City Research Online

City, University of London Institutional Repository

Citation: Asimit, A.V. & Li, J. (2016). Extremes for coherent risk measures. Insurance: Mathematics and Economics, 71, pp. 332-341. doi: 10.1016/j.insmatheco.2016.10.003

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/16268/

Link to published version: https://doi.org/10.1016/j.insmatheco.2016.10.003

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online: http://openaccess.city.ac.uk/ publications@city.ac.uk

Extremes for Coherent Risk Measures

Alexandru V. Asimit
Cass Business School, City University London,
London EC1Y 8TZ, United Kingdom.
E-mail: asimit@city.ac.uk

Jinzhu Li[1]
School of Mathematical Science and LPMC,
Nankai University, Tianjin 300071, P.R. China.
E-mail: lijinzhu@nankai.edu.cn

September 9, 2016

Abstract. Various concepts appeared in the existing literature to evaluate the risk exposure of a financial or insurance firm/subsidiary/line of business due to the occurrence of some extreme scenarios. Many of those concepts, such as Marginal Expected Shortfall or Tail Conditional Expectation, are simply some conditional expectations that evaluate the risk in adverse scenarios and are useful for signaling to a decision-maker the poor performance of its risk portfolio or to identify which sub-portfolio is likely to exhibit a massive downside risk. We investigate the latter risk under the assumption that it is measured via a coherent risk measure, which obviously generalizes the idea of only taking the expectation of the downside risk. Multiple examples are given and our numerical illustrations show how the asymptotic approximations can be used in the capital allocation exercise. We have concluded that the expectation of the downside risk does not fairly take into account the individual risk contribution when allocating the VaR-based regulatory capital, and thus, more conservative risk measurements are recommended. Finally, we have found that more conservative risk measurements do not improve the fairness of the cost of capital allocation when the uncertainty with parameter estimation is present, even at a very high level.

Keywords and phrases: Capital allocation; Coherent/Distortion risk measure; Conditional Tail Expectation; Extreme Value Theory; Marginal Expected Shortfall; Rapid Variation; Regular Variation.

Mathematics Subject Classification: Primary 62P05; Secondary 62H20, 60E05.

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and denote by $L_+(\mathbb{P})$ the set of non-negative random variables. Consider $X, Y \in L_+(\mathbb{P})$ two random insurance risks possessing *distribution functions (df)* F and G, respectively. The corresponding survival functions are $\overline{F} := 1 - F$ and $\overline{G} := 1 - G$. It is assumed

[1]Corresponding author. Phone: +86-22-2350 1233.

that the decision-maker orders its preferences via a *risk measure*, ρ, which is simply a functional, i.e. $\rho : L_+(\mathbb{P}) \to \Re$.

There is a massive literature on risk measures with many excellent references, but comprehensive discussions may be found in Denuit *et al.* (2005) and Föllmer and Schied (2011). The choice of a risk measure is usually subjective, but *Value-at-Risk (VaR)* and *Tail-Value-at-Risk (TVaR)* represent the most known risk measures used in insurance regulation. Solvency II and Swiss Solvency Test are the regulatory regimes for all (re)insurance companies that operate within the European Union and Switzerland, respectively, and their capital requirements are solely based on VaR and TVaR. For these reasons and not only, these standard risk measures have received special attention by academics, practitioners and regulators, and therefore, vivid discussions have risen among them. VaR is criticized for its lack of sub-additivity, while VaR may create regulatory arbitrage in an insurance group (for details, see Asimit *et al.*, 2013). A detailed discussion on possible regulatory arbitrages in a TVaR-based regime is provided in Koch-Medina and Munari (2016). An interesting discussion about the pros and cons of the two risk measures is detailed in the recent paper of Emmer *et al.* (2015), but the general conclusion is that there is no evidence for global advantage of one risk measure against the other.

While the basic understanding of a risk measure has been well-accepted by academics and practitioners for many decades, the mathematical formulation has become a major topic in mathematical finance literature for almost two decades. An important contribution in this direction is given in the seminal paper by Artzner *et al.* (1999), which introduces the concept of *coherent risk measures* by providing core economical and mathematical reasoning to support their formulation. The class of coherent risk measures requires that their members are *monotone, translation invariant, positive homogeneous* and *sub-additive* (for details, see Artzner et al., 1999). The subclass of coherent risk measures that possesses the *law-invariant* and *comonotonic* property is known as the *distortion risk measures* (see Wang *et al.*, 1997). While law-invariance is desirable in practice as it is required for the risk measure to be identifiable via the empirical data when designing statistical inference methodology, the presence of the comonotonic property could be debatable. Fischer (2003) defines a large subclass of coherent risk measures that are not comonotonic. An example of a coherent risk measure, namely *Worst Conditional Expectation*, that is not law-invariant is given by Artzner *et al.* (1999).

Academics, practitioners and regulators have considered various ways in measuring the sensitivity of the financial well-being of a financial or insurance firm/subsidiary/line of business as a result of an "observable" factor, especially when this factor is located in an extreme region. One popular example is the widely popular indicator *Marginal Expected Shortfall (MES)*, which in the financial literature is a popular statistical measure of systemic resilience in financial markets. The literature is quite rich and a comprehensive description is given by Idierb *et al.* (2014), where the practical advantages of this risk measure in detecting extreme risk exposures of financial firms are empirically explained. In mathematical terms, MES is a conditional expectation, $\mathbb{E}\big(X|Y > t\big)$, for large values of t. Clearly, extreme cases are viewed here by considering X and Y to be some future liabilities (consisting with our prior definitions of X and Y), rather than measuring the wealth, but changing to the left hand side of the real line is a simple exercise.

Non-parametric inferences for MES are present in Cai *et al.* (2015) via statistical extreme methodologies. See also Kulik and Soulier (2015), which shows that the model of Cai *et al.* (2015) can be enlarged significantly. It is useful to note that a variant of *Conditional Tail Expectation (CTE)* has the same representation like MES. CTE asymptotic approximations have appeared in various forms in the insurance and actuarial literature; Joe and Li (2011) focuses on distributions satisfying the multivariate regularly varying property, Hua and Joe (2011) and Zhu and Li (2012) investigate the same problem for scaled mixtures and multivariate elliptical distributions, while Hashorva *et al.* (2014) consider dependence models that exhibit the second order regularly varying tail property. The same problem is discussed in Asimit *et al.* (2011) for a variety of asymptotic dependence models, emphasizing that these extreme CTEs are useful when the total regulatory capital (based on TVaR) is allocated amongst many subsidiaries/lines of business (LOBs). The purpose of this paper is to evaluate asymptotic approximations for coherent risk measures of $\big(X|Y>t\big)$ in extreme regions, which incidentally generalizes the work of Asimit *et al.* (2011). It is quite apparent that our formulation may help the decision-maker to much better understand the downside risk and for this reason, we choose to investigate this problem.

The structure of the paper is as follows: the next section contains the necessary background and the mathematical formulation of our main aims, Section 3 provides the main results of the paper, Section 4 outlines various examples and provides some numerical illustrations of our findings, while all proofs are relegated to Section 5.

2. Preliminaries

It has been anticipated in Section 1 that we are interested in finding asymptotic approximations for coherent risk measures of conditional distributions around extreme regions. Therefore, we first define the two classes of risk measures that allow us to produce such computations. The first class is known as distortion risk measures which has the following mathematical formulation:

$$\rho(X;g)=\int_0^\infty g\big(\mathbb{P}(X>x)\big)\,\mathrm{d}x, \tag{2.1}$$

where $g:[0,1]\to[0,1]$ is a non-decreasing function such that $g(0)=0$ and $g(1)=1$, known as the *distortion function* (for example, see Denuit *et al.*, 2005 or Dhaene *et al.*, 2012). The second class is given in Fischer (2003) and is as follows

$$\rho(X;\theta,p)=\mathbb{E}X+\theta\left(\mathbb{E}\Big[\big(X-\mathbb{E}X\big)_+^p\Big]\right)^{1/p}, \tag{2.2}$$

where $\theta\in[0,1]$, $p\in[1,\infty)$ and $(x)_+=\max\{x,0\}$. Recall that distortion risk measures belong to the large class of coherent risk measures if g is concave. Our main aim is to approximate $\rho(\cdot;g)$ and $\rho(\cdot;\theta,p)$ for the conditional random variable $X|Y>t$ for large t. Therefore, we shall investigate the asymptotic approximation as $t\to\infty$ for

$$\rho(t;g)=\rho\big(X|Y>t;g\big) \tag{2.3}$$

and

$$\rho(t;\theta,p)=\rho\big(X|Y>t;\theta,p\big), \tag{2.4}$$

provided that the quantities defined in (2.3) and (2.4) are finite.

There are many choices for the distortion function (for example, see Jones and Zitikis, 2003 and 2007) and some of the well-known examples are as follows:

i) Dual-power: $g(s) = 1 - (1-s)^\beta, \quad \beta > 1$;
ii) TVaR: $g(s) = \min(s/(1-\beta), 1), \quad 0 < \beta < 1$;
iii) Gini: $g(s) = (1+\beta)s - \beta s^2, \quad 0 \leq \beta \leq 1$;
iv) Proportional hazard transform (PHT): $g(s) = s^{1-\beta}, \quad 0 \leq \beta < 1$;
v) Wang Transform: $g(s) = F_N\big(F_N^{-1}(s) + \lambda\big), \quad \lambda > 0$, where $F_N(\cdot)$ and $F_N^{-1}(\cdot)$ represent the standard normal df and its inverse, respectively.

Finding asymptotic approximations of any quantity of interest, including our specific tail risk measures defined in equations (2.3) and (2.4), requires knowledge about the behavior in the extreme region for which necessary background is provided next.

