WORKING PAPERS

N° TSE - 953

September 2018

“ExpectHill estimation, extreme risk and heavy tails"

Abdelaati Daouia, Stéphane Girard and Gilles Stupffer

Toulouse School of Economics

ExpectHill estimation, extreme risk and heavy tails

Abdelaati Daouia[a], Stéphane Girard[b] and Gilles Stupfler[c]

[a] Toulouse School of Economics, University of Toulouse Capitole, France

[b] Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France

[c] School of Mathematical Sciences, University of Nottingham, United Kingdom

Abstract: Risk measures of a financial position are traditionally based on quantiles. Replacing quantiles with their least squares analogues, called expectiles, has recently received increasing attention. The novel expectile-based risk measures satisfy all coherence requirements. We revisit their extreme value estimation for heavy-tailed distributions. First, we estimate the underlying tail index via weighted combinations of top order statistics and asymmetric least squares estimates. The resulting expectHill estimators are then used as the basis for estimating tail expectiles and Expected Shortfall. The asymptotic theory of the proposed estimators is provided, along with numerical simulations and applications to actuarial and financial data.

JEL Classifications: C13, C14.

Keywords: Asymmetric least squares, Coherent risk measures, Expected shortfall, Expectile, Extrapolation, Extremes, Heavy tails, Tail index.

1 Introduction

The risk of a financial position Y is usually summarized by a risk measure. Value at Risk (VaR) is arguably the most common risk measure used in practice. The VaR at probability level $\tau \in (0,1)$ is given by the τ-quantile $q_\tau := F_Y^{\leftarrow}(\tau) = \inf\{y \in \mathbb{R} : F(y) \geqslant \tau\}$, where F is the distribution function of Y. Koenker and Bassett [22] elaborated an absolute error loss minimization framework extending this definition of quantiles as left continuous inverse functions to the minimizers

$$q_\tau \in \underset{\theta \in \mathbb{R}}{\arg\min}\, \mathbb{E}\left\{\rho_\tau(Y-\theta) - \rho_\tau(Y)\right\},$$

with equality if F is increasing, where $\rho_\tau(y) = |\tau - \mathbb{1}(y \leqslant 0)|\,|y|$ and $\mathbb{1}(\cdot)$ is the indicator function. There are different sign conventions for VaR which co-exist in the literature. In this paper, the position Y is a real-valued random variable whose values are the negative of financial returns. The right-tail of the distribution of Y, for levels τ close to one, then corresponds to the negative of extreme losses. In actuarial science where Y is typically a non-negative loss variable, the sign convention we have chosen implies that extreme losses also correspond to levels τ close to one. The position Y is therefore considered riskier as its risk measure gets higher.

One of the major criticisms on VaR q_τ is its failure to fulfill the subadditivity property in general (Acerbi [1]), and hence it is not a coherent risk measure according to the axiomatic foundations in Artzner *et al.* [2]. Furthermore, it fails to account for the size of losses beyond the level τ, since quantiles only depend on the frequency of tail losses and not on their values (Daníelsson *et al.* [8]). In both of these aspects, expectiles are a perfectly reasonable alternative to quantiles as they depend on both the tail realizations and their probability (Kuan *et al.* [24]) and define a coherent risk measure (Bellini *et al.* [4]). This is mainly due to their conception as a least squares analogue of quantiles. More precisely, by substituting the absolute deviations in the asymmetric loss function ρ_τ with squared deviations, Newey and Powell [25] obtain the τth expectile of the distribution of Y as the minimizer

$$\xi_\tau := \underset{\theta\in\mathbb{R}}{\arg\min}\, \mathbb{E}\left\{\eta_\tau(Y-\theta)-\eta_\tau(Y)\right\}, \tag{1}$$

with $\eta_\tau(y) = |\tau - \mathbb{I}(y \leqslant 0)|\, y^2$. The additional term $\eta_\tau(Y)$ ensures the existence of a unique solution ξ_τ for distributions with finite absolute first moment. Expectiles are determined by tail expectations rather than tail probabilities, which allows for more prudent and reactive risk management. Altering the shape of extreme losses may not change the quantile-VaR, but it does impact all the expectiles (Taylor [31]). Another advantage of expectiles is that they make more efficient use of the available data since they rely on the distance to all observations and not only on the frequency of tail losses (Sobotka and Kneib [30]). Moreover, using expectiles has the appeal of avoiding recourse to regularity conditions on the underlying distribution (see *e.g.* Holzmann and Klar [21], Krätschmer and Zähle [23]). Perhaps most importantly, expectiles induce the only coherent law-invariant risk measure that is elicitable (Ziegel [33]). The property of elicitability corresponds to the existence of a natural backtesting methodology. Also, expectiles are the only M-quantiles (Breckling and Chambers [6]) that are coherent risk measures (Bellini *et al.* [4]). Further theoretical and numerical merits in favor of the adoption of expectiles in risk management can be found in Ehm *et al.* [14] and Bellini and Di Bernardino [5].

In this article we first investigate the problem of estimating tail expectiles from the perspective of extreme value theory. This translates into considering both *intermediate* and *extreme* asymmetry levels, respectively, $\tau = \tau_n \to 1$ such that $n(1-\tau_n) \to \infty$ and $\tau = \tau_n' \to 1$ such that $n(1-\tau_n') \to c < \infty$, as $n \to \infty$. We focus on the Fréchet maximum domain of attraction of heavy-tailed distributions that perfectly describe the tail structure of most actuarial and financial data (see, *e.g.*, Embrechts *et al.* [18] and Resnick [26]). This problem is, in comparison to extreme quantile estimation, still in full development. The absence of a closed form expression for expectiles makes the extreme value analysis of their asymmetric least squares estimators a much harder mathematical problem than for order statistics. Yet, we have initiated a satisfactory solution to this problem in an earlier paper [10] by proposing intermediate and extreme expectile estimators and developing their asymptotic theory. Very recently, we have come up in [11] with powerful approximations of the tail empirical expectile process. First, Theorem 1 in Daouia *et al.* [11] derives an explicit joint asymptotic Gaussian representation of the tail expectile and quantile processes. Second, Theorem 2 in [11] unravels the discrepancy between the tail empirical expectile process and its population counterpart. As these two theorems constitute the basic theoretical tools for our asymptotic analysis in the present paper, they are briefly described below in Theorem 1 along with the statistical model in Section 2.

