On tail-risk measures for non integrable heavy-tailed random variables

Laurent Gardes

Université de Strasbourg & CNRS, IRMA, UMR 7501,

7 rue René Descartes, 67084 Strasbourg Cedex, France

Abstract

The assessment of risk for heavy-tailed distributions is a crucial question in various fields of application. An important family of risk measures is provided by the class of distortion risk (DR) measures which encompasses the Value-at-Risk and the Tail-Value-at-Risk measures. The Tail-Value-at-Risk is a coherent risk measure (which is not the case for the Value-at-Risk) but it is defined only for integrable quantile functions that is to say for heavy-tailed distributions with a tail index smaller than 1. Moreover, it is a matter of fact that the performance of the empirical estimator is strongly deteriorated when the tail index becomes close to 1. The main contribution of this work is the introduction and the estimation of a new risk measure which is defined for all heavy-tailed distributions and which is tail-equivalent to a coherent DR measure when the tail of the underlying distribution is not too heavy. Its finite sample performance is discussed on a fire claims dataset.
Keywords: Risk measure; heavy-tailed distribution; extreme value theory; finite sample.

1 Introduction

Nowadays, a crucial question for many companies is to assess the risk of a given phenomenon. For a financial analyst, the phenomenon of interest can be the potential loss of a portfolio (see, e.g., Rockafellar and Uryasev [30]). An accurate control of the risk associated to extreme weather events such as hurricane, heat wave, flash

flood among many others is of course essential for insurance companies (see for instance Brazauskas *et al.* [8] and Read and Vogel [28]). From a mathematical point of view, the phenomenon of interest is represented by a real-valued random variable. For the underlying distribution, heavy-tailed models are often considered due to their ability to model extremal events. Such models will be the focus of this paper. Among the literature devoted to the measure of risk for heavy-tailed distributions, one can cite the works of Gardes *et al.* [21] on L^p quantiles, Bellini and Di Bernardino [6] and Daouia *et al.* [10] on extreme expectiles, Daouia *et al.* [11] on extreme M-quantiles and Daouia *et al.* [9] on extremiles. Before presenting the motivation and the contribution of this work, we give some generalities about risk measures.

Generalities about risk measures Let $\mathcal{X}$ be a set of random variables defined on an atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. According to Artzner *et al.* [3], a monetary risk measure is a mapping from $\mathcal{X}$ to $\mathbb{R}$ which is monotone, translation invariant and homogeneous. In what follows, the term "monetary" is omitted. Another desirable property for a risk measure is to be law-invariant, i.e., same value of the risk measure is given for two random variables sharing the same distribution. As pointed out in the monograph of Rüschendorf [31, Remark 7.1.c)], for an atomless probability space, a law-invariant risk measure can be considered as a mapping on the set $\mathcal{Q} := \{Q_X \mid X \in \mathcal{X}\}$ of quantiles functions defined for all $\beta \in [0,1]$ by $Q_X(\beta) := \inf\{x \in \mathcal{S}_X \mid \mathbb{P}(X \leq x) \geq \beta\}$ with the convention $\inf(\emptyset) = +\infty$ and where $\mathcal{S}_X$ is the support of the random variable X. From now on, only law-invariant risk measure $\varrho : \mathcal{Q} \to \mathbb{R}$ are considered. The mapping ϱ is monotone (if $Q_{X_1}(u) \geq Q_{X_2}(u)$ for all $u \in [0,1]$ then $\varrho(Q_{X_1}) \geq \varrho(Q_{X_2})$), translation-invariant and homogeneous (for all $m \in \mathbb{R}$ and $\lambda > 0$, $\varrho(\lambda Q_X + m) = \lambda\varrho(Q_X) + m$).
An important class of risk measure is provided by the set of distortion risk (DR) measures introduced by Wang [32]. A DR measure with distortion probability measure $\mu : \mathcal{B}([0,1]) \to [0,1]$, where $\mathcal{B}([0,1])$ is the Borel set on $[0,1]$, is the mapping $Q_X \in \mathcal{Q} \mapsto \int_{[0,1]} Q_X d\mu \in \mathbb{R}$ if the integral exists. This class encompasses several popular risk measures. When μ is the Dirac measure δ_β centered at $\beta \in (0,1)$, the DR measure is the Value-at-Risk of level $\beta \in (0,1)$ given for $Q_X \in \mathcal{Q}$ by $\mathrm{VaR}_\beta(Q_X) = Q_X(\beta)$. The Tail-Value-at-Risk of level $\beta \in (0,1)$ defined for in-

stance in Kaas *et al.* [24, Definition 5.6.6] for all integrable quantile function Q_X by

$$\mathrm{TVaR}_\beta(Q_X) := \frac{1}{1-\beta}\int_\beta^1 Q_X(u)du,$$

is also a DR measure obtained by taking for μ the uniform distribution on $[\beta, 1]$. One drawback of the Tail-Value-at-Risk is to be defined only for integrable quantile functions while no integrability condition is required fo the Value-at-Risk. On the other hand, as mentioned for instance by Artzner *et al.* [3], the Value-at-Risk is not a coherent risk measure since it is not sub-additive. A measure ϱ is said to be sub-additive if $\varrho(Q_{X_1+X_2}) \leq \varrho(Q_{X_1}) + \varrho(Q_{X_2})$. Sub-additivity is a desirable property of risk measures especially in finance where the risk of a diversified portfolio is expected to be smaller than the sum of the individual risks (see for instance Artzner [2]). Note that a DR measure is coherent if and only if its distortion probability measure μ is convex (i.e., $x \mapsto \mu([0, x])$ is convex, see Acerbi [1] and Yaari [34]). The Tail-Value-at-Risk is thus a coherent measure.

Motivation and contribution In this work, we focus on heavy-tailed models, i.e., we assume that for all $X \in \mathcal{X}$, there exists $\gamma_X > 0$ such that for all $t > 0$, the associated quantile function Q_X satisfies

$$\lim_{u\to 1} \frac{Q_X(1-t(1-u))}{Q_X(u)} = t^{-\gamma_X}. \tag{1}$$

The parameter γ_X is referred to as the tail index. It controls the heaviness of the tail distribution of X: larger the tail index, heavier the tail. Overviews on heavy-tail distributions can be found in the monographs of Beirlant *et al.* [4] and de Haan and Ferreira [22]. Measuring the risk for heavy-tailed distributions can be a difficult task. Indeed, for a given convex distortion probability measure μ, it is common to observe that the corresponding DR measure is defined only when the tail index γ_X is smaller than some positive value. Moreover, as pointed out by El Methni and Stupfler [17, Section 3.1], the variability of the empirical estimator of the DR measure drastically increases as γ_X gets large. This is due to the fact that a large variability in the tail is observed when the tail index increases. As a consequence, coherent DR measures (such as the Tail-Value-at-Risk) are not well adapted to measure the risk in applications where large tail indices can be encountered (e.g., in finance, see for instance Moscadelli [26] and Neslehova et al. [27]).

Of course, for very heavy-tailed distributions, the simplest idea is to measure the risk with the Value-at-Risk. This is clearly not a fully satisfactory solution since the Value-at-Risk is not sub-additive. Ideally, we would like to find a coherent risk measure taking finite values on the class of heavy-tailed distributions but, as mentioned by Delbaen [14], 'there is not immediate solution for this problem', and, to our knowledge, there is indeed no solution yet.

The objective of this paper is to take a step towards the solution by introducing and estimating a new risk measure that realizes a good compromise between a coherent DR measure and a finite risk measure on the class of heavy-tailed distributions. More specifically, for a given convex distortion probability measure μ, the risk measure proposed in this paper is an infimum over a class of DR measures, this class being chosen in such a way that the obtained risk measure is finite for all heavy-tailed distributions and tail-equivalent to the DR measure with distortion probability measure μ, provided that the tail index is smaller than a pre-specified positive value. In this situation, the new measure is also shown to be sub-additive in the tail of the distribution.

Moreover, when we focus on the tail of the distribution, the proposed risk measure can be easily estimated by taking advantage of its tail-equivalent. In a real dataset study, it is shown that, when the tail index becomes large, this estimator is less sensitive to the sample fluctuations than the empirical counterpart of the DR measure.

The paper is organized as follows. The new risk measure is introduced and commented in Section 2. Its asymptotic properties are established in Section 3 and its estimation is considered in Section 4. An application to an insurance dataset (fire claims of a Norwegian insurance company) is also discussed in Section 5. All the proofs are gathered in Section 6.

2 A new risk measure for heavy-tailed distributions

As already mentioned, this paper focus on heavy-tailed distributions. We denote by $\mathcal{Q}^{(HT)}$ the set of heavy-tailed quantile functions Q_X, i.e, satisfying (1), which are continuous, strictly increasing and with a finite left endpoint ($Q_X(0) = c_X \in \mathbb{R}$).