Let $\{X_i; i \geq 1\}$ be a sequence of independent and identically distributed random variables with common df F with an ultimate right tail, i.e. $\inf_{x \in \Re}\{F(x) = 1\} = \infty$. *Extreme Value Theory* (EVT) assumes that there are constants $a_n > 0$ and $b_n \in \Re$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(a_n\left(\max_{1\leq i\leq n} X_i - b_n\right) \leq x\right) = H(x), \quad x \in \Re.$$

In this case, H is called an *Extreme Value Distribution* and F is said to belong to the *max-domain of attraction of* H, denoted by $F \in \mathrm{MDA}(H)$. By the Fisher-Tippett Theorem (see Fisher and Tippett, 1928), if the limit distribution H is non-degenerate, then it is of one of the following two types: $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$ for all $x > 0$ with $\alpha > 0$, or $\Lambda(x) = \exp\big\{-\mathrm{e}^{-x}\big\}$ for all $x \in \Re$. In the first case, X has a *Fréchet* tail which is *regularly varying* at ∞ with index $-\alpha$, i.e.

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha}, \quad x > 0. \tag{2.5}$$

We signify the above by $F \in \mathcal{R}_{-\alpha}$. Thus, $F \in \mathrm{MDA}(\Phi_\alpha)$ if and only if $F \in \mathcal{R}_{-\alpha}$. In the second case, X has a Gumbel tail and it is well-known (for example, see Embrechts *et al.*, 1997) that there exists a positive measurable function $a(\cdot)$ such that

$$\lim_{t\to\infty} \frac{\overline{F}(t + xa(t))}{\overline{F}(t)} = \mathrm{e}^{-x}, \quad x \in \Re. \tag{2.6}$$

Equation (2.6) implies that X has a rapidly varying tail, written as $F \in \mathcal{R}_{-\infty}$, which by definition means that

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = 0, \quad x > 1.$$

Moreover, we say that X has a dominatedly varying tail, denoted by $F \in \mathcal{D}$, if

$$\liminf_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} > 0, \quad x > 1.$$

For further details, we refer the reader to Bingham *et al.* (1987) or Embrechts *et al.* (1997). Further, for a df F with an ultimate right tail, we define its lower Matuszewska index as

$$\alpha_F^* = \sup\left\{-\frac{\log \overline{F}^*(x)}{\log x} : x > 1\right\} \in [0, \infty],$$

where $\overline{F}^*(x) = \limsup_{t\to\infty} \overline{F}(tx)/\overline{F}(t)$. It is clear that $0 < \alpha_F^* \le \infty$ if and only if $\overline{F}^*(x) < 1$ for some $x > 1$. In this case, Proposition 2.2.1 of Bingham *et al.* (1987) tells us that for every $0 < \alpha' < \alpha_F^*$, there are some $K > 1$ and $t_0 > 0$ such that

$$\frac{\overline{F}(tx)}{\overline{F}(t)} \le K x^{-\alpha'} \tag{2.7}$$

holds for all $tx > t > t_0$. It is not difficult to see that if $F \in \mathcal{R}_{-\alpha}$ for some $\alpha \in [0,\infty]$ then $\alpha_F^* = \alpha$.

We end this section with a summary of notations used in this paper. Let $f_1(\cdot)$ and $f_2(\cdot)$ be two positive functions and let $\delta \in \{0,\infty\}$. We write $f_1(z) \sim f_2(z)$ to mean *strong equivalence as* $z \to \delta$, i.e. $\lim_{z\to\delta} f_1(z)/f_2(z) = 1$. Moreover, we write $f_1(z) = O\big(f_2(z)\big)$ and $f_1(z) = o\big(f_2(z)\big)$ if $\limsup_{z\to\delta} f_1(z)/f_2(z) < \infty$ and $\lim_{z\to\delta} f_1(z)/f_2(z) = 0$, respectively. For two real numbers a_1 and a_2, we write $a_1 \wedge a_2 = \min\{a_1, a_2\}$ and $a_1 \vee a_2 = \max\{a_1, a_2\}$. Finally, $\mathbf{1}_{\{\cdot\}}$ represents the indicator function.

3. Main Results

The aim of this section is to provide our main theoretical results. The initial step is to explain the assumptions under which the asymptotic approximations hold. The fist set of conditions is for the distortion function g defined in (2.3) and is stated as Assumption 3.1.

Assumption 3.1. *The distortion function g from* (2.3) *is such that*

$$\Omega_g = \big\{\beta > 0 : g(s) = O\big(s^{\beta}\ \big) \, as \ s \to 0\big\} \neq \varnothing.$$

Note that Assumption 3.1 implies the right continuity of g at 0. Moreover, Assumption 3.1 is very mild, since it is satisfied by all power functions s^β with $\beta > 0$ and all distortion functions analytic in a right neighborhood of 0 such as $g(s) = \big(1 - \mathrm{e}^{-s}\big)/\big(1 - \mathrm{e}^{-1}\big)$ and $g(s) = \log(1+s)/\log 2$. It is a simple exercise to check that each of Examples i)–v) given in Section 2 satisfies Assumption 3.1. In fact, one only needs to verify that Ω_g is non-empty, which is straightforward for Examples i)–iv). Example v) holds this property as well and it can be shown by taking $z = F_N^{-1}(s)$ and noting that for every $0 < \beta < 1$ we have that

$$\lim_{s\to 0} \frac{F_N\big(F_N^{-1}(s) + \lambda\big)}{s^\beta} = \lim_{z\to-\infty} \frac{F_N(z+\lambda)}{\big(F_N(z)\big)^\beta} = \lim_{z\to\infty} \frac{\overline{F}_N(z-\lambda)}{\Big(\overline{F}_N(z)\Big)^\beta}, \tag{3.1}$$

where the last step is due to $F_N(\cdot) = \overline{F}_N(-\cdot)$. The survival function of the standard normal distribution, $\overline{F}_N(\cdot) = 1 - F_N(\cdot)$, satisfies

$$\overline{F}_N(z) \sim \frac{1}{\sqrt{2\pi} z} \mathrm{e}^{-\frac{z^2}{2}}, \quad z \to \infty.$$

Plugging this result into (3.1) leads to $F_N\big(F_N^{-1}(s) + \lambda\big) = o\big(s^\beta\big)$ as $s \to 0$ for every $0 < \beta < 1$. Therefore, Example v) satisfies Assumption 3.1.

The next set of conditions summaries the required mild conditions for the joint asymptotic behavior of (X, Y) and is given as Assumption 3.2.

Assumption 3.2. *Let $X, Y \in L_+(\mathbb{P})$ such that their dfs, F and G, have an ultimate right tail. In addition, $\overline{F}(t) = O\big(\overline{G}(t)\big)$ as $t \to \infty$ and the limit*

$$\lim_{t\to\infty} \mathbb{P}(X > tx | Y > t) =: h(x) \in [0, 1] \tag{3.2}$$

exists almost everywhere for $x > 0$.

Define now $\Delta_a = \{x > a : h \text{ exists at } x\}$ for any $a \geq 0$. Clearly, Δ_0 is nothing but the definitional domain of h. In view of Assumption 3.2, $(a, \infty)\backslash\Delta_a$ is a null Lebesgue measure set and hence, Δ_a is dense in (a, ∞). Additionally, it follows immediately from (3.2) that $h(x)$ is non-increasing in Δ_0. These facts and the monotonicity of $\mathbb{P}\big(X > tx|Y > t\big)$ with respect to x imply that

(i) if $h(a) = 0$ for some $a \in \Delta_0$, then $\Delta_a = (a, \infty)$ with $h(x) = 0$ for all $x \in (a, \infty)$;
(ii) if $h(a) = 1$ for some $a \in \Delta_0$, then $(0, a) \subseteq \Delta_0$ with $h(x) = 1$ for all $x \in (0, a)$.

The next lemma gives some further properties of the function h, which is crucial for demonstrating all of our main results.

Lemma 3.1. *Let Assumption 3.2 hold.*

(i) *If $0 < \alpha_F^* \leq \infty$ then $\lim_{x\to\infty,\, x\in\Delta_0} h(x) = 0$.*
(ii) *We have that $h(x) = 0$ for all $x > 0$ if either*
  (a) *$\mathbb{P}(X > tx, Y > t) = O\big(\overline{F}(tx)\overline{G}(t)\big)$ as $t \to \infty$ for every $x > 0$ or*
  (b) *$1 \in \Delta_0$ with $h(1) = 0$ and $G \in \mathcal{D}$.*
(iii) *Assume $1 \in \Delta_0$ with $h(1) > 0$, then*
  (a) *$h(x_1) = 0$ for some $x_1 \in \Delta_1$, $h(x) = 0$ for all $x > 1$ and $F \in \mathcal{R}_{-\infty}$ are equivalent.*
  (b) *$h(x_1) > 0$ for some $x_1 \in \Delta_1$, $h(x) > 0$ for all $x \in \Delta_0$ and $F \in \mathcal{D}$ are equivalent.*
(iv) *If $1 \in \Delta_0$ with $h(1) = \lim_{t\to\infty} F(t)/G(t) > 0$ and $h(x_1) > 0$ for some $x_1 \in \Delta_1$, then $F \in \mathcal{R}_{-\alpha}$ for some $0 \leq \alpha < \infty$ and $h(x) = h(1)x^{-\alpha}$ for all $x \geq 1$.*

We are now ready to provide the first main result, stated as Theorem 3.1, which gives asymptotic approximations for distortion risk measures in extreme regions as defined in (2.3).

Theorem 3.1. *Consider the distortion risk measure $\rho(t; g)$ defined in (2.3) and let Assumptions 3.1 and 3.2 hold. If $1/\beta^* < \alpha_F^* \leq \infty$, where $\beta^* = \sup\big\{\Omega_g\big\}$ and $1/\infty = 0$ by convention, then*

$$\lim_{t\to\infty} \frac{\rho(t; g)}{t} = \int_0^\infty g\big(h(x)\big)\mathrm{d}x. \tag{3.3}$$

It is interesting to point out that a very weak asymptotic dependence amongst X and Y may lead to a non-informative approximation in Theorem 3.1. Specifically, under the condition of Lemma 3.1(ii), the limit displayed in (3.3) is 0, which is definitely not informative, unless a second order condition is imposed in (3.2) in order to better understand the rate of convergence of $\rho(t; g)/t$. This weak dependence occurs especially when concomitant extreme events for X and Y are not possible, which in the EVT lingo is known as *asymptotic independence*.