Built on these recent advances, Section 3 shows that the tail index of the underlying Pareto-type distribution can be estimated in a novel and more general manner. This index tunes the tail heaviness of F and its knowledge is of utmost interest since it makes the estimation of extreme quantiles and expectiles possible by means of appropriate extrapolation techniques. We first construct asymmetric least squares estimators of the tail index and derive their asymptotic normality in Theorem 2. We then construct a more general class of weighted estimators by computing a linear combination of these pure expectile-based estimators and of the popular Hill estimator (Hill [20]). This inspired the name *expectHill* estimators for this class. Thanks to the joint weighted Gaussian approximations of the tail expectile and quantile processes in Theorem 1, we get the asymptotic normality of the *expectHill* estimators and derive their joint convergence with both intermediate quantile and expectile estimators in Theorem 3.

Built on the *expectHill* estimators themselves, we propose in Section 4 general weighted estimators for intermediate expectiles ξ_{τ_n} whose asymptotic normality, obtained in Theorem 4, follows as a corollary of Theorem 3. Based on the ideas of Daouia *et al.* [10, 11], the weighted intermediate expectile estimators are then extrapolated to the very extreme expectile level τ_n' that may approach one at an arbitrarily fast rate. The asymptotic properties of the extrapolated $\xi_{\tau_n'}$ estimators are established in Theorem 5.

An important alternative to the VaR q_τ and its coherent least squares analogue ξ_τ is Expected Shortfall (ES). It is favored by practitioners who are more concerned with the risk exposure to a catastrophic event that may wipe out an investment in terms of the size of potential losses. The conventional quantile-based ES at level τ equals

$$\mathrm{QES}_\tau := \frac{1}{1-\tau}\int_\tau^1 q_t\,dt.$$

It is coherent (Acerbi [1]) and identical, when the financial position Y is continuous, to the so-called Conditional Value at Risk $\mathbb{E}[Y|Y > q_\tau]$ (Rockafellar and Uryasev [28, 29]). Similarly to this intuitive tail conditional expectation, Taylor [31] has introduced and used the expectile-based form $\mathbb{E}[Y|Y > \xi_\tau]$ as the basis for estimating the standard quantile-based measure $\mathbb{E}[Y|Y > q_\tau]$. Given that both conditional expectations $\mathbb{E}[Y|Y > q_\tau]$ and

$\mathbb{E}[Y|Y > \xi_\tau]$ are not coherent risk measures in general, Daouia *et al.* [11] have suggested to estimate the coherent ES form QES_τ on the basis of its expectile-based analogue

$$\text{XES}_\tau := \frac{1}{1-\tau}\int_\tau^1 \xi_t \, dt,$$

obtained by substituting the expectile ξ_t in place of the quantile q_t in QES_τ. This definition is more convenient than $\mathbb{E}[Y|Y > \xi_\tau]$ as it induces a proper coherent risk measure (see Proposition 2 in [11]), while keeping the intuitive meaning of the conditional expectation, when $\tau \to 1$, since $\text{XES}_\tau \sim \mathbb{E}[Y|Y > \xi_\tau]$ (see Proposition 3 in [11]). In addition to this asymptotic equivalence, the tail values XES_τ and $\mathbb{E}[Y|Y > \xi_\tau]$ share exactly the same estimators, for both intermediate and extreme expectile levels $\tau = \tau_n$ and $\tau = \tau_n'$.

The proposed estimation procedures in Daouia *et al.* [11] for both extreme values $\text{XES}_{\tau_n'}$ and $\text{QES}_{\tau_n'}$ are mainly based on the classical Hill estimator of the tail index. In Section 5, we extend their extrapolation devices by using the generalized weighted *expectHill* estimator; see Theorems 6-7. In particular, when the ultimate interest is in estimating the traditional form $\text{QES}_{\tau_n'}$ in the case of real-valued profit-loss distributions, our composite asymmetric least squares estimators perform better than the rival estimators of Daouia *et al.* [11] and El Methni *et al.* [15]. Section 6 contains our experiments with simulated data and Section 7 presents applications to medical insurance data and financial returns data. The proofs and auxiliary results are deferred to the Supplementary Material document.

2 Statistical model and basic tools

In this paper we consider the class of heavy-tailed distributions, referred to as the Fréchet maximum domain of attraction, with tail index $0 < \gamma < 1$. The survival function of these Pareto-type distributions has the form

$$\overline{F}(y) := 1 - F(y) = y^{-1/\gamma}\ell(y), \tag{2}$$

for $y > 0$ large enough, where ℓ is a slowly varying function at infinity, *i.e.*, a positive function on $(0, \infty)$ satisfying $\ell(ty)/\ell(t) \to 1$, as $t \to \infty$, for any $y > 0$. The index γ tunes the tail heaviness of $\overline{F}$: the larger the index, the heavier the right tail. Let Y be the actuarial or financial position of interest having survival function $\overline{F}$, and let $Y_- = \min(Y, 0)$ denote the negative part of Y. Then, together with condition $\mathbb{E}|Y_-| < \infty$, the assumption $\gamma < 1$ ensures the existence of the first moment of Y, and hence the existence of expectiles. By Corollary 1.2.10 in de Haan and Ferreira [12], the model assumption (2) is equivalent to

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\gamma} \quad \text{for all} \quad x > 0, \tag{3}$$

where $U(t) := q_{1-t^{-1}} \equiv \inf\{y \in \mathbb{R} : 1/\overline{F}(y) \geqslant t\}$ stands for the tail quantile function of Y. Under (2) or equivalently (3), it has been found that

$$\frac{\xi_\tau}{q_\tau} \sim (\gamma^{-1} - 1)^{-\gamma} \quad \text{as} \quad \tau \to 1 \tag{4}$$

(Bellini and Di Bernardino [5]). A refined asymptotic expansion of ξ_τ/q_τ with a precise quantification of the bias term is obtained in Proposition 1(i) of Daouia *et al.* [11] under the following second-order regular variation condition:

$\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$ For all $x > 0$,

$$\lim_{t\to\infty} \frac{1}{A(t)} \left[\frac{U(tx)}{U(t)} - x^{\gamma} \right] = x^{\gamma} \frac{x^{\rho} - 1}{\rho}$$

where $\rho \leqslant 0$ is a constant parameter and A is an auxiliary function converging to 0 at infinity and having ultimately constant sign. Hereafter, $(x^{\rho} - 1)/\rho$ is to be understood as $\log x$ when $\rho = 0$.

Assumption $\mathcal{C}_2(\gamma, \rho, A)$ is a standard condition in extreme value theory, which controls the rate of convergence in (3). The monographs of Beirlant *et al.* [3] and de Haan and Ferreira [12] give abundant examples of commonly used continuous distributions satisfying $\mathcal{C}_2(\gamma, \rho, A)$, along with thorough discussions on the interpretation and the rationale behind this second-order condition.

