlands.

Svindland, G. (2009): Convex risk measures beyond bounded risks. *Doctoral dissertation*, Ludwig Maximilian Univeristy.

Wang, B. and Wang, R. (2011): The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis*, **102**(10), 1344–1360.

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. 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.