By Lemma 3.1 (iii)(a) and (iv), under Assumption 3.2, the case of $h(1) = \lim_{t\to\infty} \overline{F}(t)/\overline{G}(t) > 0$ can only occur when $F \in \mathcal{R}_{-\alpha}$ for some $0 \leq \alpha < \infty$ or $F \in \mathcal{R}_{-\infty}$, according to whether or not there is some $x_1 \in \Delta_1$ such that $h\big(x_1\big) > 0$. In such a case, the value of $h(x)$ for $x > 1$ has been explicitly given by Lemma 3.1(iii)(a) and (iv). For an example of this case, consider that $X \le Y$ and $\lim_{t\to\infty} \overline{F}(t)/\overline{G}(t) = c \in (0,1]$ and in turn, we have $h(1) = \lim_{t\to\infty} \overline{F}(t)/\overline{G}(t) = c > 0$.

A more extreme case is the one in which $X \le Y$ and $\overline{F} \sim \overline{G}$ (in particular when $X \equiv Y$). Clearly, $h(1) = \lim_{t\to\infty} \overline{F}(t)/\overline{G}(t) = 1$, and in turn $h(x) = 1$ for all $0 < x < 1$ in view of comment (ii) displayed after Assumption 3.2. By the above analysis, it is not difficult to verify that Assumption 3.2 holds in this setting if and only if $F \in \mathcal{R}_{-\alpha}$ with some $0 \le \alpha \le \infty$, and if either of them holds then

$$h(x) = \begin{cases} 1, & 0 < x \le 1, \\ x^{-\alpha}\mathbf{1}_{\{\alpha<\infty\}} + 0\cdot\mathbf{1}_{\{\alpha=\infty\}}, & x > 1. \end{cases}$$

Consequently, Theorem 3.1 is applicable to $\rho(t; g)$ if $1/\beta^* < \alpha \le \infty$. We refine these discussions and establish the following corollary.

Corollary 3.1. *Consider the distortion risk measure $\rho(t; g)$ defined in (2.3) with $X \le Y$ and $\overline{F}(t) \sim \overline{G}(t)$ as $t \to \infty$. If Assumption 3.1 holds and $F \in \mathcal{R}_{-\alpha}$ with $1/\beta^* < \alpha \le \infty$, then we have*

$$\lim_{t\to\infty} \frac{\rho(t; g)}{t} = \begin{cases} 1 + \int_1^\infty g\big(x^{-\alpha}\big)\mathrm{d}x, & 1/\beta^* < \alpha < \infty, \\ 1, & \alpha = \infty. \end{cases}$$

Our second main result is given in Theorem 3.2 and provides asymptotic approximations for non-comonotonic coherent risk measures in extreme regions as defined in (2.4).

Theorem 3.2. *Consider the risk measure $\rho(t; \theta, p)$ defined in (2.4) and let Assumption 3.2 hold. If $p < \alpha_F^* \le \infty$ then*

$$\lim_{t\to\infty} \frac{\rho(t; \theta, p)}{t} = \widetilde{h} + \theta p^{1/p} \left(\int_{\widetilde{h}}^\infty h(x)\Big(x - \widetilde{h}\Big)^{p-1}\mathrm{d}x \right)^{1/p} \quad \text{where} \quad \widetilde{h} = \int_0^\infty h(x)\mathrm{d}x. \tag{3.4}$$

4. Examples and Applications

The current section begins with some examples given in Section 4.1 under which Assumption 3.2 holds. As a result, Theorems 3.1 and 3.2 could be used to approximate various coherent risk measures in extreme regions as given in (2.3) and (2.4). Clearly, many examples could be chosen and we focus mainly on those that are related to our numerical illustrations provided in Section 4.2. Even though the examples are simply based on the aggregate risk or maximal risk, other examples could be provided; for instance, various large layers (based on weighted order statistics) could be considered, as discussed in Asimit *et al.* (2016).

Assume that $X_i \in L_+(\mathbb{P})$ with survival functions $\overline{F}_i$ for all $i \in \{1, \dots, d\}$, where d is a positive integer. The following three settings are further considered:

(A) $X = X_k$ and $Y = S_d := \sum_{i=1}^d X_i$, where $k \in \{1, \dots, d\}$;

(B) $X = X_k$ and $Y = M_d := \bigvee_{i=1}^d X_i$, where $k \in \{1, \dots, d\}$;

(C) $X = M_d$ and $Y = S_d$.

Recall that all random variables are considered to have ultimate right tails and hence they could be either of Fréchet or Gumbel type.

4.1. **Examples.** An important notion for detailing our examples is the *vague convergence*. Let $\{\mu_n; n \geq 1\}$ be a sequence of measures on a locally compact Hausdorff space $\mathbb{B}$ with countable base. Then, μ_n converges vaguely to some measure μ, written as $\mu_n \xrightarrow{v} \mu$, if for all continuous functions f with compact support we have

$$\lim_{n\to\infty} \int_{\mathbb{B}} f \, \mathrm{d}\mu_n = \int_{\mathbb{B}} f \, \mathrm{d}\mu.$$

Note that we deal only with Radon measures, i.e. measures that are finite for every compact set in $\mathbb{B}$. A thorough background on vague convergence is given by Kallenberg (1983) and Resnick (1987). Two specific choices for $\mathbb{B}$ are considered in this section: $\mathbb{B}_\Psi := [0,\infty]^d \backslash \{\mathbf{0}\}$ for Fréchet tails with a metric for which relatively compact sets are those that are bounded away from $\mathbf{0}$, and $\mathbb{B}_\Lambda := [-\infty,\infty]^d \backslash \{-\boldsymbol{\infty}\}$ for Gumbel tails with a metric for which relatively compact sets are those that are bounded away from $-\boldsymbol{\infty}$. Unless otherwise stated, all vectors appear in the following are of dimension d.

We now show that Assumption 3.2 holds for the next three examples, namely Examples 4.1-4.3. Since the other conditions of Assumption 3.2 are obvious, we only focus on the verification of relation (3.2). The first example is fairly general and the underlying assumption implies the property of multivariate regular variation (for details, see Resnick, 1987 and 2007).

Example 4.1. *Assume that there is some function $H_\Psi(\cdot)$ such that the relation*

$$\lim_{t\to\infty} \frac{\mathbb{P}\big(X_1 > tx_1, \ldots, X_d > tx_d\big)}{\overline{F}_1(t)} = H_\Psi(\mathbf{x}) \tag{4.1}$$

holds for every $\mathbf{x} \in \mathbb{B}_\Psi$.

Relation (4.1) *implies that the relation*

$$\frac{\mathbb{P}\Big(\big(X_1/t, \ldots, X_d/t\big) \in \cdot\Big)}{\overline{F}_1(t)} \xrightarrow{v} \mu_\Psi(\cdot) \quad \text{as} \quad t \to \infty, \tag{4.2}$$

holds on $\mathbb{B}_\Psi$ with a measure μ_Ψ such that $\mu_\Psi\big(\boldsymbol{y}\colon y_i > x_i, \text{ for all } i \in \{1,\ldots,d\}\big) = H_\Psi(\boldsymbol{x})$. It is clear that $\mu_\Psi\big(\boldsymbol{y} : y_1 > 1\big) = 1$ and hence, for every $x > 0$,

$$\lim_{t\to\infty} \frac{\overline{F}_1(tx)}{\overline{F}_1(t)} = \mu_\Psi\big(\mathbf{y} : y_1 > x\big) > 0,$$

which indicates that $F_1 \in \mathcal{R}_{-\alpha}$ for some $\alpha \in [0,\infty)$ by Theorem 1.4.1 of Bingham et al. (1987).

A more specific example for (4.1) *is the so called scale mixing that has been investigated many times in the literature; examples that are very much related to our paper are Hua and Joe (2011) and Zhu and Li (2012). Under scale mixing, we have that $\big(X_1,\ldots,X_d\big) = \big(ZU_1,\ldots,ZU_d\big)$, where Z is a non-negative random variable with df satisfying $F_Z \in \mathcal{R}_{-\alpha}$ for some $\alpha \in (0,\infty)$ and $\big(U_1,\ldots,U_d\big)$ is a non-negative random vector independent of Z with all components not degenerate at* 0. *Our asymptotic calculations require that there exists some $\varepsilon > 0$ such that*

$$\mathbb{E}U_i^{\alpha+\varepsilon} < \infty, \quad \textit{for all} \quad i \in \{1,\ldots,d\}.$$

Then, by Breiman's Lemma (see Breiman, 1965), we have that

$$\lim_{t\to\infty} \frac{\mathbb{P}\big(ZU_i > t\big)}{\overline{F}_Z(t)} = \mathbb{E}\big(U_i^\alpha\big), \quad i \in \{1,\ldots,d\},$$

i.e. $X_i \in \mathrm{MDA}\big(\Psi_\alpha\big)$. *Moreover, Breiman's Lemma tells us that for every* $\boldsymbol{x} \in \mathbb{B}_\Psi$ *we have that*

$$\begin{aligned} H_\Psi(\boldsymbol{x}) &= \lim_{t\to\infty} \frac{\mathbb{P}\big(ZU_1 > tx_1, \ldots, ZU_d > tx_d\big)}{\mathbb{P}\big(ZU_1 > t\big)} \\ &= \lim_{t\to\infty} \frac{\mathbb{P}\left(Z \bigwedge_{i=1}^{d} \big(U_i/x_i\big) > t\right)}{\overline{F}_Z(t)} \frac{\overline{F}_Z(t)}{\mathbb{P}\big(ZU_1 > t\big)} \\ &= \frac{\mathbb{E}\left(\bigwedge_{i=1}^{d} \big(U_i/x_i\big)^\alpha\right)}{\mathbb{E}\big(U_1^\alpha\big)}, \end{aligned}$$

where U_i/x_i *is understood as* ∞ *if* $x_i = 0$ *for some* $i \in \{1, \ldots, d\}$. *Further examples for* (4.1) *beyond scale mixing are based on copula assumptions and are detailed in Kortschak and Albrecher (2009) and Asimit et al. (2011).*