Suppose we observe independent copies $\{Y_1, \ldots, Y_n\}$ of the random variable Y and denote by $Y_{1,n} \leqslant Y_{2,n} \leqslant \cdots \leqslant Y_{n,n}$ their nth order statistics. Let the expectile level $\tau = \tau_n$ approach one at an *intermediate* rate in the sense that $n(1 - \tau_n) \to \infty$ as $n \to \infty$. A natural estimator of the corresponding intermediate expectile ξ_{τ_n} is given by its empirical version

$$\widetilde{\xi}_{\tau_n} = \arg\min_{u\in\mathbb{R}} \sum_{i=1}^{n} \eta_{\tau_n}(Y_i - u). \tag{5}$$

Under condition $\mathcal{C}_2(\gamma, \rho, A)$, Daouia *et al.* [11] prove in their Theorem 1 that the tail empirical expectile process

$$(0, 1] \to \mathbb{R},\ \ s \mapsto \widetilde{\xi}_{1-(1-\tau_n)s}$$

can be approximated by a sequence of Gaussian processes with drift and derive its joint asymptotic behavior with the tail empirical quantile process

$$(0,1] \to \mathbb{R},\ s \mapsto \widehat{q}_{1-(1-\tau_n)s} := Y_{n-\lfloor n(1-\tau_n)s \rfloor, n},$$

where $\lfloor \cdot \rfloor$ stands for the floor function. They also analyze in their Theorem 2 the difference between the tail empirical expectile process and its population counterpart. For our purposes below, we recall these two approximations in the following result.

Theorem 1 (Daouia *et al.*, 2018b)**.** *Suppose that $\mathbb{E}|Y_-|^2 < \infty$. Assume further that condition $\mathcal{C}_2(\gamma,\rho,A)$ holds, with $0 < \gamma < 1/2$. Let $\tau_n \to 1$ be such that $n(1-\tau_n) \to \infty$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) = \mathrm{O}(1)$. Then there exists a sequence W_n of standard Brownian motions such that, for any $\varepsilon > 0$ sufficiently small,*

$$\begin{aligned}
\frac{\widehat{q}_{1-(1-\tau_n)s}}{q_{\tau_n}} &= s^{-\gamma}\left(1 + \frac{1}{\sqrt{n(1-\tau_n)}}\gamma\sqrt{\gamma^{-1}-1}\, s^{-1} W_n\left(\frac{s}{\gamma^{-1}-1}\right)\right. \\
&\quad \left. + \frac{s^{-\rho}-1}{\rho}A((1-\tau_n)^{-1}) + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{n(1-\tau_n)}}\right)\right) \\
\text{and} \quad \frac{\widetilde{\xi}_{1-(1-\tau_n)s}}{\xi_{\tau_n}} &= s^{-\gamma}\left(1 + (s^{\gamma}-1)\frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_{\tau_n}}(\mathbb{E}(Y) + \mathrm{o}_{\mathbb{P}}(1))\right. \\
&\quad + \frac{1}{\sqrt{n(1-\tau_n)}}\gamma^2\sqrt{\gamma^{-1}-1}\, s^{\gamma-1}\int_0^s W_n(t)\, t^{-\gamma-1}\, dt \\
&\quad + \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} \times \frac{s^{-\rho}-1}{\rho}A((1-\tau_n)^{-1}) \\
&\quad \left. + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{n(1-\tau_n)}}\right)\right) \quad \textit{uniformly in } s \in (0,1].
\end{aligned}$$

If in addition $\rho < 0$, then

$$\frac{\widetilde{\xi}_{1-(1-\tau_n)s}}{\xi_{1-(1-\tau_n)s}} = 1 + \frac{1}{\sqrt{n(1-\tau_n)}}\gamma^2\sqrt{\gamma^{-1}-1}\,s^{\gamma-1}\int_0^s W_n(t)\,t^{-\gamma-1}\,dt + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{n(1-\tau_n)}}\right) \quad \textit{uniformly in } s \in (0,1].$$

The assumptions that $\gamma \in (0, 1/2)$ and $\mathbb{E}|Y_-|^2 < \infty$ essentially guarantee that the loss variable has a finite variance. This is the case in most studies on actuarial and financial data where the realized values of γ have been found to lie well below 1/2; see, *e.g.*, the R package `CASdatasets`, Daouia *et al.* [10] and the references therein.

The extra condition $\rho < 0$, in the second part of Theorem 1, is required in most extrapolation results formulated in the extreme value literature under condition $\mathcal{C}_2(\gamma, \rho, A)$; see, *e.g.*, Chapter 4 of de Haan and Ferreira [12] regarding extreme quantile estimation and Daouia *et al.* [10] for extreme expectile estimation. Note also that, in contrast to the first part of Theorem 1, the second part avoids the error terms that are proportional to $1/q_{\tau_n}$ and $A((1-\tau_n)^{-1})$.

This theorem, already proved in Daouia *et al.* [11], constitutes the main intermediate theoretical tool for our ultimate interest in constructing general weighted estimators of the tail index and extreme expectiles, as well as of Expected Shortfall risk measures.

3 Estimation of the tail index

In this section, we first construct purely expectile-based estimators of the tail index γ and derive their asymptotic distributions. We shall then construct a more general class of estimators by combining both intermediate empirical expectiles and quantiles. The basic idea stems from Theorem 1 which suggests the following approximation:

$$\int_0^1 \log\left(\frac{\widetilde{\xi}_{1-(1-\tau_n)s}}{\xi_{\tau_n}}\right) ds \approx \int_0^1 \log(s^{-\gamma})\, ds = \gamma$$

where $\tau_n \to 1$ is such that $n(1-\tau_n) \to \infty$. One can then estimate γ by

$$\breve{\gamma}_{\tau_n} := \int_0^1 \log\left(\frac{\widetilde{\xi}_{1-(1-\tau_n)s}}{\widetilde{\xi}_{\tau_n}}\right) ds.$$