For all probability measures ν defined on $\mathcal{B}([0,1])$, we denote by ν_β with $\beta \in [0,1)$, the push-forward measure under the transformation $\mathcal{T}_\beta : [0,1] \to [\beta,1]$ defined by $\mathcal{T}_\beta(x) = (1-\beta)x + \beta$. The new risk measure depends on a probability measure μ, called the reference probability measure. From now on, we assume that μ is convex and that

$$\gamma_\mu^* := \sup\left\{\gamma > 0 \,\middle|\, \int_{[0,1]} (1-u)^{-\gamma} d\mu(u) < \infty\right\} > 0.$$

It is shown in Lemma 2 that when $\gamma_X < \gamma_\mu^*$ the quantile function $Q_X \in \mathcal{Q}^{(HT)}$ is integrable with respect to μ_β while if $\gamma_X > \gamma_\mu^*$,

$$\int_{[0,1]} |Q_X| d\mu_\beta = \infty.$$

For instance, when μ is the uniform distribution, $\gamma_\mu^* = 1$ (see, e.g., de Haan and Ferreira [22, Exercice 1.16]). When $\mu = \delta_b$ with $b \in [0,1)$ then $\gamma_\mu^* = +\infty$.
We propose to measure the risk of a quantile function $Q_X \in \mathcal{Q}^{(HT)}$ by

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0)) := \inf\left\{\int_{[0,1]} Q_X d\nu_\beta \,\middle|\, \nu \in \mathcal{P}_\mu(\kappa,\gamma_0)\right\}, \tag{2}$$

where, for some $\kappa \in [1,\infty)$ and $\gamma_0 \in (0,\gamma_\mu^*)$, the set $\mathcal{P}_\mu(\kappa,\gamma_0)$ is the class of probability measures containing the reference probability measure μ and all probability measures ν such that

$$\int_{[0,1]} (1-u)^{-\gamma} d\nu(u) \geq \int_{[0,1]} (1-u)^{-\gamma} d\mu(u) \text{ for all } \gamma \in (0,\gamma_0], \tag{3}$$

and

$$\left(\int_{[0,1]} (1-u)^{-\gamma} d\nu(u)\right)^{\gamma_0/\gamma} \leq \kappa \int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u) \text{ for all } \gamma > 0. \tag{4}$$

Comments on the definition of the risk measure

1) First note that for all $\nu \in \mathcal{P}_\mu(\kappa,\gamma_0)$,

$$\int_{[0,1]} Q_X d\nu_\beta = \int_{[0,1]} Q_{X,\beta} d\nu,$$

where $Q_{X,\beta}(\cdot) := Q_X((1-\beta)\cdot + \beta) \in \mathcal{Q}^{(HT)}$ is the conditional quantile of X given that $X > Q_X(\beta)$. As a consequence, the risk measure $\overline{\varrho}_\beta(\cdot \mid \mathcal{P}_\mu(\kappa,\gamma_0))$ gives emphasis to the right tail of the distribution when its level β is close to 1

2) For all $\kappa \in [1, \infty)$ and $\gamma_0 \in (0, \gamma_\mu^*)$, the Dirac measure $\delta_{b(\mu,\gamma_0)}$ with

$$b(\mu, \gamma_0) := 1 - \left(\int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u) \right)^{-1/\gamma_0},$$

belongs to the class $\mathcal{P}_\mu(\kappa, \gamma_0)$ (see Lemma 3). This entails in particular that the risk measure (2) is bounded for all $Q_X \in \mathcal{Q}^{(HT)}$. Indeed, since $Q_X(0) = c_X$, it is readily seen that $\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) \geq c_X$. Moreover, since $0 < \gamma_0 < \gamma_\mu^*$, one has $b(\mu, \gamma_0) \in [0, 1)$ and thus $Q_{X,\beta}(b(\mu, \gamma_0)) < \infty$ leading to

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) \leq \int_{[0,1]} Q_{X,\beta} d\delta_{b(\mu,\gamma_0)} = Q_{X,\beta}(b(\mu, \gamma_0)) < \infty.$$

3) The mapping $\overline{\varrho}_\beta(\cdot \mid \mathcal{P}_\mu(\kappa, \gamma_0))$ is a proper risk measure. The monotonicity is guaranteed by the monotonicity of the integral and for all $\lambda > 0$ and $m \in \mathbb{R}$, $\overline{\varrho}_\beta(\lambda \cdot + m \mid \mathcal{P}_\mu(\kappa, \gamma_0)) = \lambda \overline{\varrho}_\beta(\cdot \mid \mathcal{P}_\mu(\kappa, \gamma_0)) + m$.

4) The choice of the set $\mathcal{P}_\mu(\kappa, \gamma_0)$ can be motivated by looking at the class of strict Pareto distributions. If Q_X is a strict Pareto quantile function, i.e., if there exist $\gamma_X > 0$, $\lambda_X > 0$ and $m_X \in \mathbb{R}$ such that $Q_X(u) = \lambda_X (1-u)^{-\gamma_X} + m_X$, for all $u \in [0, 1]$, then, when $\gamma_X \leq \gamma_0$, condition (3) is equivalent to say that

$$\int_{[0,1]} Q_X d\mu \leq \int_{[0,1]} Q_X d\nu \text{ for all } \nu \in \mathcal{P}_\mu(\kappa, \gamma_0).$$

As a consequence, when $\gamma_X \leq \gamma_0$, the new measure $\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0))$ is equal to the DR measure with distortion probability measure μ_β. More precisely, the explicit expression of the risk measure (2) for strict Pareto distributions is given in the next result.

Proposition 1 *For all $\kappa \in [1, \infty)$ and $\gamma_0 \in (0, \gamma_\mu^*)$, if Q_X is a strict Pareto quantile function, then, if $\gamma_X \leq \gamma_0$,*

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) = \int_{[0,1]} Q_X d\mu_\beta$$

and, if $\gamma_X > \gamma_0$,

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) = \int_{[0,1]} Q_{X,\beta} d\delta_{b(\mu,\gamma_0)} = Q_X(\beta + (1-\beta) b(\mu, \gamma_0)).$$

For the set of strict Pareto quantile functions, which is included in $\mathcal{Q}^{(HT)}$, the new measure given in (2) achieves the objective announced in the introduction: it coincides with the DR measure with distortion probability measure μ provided that $\gamma_X \leq \gamma_0$. When $\gamma_X > \gamma_0$, the new measure switch continuously to the Value-at-Risk of level $\beta + (1-\beta)b(\mu, \gamma_0)$. Let us emphasis again that the transition is designed in such a way that $\overline{\varrho}_\beta\left(\cdot \mid \mathcal{P}_\mu(\kappa, \gamma_0)\right)$ is a proper risk measure on the set of strict Pareto distributions. In particular, when μ is the uniform distribution on $[0,1]$, the risk measure $\overline{\varrho}_\beta(\cdot \mid \mathcal{P}_\mu(\kappa, \gamma_0))$ coincides, on the class of strict Pareto distributions, with the Tail-Value-at-Risk of level β if $\gamma_X \leq \gamma_0$ and with the Value-at-Risk of level $1-(1-\beta)(1-\gamma_0)^{1/\gamma_0}$ otherwise.

Let us emphasis that Proposition 1 remains true if condition (4) in the definition of the set of probability measures $\mathcal{P}_\mu(\kappa, \gamma_0)$ is replaced by the less restrictive condition

$$\int_{[0,1]} (1-u)^{-\gamma} d\mu(u) < \infty \text{ for all } \gamma > 0.$$

In particular, the value of the risk measure for strict Pareto distributions does not depend on κ. However, condition (4) will be necessary in the next section to prove that Proposition 1 is asymptotically true (as $\beta \to 1$) for the set $\mathcal{Q}^{(HT)}$ of heavy-tailed quantile functions. Finally, note that Proposition 1 can be rewritten

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) = \int_{[0,1]} Q_X \, d\nu_X^*,$$

where $\nu_X^* = \mu$ if $\gamma_X \leq \gamma_0$ and $\nu_X^* = \delta_{b(\mu,\gamma_0)}$ if $\gamma_X > \gamma_0$. The new risk measure (2) is thus not a DR measure since the probability measure depends on the underlying distribution.

5) The size of the set of strict Pareto distributions for which the tail risk measure (2) is equal to the DR measure is controlled by the parameter γ_0. A large set is obtain when γ_0 is close to γ_μ^*, but, as already pointed out, the performance of the empirical counterpart of a DR measure is badly affected by values of γ_X close to γ_μ^*. A choice of a reasonable value of γ_0 is proposed in the real data study.

3 Asymptotic properties of the risk measure

This section is devoted to the asymptotic properties (as the level β goes to 1) of the new risk measure presented above. The first result is an asymptotic version of Proposition 1 when $\beta \to 1$. It will be very useful to propose an estimator of the risk measure when the level $\beta = \beta_n$ goes to 1 as the sample size n increases, see Section 4. The following condition on the reference probability measure is required.

(H.1) The reference probability measure μ is such that the function

$$\gamma \in (0, \gamma_\mu^*) \mapsto \int_{[0,1]} (1-u)^{-\gamma} d\mu(u),$$

is continuous.

Note that if μ is the uniform distribution on $[0,1]$, condition (H.1) holds since for all $\gamma \in (0,1)$, $\int_{[0,1]}(1-u)^{-\gamma}d\mu(u) = 1/(1-\gamma)$.

Proposition 2 *Let $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_X > 0$. If μ satisfies condition (H.1) then for all $\kappa \in [1,\infty)$ and $\gamma_0 \in (0, \gamma_\mu^*)$,*

$$\lim_{\beta \to 1} \frac{\overline{\varrho}_\beta \left(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0) \right)}{Q_X(\beta)} = \left[\int_{[0,1]} (1-u)^{-(\gamma_X \wedge \gamma_0)} d\mu(u) \right]^{(\gamma_X \vee \gamma_0)/\gamma_0}. \tag{5}$$

Moreover, as $\beta \to 1$,

$$\overline{\varrho}_\beta \left(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0) \right) \sim \begin{cases} \int_{[0,1]} Q_X d\mu_\beta & \text{if } \gamma_X \leq \gamma_0, \\ Q_X \left(\beta + (1-\beta) b(\mu, \gamma_0) \right) & \text{if } \gamma_X > \gamma_0. \end{cases} \tag{6}$$

Roughly speaking, these equivalences are a consequence of Proposition 1 together with the fact that for a quantile function $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_X > 0$, one has for all $u \in (0,1)$ that $Q_{X,\beta}(u)/Q_X(\beta) \to (1-u)^{-\gamma_X}$ as $\beta \to 1$.

It is worth noting that the asymptotic equivalents do not depend on κ. In fact, only the rate of convergence is influenced by the choice of κ; the larger the value of κ, the slower the convergence. The rate of convergence is also badly influenced by a large value of γ_X. More specifically, we can show that for all $\varepsilon > 0$, there exits $\beta(\varepsilon) \in (0,1)$ such that for all $\beta \geq \beta(\varepsilon)$ and $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$,

$$\left| \frac{\overline{\varrho}_\beta \left(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0) \right)}{Q_X(\beta)} - \left[\int_{[0,1]} (1-u)^{-(\gamma_X \wedge \gamma_0)} d\mu(u) \right]^{(\gamma_X \vee \gamma_0)/\gamma_0} \right| \leq 2\varepsilon g(\kappa, \gamma_X),$$

where $g(\kappa, \gamma_X)$ increases as κ and/or γ_X increases. The expression of $g(\kappa, \gamma_X)$ is given in the proof of Proposition 2.