We now show that Assumption 3.2 holds for all settings (A)-(C) if (4.1) *is true. Consider first the setting (A). Note that* $\mu_\Psi\big(\boldsymbol{y}\colon y_1 > 1\big) = 1 > 0$ *implies* $\mu_\Psi\left(\boldsymbol{y} : \sum_{i=1}^{d} y_i > 1\right) > 0$, *which makes the denominator appearing in the next equation meaningful. For every* $0 < x < 1$,

$$\lim_{t\to\infty} \mathbb{P}\left(X_k > tx | S_d > t\right) = \lim_{t\to\infty} \frac{\mathbb{P}\big(X_k > tx, S_d > t\big)}{\overline{F}_1(t)} \frac{\overline{F}_1(t)}{\mathbb{P}\big(S_d > t\big)} = \frac{\mu_\Psi\left(\boldsymbol{y} : y_k > x, \sum_{i=1}^{d} y_i > 1\right)}{\mu_\Psi\left(\boldsymbol{y} : \sum_{i=1}^{d} y_i > 1\right)},$$

where the last step is due to (4.2) *and Proposition A2.12 of Embrechts et al. (1997). Note that the latter proposition can be applied as long as*

$$\mu_\Psi\left(\partial\left\{\mathbf{y} : \sum_{i=1}^{d} y_i > 1\right\}\right) = 0 \;\; \textit{and} \;\; \mu_\Psi\left(\partial\left\{\mathbf{y} : y_k > x, \sum_{i=1}^{d} y_i > 1\right\}\right) = 0.$$

The first claim is shown in the proof of Theorem 3.2 of Kortschak and Albrecher (2009), while the second claim is true for the same reason and the fact that $\mu_\Psi\big(\partial\{\mathbf{y} : y_k > x\}\big) = 0$, *which is a consequence of the uniform convergence of* (2.5) *(see Embrechts et al., 1997 or Resnick, 1987). Similar arguments help in justifying that, for every* $x \geq 1$,

$$\lim_{t\to\infty} \mathbb{P}\big(X_k > tx | S_d > t\big) = \lim_{t\to\infty} \frac{\overline{F}_k(tx)}{\overline{F}_1(tx)} \frac{\overline{F}_1(t)}{\mathbb{P}\big(S_d > t\big)} \frac{\overline{F}_1(tx)}{\overline{F}_1(t)} = \frac{\mu_\Psi\left(\mathbf{y} : y_k > 1\right)}{\mu_\Psi\left(\mathbf{y} : \sum_{i=1}^{d} y_i > 1\right)} x^{-\alpha}.$$

Hence, (3.2) *holds with*

$$h_A(x) = \frac{\mu_\Psi\left(\mathbf{y} : y_k > x, \sum_{i=1}^{d} y_i > 1\right)}{\mu_\Psi\left(\mathbf{y} : \sum_{i=1}^{d} y_i > 1\right)} \mathbf{1}_{\{0<x<1\}} + \frac{\mu_\Psi\left(\mathbf{y} : y_k > 1\right)}{\mu_\Psi\left(\mathbf{y} : \sum_{i=1}^{d} y_i > 1\right)} x^{-\alpha} \mathbf{1}_{\{x\geq 1\}}.$$

For settings (B) and (C), similar arguments are used when showing setting (A) may lead to

$$h_B(x) = \frac{\mu_\Psi\left(\mathbf{y} : y_k > x, \bigvee_{i=1}^d y_i > 1\right)}{\mu_\Psi\left(\mathbf{y} : \bigvee_{i=1}^d y_i > 1\right)}\mathbf{1}_{\{0<x<1\}} + \frac{\mu_\Psi\left(\mathbf{y} : y_k > 1\right)}{\mu_\Psi\left(\mathbf{y} : \bigvee_{i=1}^d y_i > 1\right)}x^{-\alpha}\mathbf{1}_{\{x\geq 1\}}$$

and

$$h_C(x) = \frac{\mu_\Psi\left(\mathbf{y} : \bigvee_{i=1}^d y_i > x, \sum_{i=1}^d y_i > 1\right)}{\mu_\Psi\left(\mathbf{y} : \sum_{i=1}^d y_i > 1\right)}\mathbf{1}_{\{0<x<1\}} + \frac{\mu_\Psi\left(\mathbf{y} : \bigvee_{i=1}^d y_i > 1\right)}{\mu_\Psi\left(\mathbf{y} : \sum_{i=1}^d y_i > 1\right)}x^{-\alpha}\mathbf{1}_{\{x\geq 1\}}.$$

The second example is the mirror setting of Example 4.1 and requires all distribution functions to have a Gumbel tail.

Example 4.2. *Assume that there is some function $H_\Lambda(\cdot)$ such that the relation*

$$\lim_{t\to\infty} \frac{\mathbb{P}\big(X_1 > t + x_1 a(t), \ldots, X_d > t + x_d a(t)\big)}{\overline{F}_1(t)} = H_\Lambda(\mathbf{x}) \tag{4.3}$$

holds for every $\mathbf{x} \in \mathbb{B}_\Lambda$. Similar to Example 4.1, relation (4.3) implies that

$$\frac{\mathbb{P}\left(\left(\frac{X_1-t}{a(t)}, \ldots, \frac{X_d-t}{a(t)}\right) \in \cdot\right)}{\overline{F}_1(t)} \xrightarrow{v} \mu_\Lambda(\cdot), \quad t\to\infty \tag{4.4}$$

holds on $\mathbb{B}_\Lambda$ with a measure μ_Λ such that $\mu_\Lambda\big(\mathbf{y} : y_i > x_i,$ for all $i \in \{1,\ldots,d\}\big) = H_\Lambda(\mathbf{x})$. Once again, further examples for (4.3) that are based on the concept of copula are detailed in Kortschak and Albrecher (2009) and Asimit et al. (2011). From now on, it is imperative to impose that $\mu_\Lambda\big(\mathbf{y} : \sum_{i=1}^d y_i > 0\big) > 0$ holds without which all further results would have not been true, but a remedy to this restriction is discussed in Example 4.3. Note that such a technical condition excludes the following asymptotic independence case:

$$\mathbb{P}\big(X_1 > t, \ldots, X_d > t\big) = o\big(\overline{F}_1(t)\big), \quad t\to\infty.$$

In addition, $\mu_\Lambda\left(\mathbf{y} : \bigwedge_{i=1}^d y_i > 0\right) > 0$, since otherwise some mass would have been put on one of the axes, which is not possible due to the fact that μ_Λ is a Radon measure and keeping in mind that $\mu_\Lambda(A + x\mathbf{1}) = e^{-x}\mu_\Lambda(A)$ holds for any $x \in \Re$ and set A such that $\mu_\Lambda(\partial A) = 0$. Similarly, one may show that $\mu_\Lambda(\mathbf{y} : y_k > 0) > 0$ for all $1 \leq k \leq n$, which in turn implies that $F_k \in \mathrm{MDA}(\Lambda) \subset \mathcal{R}_{-\infty}$ with auxiliary function $a(\cdot)$ and $\alpha^*_{F_k} = \infty$ for all $k \in \{1,\ldots,d\}$.

We now show that Assumption 3.2 holds for settings (A) if (4.3) is true. Let $r = t/d$ and write

$$\mathbb{P}\big(X_k > tx | S_d > t\big) = \frac{\mathbb{P}\big(X_k > dxr, S_d > dr\big)}{\overline{F}_1(r)} \frac{\overline{F}_1(r)}{\mathbb{P}\big(S_d > dr\big)}. \tag{4.5}$$

The second quotient in (4.5) *satisfies*

$$\lim_{r\to\infty} \frac{\overline{F}_1(r)}{\mathbb{P}\big(S_d > dr\big)} = \frac{1}{\mu_\Lambda\left(\mathbf{y} : \sum_{i=1}^{d} y_i > 0\right)}, \tag{4.6}$$

where we used (4.4) *and Proposition A2.12 of Embrechts et al. (1997). Note that the latter proposition can be applied as long as*

$$\mu_\Lambda\left(\partial\left\{\mathbf{y} : \sum_{i=1}^{d} y_i > 0\right\}\right) = 0, \tag{4.7}$$

which is shown in the proof of Theorem 3.3 of Kortschak and Albrecher (2009). It only remains to estimate the first quotient in (4.5)*. For* $0 < x < 1/d$*, we further write*

$$\frac{\mathbb{P}\big(X_k > dxr, S_d > dr\big)}{\overline{F}_1(r)} = \frac{\mathbb{P}\big(S_d > dr\big) - \mathbb{P}\big(X_k \leq dxr, S_d > dr\big)}{\overline{F}_1(r)}. \tag{4.8}$$