A computationally more viable option is to use a discretized version of the integral estimator $\breve{\gamma}_{\tau_n}$ on a regular $l-$grid of points in $[0,1]$, namely:

$$\widetilde{\gamma}_{\tau_n,l} := \frac{1}{l}\sum_{i=1}^{l} \log\left(\frac{\widetilde{\xi}_{1-(1-\tau_n)(i-1)/l}}{\widetilde{\xi}_{\tau_n}}\right)$$

where $l = l(n) \to \infty$. A particularly interesting example is

$$\widetilde{\gamma}_{\tau_n} := \frac{1}{\lfloor n(1-\tau_n)\rfloor}\sum_{i=1}^{\lfloor n(1-\tau_n)\rfloor} \log\left(\frac{\widetilde{\xi}_{1-(i-1)/n}}{\widetilde{\xi}_{1-\lfloor n(1-\tau_n)\rfloor/n}}\right) \tag{6}$$

or, equivalently, $\widetilde{\gamma}_{\tau_n} = \widetilde{\gamma}_{1-\lfloor n(1-\tau_n)\rfloor/n,\lfloor n(1-\tau_n)\rfloor}$. This simple estimator has exactly the same form as the popular Hill estimator (Hill [20])

$$\widehat{\gamma}_{\tau_n} = \frac{1}{\lfloor n(1-\tau_n)\rfloor}\sum_{i=1}^{\lfloor n(1-\tau_n)\rfloor} \log\left(\frac{\widehat{q}_{1-(i-1)/n}}{\widehat{q}_{1-\lfloor n(1-\tau_n)\rfloor/n}}\right) \tag{7}$$

with the tail empirical quantile process $\widehat{q}$ in (7) replaced by its asymmetric least squares analogue $\widetilde{\xi}$. Beirlant *et al.* [3] and de Haan and Ferreira [12] provide an extensive overview of the asymptotic theory for the Hill estimator $\widehat{\gamma}_{\tau_n}$. The next theorem gives the asymptotic normality of the three new estimators $\check{\gamma}_{\tau_n}$, $\widetilde{\gamma}_{\tau_n,l}$ and $\widetilde{\gamma}_{\tau_n}$. Its proof essentially consists in writing

$$\log\left(\frac{\widetilde{\xi}_{1-(1-\tau_n)s}}{\widetilde{\xi}_{\tau_n}}\right) = \log\left(\frac{\widetilde{\xi}_{1-(1-\tau_n)s}}{\xi_{\tau_n}}\right) - \log\left(\frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}}\right)$$

before integrating and crucially using Theorem 1 twice in order to control both of the logarithms on the right-hand side.

Theorem 2. *Suppose that $\mathbb{E}|Y_-|^2 < \infty$. Assume further that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds, with $0 < \gamma < 1/2$. Let $\tau_n \to 1$ be such that $n(1-\tau_n) \to \infty$, and suppose that the bias conditions $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/q_{\tau_n} \to \lambda_2 \in \mathbb{R}$ are satisfied. Then:*

(i) $\sqrt{n(1-\tau_n)}(\check{\gamma}_{\tau_n} - \gamma)$

$$\xrightarrow{d} \mathcal{N}\left(\frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1 - \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2, \frac{2\gamma^3}{1-2\gamma}\right).$$

(ii) If $l = l(n)$ fulfills $\sqrt{n(1-\tau_n)}\log(n(1-\tau_n))/l \to 0$, then (i) holds with $\check{\gamma}_{\tau_n}$ replaced by $\widetilde{\gamma}_{\tau_n,l}$. Especially, (i) holds with $\check{\gamma}_{\tau_n}$ replaced by $\widetilde{\gamma}_{\tau_n}$.

Before using the estimator $\widetilde{\gamma}_{\tau_n}$ to construct a more general class of tail index estimators, we formulate a couple of remarks about its theoretical and practical behavior.

Remark 1. The conditions involving the auxiliary function A in Theorem 2 are also required to derive the asymptotic normality of the conventional Hill estimator $\widehat{\gamma}_{\tau_n}$ in (7), with asymptotic bias $\lambda_1/(1-\rho)$ and asymptotic variance γ^2 [see Theorem 3.2.5 in de Haan and Ferreira ([12], p.74)]. Theorem 2 also features a further bias condition involving the quantile function q; this was to be expected in view of Theorem 1, of which a consequence is that the remainder term in the approximation $\xi_{1-(1-\tau_n)s}/\xi_{\tau_n} \approx s^{-\gamma}$ depends on both A and q. Yet, it is straightforward to eliminate this bias component: note that the centered variable $Z = Y - \mathbb{E}(Y)$ is also heavy-tailed, with the same extreme value parameters as Y, and thus the estimator $\breve{\gamma}_{\tau_n}^{Z}$ constructed on the $Z_i = Y_i - \mathbb{E}(Y)$ satisfies

$$\sqrt{n(1-\tau_n)}(\breve{\gamma}_{\tau_n}^{Z} - \gamma) \xrightarrow{d} \mathcal{N}\left(\frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1, \frac{2\gamma^3}{1-2\gamma}\right).$$

This suggests to define $\widehat{Z}_i = Y_i - \overline{Y}_n$, where $\overline{Y}_n$ is the sample mean, and then to consider the estimator $\breve{\gamma}_{\tau_n}^{\widehat{Z}}$. Due to the translation equivariance of expectiles, the gap between $\breve{\gamma}_{\tau_n}^{\widehat{Z}}$ and $\breve{\gamma}_{\tau_n}^{Z}$ has the same order as $|\overline{Y}_n - \mathbb{E}(Y)| = \mathrm{O}_{\mathbb{P}}(1/\sqrt{n})$. It follows that $\breve{\gamma}_{\tau_n}^{\widehat{Z}}$ has the same asymptotic distribution as $\breve{\gamma}_{\tau_n}^{Z}$, and is therefore a bias-reduced version of $\breve{\gamma}_{\tau_n}$ which eliminates the quantile component of the bias.

Remark 2. The selection of τ_n is a difficult problem in general, since any sort of optimal choice will involve the unknown parameter ρ as well as the function A; for a discussion about the optimal choice of τ_n in the Hill estimator based on mean-squared error, see Hall and Welsh [19]. A usual practice for selecting a reasonable estimate $\widehat{\gamma}_{\tau_n}$ is, in the reparametrization $\tau_n = 1 - k/n$, to plot the graph of $k \mapsto \widehat{\gamma}_{1-k/n}$ for $k \in \{1, 2, \ldots, n-1\}$, and then to pick out a value of k corresponding to the first stable part of the plot [see, *e.g.*, de Haan and Ferreira ([12], Section 3)]. There have been a number of attempts at formalizing this procedure, including Resnick and Stărică [27], Drees *et al.* [13], and more recently El Methni and Stupfler [16, 17]. The Hill plot may be, however, so unstable that reasonable values of k (which would correspond to estimates close to the true value of γ) may be hidden in the graph. The least squares analogue $\widetilde{\gamma}_{1-k/n}$ in (6) is, in contrast to $\widehat{\gamma}_{1-k/n}$, based on expectiles that enjoy superior regularity properties compared to quantiles (see Proposition 1 in Holzmann and Klar [21]). One may thus expect that $\widetilde{\gamma}_{1-k/n}$ affords smoother and more stable plots compared to those of the Hill estimator $\widehat{\gamma}_{1-k/n}$. This advantage is illustrated in Section A of the Supplementary Material document, where we examine the behavior of $\widehat{\gamma}$ and $\widetilde{\gamma}$ on two concrete actuarial and financial data sets. It can be seen thereon that the plots of $k \mapsto \widetilde{\gamma}_{1-k/n}$ are indeed far smoother than the arguably wiggly plots of $k \mapsto \widehat{\gamma}_{1-k/n}$.