In the next result, we show that the risk measure $\overline{\varrho}_\beta(\cdot \mid \mathcal{P}_\mu(\kappa, \gamma_0))$ is tail-sub-additive.

Proposition 3 *Assume that condition (H.1) holds for μ. Let X_1 and X_2 be two random variables such that Q_{X_1} and Q_{X_2} belong to the set $\mathcal{Q}^{(\mathrm{HT})}$ with respective tail index γ_1 and γ_2. If $0 < \gamma_1 \leq \gamma_2 \leq \gamma_0$ and if*

$$\lim_{x \to \infty} \frac{\mathbb{P}(X_1 > x, X_2 > x)}{\mathbb{P}(X_2 > x)} = 0, \tag{7}$$

then for all $\kappa \in [1, \infty)$ and $\varepsilon > 0$, there exists $\tau_0 \in (0,1)$ such that for all $\beta > 1 - \tau_0$,

$$\begin{aligned} & \overline{\varrho}_\beta \left(Q_{X_1 + X_2} \mid \mathcal{P}_\mu(\kappa, \gamma_0) \right) \\ \leq \quad & \left[\overline{\varrho}_\beta \left(Q_{X_1} \mid \mathcal{P}_\mu(\kappa, \gamma_0) \right) + \overline{\varrho}_\beta \left(Q_{X_2} \mid \mathcal{P}_\mu(\kappa, \gamma_0) \right) \right] (1 + \varepsilon). \end{aligned}$$

Condition (7) is a so-called asymptotic dependence condition. It ensures that when X_1 and X_2 are heavy-tailed, the sum $X_1 + X_2$ is also heavy-tailed with tail index $\gamma_1 \vee \gamma_2$ (see e.g., Davis and Resnick [13, Lemma 2.1]).

As a conclusion, Propositions 2 and 3 ensure that the new risk measure achieves the objectif we have laid out in the introduction. The new risk measure (2) can thus be seen as a good alternative between a coherent DR measure but subject to an integrability condition and the Value-at-Risk which is not a coherent risk measure (but defined everywhere).

4 Estimation

In practice, the distribution of the random variable of interest X is unknown. The available information is contained in a realization of a sample $X_1, \cdots, X_n$ of n independent replications of X. Throughout this section, the quantile function Q_X is taken in the set $\mathcal{Q}^{(\mathrm{HT})}$.

To implement the risk measure introduced above, we are looking for an estimator

of $\overline{\varrho}_{\beta_n}(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0))$ where β_n is a sequence converging to 1 as the sample size increases and μ is a given reference probability measure. Two situations for the rate of convergence of β_n to 1 are considered:

i) the *intermediate case* where $\beta_n \to 1$ and $n(1-\beta_n) \to \infty$ as $n \to \infty$,

ii) the *extreme case* where $n(1-\beta_n) \to b \in (0,\infty)$ as $n \to \infty$.

In the first scenario, the quantile $Q_X(\beta_n)$ can be consistently estimated by its empirical counterpart $\widehat{Q}_{n,X}(\beta_n) = X_{\lceil n\beta_n \rceil:n}$ where $\lceil\cdot\rceil$ is the ceiling function and $X_{1:n} \leq \cdots \leq X_{n:n}$ are the observations arranged in ascending order (see for instance de Haan and Ferreira [22, Theorem 2.4.1]). As a consequence, we will show hereafter that a consistent estimator of $\overline{\varrho}_{\beta_n}(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0))$ can be obtained directly from the equivalence (5) in Proposition 2.
The second scenario is more challenging since $Q_X(\beta_n)$ cannot be estimated by an order statistic. We thus have to extrapolate the intermediate estimator proposed in the first scenario to an extreme level.

Estimation in the intermediate case We take advantage of the asymptotic equivalence (5) obtained in Proposition 2 to propose the following estimator

$$\widehat{\overline{\varrho}}_{n,\beta_n}(Q_X \mid \mu,\gamma_0) := X_{\lceil n\beta_n \rceil:n} \left(\int_{[0,1]} (1-u)^{-(\widehat{\gamma}_{n,X} \wedge \gamma_0)} d\mu(u) \right)^{(\widehat{\gamma}_{n,X} \vee \gamma_0)/\gamma_0}, \quad (8)$$

where $\widehat{\gamma}_{n,X}$ is any estimator of γ_X that converges in probability to γ_X (in short $\widehat{\gamma}_{n,X} \xrightarrow{\mathbb{P}} \gamma_X$). Consistency of (8) is established in the next result.

Proposition 4 *Assume that the reference probability measure μ satisfies condition (H.1). If $\beta_n \to 1$ and $n(1-\beta_n) \to \infty$ then for all $\kappa \in [1,\infty)$, $\gamma_0 \in (0,\gamma_\mu^*)$ and for any estimator $\widehat{\gamma}_{n,X} \xrightarrow{\mathbb{P}} \gamma_X$,*

$$\frac{\widehat{\overline{\varrho}}_{n,\beta_n}(Q_X \mid \mu,\gamma_0)}{\overline{\varrho}_{\beta_n}\left(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0)\right)} \xrightarrow{\mathbb{P}} 1.$$

As expected, the consistency holds for all $\gamma_X > 0$. Let us pay particular attention to the interesting situation where the reference probability measure μ is the uniform

distribution. In this case, estimator (8) is given by

$$\widehat{\varrho}_{n,\beta_n}(Q_X \mid \mu, \gamma_0) = \begin{cases} X_{\lceil n\beta_n \rceil:n}/[1-\widehat{\gamma}_{n,X}] & \text{if } \gamma_X \in (0, \gamma_0], \\ X_{\lceil n\beta_n \rceil:n}/[1-\gamma_0]^{\widehat{\gamma}_{n,X}} & \text{if } \gamma_X > \gamma_0. \end{cases}$$

When $\gamma_X \leq \gamma_0$, the new estimator coincides with the estimator of the Tail-Value-at-Risk of level β_n studied in Gardes *et al.* [21, Section 4.2] where its asymptotic normality has been established under the condition $\gamma_X < 1/2$. It is worth pointing out that when $\gamma_X < 1$ is close to 1, the tail index estimator $\widehat{\gamma}_{n,X}$ is also expected to be close to 1 (or even larger!). As a consequence, a small variation in the value of $\widehat{\gamma}_{n,X}$ leads to an important change in the value of the Tail-Value-at-Risk estimator $X_{\lceil n\beta_n \rceil:n}/[1-\widehat{\gamma}_{n,X}]$. This is no longer the case for the new estimator since when $\gamma_X > \gamma_0$, it switches to a more stable estimator of the Value-at-Risk of level $\beta_n + (1-\beta_n)[1-(1-\gamma_0)^{1/\gamma_0}]$. This point is illustrated in Section 5 where a choice of γ_0 is proposed.

A natural estimator of the tail index γ_X is the Hill estimator, see Hill [23], defined for a sequence $k_n \in \{1, \cdots, n\}$ by

$$\widehat{\gamma}^{(\mathrm{H})}_{n,X}(k_n) := \frac{1}{k_n} \sum_{i=1}^{k_n} \log \frac{X_{n-i+1:n}}{X_{n:n}}.$$

The convergence in probability of the Hill estimator is established for instance in de Haan and Ferreira [22, Theorem 3.2.2] under the condition that $k_n \to \infty$ and $n/k_n \to \infty$. As a consequence, the estimator $\widehat{\varrho}^{(\mathrm{H})}_{n,\beta_n}(Q_X \mid \mu, \gamma_0)$ obtained by using $\widehat{\gamma}^{(\mathrm{H})}_{n,X}(k_n)$ with $k_n = n\lceil 1-\beta_n \rceil$ as a tail index estimator is a relative consistent estimator of the risk measure $\overline{\varrho}_{\beta_n}(Q_X \,|\mathcal{P}_\mu(\kappa, \gamma_0))$.

To determinate the rate of convergence in Proposition 4 or to establish the asymptotic normality of (8), we need in particular to establish the rate of convergence in Proposition 2 which is a difficult task. This question is not addressed in this paper.

Estimation in the extreme case – Of course when $n(1-\beta_n) \to b \geq 0$, there is no hope for (8) to be a consistent estimator since $X_{\lceil n\beta_n \rceil:n}/Q_X(\beta_n)$ does not converge in probability to 1. We thus need to extrapolate the estimate at an intermediate level α_n to the extreme level β_n. The extrapolation procedure is based

on the approximation

$$\frac{\overline{\varrho}_{\beta_n}\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)}{\overline{\varrho}_{\alpha_n}\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)} \approx \left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\gamma_X}.$$

This approximation is a direct consequence of the definition (1) of heavy-tailed quantile functions. A similar idea was used by Weissman [33] for the estimation of extreme quantiles. Consequently, for an extreme level β_n, we propose the extrapolated estimator

$$\widehat{\overline{\varrho}}^{(\mathrm{W})}_{n,\beta_n}(Q_X \mid \mu,\gamma_0,\alpha_n) := \widehat{\overline{\varrho}}_{n,\alpha_n}(Q_X \mid \mu,\gamma_0)\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\widehat{\gamma}_{n,X}}, \tag{9}$$

where $\widehat{\gamma}_{n,X}$ is a consistent estimator of γ_X and α_n is an intermediate sequence (i.e., $\alpha_n \to 1$ and $n(1-\alpha_n) \to \infty$). Before establishing the consistency of (9), let us mention that (1) is equivalent to say that $Q_X(u) = (1-u)^{-\gamma_X}\ell_X((1-u)^{-1})$ where ℓ_X is a slowly varying function at infinity, i.e., for all $t > 0$,

$$\lim_{x\to\infty}\frac{\ell_X(tx)}{\ell_X(x)} = 1. \tag{10}$$

It is well known (see for instance Bingham *et al.* [7, Theorem 1.3.1]) that a slowly varying function can be represented for all $x > 1$ as,

$$\ell_X(x) = c_X(x)\exp\left(\int_1^x \frac{\Delta_X(t)}{t}dt\right), \tag{11}$$

where $c_X(x) \to c \geq 0$ as $x \to \infty$ and $\Delta_X(t) \to 0$ as $t \to \infty$. The function Δ_X controls the rate of convergence in (10). The consistency of (9) is established in the next result.