Now, for $0 < x < 1/d$*,*

$$\begin{aligned}
\limsup_{r\to\infty} \frac{\mathbb{P}\big(X_k \leq dxr, S_d > dr\big)}{\overline{F}_1(r)} &\leq \limsup_{r\to\infty} \frac{\mathbb{P}\left(\sum_{i=1, i\neq k}^{d} X_i > d(1-x)r\right)}{\overline{F}_1(r)} \\
&= \lim_{r\to\infty} \frac{\mathbb{P}\left(\sum_{i=1, i\neq k}^{d} X_i > (d-1)\frac{d(1-x)}{d-1} r\right)}{\overline{F}_1\left(\frac{d(1-x)}{d-1} r\right)} \frac{\overline{F}_1\left(\frac{d(1-x)}{d-1} r\right)}{\overline{F}_1(r)} \\
&= \nu_\Lambda\left(\boldsymbol{y} : \sum_{i=1, i\neq k}^{d} y_i > 0\right) \cdot 0 = 0,
\end{aligned}$$

where the last step is due to $F_1 \in \mathcal{R}_{-\infty}$*, similar arguments to those used in* (4.7)*, and the fact that* ν_Λ *is a Radon measure. A combination of the latter result,* (4.6) *and* (4.8) *leads to*

$$\lim_{r\to\infty} \frac{\mathbb{P}\big(X_k > dxr, S_d > dr\big)}{\overline{F}_1(r)} = \nu_\Lambda\left(\mathbf{y} : \sum_{i=1}^{d} y_i > 0\right), \quad 0 < x < 1/d. \tag{4.9}$$

If $x = 1/d$*, we get that*

$$\lim_{r\to\infty} \frac{\mathbb{P}\big(X_k > dxr, S_d > dr\big)}{\overline{F}_1(r)} = \mu_\Lambda\left(\mathbf{y} : y_k > 0, \sum_{i=1}^{d} y_i > 0\right), \tag{4.10}$$

as a result of (4.4) *and Proposition A2.12 of Embrechts et al. (1997). The latter proposition could be applied due to* (4.7) *and the uniform convergence of* (2.6) *(see Embrechts et al., 1997 or Resnick,* *1987). Finally, for $x > 1/d$, it follows from $F \in \mathcal{R}_{-\infty}$ and* (4.4) *that*

$$\begin{aligned}
\limsup_{r\to\infty} \frac{\mathbb{P}\big(X_k > dxr, S_d > dr\big)}{\overline{F}_1(r)} &\leq \limsup_{r\to\infty} \frac{\mathbb{P}\big(X_k > dxr\big)}{\overline{F}_1(r)} \\
&= \lim_{r\to\infty} \frac{\overline{F}_k(dxr)}{\overline{F}_1(dxr)} \frac{\overline{F}_1(dxr)}{\overline{F}_1(r)} \\
&= \nu_\Lambda\big(\mathbf{y} : y_k > 0\big) \cdot 0 = 0. \qquad (4.11)
\end{aligned}$$

Hence, plugging (4.6) *and* (4.9)*-*(4.11) *into* (4.5)*, we know that* (3.2) *holds with*

$$h_A(x) = \mathbf{1}_{\{0<x<1/d\}} + \frac{\mu_\Lambda\left(\mathbf{y} : y_k > 0, \sum_{i=1}^d y_i > 0\right)}{\mu_\Lambda\left(\mathbf{y} : \sum_{i=1}^d y_i > 0\right)} \mathbf{1}_{\{x=1/d\}} + 0 \cdot \mathbf{1}_{\{x>1/d\}}.$$

Note that the jump of h_A at $x = 1/d$ does not change the actual values of the asymptotic constants in equations (3.3) *and* (3.4)*.*

Unfortunately, (3.2) *does not hold for $0 < x < 1$ in settings (B) and (C).*

The next example is a reaction to the restriction we impose in Example 4.2, i.e. removing the asymptotic independence case with Gumbel tails. Example 4.3 provides a case in which (4.3) holds such that $\mu_\Lambda\left(\mathbf{y} : \sum_{i=1}^d y_i > 0\right) = 0$, where different arguments are needed. The set of assumptions are precisely the same as in Mitra and Resnick (2009), but a similar setup can be found in Hashorva and Li (2015).

Example 4.3. *Let $F_1 \in \mathrm{MDA}(\Lambda)$ with an auxiliary function $a(\cdot)$. Further, $\overline{F}_i(x) \sim c_i \overline{F}_1(x)$ holds as $x \to \infty$ with $c_i \geq 0$ for all $i \in \{1, \ldots, d\}$. Moreover, for all $1 \leq i \neq j \leq d$, assume that*

$$\lim_{t\to\infty} \frac{\mathbb{P}\left(X_i > t, X_j > a(t)x\right)}{\overline{F}_1(t)} = 0, \quad \textit{for all } x > 0, \qquad (4.12)$$

and

$$\lim_{t\to\infty} \frac{\mathbb{P}\left(X_i > L_{ij}t, X_j > L_{ij}t\right)}{\overline{F}_1(t)} = 0, \quad \textit{for some } L_{ij} > 0.$$

We show now that Assumption 3.2 holds for all settings (A)-(C) under the above conditions. Consider first the setting (A). Recall that Corollary 2.2 of Mitra and Resnick (2009) tells us

$$\mathbb{P}\big(S_d > t\big) \sim \sum_{i=1}^d c_i \, \overline{F}_1(t), \quad t \to \infty. \qquad (4.13)$$

For $0 < x \le 1$, one can get that

$$\begin{aligned}
\mathbb{P}\big(X_k > tx, S_d > t\big) &= \mathbb{P}\big(S_d > t\big) - \mathbb{P}\big(X_k \le tx, S_d > t\big) \\
&\le \mathbb{P}\big(S_d > t\big) - \mathbb{P}\left(X_k \le tx, \bigcup_{i=1, i\neq k}^{d} \{X_i > t\}\right) \\
&= \mathbb{P}\big(S_d > t\big) - \mathbb{P}\left(\bigcup_{i=1, i\neq k}^{d} \{X_i > t\}\right) + \mathbb{P}\left(X_k > tx, \bigcup_{i=1, i\neq k}^{d} \{X_i > t\}\right) \\
&\le \mathbb{P}\big(S_d > t\big) - \sum_{\substack{i=1\\ i\neq k}}^{d} \mathbb{P}\big(X_i > t\big) + \sum_{\substack{1\le i<j\le d\\ i\neq k\neq j}} \mathbb{P}\big(X_i > t, X_j > t\big) \\
&\qquad + \sum_{\substack{i=1\\ i\neq k}}^{d} \mathbb{P}\big(X_k > tx, X_i > t\big),
\end{aligned}$$

which together with (4.13), (4.12) *and the well-known fact $a(t) = o(t)$ as $t \to \infty$ lead to*

$$\limsup_{t\to\infty} \frac{\mathbb{P}\big(X_k > tx, S_d > t\big)}{\overline{F}_1(t)} \le \sum_{i=1}^{d} c_i - \sum_{i=1, i\neq k}^{d} c_i + 0 + 0 = c_k.$$

Clearly,

$$\liminf_{t\to\infty} \frac{\mathbb{P}\big(X_k > tx, S_d > t\big)}{\overline{F}_1(t)} \ge \liminf_{t\to\infty} \frac{\overline{F}_k(t)}{\overline{F}_1(t)} = c_k.$$

Hence, combining the two relations from above with (4.13) *leads to*

$$\lim_{t\to\infty} \mathbb{P}\big(X_k > tx | S_d > t\big) = c_k \left(\sum_{i=1}^{d} c_i\right)^{-1}, \quad 0 < x \le 1.$$

On the other hand, for $x > 1$, we have

$$\lim_{t\to\infty} \mathbb{P}\left(X_k > tx | S_d > t\right) = \lim_{t\to\infty} \frac{\overline{F}_k(tx)}{\overline{F}_1(tx)} \frac{\overline{F}_1(t)}{\mathbb{P}\big(S_d > t\big)} \frac{\overline{F}_1(tx)}{\overline{F}_1(t)} = c_k \left(\sum_{i=1}^{d} c_i\right)^{-1} \cdot 0 = 0,$$

since $F_1 \in \mathcal{R}_{-\infty}$. Hence, (3.2) *holds with*

$$h_A(x) = c_k \left(\sum_{i=1}^{d} c_i\right)^{-1} \mathbf{1}_{\{0<x\le 1\}} + 0 \cdot \mathbf{1}_{\{x>1\}}.$$

Setting (B) may be concluded in the same manner as above and one may find that $h_B(x) = h_A(x)$. Finally, for setting (C) note that (4.12) *and* (4.13) *imply*

$$\mathbb{P}\big(M_d > t\big) \sim \sum_{i=1}^{d} c_i \, \overline{F}_1(t) \sim \mathbb{P}\big(S_d > t\big), \quad t \to \infty,$$

and hence the df of M_d also belongs to $\mathcal{R}_{-\infty}$. Thus, we may directly apply Corollary 3.1 to obtain that

$$h_C(x) = \mathbf{1}_{\{0<x\le 1\}} + 0 \cdot \mathbf{1}_{\{x>1\}}.$$

It is interesting to mention that Theorems 2.1, 2.2, 3.1 and 3.3 of Asimit et al. (2011) can be retrieved by plugging the appropriate functions $h_A(\cdot)$ obtained in Examples 4.1-4.3 into (3.3) with $g(s) = s$.