It could, however, happen that $\widetilde{\gamma}$ has a higher bias than the Hill estimator. This is for instance the case if $|\rho|$ is large, since a large $|\rho|$ means that the underlying distribution is, in its right tail, very close to a multiple of the Pareto distribution for which the Hill estimator is unbiased. An efficient way to take advantage of the desirable properties of both $\widetilde{\gamma}$ and $\widehat{\gamma}$ in a large class of models is by using their linear combination for estimating γ. For $\alpha \in \mathbb{R}$, we then define the more general estimator

$$\overline{\gamma}_{\tau_n}(\alpha) := \alpha\widehat{\gamma}_{\tau_n} + (1-\alpha)\widetilde{\gamma}_{\tau_n}. \tag{8}$$

We shall call this linear combination the *expectHill estimator*. For example, the simple mean $\overline{\gamma}_{\tau_n}(1/2)$ would represent an equal balance between the use of large asymmetric least squares statistics in (6) and top order statistics in (7). The convergence of the *expectHill* estimator is, however, a highly non-trivial problem as it hinges, by construction, on both the tail expectile and quantile processes. The explicit joint asymptotic Gaussian representation of these two processes, obtained in Theorem 1, is a pivotal tool for our analysis, and enables us to address the convergence problem in its full generality. We establish below the asymptotic normality of the *expectHill* estimator, along with its joint convergence with intermediate sample quantiles and expectiles.

Theorem 3. *Suppose that the conditions of Theorem 2 hold. Then, for any $\alpha \in \mathbb{R}$,*

$$\sqrt{n(1-\tau_n)}\left(\overline{\gamma}_{\tau_n}(\alpha)-\gamma, \frac{\widehat{q}_{\tau_n}}{q_{\tau_n}}-1, \frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right) \xrightarrow{d} \mathcal{N}(\mathfrak{m}_\alpha, \mathfrak{V}_\alpha)$$

where $\mathfrak{m}_\alpha$ is the 1×3 vector $\mathfrak{m}_\alpha := (b_\alpha, 0, 0)$, with

$$b_\alpha = \frac{\lambda_1}{1-\rho}\left(\alpha + (1-\alpha)\frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho}\right) - (1-\alpha)\mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2, \tag{9}$$

and $\mathfrak{V}_\alpha$ is the 3×3 symmetric matrix with entries

$$\begin{aligned}
\mathfrak{V}_\alpha(1,1) &= \gamma^2\left(\alpha^2\left[\frac{3-4\gamma}{1-2\gamma} - 2\frac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma}\right] - 2\alpha\left[\frac{1}{1-2\gamma} - \frac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma}\right] + \frac{2\gamma}{1-2\gamma}\right), \\
\mathfrak{V}_\alpha(1,2) &= (1-\alpha)\gamma[(\gamma^{-1}-1)^{\gamma} - 1 - \gamma\log(\gamma^{-1}-1)], \\
\mathfrak{V}_\alpha(1,3) &= \frac{\gamma^3}{(1-\gamma)^2}\left[\alpha(\gamma^{-1}-1)^{\gamma} + (1-\alpha)\frac{1-\gamma}{1-2\gamma}\right], \\
\mathfrak{V}_\alpha(2,2) &= \gamma^2, \quad \mathfrak{V}_\alpha(2,3) = \gamma^2\left(\frac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma} - 1\right), \quad \mathfrak{V}_\alpha(3,3) = \frac{2\gamma^3}{1-2\gamma}.
\end{aligned}$$

As an immediate consequence, we have for any $\alpha \in \mathbb{R}$,

$$\sqrt{n(1-\tau_n)}\left(\overline{\gamma}_{\tau_n}(\alpha)-\gamma\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha) \text{ where } v_\alpha = \mathfrak{V}_\alpha(1,1). \tag{10}$$

This remains valid if $\widetilde{\gamma}_{\tau_n}$ is replaced in (8) by the continuous version $\check{\gamma}_{\tau_n}$, or any other discretized version $\widetilde{\gamma}_{\tau_n,l}$ provided $\sqrt{n(1-\tau_n)}\log(n(1-\tau_n))/l \to 0$.

Remark 3. The optimal value of the weighting coefficient α in (8), which minimizes the asymptotic variance v_α of $\overline{\gamma}_{\tau_n}(\alpha)$, only depends on the tail index γ and has the explicit expression

$$\alpha(\gamma) = \frac{(1-\gamma) - (1-2\gamma)(\gamma^{-1}-1)^{\gamma}}{(1-\gamma)(3-4\gamma) - 2(1-2\gamma)(\gamma^{-1}-1)^{\gamma}}.$$

Its plot against $\gamma \in (0, 1/2)$ is given in Section B of the Supplementary Material document. It can be seen thereon that the simple mean $\overline{\gamma}_{\tau_n}(1/2)$ of $\widehat{\gamma}_{\tau_n}$ and $\widetilde{\gamma}_{\tau_n}$, with $\alpha = 1/2$, affords a middle course between $\widehat{\gamma}_{\tau_n} \equiv \overline{\gamma}_{\tau_n}(1)$ and $\widetilde{\gamma}_{\tau_n} \equiv \overline{\gamma}_{\tau_n}(0)$ in terms of asymptotic variance. In terms of smoothness, $\overline{\gamma}_{\tau_n}(1/2)$ offers a middle course as well, as shown in Section A of the Supplementary Material document.

4 Extreme expectile estimation

In this section, we first return to intermediate expectile estimation by making use of the general class of γ estimators $\{\overline{\gamma}_{\tau_n}(\alpha)\}_{\alpha\in\mathbb{R}}$ to construct alternative estimators for high expectiles ξ_{τ_n} such that $\tau_n \to 1$ and $n(1-\tau_n) \to \infty$ as $n \to \infty$. Then we extrapolate the obtained estimators to the very high expectile levels that may approach one at an arbitrarily fast rate.