Proposition 5 *Assume that the reference probability measure μ satisfies condition (H.1) and that $|\Delta_X|$ is asymptotically decreasing. Let α_n and β_n be sequences converging to 1 and such that $n(1-\alpha_n) \to \infty$, $n(1-\beta_n) \to b \geq 0$ and*

$$\lim_{n\to\infty}\ln\left(\frac{1-\alpha_n}{1-\beta_n}\right)\left|\Delta_X\left(\frac{1}{1-\alpha_n}\right)\right| = 0. \tag{12}$$

If the estimator $\widehat{\gamma}_{n,X}$ satisfies the condition

$$\ln\left(\frac{1-\alpha_n}{1-\beta_n}\right)(\widehat{\gamma}_{n,X}-\gamma_X) \xrightarrow{\mathbb{P}} 0, \tag{13}$$

then for all $\kappa \in [1,\infty)$ *and* $\gamma_0 \in (0,\gamma_\mu^*)$,

$$\frac{\widehat{\overline{\varrho}}^{(\mathrm{W})}_{n,\beta_n}(Q_X \mid \mu,\gamma_0,\alpha_n)}{\overline{\varrho}_{\beta_n}\left(Q_X \,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)} \xrightarrow{\mathbb{P}} 1.$$

Note that conditions (12) and (13) are required to ensure that the Weissman estimator

$$X_{\lceil n\alpha_n\rceil:n}\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\widehat{\gamma}_{n,X}},$$

is a relative consistent estimator of the extreme quantile $Q_X(\beta_n)$.

When the reference probability measure μ is the uniform distribution, the extrapolated estimator (9) coincides, when $\gamma_X \leq \gamma_0$, to the so-called indirect extrapolated estimator of the Tail-Value-at-Risk proposed in Gardes *et al.* [21]. The latter is equivalent in probability to the Tail-Value-at-Risk when $\gamma_X < 1$ and asymptotically Gaussian when $\gamma_X < 1/2$. Again, (9) is expected to be more stable than the indirect extrapolated estimator of Gardes *et al.* [21] when the tail index $\gamma_X < 1$ is close to 1.

Let us finally take a look to the particular case where the Hill estimator is used for the estimation of the tail index. It can be shown (see Lemma 5) that under additional technical conditions, the Hill estimator $\widehat{\gamma}^{(\mathrm{H})}_{n,X}(k_n)$ with $k_n = \lceil n(1-\alpha_n)\rceil$ satisfies (13) and thus can be used in (9) as an estimator of the tail index leading to a relative consistent estimator $\widehat{\overline{\varrho}}^{(\mathrm{W,H})}_{n,\beta_n}(Q_X \mid \mu,\gamma_0,\alpha_n)$ of $\overline{\varrho}_{\beta_n}(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0))$.

5 Real data study

We illustrate the estimation procedure on a fire insurance dataset studied by Beirlant *et al.* [5]. This dataset provides the sizes of 9181 fire insurance claims for a Norwegian insurance company for the period 1972 to 1992. These claim sizes are corrected for inflation using the Norwegian CPI and are expressed in thousands Norwegian Krone (NKR). The dataset is available for instance in the R package `CASdatasets` that can be downloaded at the address `http://dutangc.perso.math.cnrs.fr/RRepository/pub/`.

For the period 1985 to 1992, the annual numbers of claim sizes are similar. We

thus focus on this period for this study. For each year $j \in \{1985, \cdots, 1992\}$, we denote by $x_1^{(j)}, \cdots, x_{n_j}^{(j)}$ the observed n_j fire losses. As in Gardes and Girard [20], it is assumed that these observations are realizations of a sample $X_1^{(j)}, \cdots, X_{n_j}^{(j)}$ of n_j independent copies of a heavy-tailed random variable $X^{(j)}$.

Our goal is to compare the annual risks of the fire losses over the considered period. More specifically, for each year j, we estimate the risk measure $\overline{\varrho}_\beta\left(Q_{X^{(j)}} \mid \mathcal{P}_\mu(\kappa, \gamma_0)\right)$ when μ is the uniform probability measure. For the level β, we take the value $1-1/\tilde{n}$ where $\tilde{n} = 637.5$ is the median of the set $\{n_{1985}, \cdots, n_{1992}\}$. When the sample size n_j is lower than $\tilde{n}$, the level β can be considered as extreme since $n_j(1-\beta) = n_j/\tilde{n} < 1$. This is the case for the years 1985, 1990, 1991 and 1992. For each year j, we thus chose to use the extrapolated estimator $\widehat{\overline{\varrho}}_{n_j,\beta}^{(\mathrm{W},\mathrm{H})}(Q_{X^{(j)}} \mid \mu, \gamma_0, \alpha_{n_j})$. Its expression is given by

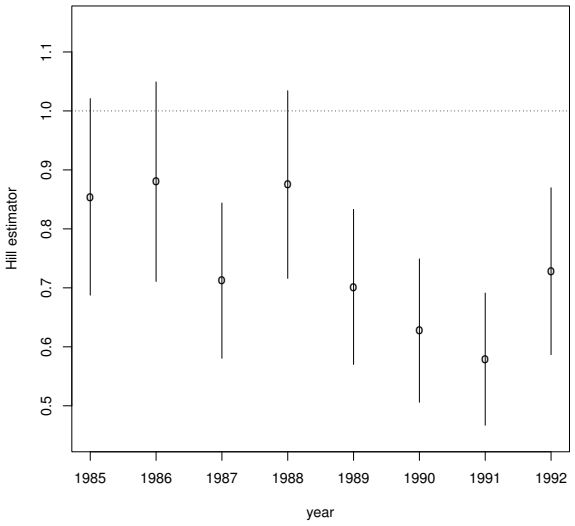
$$\left(\frac{1-\alpha_{n_j}}{1-\beta}\right)^{\widehat{\gamma}_j^{(H)}} X^{(j)}_{\lceil n\alpha_{n_j}\rceil:n_j} \left(\frac{1}{1-\widehat{\gamma}_j^{(H)} \wedge \gamma_0}\right)^{(\widehat{\gamma}_j^{(H)} \vee \gamma_0)/\gamma_0}, \tag{14}$$

where $\alpha_{n_j} = 1 - \ln^2(n_j)/(2n_j)$ and $\widehat{\gamma}_j^{(H)} := \widehat{\gamma}^{(\mathrm{H})}_{n_j, X^{(j)}}(k_{n,j})$ is the Hill estimator computed with $k_{n_j} = \lceil n_j(1-\alpha_{n_j})\rceil$ order statistics. For the parameter γ_0, we take the value $1/2$. This choice is motivated by the fact that the second moment of a heavy-tailed distribution exits as soon as the tail index is lower than $1/2$. The risk measure (14) is then asymptotically equivalent (as $\beta \to 1$) to the Tail-Value-at-Risk only in this situation. This limitation was suggested by Gardes *et al.* [21] in order to obtain a stable estimation of the Tail-Value-at-Risk. In Figure 1, we represent for each year j the value of the Hill estimator $\widehat{\gamma}_j^{(H)}$ together with its asymptotic confidence interval of level 0.95 given by

$$\left[\widehat{\gamma}_j^{(H)}\left(1 - k_n^{-1/2} u_{0.975}\right) \; ; \; \widehat{\gamma}_j^{(H)}\left(1 - k_n^{-1/2} u_{0.025}\right)\right], \tag{15}$$

where u_α is the quantile of order α of a standard normal distribution. This asymptotic confidence interval is constructed on the base of the asymptotic normality of the Hill estimator (see de Haan and Ferreira [22, Theorem 3.2.5]).

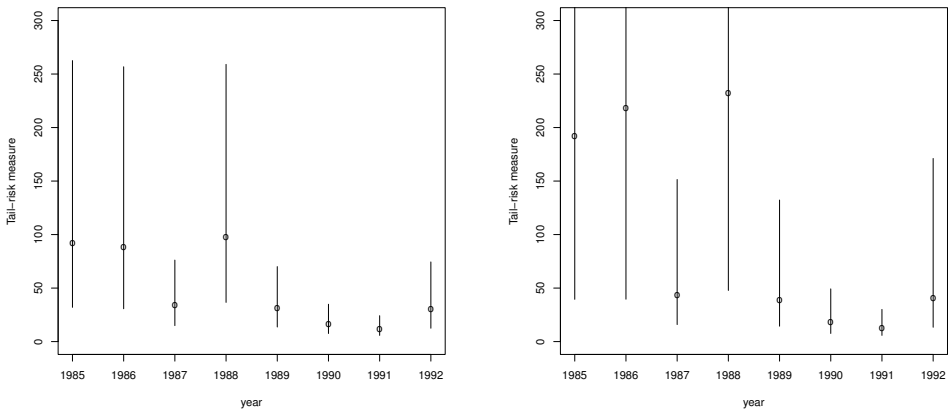
Figure 1: For each year $j \in \{1985, \cdots, 1992\}$, the point o represents the value of the Hill estimator and the vertical full line its confidence interval of level 0.95. The horizontal line is the critical value $\gamma_\mu^* = 1$.



It appears that for the years 1985, 1986 and 1988, the critical value $\gamma_\mu^* = 1$ is contained in the asymptotic confidence interval (15). Consequently, since the quantile function Q_X is not integrable when $\gamma_X > 1$, we can reasonably have some doubt about the existence of the Tail-Value-at-Risk for these years. In the left panel of Figure 2, the values of the extrapolated tail risk measure estimator (14) are depicted for each year j. The interval of variation of the estimator when the tail index estimator varies in the confidence interval (15) is also represented. The lower (resp. upper) bound of this interval of variation is thus the estimator (14) where the Hill estimator is replaced by the lower (resp. upper) bound of the confidence interval (15). Same thing is done in the right panel with the indirect extrapolated estimator of the Tail-Value-at-Risk introduced in Gardes *et al.* [21] and given by

$$\left(\frac{1-\alpha_{n_j}}{1-\beta}\right)^{\widehat{\gamma}_j^{(H)}} \frac{X^{(j)}_{\lceil n_j \alpha_{n_j} \rceil : n_j}}{1-\widehat{\gamma}_j^{(H)}}. \tag{16}$$

Figure 2: Left panel: for each year $j \in \{1985, \cdots, 1992\}$, value of the estimator (14) together with its interval of variation when the Hill estimator varies in the confidence interval (15). Right panel: idem with the Tail-Value-at-Risk estimator (16).