4.2. **Numerical Illustrations.** Consider a European-based insurance company, i.e. Solvency II Regime is in force, that holds a portfolio consisting of two LOBs with random future liabilities (corresponding to business written in the coming year) denoted by X_1 and X_2, respectively. Under this assumption, the total capital is VaR-based and is calculated at the 99.5% level. By definition, the VaR of a generic loss variable Z at a confidence level a, $VaR_a(Z)$, is the a%-quantile, i.e.

$$VaR_a(Z) := \inf\{z \geq z_0 : \mathbb{P}(Z \leq z) \geq a\},$$

where by convention, $\inf \emptyset := +\infty$. Specifically, the total cost of meeting the capital requirements, according to the Solvency II Regime, is

$$\mathbb{E}X_1 + \mathbb{E}X_2 + \lambda_{CoC}\Big(VaR_{99.5\%}\big(X_1 + X_2\big) - \mathbb{E}X_1 - \mathbb{E}X_2\Big), \tag{4.14}$$

where λ_{CoC} is a constant representing the return on capital that the shareholders expect to receive for making their capital available to run this specific business. This constant is usually greater than 6% and depends on the level of taxation of the economic area that the insurance business takes place. The random liabilities are assumed to be Pareto distributed with survival function $\overline{F}_k(x) = \left(\frac{\lambda_k}{\lambda_k + x}\right)^{\alpha}$ with $x, \lambda_k > 0$ and $\alpha > 1$ for all $k \in \{1, 2\}$, which implies that $\mathbb{E}X_k = \lambda_k/(\alpha - 1)$. Thus, $\overline{F}_2(x) \sim c\overline{F}_1(x)$, where $c = \big(\lambda_2/\lambda_1\big)^{\alpha}$, as $t \to \infty$ and $F_1, F_2 \in \mathcal{R}_{-\alpha}$. Assume that (4.1) holds such that

$$H_\Psi(\mathbf{x}) = \left(x_1^{\alpha\beta} + c^{-\beta} x_2^{\alpha\beta}\right)^{-1/\beta}, \quad \text{with} \quad \beta > 0. \tag{4.15}$$

The structure shown in (4.15) is derived from an Archimedean survival copula with a regularly varying generator function (for details, see Juri and Wüthrich, 2003 or Section 4 of Asimit *et al.*, 2011). Recall that Lemma 2.1 of Asimit *et al.* (2011) implies that $VaR_{99.5\%}\big(X_1 + X_2\big) \approx C_+^{1/\alpha} VaR_{99.5\%}\big(X_1\big)$ with $C_+ := \mu_\Psi\big(\mathbf{y} : y_1 + y_2 > 1\big)$, where μ_Ψ is the measure induced by the limit from (4.15).

As it has been anticipated, the main exercise is to allocate the total cost of meeting the capital requirements given in (4.14). The first step is to evaluate the total cost, which is computable via an asymptotic approximation that has been detailed above. The second step is to distribute the total cost among the LOBs in the most fair way that takes into account the individual risk profile and the dependence amongst risks; this step is essential when evaluating the financial performance of the two LOBs, making the cost allocation to be a genuine measure of performance in the evaluation process of the actual emerging profits. The cost allocation may be done in many ways and one possibility is the so-called proportional allocation, which is the case when asymptotic approximations are considered (for details, see Asimit *et al.*, 2011). Given the nature of this paper, proportional allocations are sought in further numerical illustrations. Specifically, take $g(s; \xi) = (1 + \xi)s - \xi s^2$ with $0 \leq \xi \leq 1$ in Theorem 3.1, which implies that equation (3.3) becomes

$$\lim_{t\to\infty} \frac{\rho(X_k | X_1 + X_2 > t; \xi)}{t} = \int_0^\infty g\big(h_k(x); \xi\big) \mathrm{d}x := C_k, \quad k \in \{1, 2\}.$$

Note that we choose the Gini risk measure (see Example iii) from Section 2) for the reason of simplifying the numerical calculations when evaluating multiple integrals, but any other choice is doable via more advanced Monte-Carlo integration. Moreover, as in Example 4.1, we have that

$$h_1(x)=\frac{\mu_\Psi\big(\mathbf{y}:y_1>x,y_1+y_2>1\big)}{C_+}\mathbf{1}_{\{0<x<1\}}+\frac{x^{-\alpha}}{C_+}\mathbf{1}_{\{x\geq 1\}}$$

and

$$h_2(x)=\frac{\mu_\Psi\big(\mathbf{y}:y_2>x,y_1+y_2>1\big)}{C_+}\mathbf{1}_{\{0<x<1\}}+\frac{cx^{-\alpha}}{C_+}\mathbf{1}_{\{x\geq 1\}}.$$

Having all of these in mind, the total capital cost from (4.14) is allocated (amongst the two LOBs) as follows:

$$COC_k:=\mathbb{E}X_k+\lambda_{CoC}\left(\frac{C_k}{C_1+C_2}C_+^{1/\alpha}VaR_{99.5\%}\big(X_1\big)-\mathbb{E}X_k\right),\quad k\in\{1,2\}.$$

Now, we further simplify the expressions for C_+, C_1 and C_2 in a way that the Monte Carlo methods are implementable in a straightforward manner when integrals are numerically computed. Denote $H^{(1)}_\Psi(y_1,y_2)=-\frac{\partial H_\Psi(y_1,y_2)}{\partial y_1}$ and $H^{(2)}_\Psi(y_1,y_2)=-\frac{\partial H_\Psi(y_1,y_2)}{\partial y_2}$. Clearly,

$$C_+=1+\int_0^1\mu_\Psi\big(\mathrm{d}x\times(1-x,\infty]\big)=1+\int_0^1 H^{(1)}_\Psi(x,1-x)\mathrm{d}x$$

and

$$\begin{aligned}
C_1 &= (1+\xi)\int_0^\infty h_1(x)\mathrm{d}x-\xi\int_0^\infty h_1^2(x)\mathrm{d}x\\
&= \frac{1+\xi}{C_+}\left(\int_0^1\mu_\Psi\big(\mathbf{y}:y_1>x,y_1+y_2>1\big)\mathrm{d}x+\int_1^\infty x^{-\alpha}\mathrm{d}x\right)\\
&\quad -\frac{\xi}{C_+^2}\left(\int_0^1\mu_\Psi^2\big(\mathbf{y}:y_1>x,y_1+y_2>1\big)\mathrm{d}x+\int_1^\infty x^{-2\alpha}\mathrm{d}x\right)\\
&= \frac{1+\xi}{C_+}\left\{\frac{1}{\alpha-1}+\int_0^1\left(1+\int_x^1 H^{(1)}_\Psi(z,1-z)\mathrm{d}z\right)\mathrm{d}x\right\}\\
&\quad -\frac{\xi}{C_+^2}\left\{\frac{1}{2\alpha-1}+1+2\int_0^1 x\cdot\mu_\Psi\big(\mathbf{y}:y_1>x,y_1+y_2>1\big)H^{(1)}_\Psi(x,1-x)\mathrm{d}x\right\}\\
&= \frac{1+\xi}{C_+}\left\{\frac{\alpha}{\alpha-1}+\int_0^1\int_x^1 H^{(1)}_\Psi(z,1-z)\mathrm{d}z\mathrm{d}x\right\}\\
&\quad -\frac{\xi}{C_+^2}\left\{\frac{2\alpha}{2\alpha-1}+2\int_0^1 xH^{(1)}_\Psi(x,1-x)\mathrm{d}x+2\int_0^1\int_x^1 xH^{(1)}_\Psi(x,1-x)H^{(1)}_\Psi(z,1-z)\mathrm{d}z\mathrm{d}x\right\},
\end{aligned}$$

where the second last step is a simple consequence of integration by parts. Similarly,

$$\begin{aligned}
C_2 &= \frac{1+\xi}{C_+}\left\{\frac{c\alpha}{\alpha-1}+\int_0^1\int_x^1 H^{(2)}_\Psi(1-z,z)\mathrm{d}z\mathrm{d}x\right\}\\
&\quad -\frac{\xi}{C_+^2}\left\{\frac{2c^2\alpha}{2\alpha-1}+2c\int_0^1 xH^{(2)}_\Psi(1-x,x)\mathrm{d}x+2\int_0^1\int_x^1 xH^{(2)}_\Psi(1-x,x)H^{(2)}_\Psi(1-z,z)\mathrm{d}z\mathrm{d}x\right\}.
\end{aligned}$$

Assume now that $\alpha=3$, $\beta=2$, $c=0.8$, $\lambda_{CoC}=10\%$ and $\lambda_1=5,000$, which implies that a total cost of capital of $3,858.72$ needs to be allocated. Using the previously-mentioned derivations for C_+, C_1 and C_2, proportional allocations based on Gini risk measures are displayed in Table 4.1. Recall that the lower the value of ξ, the more conservative the risk measure is. The results are sensible

Allocations	$\xi = 1$	$\xi = 0.75$	$\xi = 0.5$
COC_1	$1,931.11$	$1,964.60$	$1,995.50$
COC_2	$1,927.61$	$1,894.12$	$1,863.22$
Total	$3,858.72$	$3,858.72$	$3,858.72$

TABLE 4.1. Cost of capital allocations for various values of ξ, when $\alpha = 3$, $\beta = 2$, $c = 0.8$, $\lambda_{CoC} = 10\%$ and $\lambda_1 = 5,000$.

and it can be observed that the riskier LOB, i.e. the first LOB, requires a large amount of capital cost when the risk measure becomes more conservative. The increase in capital cost for LOB1 is 3.33% when the value for ξ is reduced from 1 to 0.5. Thus, we may conclude that the expectation of the downside risk (when $\xi = 1$) does not reasonably account for the individual risk contribution when the costs of the VaR-based regulatory capital is allocated, and thus, more conservative risk measurements would be recommended.

We finally investigate the effect of parameter uncertainty for the cost of capital proportions from the total cost. Thus, we alter the previous toy model, where it was assumed that $\alpha = 3$, $\beta = 2$, $c = 0.8$, $\lambda_{CoC} = 10\%$ and $\lambda_1 = 5,000$. Numerical illustrations are displayed in Table 4.2, where the parameter α changes by $\{-10\%, -5\%, -2.5\%, 2.5\%, 5\%, 10\%\}$ and the actual costs of capital for each LOB are compared to the toy model. We observe that there is no gain (in the sense that the allocations take into account the change of the individual risk profiles) for allocating the costs via a more conservative risk measure. Similar sensitivity analyses were performed for all other parameters and the conclusion remain the same, i.e. conservative risk measures do not make the allocations fairer when the risk parameter is present. Therefore, we may conclude that more conservative risk measurements do not improve the fairness of the proportions for cost of capital allocations when the uncertainty with parameter estimation is present, even at a very high level. We simply infer that the most influential estimation for accurate cost of capital allocations is given by the estimate of $VaR_{99.5\%}\big(X_1 + X_2\big)$, which in our case is reduced to estimating a high quantile of the LOB1, i.e. $VaR_{99.5\%}\big(X_1\big)$.