Alternatively to the asymmetric least squares estimator $\widetilde{\xi}_{\tau_n}$ defined in (5), one may use the asymptotic connection $\xi_{\tau_n} \sim (\gamma^{-1}-1)^{-\gamma} q_{\tau_n}$, described in (4), to define the following semiparametric estimator of ξ_{τ_n}:

$$\widehat{\xi}_{\tau_n}(\alpha) := \left(\overline{\gamma}_{\tau_n}(\alpha)^{-1} - 1\right)^{-\overline{\gamma}_{\tau_n}(\alpha)} \widehat{q}_{\tau_n}.$$

Even more generally, one may combine the two estimators $\widehat{\xi}_{\tau_n}(\alpha)$ and $\widetilde{\xi}_{\tau_n}$ to define, for $\beta \in \mathbb{R}$, the weighted estimator

$$\overline{\xi}_{\tau_n}(\alpha,\beta) := \beta\,\widehat{\xi}_{\tau_n}(\alpha) + (1-\beta)\,\widetilde{\xi}_{\tau_n}.$$

When $\alpha = 1$, we recover the particular expectile estimator $\overline{\xi}_{\tau_n}(\beta) := \overline{\xi}_{\tau_n}(1,\beta)$ introduced in Daouia *et al.* [11]. The limit distribution of the more general variant $\overline{\xi}_{\tau_n}(\alpha,\beta)$ crucially relies on the asymptotic dependence structure in Theorem 3 between $\overline{\gamma}_{\tau_n}(\alpha)$, $\widehat{q}_{\tau_n}$ and $\widetilde{\xi}_{\tau_n}$.

Theorem 4. *Suppose that the conditions of Theorem 2 hold. Then, for any $\alpha,\beta \in \mathbb{R}$,*

$$\sqrt{n(1-\tau_n)}\left(\frac{\overline{\xi}_{\tau_n}(\alpha,\beta)}{\xi_{\tau_n}} - 1\right) \xrightarrow{d} \beta\left(\mathfrak{b}_\alpha + [(1-\gamma)^{-1} - \log(\gamma^{-1}-1)]\Psi_\alpha + \Theta\right) + (1-\beta)\Xi$$

where the bias component $\mathfrak{b}_\alpha$ is $\mathfrak{b}_\alpha = \lambda_1\mathfrak{b}_{1,\alpha} + \lambda_2\mathfrak{b}_{2,\alpha}$ with

$$\begin{aligned}
\mathfrak{b}_{1,\alpha} &= \frac{(1-\gamma)^{-1} - \log(\gamma^{-1}-1)}{1-\rho}\left[\alpha + (1-\alpha)\frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho}\right] \\
&\quad - \frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} - \frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho}, \\
\mathfrak{b}_{2,\alpha} &= -\gamma(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y)\left(1 + (1-\alpha)[(1-\gamma)^{-1} - \log(\gamma^{-1}-1)]\frac{\gamma}{\gamma+1}\right),
\end{aligned}$$

and $(\Psi_\alpha, \Theta, \Xi)$ is a trivariate Gaussian centered random vector with covariance matrix $\mathfrak{V}_\alpha$ as in Theorem 3.

Let us now extend the estimation procedure far into the right tail, where few or no observations are available. This translates into considering the expectile level $\tau = \tau_n' \to 1$ such that $n(1-\tau_n') \to c \in [0,\infty)$, as $n \to \infty$. To estimate the extreme expectile $\xi_{\tau_n'}$, the basic idea is to extrapolate a consistent expectile estimator of intermediate order τ_n to the very high level τ_n'. To do so, note that on the one hand we have $\xi_{\tau_n'}/\xi_{\tau_n} \sim q_{\tau_n'}/q_{\tau_n}$ in view of (4).

On the other hand, we have the classical Weissman extrapolation formula

$$\frac{q_{\tau_n'}}{q_{\tau_n}} = \frac{U((1-\tau_n')^{-1})}{U((1-\tau_n)^{-1})} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma}$$

as τ_n and τ_n' approach one (Weissman [32]). Thus, we arrive at the expectile approximation

$$\xi_{\tau_n'} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma} \xi_{\tau_n}. \tag{11}$$

By substituting our *expectHill* estimator $\overline{\gamma}_{\tau_n}(\alpha)$ and the general weighted intermediate estimator $\overline{\xi}_{\tau_n}(\alpha,\beta)$, respectively, in place of γ and ξ_{τ_n}, we get the extrapolated expectile estimator

$$\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\overline{\gamma}_{\tau_n}(\alpha)} \overline{\xi}_{\tau_n}(\alpha,\beta). \tag{12}$$

The special case $\alpha = 1$ corresponds to the estimator $\overline{\xi}^{\star}_{\tau_n'}(\beta) := \overline{\xi}^{\star}_{\tau_n'}(1,\beta)$ introduced by Daouia *et al.* [11]. We extend this estimator by using the generalized *expectHill* estimator $\overline{\gamma}_{\tau_n}(\alpha)$ instead of the Hill estimator $\widehat{\gamma}_{\tau_n}$. The next theorem gives the asymptotic behavior of $\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta)$.

Theorem 5. *Suppose that the conditions of Theorem 2 hold. Assume also that $\rho < 0$ and $n(1-\tau_n') \to c < \infty$ with $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau_n')] \to \infty$. Then, for any $\alpha, \beta \in \mathbb{R}$,*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]} \left(\frac{\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta)}{\xi_{\tau_n'}} - 1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha)$$

with (b_α, v_α) as in (9) and (10).

One can observe that the limiting distribution of $\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta)$ is controlled by the asymptotic distribution of $\overline{\gamma}_{\tau_n}(\alpha)$. This is a consequence of the fact that the convergence of $\overline{\xi}^{\star}_{\tau'_n}(\alpha,\beta)$ is governed by that of the extrapolation factor $[(1-\tau'_n)/(1-\tau_n)]^{-\overline{\gamma}_{\tau_n}(\alpha)}$. The latter approximates the theoretical factor $[(1-\tau'_n)/(1-\tau_n)]^{-\gamma}$ in the extrapolation (11) at a slower rate than both the speed of convergence of $\overline{\xi}_{\tau_n}(\alpha,\beta)$ to ξ_{τ_n}, given by Theorem 4, and the speed of convergence to 0 of the bias term that is incurred by the use of (11) and that can be controlled by Theorem 1.