This estimator is not valid when the tail index is larger than 1. As a consequence, for the years 1985, 1986 and 1988, the upper bound of the interval of variation is infinite since for these years, the right bound of the confidence interval (15) is larger than 1. As expected, the estimator (14) is less sensitive to the variations of the tail index estimator. For the period 1989 to 1992, the estimators (14) and (16) are similar since for these years, the tail index is not too large. At the opposite, for the years 1985, 1986 and 1988, the range of variation of the indirect extrapolated estimator (16) is important which is not surprising since the estimation of the tail index is close to 1.

Conclusion

The main contribution of this work is the introduction and the estimation of a new risk measure dedicated to heavy-tailed distributions. Contrary to the Tail-Value-at-Risk, it presents the advantage to be defined whatever the value of the tail index. It is preferred to use the new risk measure instead of the classical Value-at-Risk since, for reasonable values of the tail index, it is shown to be equivalent to a coherent DR measure. From a finite sample point of view, when the tail index is large,

the proposed estimator is less sensitive to sample fluctuations than the empirical counterpart of a coherent DR measure. For these reasons, the risk measure proposed in this paper seems to be a good alternative to the use of Value-at-Risk for very heavy-tailed distributions.

6 Proofs

In this section, we use the following standard notation in extreme value theory: U_X is the function defined for all $x \geq 1$ by $U_X(x) = Q_X(1 - 1/x)$. When $Q_X \in \mathcal{Q}^{(HT)}$, the function U_X is regularly varying with index γ_X that is to say that for all $t > 0$,

$$\lim_{x\to\infty} \frac{U_X(tx)}{U_X(x)} = t^{\gamma_X}.$$

In particular, we have the following result (proved for instance in de Haan and Ferreira [22, Proposition B.1.10]) which will be used several times in this section.

Lemma 1 *For $u \in (0,1)$ and $\beta \in (0,1)$, let*

$$R_X(u,\beta) := \frac{U_X((1-\beta)^{-1}(1-u)^{-1})}{U_X((1-\beta)^{-1})} - (1-u)^{-\gamma_X}. \tag{17}$$

For all $\varepsilon > 0$ and $\delta > 0$, there exists $\tau_0 = \tau_0(\varepsilon,\delta) \in (0,1)$ such that for all $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ and $\beta > 1 - \tau_0$ one has $|R_X(u,\beta)| \leq \varepsilon(1-u)^{-(\gamma_X+\delta)}$.

6.1 Preliminary results

We first show that $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ is integrable with respect to μ if its tail index γ_X is lower than γ_μ^*.

Lemma 2 *For all quantile functions $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_X > 0$ one has*

$$\begin{cases} \int_{[0,1]} |Q_X| d\mu_\beta < \infty & \text{if } \gamma_X \in (0, \gamma_\mu^*), \\ \int_{[0,1]} |Q_X| d\mu_\beta = \infty & \text{if } \gamma_X > \gamma_\mu^*. \end{cases}$$

Proof – Without loss of generality, we assume that $Q_X(u) \geq 0$ for all $u \in [0,1]$ (add c_X if it is not the case). We thus have $Q_{X,\beta}(u) \geq 0$ for all $u \in [0,1]$. It thus suffices to work with the integral

$$\int_{[0,1]} Q_X d\mu_\beta = \int_{[0,1]} Q_{X,\beta} d\mu.$$

First assume that $\gamma_X \in (0, \gamma_\mu^*)$. For all $x \geq 1$, we have

$$\int_{[0,1]} Q_X d\mu_\beta = Q_X(\beta) \int_{[0,1]} \frac{U_X((1-\beta)^{-1}(1-u)^{-1})}{U_X((1-\beta)^{-1})} d\mu(u).$$

Hence,

$$\int_{[0,1]} Q_X d\mu_\beta \leq Q_X(\beta) \left[\int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u) + \int_{[0,1]} |R_X(u,\beta)| d\mu(u) \right],$$

where $R_X(u,\beta)$ is defined in Lemma 1, equation (17). Taking $\delta = (\gamma_\mu^* - \gamma)/2 > 0$ in Lemma 1, for all $\varepsilon > 0$, there exists $\tau_0 \in (0,1)$ such that for $\beta > 1 - \tau_0$,

$$\int_{[0,1]} Q_X d\mu_\beta < Q_X(\beta) \left[\int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u) + \varepsilon \int_{[0,1]} (1-u)^{-(\gamma_X+\delta)} d\mu(u) \right].$$

We conclude the first part of the proof by remarking that

$$\int_{[0,1]} (1-u)^{-(\gamma_X+\delta)} d\mu < \infty,$$

since $\gamma_X + \delta < \gamma_\mu^*$.

Now assume that $\gamma_X > \gamma_\mu^*$. We start with

$$\int_{[0,1]} Q_X d\mu_\beta = \int_{[0,1]} \frac{U_X((1-\beta)^{-1}(1-u)^{-1})}{U_X((1-\beta)^{-1})} d\mu(u).$$

Using Potter's bounds (see for instance de Haan and Ferreira [22, eq. (B.1.19)]) for $\delta < \gamma_X - \gamma_\mu^*$, there exists β_0 such that for all $\beta > \beta_0$ and $u \in (0,1)$,

$$\frac{1}{2}(1-u)^{-(\gamma_X-\delta)} \leq \frac{U_X((1-\beta)^{-1}(1-u)^{-1})}{U_X((1-\beta)^{-1})}.$$

Hence,

$$\int_{[0,1]} Q_X d\mu_\beta \geq \frac{1}{2} \int_{[0,1]} (1-u)^{-(\gamma_X-\delta)} d\mu(u),$$

which is infinite since $\gamma_X - \delta > \gamma_\mu^*$. ■

Lemma 3 *For all $\kappa \in [1, \infty)$ and $\gamma_0 \in (0, \gamma_\mu^*)$, the Dirac measure $\delta_{b(\mu,\gamma_0)}$ belongs to the set $\mathcal{P}_\mu(\kappa, \gamma_0)$.*

Proof − When $\gamma \leq \gamma_0$, an application of the Jensen's inequality leads to

$$1 - \left(\int_{[0,1]} (1-u)^{-\gamma} d\mu(u) \right)^{-1/\gamma} \leq b(\mu, \gamma_0).$$

As a consequence, for all $\gamma \in (0, \gamma_0]$,

$$\int_{[0,1]} (1-u)^{-\gamma} d\mu(u) \leq (1 - b(\mu, \gamma_0))^{-\gamma} = \int_{[0,1]} (1-u)^{-\gamma} d\delta_{b(\mu,\gamma_0)}(u),$$

and condition (3) holds. Moreover, for all $\gamma > 0$,

$$\int_{[0,1]} (1-u)^{-\gamma} d\delta_{b(\mu,\gamma_0)}(u) = \left(\int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u) \right)^{\gamma/\gamma_0}.$$

Condition (4) is thus satisfied with $\kappa = 1$ and thus for all $\kappa \in [1, \infty)$ ensuring that $\delta_{b(\mu,\gamma_0)} \in \mathcal{P}_\mu(\kappa, \gamma_0)$. ■

In the next result, an asymptotic equivalent, as $\beta \to 1$, of the DR measure

$$\int_{[0,1]} Q_X d\mu_\beta,$$

is given.

Lemma 4 *For all probability measure μ such that $\gamma_\mu^* > 0$ and all quantile functions $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_X \in (0, \gamma_\mu^*)$*

$$\lim_{\beta \to 1} \int_{[0,1]} \frac{Q_X(u)}{Q_X(\beta)} d\mu_\beta(u) = \int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u).$$

Proof − We have

$$\begin{aligned} \int_{[0,1]} \frac{Q_X(u)}{Q_X(\beta)} d\mu_\beta(u) &= \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu(u) \\ &= \int_{[0,1]} \frac{U_X((1-\beta)^{-1}(1-u)^{-1})}{U_X((1-\beta)^{-1})} d\mu(u) \\ &= \int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u) + \int_{[0,1]} |R_X(u,\beta)| d\mu(u), \end{aligned}$$

where $R_X(u, \beta)$ is defined in equation (17). From Lemma 1 with $\delta > 0$ such that $\gamma_X + \delta < \gamma_\mu^*$, for all $\varepsilon > 0$, there exists $\tau_0 \in (0, 1)$ such that for all $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ and

$\beta > 1 - \tau_0$,

$$\int_{[0,1]} |R_X(u,\beta)| d\mu(u) \leq \varepsilon \int_{[0,1]} (1-u)^{-(\gamma_X+\delta)} d\mu(u).$$

Since $\int_{[0,1]} (1-u)^{-(\gamma_X+\delta)} d\mu(u) < \infty$ the desired result is proved. ■

The last result deals with the Hill estimator of the tail index. It is obtained under classical restrictions on the representation (11) of the slowly varying function ℓ_X.

(H.2) In representation (11), the function c_X is constant, the sign of the function Δ_X is asymptotically constant and $|\Delta_X|$ is a regularly varying function of index $\rho < 0$.

When c_X is a constant function, the slowly varying function is said to be normalized. As mentioned by Bingham *et al.* [7, Page 15], "we lose nothing by restricting attention to the case of constant c-function in (11)". The parameter $\rho < 0$, called the second-order parameter, tunes the rate of convergence in (10).