5. PROOFS

Proof of Lemma 3.1. (i) By (2.7), for every $0 < \alpha' < \alpha^*_F$, there are some $K_1 > 1$ and $t_0 > 0$ such that $\overline{F}(tx)/\overline{F}(t) \le K_1 x^{-\alpha'}$ holds for all $tx > t > t_0$. Since $\overline{F}(t) = O\big(\overline{G}(t)\big)$ as $t \to \infty$, there is some $K_2 > 1$ such that $\overline{F}(t)/\overline{G}(t) \le K_2$ is true for all $t \ge 0$. Thus, for any $tx > t > t_0$ we have that

$$\mathbb{P}(X > tx|Y > t) = \frac{\mathbb{P}(X > tx, Y > t)}{\overline{F}(t)}\frac{\overline{F}(t)}{\overline{G}(t)} \le \frac{\overline{F}(tx)}{\overline{F}(t)}\frac{\overline{F}(t)}{\overline{G}(t)} \le K_1 K_2 x^{-\alpha'}. \tag{5.1}$$

Hence, the following holds for all $x \in \Delta_1$:

$$0 \le h(x) = \lim_{t\to\infty} \mathbb{P}(X > tx|Y > t) \le K_1 K_2 x^{-\alpha'},$$

and in turn we get that $\lim\limits_{x\to\infty,\, x\in\Delta_0} h(x) = 0$. This completes the argumentation for part (i).

(ii) If $\mathbb{P}(X > tx, Y > t) = O\big(\overline{F}(tx)\overline{G}(t)\big)$ as $t \to \infty$ for every $x > 0$, then it is clear that $\mathbb{P}(X > tx|Y > t) = O\big(\overline{F}(tx)\big)$ as $t \to \infty$. Thus, $h(x) = 0$.

Allocations	$\xi = 1$					
	-10%	-5%	-2.5%	2.5%	5%	10%
COC_1	-22.77%	-11.65%	-5.91%	6.00%	12.12%	24.64%
COC_2	-20.95%	-10.71%	-5.39%	5.53%	11.18%	22.86%
Total	-21.89%	-11.18%	-5.65%	5.77%	11.65%	23.75%

Allocations	$\xi = 0.75$					
	-10%	-5%	-2.5%	2.5%	5%	10%
COC_1	-22.40%	-11.43%	-5.79%	5.87%	11.85%	24.05%
COC_2	-21.35%	-10.92%	-5.50%	5.66%	11.45%	23.44%
Total	-21.89%	-11.18%	-5.65%	5.77%	11.65%	23.75%

Allocations	$\xi = 0.5$					
	-10%	-5%	-2.5%	2.5%	5%	10%
COC_1	-22.10%	-11.27%	-5.69%	5.77%	11.64%	23.60%
COC_2	-21.66%	-11.09%	-5.60%	5.76%	11.66%	23.90%
Total	-21.89%	-11.18%	-5.65%	5.77%	11.65%	23.75%

TABLE 4.2. Sensitivity analysis for the cost of capital allocations with respect to parameter α.

Assume now that $1 \in \Delta_0$ with $h(1) = 0$ and $G \in \mathcal{D}$. The comment (i) after Assumption 3.2 gives that $h(x) = 0$ for all $x > 1$, since $h(1) = 0$. In addition, for every $0 < x \le 1$,

$$\limsup_{t\to\infty} \mathbb{P}(X > tx | Y > t) \le \limsup_{t\to\infty} \frac{\mathbb{P}(X > tx, Y > tx)}{\overline{G}(tx)} \frac{\overline{G}(tx)}{\overline{G}(t)} \le h(1) \limsup_{t\to\infty} \frac{\overline{G}(tx)}{\overline{G}(t)} = 0,$$

since $h(1) = 0$ and $G \in \mathcal{D}$. Therefore, $h(x) = 0$ for all $x > 0$, which concludes part (ii).

(iii) We first prove part (a). Clearly, $\liminf_{t\to\infty} \overline{F}(t)/\overline{G}(t) \ge h(1) > 0$. This fact and $\overline{F}(t) = O\big(\overline{G}(t)\big)$ as $t \to \infty$ imply that there exist c_1 and c_2 such that

$$0 < c_1 := \liminf_{t\to\infty} \frac{\overline{F}(t)}{\overline{G}(t)} \le \limsup_{t\to\infty} \frac{\overline{F}(t)}{\overline{G}(t)} =: c_2 < \infty.$$

Thus, for every $x \in \Delta_0$,

$$\liminf_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} \ge \liminf_{t\to\infty} \frac{\mathbb{P}(X > tx, Y > t)}{\overline{G}(t)} \frac{\overline{G}(t)}{\overline{F}(t)} \ge \frac{1}{c_2} h(x),$$

and for every $x \in \Delta_1$,

$$\limsup_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} \le \limsup_{t\to\infty} \frac{\mathbb{P}(X > tx, Y > t)}{\overline{G}(t)} \frac{\overline{G}(tx)}{\mathbb{P}(X > tx, Y > tx)} \frac{\overline{F}(tx)}{\overline{G}(tx)} \frac{\overline{G}(t)}{\overline{F}(t)} \le \frac{c_2}{c_1 h(1)} h(x).$$

Hence, for every $x \in \Delta_1$ we have that

$$\frac{1}{c_2} h(x) \le \liminf_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} \le \limsup_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} \le \frac{c_2}{c_1 h(1)} h(x). \tag{5.2}$$

Keeping (5.2) in mind, the only non-trivial task in proving part (a) is to verify that $h(x) = 0$ for all $x > 1$ if $h(x_1) = 0$ for some $x_1 \in \Delta_1$. The comment (i) following Assumption 3.2 tells us that if $h(x_1) = 0$ for some $x_1 \in \Delta_1$ then $h(x) = 0$ for all $x > x_1$. Note also that $h(x_1) = 0$, which together with (5.2), imply

$$\lim_{t\to\infty} \frac{\overline{F}(tx_1)}{\overline{F}(t)} = 0. \tag{5.3}$$

Suppose now that there is some $x_0 \in \Delta_1 \cap (1, x_1)$ such that $h(x_0) > 0$. Since $x_0 > 1$, there is some positive integer $n \geq 2$ such that $x_0^n \geq x_1$, and therefore, we have that

$$\frac{\overline{F}(tx_1)}{\overline{F}(t)} \geq \frac{\overline{F}(tx_0^n)}{\overline{F}(t)} = \frac{\overline{F}(tx_0^n)}{\overline{F}(tx_0^{n-1})} \frac{\overline{F}(tx_0^{n-1})}{\overline{F}(tx_0^{n-2})} \cdots \frac{\overline{F}(tx_0)}{\overline{F}(t)}.$$

Applying (5.2) to the above yields

$$\liminf_{t\to\infty} \frac{\overline{F}(tx_1)}{\overline{F}(t)} \geq \left(\liminf_{t\to\infty} \frac{\overline{F}(tx_0)}{\overline{F}(t)}\right)^n \geq \left(\frac{1}{c_2} h(x_0)\right)^n > 0,$$

which contradicts (5.3). Hence, $h(x) = 0$ for all $x \in \Delta_1$. Now, since Δ_1 is dense in $(1, \infty)$, for every $x > 1$ there exists some $x_2 \in \Delta_1$ such that $1 < x_2 < x$. Then,

$$\limsup_{t\to\infty} \mathbb{P}(X > tx | Y > t) \leq \limsup_{t\to\infty} \mathbb{P}(X > tx_2 | Y > t) = h(x_2) = 0,$$

which implies that $h(x)$ exists for every $x > 1$ with $h(x) = 0$.

We now outline the proof for part (b). The first equivalence is an immediate consequence of part (a). The second equivalence can be obtained by using (5.2), the monotonicity of $\overline{F}(tx)$ with respect to $x > 0$ and the denseness of Δ_1 in $(1, \infty)$. This completes the argumentation for part (iii).

(iv) The result from (iii)(b) tells us that $h(x) > 0$ for all $x \in \Delta_0$. Now, $\lim_{t\to\infty} F(t)/G(t) = h(1)$ leads to $c_1 = c_2 = h(1)$ in equation (5.2), which implies that

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = \frac{h(x)}{h(1)} \in (0, \infty) \tag{5.4}$$

holds for every $x \in \widetilde{\Delta}_1 =: \Delta_1 \cup \{1\}$. Hence, the set of all $x > 0$ such that $\lim_{t\to\infty} \overline{F}(tx)/\overline{F}(t)$ exists and belongs to $(0, \infty)$ has a positive measure. Then, by Theorem 1.4.1 of Bingham et al. (1987), we have $F \in \mathcal{R}_{-\alpha}$ for some $0 \leq \alpha < \infty$. This fact and (5.4) imply that $h(x) = h(1)x^{-\alpha}$ for every $x \in \widetilde{\Delta}_1$. Further, since $\widetilde{\Delta}_1$ is dense in $[1, \infty)$, for every $x \geq 1$ there are two sequences $\{y_n; n \geq 1\}$ and $\{z_n; n \geq 1\}$ from $\widetilde{\Delta}_1$ such that $y_n \nearrow x \swarrow z_n$ as $n \to \infty$. Then, for every $n \geq 1$,

$$h(1)z_n^{-\alpha} = h(z_n) \leq \liminf_{t\to\infty} \mathbb{P}(X > tx | Y > t) \leq \limsup_{t\to\infty} \mathbb{P}(X > tx | Y > t) \leq h(y_n) = h(1)y_n^{-\alpha}.$$

Letting $n \to \infty$ in the above relation, we obtain that $h(x)$ exists for all $x \geq 1$ and $h(x) = h(1)x^{-\alpha}$, which concludes part (iv). The proof is now complete. □