5 Estimation of tail Expected Shortfall

This section aims to estimate both expectile- and quantile-based forms of Expected Shortfall,

$$\text{XES}_\tau := \frac{1}{1-\tau}\int_\tau^1 \xi_t\,dt, \quad \text{QES}_\tau := \frac{1}{1-\tau}\int_\tau^1 q_t\,dt, \tag{13}$$

at a very extreme security level τ that may approach one at an arbitrarily fast rate. To do so, Daouia *et al.* [11] have already suggested to start by estimating these risk measures at an intermediate level $\tau_n \to 1$ such that $n(1-\tau_n)\to\infty$, before extrapolating the resulting estimates to the far tail by making use of the traditional Hill estimator $\widehat{\gamma}_{\tau_n}$ of the tail index γ. Here, we extend their device by using the generalized *expectHill* estimator $\overline{\gamma}_{\tau_n}(\alpha)$ in place of $\widehat{\gamma}_{\tau_n}$. The following asymptotic connections, established in Proposition 3 of Daouia *et al.* [11], will prove instrumental in the estimation procedure.

Proposition 1 (Daouia *et al.*, 2018b). *Assume that* $\mathbb{E}|Y_-| < \infty$ *and that* Y *has a Pareto-type distribution* (2) *with tail index* $0 < \gamma < 1$. *Then*

$$\frac{XES_\tau}{QES_\tau} \sim \frac{\xi_\tau}{q_\tau} \sim \frac{\mathbb{E}[Y|Y > \xi_\tau]}{\mathbb{E}[Y|Y > q_\tau]} \quad \textit{and} \quad \frac{XES_\tau}{\xi_\tau} \sim \frac{1}{1-\gamma} \sim \frac{\mathbb{E}[Y|Y > \xi_\tau]}{\xi_\tau}, \quad \tau \to 1.$$

5.1 Expectile-based Expected Shortfall

Under the model assumptions that $\mathbb{E}|Y_-| < \infty$ and Y has a heavy-tailed distribution (2), we wish to estimate an extreme value of the expectile-based form $\mathrm{XES}_{\tau_n'}$, where $\tau_n' \to 1$ and $n(1-\tau_n') \to c < \infty$. By Proposition 1, we have

$$\frac{\mathrm{XES}_{\tau_n'}}{\mathrm{XES}_{\tau_n}} \sim \frac{\xi_{\tau_n'}}{\xi_{\tau_n}} \quad \text{as} \quad n \to \infty.$$

It follows from the approximation (11) that $\mathrm{XES}_{\tau_n'} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma} \mathrm{XES}_{\tau_n}$. Then, by replacing γ with $\overline{\gamma}_{\tau_n}(\alpha)$ and XES_{τ_n} with its empirical counterpart

$$\widetilde{\mathrm{XES}}_{\tau_n} := \frac{1}{1-\tau_n}\int_{\tau_n}^1 \widetilde{\xi}_t \, dt,$$

we obtain the extrapolated $\mathrm{XES}_{\tau_n'}$ estimator

$$\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\overline{\gamma}_{\tau_n}(\alpha)} \widetilde{\mathrm{XES}}_{\tau_n}. \tag{14}$$

One may also estimate $\mathrm{XES}_{\tau_n'}$ by using the asymptotic equivalence $\mathrm{XES}_{\tau_n'} \sim (1-\gamma)^{-1}\xi_{\tau_n'}$ in Proposition 1. By substituting γ and $\xi_{\tau_n'}$ with their estimators $\overline{\gamma}_{\tau_n}(\alpha)$ and $\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta)$, described respectively in (8) and (12), we define the alternative $\mathrm{XES}_{\tau_n'}$ estimator

$$\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta) := [1-\overline{\gamma}_{\tau_n}(\alpha)]^{-1}\,\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta) \tag{15}$$

for the weights $\alpha,\beta\in\mathbb{R}$. A last option for estimating $\mathrm{XES}_{\tau_n'}$ is motivated by the different asymptotic equivalence $\mathrm{XES}_{\tau_n'}\sim\frac{\xi_{\tau_n'}}{q_{\tau_n'}}\,\mathrm{QES}_{\tau_n'}$ in Proposition 1. This yields the $\mathrm{XES}_{\tau_n'}$ estimator

$$\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta) := \frac{\widehat{\mathrm{QES}}^{\star}_{\tau_n'}(\alpha)}{\widehat{q}^{\star}_{\tau_n'}(\alpha)}\,\overline{\xi}^{\star}_{\tau_n'}(\alpha,\beta) \tag{16}$$

for the estimators $\widehat{q}^{\star}_{\tau_n'}(\alpha)$ of $q_{\tau_n'}$ and $\widehat{\mathrm{QES}}^{\star}_{\tau_n'}(\alpha)$ of $\mathrm{QES}_{\tau_n'}$ defined as

$$\widehat{q}^{\star}_{\tau_n'}(\alpha) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\overline{\gamma}_{\tau_n}(\alpha)}\widehat{q}_{\tau_n}, \tag{17}$$

$$\widehat{\mathrm{QES}}^{\star}_{\tau_n'}(\alpha) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\overline{\gamma}_{\tau_n}(\alpha)}\frac{1}{\lfloor n(1-\tau_n)\rfloor}\sum_{i=1}^{\lfloor n(1-\tau_n)\rfloor}Y_{n-i+1,n}. \tag{18}$$

In the special case $\alpha=1$, the latter estimators are identical to the popular $q_{\tau_n'}$ estimator of Weissman [32] and to the extrapolated $\mathrm{QES}_{\tau_n'}$ estimator of El Methni *et al.* [15], respectively.

The next result provides the convergence of the three estimators $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$, $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ and $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ of $\mathrm{XES}_{\tau_n'}$.

Theorem 6. *Assume that the conditions of Theorem 5 hold. Then, for any* $\alpha, \beta \in \mathbb{R}$,

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)}{\mathrm{XES}_{\tau_n'}}-1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha),$$

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)}{\mathrm{XES}_{\tau_n'}}-1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha),$$

$$\text{and} \qquad \frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)}{\mathrm{XES}_{\tau_n'}}-1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha)$$

with (b_α, v_α) *as in* (9) *and* (10).

The three estimators share the same asymptotic behavior from a theoretical point of view. However, our experience with simulated data in Section 6.2.1 indicates that $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$ is more efficient in the case of real-valued profit-loss distributions with heavy left and right tails, while $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ affords advantageous estimates in the case of non-negative heavy-tailed loss distributions.