Lemma 5 *Assume that the reference probability measure μ satisfies condition (H.1) and that condition (H.2) holds. If α_n and β_n are sequences converging to 1 and such that $n(1-\alpha_n) \to \infty$, $n(1-\beta_n) \to b \geq 0$,*

$$[n(1-\alpha_n)]^{-1/2} \ln\left(\frac{1-\alpha_n}{1-\beta_n}\right) \to 0 \text{ and } [n(1-\alpha_n)]^{1/2} \Delta_X\left(\frac{1}{1-\alpha_n}\right) \to \lambda \in \mathbb{R},$$

as $n \to \infty$, then

$$\ln\left(\frac{1-\alpha_n}{1-\beta_n}\right)\left(\widehat{\gamma}_{n,X}^{(H)} - \gamma_X\right) \xrightarrow{\mathbb{P}} 0.$$

Proof − Our goal is to prove that

$$(n(1-\alpha_n))^{1/2}\left(\widehat{\gamma}_{n,X}^{(\mathrm{H})} - \gamma_X\right) \xrightarrow{d} \mathcal{N}(0, \gamma_X^2). \tag{18}$$

To prove (18), we use de Haan and Ferreira [22, Theorem 3.2.5] in which the asymptotic normality of the Hill estimator is established. We thus have to check that under the assumptions of Lemma 5, the conditions of de Haan and Ferreira [22, Theorem 3.2.5] are satisfied. After a comparison between the set of conditions, we notice that we only have to check that the second order condition

$$\lim_{u\to 1} \frac{1}{\Delta_X((1-u)^{-1})}\left(\frac{Q_X(1-t(1-u))}{Q_X(u)} - t^{-\gamma_X}\right) = t^{-\gamma_X}\int_1^{t^{-1}} v^{\rho-1} dv.$$

holds or equivalently that for all $t > 1$,

$$\lim_{x\to\infty} \frac{\ell_X(tx)/\ell_X(x) - 1}{\Delta_X(x)} = \int_1^t u^{\rho-1} du = \frac{t^\rho - 1}{\rho}.$$

Under the conditions of the lemma,

$$\frac{\ell_X(tx)}{\ell_X(x)} = \exp\left(\int_x^{tx} \frac{\Delta_X(u)}{u} du\right) = \exp\left(\Delta_X(x) \int_1^t \frac{\Delta_X(ux)}{\Delta_X(x)} \frac{1}{u} du\right).$$

Since the sign of Δ_X is asymptotically constant and $|\Delta_X|$ is regularly varying of index $\rho \leq 0$, we have from Bingham *et al.* [7, Theorem 1.2.1] that

$$\lim_{x\to\infty} \frac{\Delta_X(ux)}{\Delta_X(x)} = u^\rho,$$

uniformly on u in a compact set. Hence,

$$\lim_{x\to\infty} \int_1^t \frac{\Delta_X(ux)}{\Delta_X(x)} \frac{1}{u} du = \int_1^t u^{\rho-1} du,$$

and the conclusion follows since $\Delta_X(x) \to 0$ as $x \to \infty$ and $(\exp(u) - 1)/u \to 1$ as $u \to 0$. ■

6.2 Proofs of main results

Proof of Proposition 1 − For a strict Pareto quantile function given for $u \in [0, 1]$ by $Q_X(u) = \lambda_X(1-u)^{-\gamma_X} + m_X$, one has

$$\begin{aligned} & \overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) = \lambda_X \inf\left\{ \int_{[0,1]} (1-u)^{-\gamma_X} d\nu_\beta(u) \middle| \nu \in \mathcal{P}_\mu(\kappa, \gamma_0) \right\} + m_X \\ = \quad & \lambda_X (1-\beta)^{-\gamma_X} \inf\left\{ \int_{[0,1]} (1-u)^{-\gamma_X} d\nu(u) \middle| \nu \in \mathcal{P}_\mu(\kappa, \gamma_0) \right\} + m_X \end{aligned}$$

Since $\gamma_X \in (0, \gamma_0]$, from (3) we know that for all $\nu \in \mathcal{P}_\mu(\kappa, \gamma_0)$,

$$\int_{[0,1]} (1-u)^{-\gamma_X} d\nu(u) \geq \int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u).$$

Hence, since $\mu \in \mathcal{P}_\mu(\kappa, \gamma_0)$, one has for all strict Pareto quantile function Q_X with $\gamma_X \in (0, \gamma_0]$ that

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa, \gamma_0)) = \lambda_X (1-\beta)^{-\gamma_X} \int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u) + m_X = \int_{[0,1]} Q_X d\mu_\beta.$$

It remains to prove that for all $\gamma_X > \gamma_0$,

$$\int_{[0,1]} (1-u)^{-\gamma_X} d\delta_{b(\mu,\gamma_0)}(u) \leq \int_{[0,1]} (1-u)^{-\gamma_X} d\nu(u), \tag{19}$$

for all $\nu \in \mathcal{P}_\mu(\kappa, \gamma_0)$. Indeed, if (19) holds, then

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0)) = \int_{[0,1]} Q_{X,\beta} d\delta_{b(\mu,\gamma_0)} = Q_X\left(\beta + (1-\beta) b(\mu,\gamma_0)\right).$$

Let us use an indirect proof to prove (19). Assume that there exists a probability measure $\nu \in \mathcal{P}_\mu(\kappa,\gamma_0)$ such that

$$\left(\int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u)\right)^{\gamma_X/\gamma_0} > \int_{[0,1]} (1-u)^{-\gamma_X} d\nu(u) \tag{20}$$

By Jensen's inequality,

$$\int_{[0,1]} (1-u)^{-\gamma_0} d\nu(u) < \left(\int_{[0,1]} (1-u)^{-\gamma_X} d\nu(u)\right)^{\gamma_0/\gamma_X},$$

Then, if (20) holds,

$$\int_{[0,1]} (1-u)^{-\gamma_0} d\nu(u) < \int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u),$$

which is impossible for a measure $\nu \in \mathcal{P}_\mu(\kappa,\gamma_0)$. Hence

$$\begin{aligned} \int_{[0,1]} (1-u)^{-\gamma_X} d\delta_{b(\mu,\gamma_0)}(u) &= \left(\int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u)\right)^{\gamma_X/\gamma_0} \\ &\leq \int_{[0,1]} (1-u)^{-\gamma_X} d\nu(u), \end{aligned}$$

and the proof is complete. ■

Proof of Proposition 2 – Let us first prove equivalence (5). Let Q_X be a quantile function in $\mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_X > 0$. For all $\varepsilon > 0$ and $\beta \in (0,1)$, there exists a measure $\nu^{(\beta)}_{X,\varepsilon} \in \mathcal{P}_\mu(\kappa,\gamma_0)$ such that

$$0 \leq \int_{[0,1]} Q_{X,\beta} d\nu^{(\beta)}_{X,\varepsilon} - \overline{\varrho}_\beta\left(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0)\right) \leq \varepsilon. \tag{21}$$

Let us now introduce the set

$$B_\varepsilon = \left\{ \beta \in (0,1) \left| \int_{[0,1]} Q_\beta d\nu_{X,\varepsilon}^{(\beta)} < \int_{[0,1]} Q_\beta d\nu_X^* \right. \right\},$$

where the probability measure ν_X^* is given by $\nu_X^* = \mu$ if $\gamma_X \leq \gamma_0$ and $\nu_X^* = \delta_{b(\mu,\gamma_0)}$ if $\gamma_X > \gamma_0$.

If $\beta_1 := \sup B_\varepsilon < 1$, then for all $\beta > \beta_1$ or for all $\beta \notin B_\varepsilon$,

$$\int_{[0,1]} Q_{X,\beta} d\nu_{X,\varepsilon}^{(\beta)} \geq \int_{[0,1]} Q_{X,\beta} d\nu_X^*,$$

and thus, for all $\beta > \beta_1$,

$$0 \leq \int_{[0,1]} Q_{X,\beta} d\nu_X^* - \overline{\varrho}_\beta \left(Q_X \,|\mathcal{P}_\mu(\kappa,\gamma_0) \right) \leq \varepsilon. \tag{22}$$

Now, if $\sup B_\varepsilon = 1$, using Lemma 1, for all $\delta > 0$ (with $0 < \delta < \gamma_\mu^* - \gamma_X$ if $\gamma_X < \gamma_\mu^*$), there exists $\beta_2 \in (0,1)$ such that for all $\beta \in B_\varepsilon$ with $\beta > \beta_2$,

$$\begin{aligned} 0 &\leq \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu_X^*(u) - \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu_{X,\varepsilon}^{(\beta)}(u) \\ &\leq \int_{[0,1]} (1-u)^{-\gamma_X} d\nu_X^*(u) - \int_{[0,1]} (1-u)^{-\gamma_X} d\nu_{X,\varepsilon}^{(\beta)}(u) \\ &\quad + \varepsilon \left[\int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\nu_X^*(u) + \int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\nu_{X,\varepsilon}^{(\beta)}(u) \right]. \end{aligned}$$

From Proposition 1,

$$\int_{[0,1]} (1-u)^{-\gamma_X} d\nu_X^*(u) - \int_{[0,1]} (1-u)^{-\gamma_X} d\nu_{X,\varepsilon}^{(\beta)}(u) \leq 0,$$

and thus

$$\begin{aligned} 0 &\leq \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu_X^*(u) - \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu_{X,\varepsilon}^{(\beta)}(u) \\ &\leq \varepsilon \left[\int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\nu_X^*(u) + \int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\nu_{X,\varepsilon}^{(\beta)}(u) \right]. \end{aligned} \tag{23}$$

Let us now show that there exits $\beta_3 \in (0,1)$ such that

$$\int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\nu_X^*(u) + \sup_{\substack{\beta \geq \beta_3 \\ \beta \in B_\varepsilon}} \int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\nu_{X,\varepsilon}^{(\beta)}(u) \leq c(\kappa, \gamma_X), \quad (24)$$

where $c(\kappa, \gamma_X)$ is a positive constant independent of ε.

$\hookrightarrow$ We start with the first term.

- If $\gamma_X \leq \gamma_0$, one can pick δ such that $\gamma_X + \delta < \gamma_\mu^*$ and thus

$$\int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\nu_X^*(u) = \int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\mu(u) < \infty. \quad (25)$$

- Recall that if $\gamma_X > \gamma_0$ we have $\nu_X^*(u) = \delta_{b(\mu,\gamma_0)}$. Hence,

$$\int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\nu_X^*(u) = \left(\int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u) \right)^{(\gamma_X+\delta)/\gamma_0} < \infty. \quad (26)$$

$\hookrightarrow$ Now, for the second term,

- if $\gamma_X > \gamma_\mu^*$, we know from Lemma 2 that

$$\int_{[0,1]} Q_{X,\beta} d\mu = \infty.$$

Hence, (21) cannot be true if $\nu_{X,\varepsilon}^{(\beta)} = \mu$ and thus $\nu_{X,\varepsilon}^{(\beta)} \in \tilde{\mathcal{P}}_\mu(\kappa, \gamma_0)$ where $\tilde{\mathcal{P}}_\mu(\kappa, \gamma_0)$ is the set of probability measures satisfying (3) and (4). As a consequence, if $\gamma_X > \gamma_\mu^*$, one has

$$\int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\nu_{X,\varepsilon}^{(\beta)}(u) \leq \left(\kappa \int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\mu(u) \right)^{(\gamma_X+\delta)/\gamma_0}, \quad (27)$$

for all $\beta \in B_\varepsilon$.