Proof of Theorem 3.1. Recalling (2.1) and (2.3), we get that

$$\frac{\rho(t; g)}{t} = \int_0^\infty g\big(\mathbb{P}(X > tx | Y > t)\big) \mathrm{d}x \tag{5.5}$$

holds for every $t > 0$. Since g is non-decreasing, equation (5.1) implies that for every $1/\beta^* < \alpha' < \alpha_F^*$ there are some $K_1 > 1$ and $t_0 > 0$ such that

$$g\big(\mathbb{P}(X > tx | Y > t)\big) \leq g\left(K_1 x^{-\alpha'}\right)$$

holds for $tx > t > t_0$. It is easy to see that $\beta' \in \Omega_g$ for every $0 < \beta' < \beta^*$. Hence, by Assumption 3.1, for every $0 < \beta' < \beta^*$ there are some $K_2 > 1$ and $x_0 > 1$ such that

$$g\big(\mathbb{P}(X > tx|Y > t)\big) \le K_2 x^{-\alpha'\beta'}$$

holds for all $x > x_0$ and $t > t_0$. Clearly, $g\big(\mathbb{P}(X > tx|Y > t)\big) \le g(1) = 1$ for all $x > 0$. Hence, for all $t > t_0$ and $x > 0$ we have that

$$g\big(\mathbb{P}(X > tx|Y > t)\big) \le \mathbf{1}_{\{0<x\le x_0\}} + K_2 x^{-\alpha'\beta'}\mathbf{1}_{\{x>x_0\}}.$$

Noting that $\alpha_F^*\beta^* > 1$, we may choose α' and β' satisfying $\alpha'\beta' > 1$, which in turn implies that $\mathbf{1}_{\{0<x\le x_0\}} + K_2 x^{-\alpha'\beta'}\mathbf{1}_{\{x>x_0\}}$ is integrable over $(0,\infty)$. The latter and (3.2) enable us to apply the Dominated Convergence Theorem to (5.5) to justify our claim stated in (3.3), since the function g is monotonic and hence continuous almost everywhere. The proof is now complete. □

Proof of Theorem 3.2. Next, denote $\xi_t := X|Y > t$ and $\mu_t := \mathbb{E}\xi_t$. Then, recalling (2.2), (2.4) can be rewritten as follows:

$$\rho(t;\theta,p) = \mu_t + \theta\left(\mathbb{E}\Big[\big(\xi_t - \mu_t\big)_+^p\Big]\right)^{1/p}. \tag{5.6}$$

Note that $1 \le p < \alpha_F^*$. Hence, applying Theorem 3.1 with $g(s) = s$ leads to

$$\lim_{t\to\infty}\frac{\mu_t}{t} = \int_0^\infty h(x)\mathrm{d}x = \widetilde{h}. \tag{5.7}$$

Next, we focus on the asymptotic approximate of $\mathbb{E}\Big[\big(\xi_t - \mu_t\big)_+^p\Big]$ as $t \to \infty$. Clearly,

$$\begin{aligned}\mathbb{E}\Big[\big(\xi_t - \mu_t\big)_+^p\Big] &= p\int_{\mu_t}^\infty \mathbb{P}(X > x|Y > t)\big(x - \mu_t\big)^{p-1}\mathrm{d}x \\ &= p\, t^p \int_{\frac{\mu_t}{t}}^\infty \mathbb{P}(X > tx|Y > t)\left(x - \frac{\mu_t}{t}\right)^{p-1}\mathrm{d}x,\end{aligned}$$

which is equivalent to

$$\frac{\mathbb{E}\Big[\big(\xi_t - \mu_t\big)_+^p\Big]}{p\, t^p} = \int_0^\infty \mathbb{P}(X > tx|Y > t)\left(x - \frac{\mu_t}{t}\right)^{p-1}\mathbf{1}_{\left\{x>\frac{\mu_t}{t}\right\}}\mathrm{d}x. \tag{5.8}$$

Due to relation (5.7), there is some large $t_0 > 0$ such that $\mu_t/t \ge \widetilde{h}/2$ holds for all $t > t_0$. Recalling (5.1), we can choose t_0 large enough such that $\mathbb{P}(X > tx|Y > t) \le Kx^{-\alpha'}$ is true for some $K > 1$, $p < \alpha' < \alpha_F^*$ and all $tx > t > t_0$. Hence, for all $t > t_0$, the integrand of (5.8) is not greater than

$$\mathbb{P}(X > tx|Y > t)\left(x - \frac{\widetilde{h}}{2}\right)^{p-1}\mathbf{1}_{\left\{x>\frac{\widetilde{h}}{2}\right\}} \le \left(x - \frac{\widetilde{h}}{2}\right)^{p-1}\mathbf{1}_{\left\{\frac{\widetilde{h}}{2}<x\le\frac{\widetilde{h}}{2}\vee 1\right\}} + Kx^{-\alpha'}\left(x - \frac{\widetilde{h}}{2}\right)^{p-1}\mathbf{1}_{\left\{x>\frac{\widetilde{h}}{2}\vee 1\right\}},$$

which is obviously integrable over $(0,\infty)$. Thus, applying (3.2), (5.7) and the Dominated Convergence Theorem to (5.8) yields

$$\lim_{t\to\infty}\frac{\mathbb{E}\Big[\big(\xi_t - \mu_t\big)_+^p\Big]}{p\, t^p} = \int_{\widetilde{h}}^\infty h(x)\big(x - \widetilde{h}\big)^{p-1}\mathrm{d}x.$$

Combining the above relation with (5.6) and (5.7), one may conclude (3.4). □

Acknowledgments. The authors are very grateful to two anonymous reviewers for their thorough reading of the paper and constructive suggestions. The second author acknowledges the support from the National Natural Science Foundation of China (Grant No. 11371020) and the project RARE–318984 (an FP7 Marie Curie IRSES Fellowship).

References

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

Asimit, A. V., Badescu, A.M. and Tsanakas, A. 2013. "Optimal risk transfers in insurance groups," *European Actuarial Journal*, 3(1), 159–190.

Asimit, A. V., Furman, E., Tang, Q. and Vernic, R. 2011. "Asymptotics for risk capital allocations based on conditional tail expectation," *Insurance: Mathematics and Economics*, 49(3), 310–324.

Asimit, A. V., Hashorva, E. and Kortschak, D. 2016. "Aggregation of randomly weighted large risks," *IMA Journal of Management Mathematics, Forthcoming.*

Bingham, N.H., Goldie, C.M., and Teugels, J.L. 1987. *Regular Variation*, Cambridge University Press, Cambridge.

Breiman, L. 1965. "On some limit theorems similar to the arc-sin law," *Theory of Probability and its Applications*, 10(2), 323–331.

Cai, J.-J., Einmahl, J.H.J, de Haan, L. and Zhou, C. 2015. "Estimation of the marginal expected shortfall: the mean when a related variable is extreme," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(2), 417–442.

Denuit, M., Dhaene, J., Goovaerts, M.J. and Kaas, R. 2005. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*, Wiley, Chichester.

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

Embrechts, P., Klüppelberg, C. and Mikosch, T. 1997. *Modeling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.

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.

Fischer, T. 2003. "Risk capital allocation by coherent risk measures based on one-sided moments," *Insurance: Mathematics and Economics*, 32(1), 135–146.

Fisher, R.A. and Tippett, L.H.C. 1928. "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2), 180–190.

Föllmer, H. and Schied, A. 2011. *Stochastic finance: An introduction in discrete time*, Walter de Gruyter & Co., Berlin.

Hashorva, E. and Li, J. 2015. "Tail behavior of weighted sums of order statistics of dependent risks," *Stochastic Models*, 31(1), 1–19.

Hashorva, E., Ling, C. and Peng, Z. 2014. "Tail asymptotic expansions for L-statistics," *Science China Mathematics*, 57(10), 1993–2012.

Hua, L. and Joe, H. 2011. "Second order regular variation and conditional tail expectation of multiple risks," *Insurance: Mathematics and Economics*, 49(3), 537–546.

Idierb, J., Laméa, G. and Mésonnierb, J.-S. 2014. "How useful is the marginal expected shortfall for the measurement of systemic exposure? A practical assessment," *Journal of Banking & Finance*, 47, 134–146.

Joe, H. and Li, H. 2011. "Tail risk of multivariate regular variation," *Methodology and Computing in Applied Probability*, 13(4), 671–693.

Jones, B.L. and Zitikis, R. 2003. "Empirical estimation of risk measures and related quantities," *North American Actuarial Journal*, 7(4), 44–54.

Jones, B.L. and Zitikis, R. 2007. "Risk measures, distortion parameters, and their empirical estimation," *Insurance: Mathematics and Economics*, 41(2), 279–297.

Juri, A. and Wüthrich, M.V. 2003. "Tail dependence from a distributional point of view," *Extremes* 6(3), 213–246.

Kallenberg, O. 1983. *Random Measures*, 3rd edition Akademie-Verlag, Berlin.

Koch-Medina, P. and Munari, C. 2016. "Unexpected shortfalls of expected shortfall: extreme default profiles and regulatory arbitrage," *Journal of Banking & Finance,* 62, 141–151.

Kortschak, D. and Albrecher, H. 2009. "Asymptotic results for the sum of dependent non-identically distributed random variables," *Methodology and Computing in Applied Probability* 11(3), 279–306.

Kulik, R. and Soulier, P. 2015. "Heavy tailed time series with extremal independence", *Extremes* 18(2), 273–299.

Mitra, A. and Resnick, S.I. 2009. "Aggregation of rapidly varying risks and asymptotic independence," *Advances in Applied Probability*, 41(3), 797–828.

Resnick, S.I. 1987. *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York.

Resnick, S.I. 2007. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer-Verlag, New York.

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

Zhu, L. and Li, H. 2012. "Asymptotic analysis of multivariate tail conditional expectations," *North American Actuarial Journal*, 16(3), 350–363.