- If $\gamma_X \leq \gamma_\mu^*$, we need first to prove the following assertion.

(A) There exist $\tilde{\gamma} \in (0, \gamma_\mu^*)$ and $\beta_3 \in (0,1)$ such that, if $\gamma_X > \tilde{\gamma}$, for all $\beta \in B_\varepsilon$ with $\beta > \beta_3$, the probability measure $\nu_{X,\varepsilon}^{(\beta)}$ is not equal to μ.

As a direct consequence of this assertion, if $\gamma_X \leq \gamma_\mu^*$,

$$\sup_{\substack{\beta \geq \beta_3 \\ \beta \in B_\varepsilon}} \int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\nu_{X,\varepsilon}^{(\beta)}(u) < \infty. \quad (28)$$

Indeed, if $0 < \gamma_X \leq \tilde{\gamma}$, since $\gamma_X + \delta < \gamma_\mu^*$,

$$\begin{aligned} & \int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\nu_{X,\varepsilon}^{(\beta)}(u) \\ \leq & \int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\mu(u) \bigwedge \left(\kappa \int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\mu(u) \right)^{(\gamma_X + \delta)/\gamma_0} < \infty, \end{aligned}$$

and if $\gamma_X > \tilde{\gamma}$, assertion (A) entails that for all $\beta \in B_\varepsilon$ with $\beta \geq \beta_3$,

$$\int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\nu_{X,\varepsilon}^{(\beta)}(u) \leq \left(\kappa \int_{[0,1]} (1-u)^{-\gamma_X - \delta} d\mu(u) \right)^{(\gamma_X + \delta)/\gamma_0} < \infty.$$

Let us prove assertion (A). Since the function

$$\gamma \in (0, \gamma_\mu^*) \mapsto \int_{[0,1]} (1-u)^{-\gamma} d\mu(u)$$

is increasing, continuous and converges to ∞ as $\gamma \to \gamma_\mu^*$, for all $\chi > 0$, there exists $\tilde{\gamma}(\chi) \in (0, \gamma_\mu^*)$ such that for all $\gamma > \tilde{\gamma}(\chi)$,

$$\int_{[0,1]} (1-u)^{-\gamma} d\mu(u) \geq \chi.$$

Now, for some $\delta_1 > 0$, let $\chi(\varepsilon) = \varepsilon + \tilde{\chi}$ where

$$\tilde{\chi} = 3 \max_{\gamma \in (0, \gamma_\mu^* + \delta_1]} \left(\kappa \int_{[0,1]} (1-u)^{-\gamma} d\mu(u) \right)^{\gamma/\gamma_0}.$$

Using Potter's bounds (see for instance de Haan and Ferreira [22, eq. (B.1.19)]) there exists $\beta_4 \in (0,1)$ such that for all $\beta \in B_\varepsilon$, $\beta > \beta_4$,

$$\int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu(u) \geq \frac{1}{2} \int_{[0,1]} (1-u)^{-\overline{\gamma}_X(\chi(\varepsilon),\mu)} d\mu(u),$$

where $\overline{\gamma}_X(\chi(\varepsilon), \mu) = \gamma_X - [\gamma_\mu^* - \tilde{\gamma}(\chi(\varepsilon))]/2$. Hence, if $\gamma_X \geq [\tilde{\gamma}(\chi(\varepsilon)) + \gamma_\mu^*]/2$ then $\overline{\gamma}_X(\chi(\varepsilon), \mu) > \tilde{\gamma}(\chi(\varepsilon))$ and thus,

$$\int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu(u) \geq \frac{\chi(\varepsilon)}{2}.$$

Moreover, for $\delta_1 > 0$, there exists $\beta_5 \in (0,1)$ such that for all $\beta \in B_\varepsilon$ with $\beta > \beta_5$

and for all probability measure $\nu \in \tilde{\mathcal{P}}_\mu(\kappa, \gamma_0)$,

$$\begin{aligned}\frac{\overline{\varrho}_\beta\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)}{Q_X(\beta)} &\leq \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu(u) \\ &\leq \frac{3}{2}\int_{[0,1]} (1-u)^{-\gamma_X-\delta_1} d\nu(u) \leq \frac{\tilde{\chi}}{2}.\end{aligned}$$

Since $Q_X(\beta) \to \infty$ as $\beta \to 1$, there exists $\beta_6 \in (0,1)$ such that $Q_X(\beta) > 2$. Hence, for all $\beta \in B_\varepsilon$ with $\beta > \beta_3 := \beta_4 \vee \beta_5 \vee \beta_6$, if $\gamma_X > [\tilde{\gamma}(\chi(\varepsilon)) + \gamma_\mu^*]/2 =: \tilde{\gamma}$,

$$\int_{[0,1]} Q_{X,\beta} d\mu - \overline{\varrho}_\beta\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right) \geq Q_X(\beta)\left[\frac{\chi(\varepsilon)}{2} - \frac{\tilde{\chi}}{2}\right] = \frac{Q_X(\beta)\varepsilon}{2} \geq \varepsilon,$$

and thus necessarily, $\nu_{X,\varepsilon}^{(\beta)} \neq \mu$. Assertion (A) is then proved.

Gathering (25) to (28) entails that (24) holds with

$$c(\kappa,\gamma_X) = \int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\mu(u) \bigwedge \left(\kappa \int_{[0,1]} (1-u)^{-\gamma_X-\delta} d\mu(u)\right)^{(\gamma_X+\delta)/\gamma_0},$$

if $\gamma_X \leq \tilde{\gamma}$ and if $\gamma_X > \tilde{\gamma}$,

$$c(\kappa,\gamma_X) = \left(\int_{[0,1]} (1-u)^{-[(\gamma_X+\delta)\wedge\gamma_0]} d\mu(u)\right)^{[(\gamma_X+\delta)\vee\gamma_0]/\gamma_0}.$$

Note that as expected, $c(\kappa,\gamma_X)$ does not depend on ε. Moreover when κ and/or γ_X increase, so does $c(\kappa,\gamma_X)$.

For all $\beta \in B_\varepsilon$ with $\beta > \beta_3$, inequalities (21) and (23) and the fact that $Q_X(\beta) \geq 2$ lead to

$$\begin{aligned}0 &\leq \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu_X^*(u) - \frac{\overline{\varrho}_\beta\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)}{Q_X(\beta)} \\ &\leq \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu_X^*(u) - \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu_{X,\varepsilon}^{(\beta)}(u) + \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\nu_{X,\varepsilon}^{(\beta)}(u) \\ &\quad - \frac{\overline{\varrho}_\beta\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)}{Q_X(\beta)} \leq \varepsilon\left(c(\kappa,\gamma_X) + \frac{1}{2}\right). \qquad (29)\end{aligned}$$

From (22) and (29), we finally get that for all $\varepsilon > 0$ and for $\beta > \beta_3$

$$0 \leq \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu_X^*(u) - \frac{\overline{\varrho}_\beta\left(Q_X\,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)}{Q_X(\beta)} \leq \varepsilon g(\kappa,\gamma_X), \qquad (30)$$

with

$$g(\kappa,\gamma_X) := \left[1 \bigvee \left(c(\kappa,\gamma_X) + \frac{1}{2}\right)\right].$$

Since from Lemma 4, one has for β close enough to 1,

$$\left| \int_{[0,1]} \frac{Q_{X,\beta}(u)}{Q_X(\beta)} d\mu_X^*(u) - \int_{[0,1]} (1-u)^{-\gamma_X} d\mu_X^*(u) \right| < \varepsilon g(\kappa,\gamma_X),$$

and thus, from (30),

$$\left| \frac{\overline{\varrho}_\beta\left(Q_X \,|\mathcal{P}_\mu(\kappa,\gamma_0)\right)}{Q_X(\beta)} - \int_{[0,1]} (1-u)^{-\gamma_X} d\mu_X^*(u) \right| \leq 2\varepsilon g(\kappa,\gamma_X),$$

which entails equivalence (5).

Let us now focus on equivalence (6). From Lemma 4, we have for all quantile functions $Q_X \in \mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_X \in (0,\gamma_0]$,

$$\lim_{\beta\to 1} \int_{[0,1]} \frac{Q_X(u)}{Q_X(\beta)} d\mu_\beta(u) = \int_{[0,1]} (1-u)^{-\gamma_X} d\mu(u).$$

The first part of (6) is then a direct consequence of (5). If $\gamma_X > \gamma_0$, from (5) we have

$$\begin{aligned}\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0)) &\sim Q_X(\beta) \left(\int_{[0,1]} (1-u)^{-\gamma_0} d\mu(u) \right)^{\gamma_X/\gamma_0} \\ &= Q_X(\beta)\left(1 - b(\mu,\gamma_0)\right)^{-\gamma_X}.\end{aligned}$$

From (1), we obtain

$$\overline{\varrho}_\beta(Q_X \mid \mathcal{P}_\mu(\kappa,\gamma_0)) \sim Q_X\left(1 - (1-b(\mu,\gamma_0))(1-\beta)\right) = Q_X\left(\beta + b(\mu,\gamma_0)(1-\beta)\right),$$

and the proof is complete. ■

Proof of Proposition 3 − Under (7), the sum $X_1 + X_2$ is such that $Q_{X_1+X_2} \in \mathcal{Q}^{(\mathrm{HT})}$ with tail index $\gamma_2 \leq \gamma_0$ (see for instance Davis and Resnick [13, Lemma 2.1]).

Hence, as $\beta \to 1$, we have from Lemma 4

$$\begin{aligned}\overline{\varrho}_\beta\left(Q_{X_1+X_2} \mid \mathcal{P}_\mu(\kappa,\gamma_0)\right) &\sim Q_{X_1+X_2}(\beta)\int_{[0,1]}(1-u)^{-\gamma_2}d\mu(u)\\ &\sim \int_{[0,1]}Q_{X_1+X_2}d\mu_\beta \leq \int_{[0,1]}Q_{X_1}d\mu_\beta + \int_{[0,1]}Q_{X_2}d\mu_\beta.\end{aligned}$$

The conclusion follows with another use of Lemma 4. ■

Proof of Proposition 4 – Let us first introduce the function $\Phi_{\mu,\gamma_0} : (0,\infty) \to (0,\infty)$, defined for all $\gamma > 0$ by

$$\Psi_{\mu,\gamma_0}(\gamma) = \left[\int_{[0,1]}(1-u)^{-(\gamma\wedge\gamma_0)}d\mu(u)\right]^{(\gamma\vee\gamma_0)/\gamma_0}.$$

We thus have

$$\widehat{\overline{\varrho}}_{n,\beta_n}(Q_X \mid \mu,\gamma_0) = X_{\lceil n\beta_n\rceil:n}\Psi_{\mu,\gamma_0}(\widehat{\gamma}_{n,X}).$$

Using the asymptotic equivalence obtained in Proposition 2, it suffices to prove that

$$\frac{X_{\lceil n\beta_n\rceil:n}}{Q_X(\beta_n)} \xrightarrow{\mathbb{P}} 1 \text{ and } \Psi_{\mu,\gamma_0}(\widehat{\gamma}_{n,X}) \xrightarrow{\mathbb{P}} \Psi_{\mu,\gamma_0}(\gamma_X).$$

Let's start with the first convergence in probability. If $V_1,\cdots,V_n$ is a sample of n independent copies of a standard uniform random variables, it is well known that

$$X_{\lceil n\beta_n\rceil:n} \stackrel{d}{=} Q_X\left(1-V_{n-\lceil n\beta_n\rceil+1:n}\right) = U_X\left(V^{-1}_{n-\lceil n\beta_n\rceil+1:n}\right),$$

where $V_{1:n} \leq \cdots \leq V_{n:n}$ is the sample $V_1,\cdots,V_n$ arranged in ascending order. From the representation of uniform order statistics (see David [12] and Reiss [29]),

$$V_{n-\lceil n\beta_n\rceil+1:n} \stackrel{d}{=} \sum_{i=1}^{n-\lceil n\beta_n\rceil+1}E_i \Big/ \sum_{i=1}^{n}E_i,$$

where $E_1,\cdots,E_n$ are independent copies of a standard exponential random variable. Since $n(1-\beta_n) \to \infty$, it is readily seen that $(n-\lceil n\beta_n\rceil+1)/[n(1-\beta_n)] \to 1$. An application of the law of large numbers leads to

$$\frac{V_{n-\lceil n\beta_n\rceil+1:n}}{1-\beta_n} \xrightarrow{\mathbb{P}} 1.$$

Since U_X is a regularly varying function,

$$U_X\left(V^{-1}_{n-\lceil n\beta_n\rceil+1:n}\right)\Big/ U_X\left((1-\beta_n)^{-1}\right) \xrightarrow{\mathbb{P}} 1$$

showing the first convergence. The proof of the second convergence in probability is straightforward since $\widehat{\gamma}_{n,X} \xrightarrow{\mathbb{P}} \gamma_X$ and since under (H.1) the function Ψ_{μ,γ_0} is continuous. ■

Proof of Proposition 5 – Using the function Ψ_{μ,γ_0} introduced in the proof of Proposition 4, we have

$$\widehat{\varrho}^{(\mathrm{W})}_{n,\beta_n}(Q_X \mid \mu,\gamma_0,\alpha_n) := X_{\lceil n\alpha_n\rceil:n}\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\widehat{\gamma}_{n,X}} \Psi_{\mu,\gamma_0}\left(\widehat{\gamma}_{n,X}\right).$$

According to the asymptotic equivalence obtained in Proposition 2, it thus suffices to show that

$$\frac{X_{\lceil n\alpha_n\rceil:n}}{Q_X(\beta_n)}\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\widehat{\gamma}_{n,X}} \xrightarrow{\mathbb{P}} 1.$$

Since $X_{\lceil n\alpha_n\rceil:n}/Q(\alpha_n) \xrightarrow{\mathbb{P}} 1$, see the proof of Proposition 4, it is equivalent to prove that

$$\frac{Q_X(\alpha_n)}{Q_X(\beta_n)}\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\widehat{\gamma}_{n,X}} = \frac{\ell_X((1-\alpha_n)^{-1})}{\ell_X((1-\beta_n)^{-1})}\exp\left((\widehat{\gamma}_{n,X}-\gamma_X)\ln\left(\frac{1-\alpha_n}{1-\beta_n}\right)\right) \xrightarrow{\mathbb{P}} 1.$$

Under (13),

$$\exp\left((\widehat{\gamma}_{n,X}-\gamma_X)\ln\left(\frac{1-\alpha_n}{1-\beta_n}\right)\right) \xrightarrow{\mathbb{P}} 1.$$

Now, using the representation (11) of a slowly varying function,

$$\frac{\ell_X((1-\alpha_n)^{-1})}{\ell_X((1-\beta_n)^{-1})} = \frac{c_X((1-\alpha_n)^{-1})}{c_X((1-\beta_n)^{-1})}\exp\left(-\int_{(1-\alpha_n)^{-1}}^{(1-\beta_n)^{-1}}\frac{\Delta_X(t)}{t}dt\right).$$

Since α_n and β_n converge to 1 as $n\to\infty$,

$$\lim_{n\to\infty}\frac{c_X((1-\alpha_n)^{-1})}{c_X((1-\beta_n)^{-1})} = 1.$$

Moreover, since $|\Delta_X|$ is asymptotically decreasing,

$$\left|\int_{(1-\alpha_n)^{-1}}^{(1-\beta_n)^{-1}}\frac{\Delta_X(t)}{t}dt\right| \le \left|\Delta_X\left(\frac{1}{1-\alpha_n}\right)\right|\ln\left(\frac{1-\alpha_n}{1-\beta_n}\right) \to 0,$$

and the proof is complete. ■

References

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

[2] Artzner, P. (1999). Application of coherent risk measures to capital requirements in insurance, *North American Actuarial Journal*, **3:2**, 11–25.

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

[4] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J.L. (2004). *Statistics of Extremes: Theory and Applications*, Wiley Series in Probability and Statistics.

[5] Beirlant, J., Teugels, J.L. and Vynckier, P. (1996). *Practical Analysis of Extreme Values*, Leuven University Press.

[6] Bellini, F. and Di Bernardino, E. (2017). Risk management with expectiles, *The European Journal of Finance*, **23**, 487–506.

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.

[8] Brazauskas, V., Jones, B.L., Puri, M.L. and Zitikis, R. (2008). Estimating conditional tail expectation with actuarial applications in view, *Journal of Statistical Planning and Inference*, **138**, 3590–3604.

[9] Daouia, A., Girard, S. and Stupfler, G. (2018). Extremiles: a new perspective on asymmetric least squares, *Journal of the American Statistical Association*, **114**, 1366–1381.

[10] Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles, *Journal of the Royal Statistical Society: Series B*, **80**, 263–292.

[11] Daouia, A., Girard, S. and Stupfler, G. (2019). Extreme M-quantiles as risk measures: From L1 to Lp optimization, *Bernoulli*, **25**, 264–309.

[12] David, H. (1980). *Order Statistics*, Wiley, New York.

[13] Davis, R.A. and Resnick, S.I. (1996). Limit theory for bilinear processes with heavy-tailed noise, *Annals of applied Probability*, **6(4)**, 1191–1210.

[14] Delbaen, F. (2009). Risk measures for non-integrable random variables, *Mathematical Finance*, **19**, 329–333.

[15] Denneberg, D. (1990). Premium calculation: why standard deviation should be replaced by absolute deviation, *ASTIN Bulletin: Journal of the International Actuarial Association*, **20**, 181–190.

[16] El Methni, J. and Stupfler, G. (2017). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica*, **27**, 907–930.

[17] El Methni, J. and Stupfler, G. (2018). Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions, *Econometrics and Statistics*, **6**, 129–148.

[18] Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer Berlin.

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

[20] Gardes, L. and Girard, S. (2021). On the estimation of the variability in the distribution tail, *Test Spanish Society of Statistics and Operations Research*, **30**, 884–907.

[21] Gardes, L., Girard, S. and Stupfler, G. (2020). Beyond tail median and conditional tail expectation: extreme risk estimation using tail Lp-optimisation, *Scandinavian Journal of Statistics*, **47(3)**, 922–949.

[22] de Haan L. and Ferreira A. (2006). *Extreme Value Theory: An Introduction*, Springer, New York.

[23] Hill, B. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1163–1173.

[24] Kaas, R., Goovaerts, M., Dhaene, J. and Denuit, M. (2008). *Modern Actuarial Theory Using R*, Springer, New York.

[25] Konstantinides, D. (2018). *Risk Theory. A Heavy Tail Approach*, World Scientific Publishing.

[26] Moscadelli, M. (2004). The modelling of operational risk: experience with the analysis of the data collected by the Basel committee, Technical Report 517, Banca d'Italia.

[27] Neslehova, J., Embrechts, P. and Chavez-Demoulin, V. (2006). Infinite mean models and the LDA for operational risk, *Journal of Operational Risk*, **1(1)**, 3–25.

[28] Read, L.K. and Vogel, R.M. (2015). Reliability, return periods, and risk under nonstationarity, *Water Resources Research*, **51(8)**, 6381–6398.

[29] Reiss, R. (1989). *Approximation Theorems of Order Statistics*, Springer-Verlag, New York.

[30] Rockafellar, R.T. and Uryasev, S. (2002). Conditional Value-at-Risk for general loss distributions, *Journal of Banking and Finance*, **26**, 1443–1471.

[31] Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*, Springer Series in Operations Research and Financial Ingineering.

[32] Wang, S. (1996). Premium calculation by transforming the layer premium density, *Astin Bulletin*, **26**, 71–92.

[33] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association*, **73**, 812–815.

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

[35] Yitzhaki, S. and Schechtman, E. (2013). *The Gini Methodology: A Primer on a Statistical Methodology*, Springer, New York.