5.2 Quantile-based Expected Shortfall

In this section, we return to the estimation of the usual form QES_{p_n} of tail Expected Shortfall, for a pre-specified tail probability $p_n \to 1$ with $n(1-p_n) \to c < \infty$. The generalized Weissman-type estimators $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha)$, defined in (18), already provide a first family of weighted estimators. Here, we wish to derive alternative families of composite expectile-based estimators from the three $\mathrm{XES}_{\tau_n'}$ estimators introduced above, where $\tau_n' = \tau_n'(p_n)$ is to be determined. The starting point is the asymptotic equivalences $\mathrm{QES}_{p_n} \sim \mathbb{E}[Y|Y > q_{p_n}]$ and $\mathrm{XES}_{\tau_n'} \sim \mathbb{E}[Y|Y > \xi_{\tau_n'}]$ in Proposition 1. The basic idea is then to pick out τ_n' so that $\xi_{\tau_n'} \equiv q_{p_n}$, and hence $\mathrm{QES}_{p_n} \sim \mathrm{XES}_{\tau_n'}$. In this way, QES_{p_n} inherits the extreme value estimators of $\mathrm{XES}_{\tau_n'}$ itself, namely $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$, $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ and $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ described in (14), (15) and (16). Yet, it remains to estimate the extreme expectile level $\tau_n'(p_n) := \tau_n'$ such that $\xi_{\tau_n'} = q_{p_n}$. It has been found in Proposition 3 of Daouia *et al.* [10] that such a level satisfies

$$1 - \tau_n'(p_n) \sim (1 - p_n)\,\frac{\gamma}{1-\gamma} \quad \text{as} \quad n \to \infty,$$

under the model assumption of heavy tails (2) with tail index $0 < \gamma < 1$. Built on our novel *expectHill* estimator $\overline{\gamma}_{\tau_n}(\alpha)$ of γ, we can then estimate $\tau_n'(p_n)$ by

$$\widehat{\tau}_n'(p_n) := 1 - (1 - p_n)\,\frac{\overline{\gamma}_{\tau_n}(\alpha)}{1 - \overline{\gamma}_{\tau_n}(\alpha)}. \tag{19}$$

By substituting this estimated value in place of $\tau_n'(p_n) \equiv \tau_n'$ in the extrapolated estimators $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$, $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ and $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$, we obtain composite estimators that estimate $\mathrm{XES}_{\tau_n'(p_n)} \sim \mathrm{QES}_{p_n}$. Note that the composite expectile-based estimator $\widehat{\mathrm{XES}}^{\star}_{\widehat{\tau}_n'(p_n)}(\alpha,1)$, obtained for the special weight $\beta = 1$, is actually identical to the quantile-based estimator $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha)$ defined in (18).

The asymptotic properties of the extrapolated estimators $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$, $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$ and $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha,\beta)$, stated in Theorem 6, still hold true for their composite versions as estimators of QES_{p_n}, with the same conditions.

Theorem 7. *Suppose the conditions of Theorem 5 hold with p_n in place of τ_n'. Then, for any $\alpha, \beta \in \mathbb{R}$,*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-p_n)]}\left(\frac{\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}_n'(p_n)}(\alpha)}{\mathrm{QES}_{p_n}}-1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha),$$
$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-p_n)]}\left(\frac{\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}_n'(p_n)}(\alpha,\beta)}{\mathrm{QES}_{p_n}}-1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha),$$
$$\text{and}\quad \frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-p_n)]}\left(\frac{\widehat{\mathrm{XES}}^{\star}_{\widehat{\tau}_n'(p_n)}(\alpha,\beta)}{\mathrm{QES}_{p_n}}-1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha)$$

with (b_α, v_α) as in (9) and (10).

6 Numerical simulations

In order to illustrate the behavior of the presented estimation procedures of the tail index γ and the two expected shortfall forms $\mathrm{XES}_{\tau_n'}$ and QES_{p_n}, we consider the Student t-distribution with $1/\gamma$ degrees of freedom, the Fréchet distribution $F(x) = e^{-x^{-1/\gamma}},\ x > 0$, and the Pareto distribution $F(x) = 1 - x^{-1/\gamma},\ x > 1$. The finite-sample performance of the different estimators is evaluated through their relative Mean-Squared Error (MSE) and bias, computed over 200 replications. All the experiments have sample size $n = 500$ and true tail index $\gamma \in \{0.35, 0.45\}$ (motivated by our real data applications where the realized values of γ were found to vary between 0.35 and 0.45). In our estimators we used the extreme levels $\tau_n' = p_n = 1 - 1/n$ and the intermediate level $\tau_n = 1 - k/n$, where the integer k can be viewed as the effective sample size for tail extrapolation. To save space, all figures illustrating our simulation results are deferred to Section C of the Supplementary Material document.

6.1 Estimation of the tail index

Our Monte-Carlo simulations in Supplement C.1 indicate that the *expectHill* estimator $\overline{\gamma}_{1-k/n}(\alpha)$, introduced in (8) with the weight $\alpha = 1/2$, is more efficient relative to the standard Hill estimator $\widehat{\gamma}_{1-k/n}$, given in (7), for both Student and Fréchet distributions. In the case of the real-valued Student distribution, it may be seen therein that $\overline{\gamma}_{1-k/n}(\frac{1}{2})$ performs better than $\widehat{\gamma}_{1-k/n}$ in terms of MSE, for all values of k, without sacrificing too much quality in terms of bias, especially for the larger value of γ. We arrive at the same tentative conclusion in the case of the Fréchet distribution. By contrast, in the special case of the Pareto distribution, the Hill estimator $\widehat{\gamma}_{1-k/n}$ is exactly the maximum likelihood estimator of γ and is unbiased, whereas the *expectHill* estimator $\overline{\gamma}_{1-k/n}(\frac{1}{2}) = \frac{1}{2}(\widehat{\gamma}_{1-k/n} + \widetilde{\gamma}_{1-k/n})$ is biased in this case. Unsurprisingly, the Monte Carlo results obtained here indicate that $\widehat{\gamma}_{1-k/n}$ is the winner.

6.2 Expected Shortfall estimation

6.2.1 Estimates of $\mathrm{XES}_{\tau_n'}$

Before comparing the finite-sample performance of $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$ described in (14), $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha, \beta)$ in (15) and $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha, \beta)$ in (16), as estimators of $\mathrm{XES}_{\tau_n'}$, we first investigated the accuracy of each estimator in terms of the associated weights α and β. Then we compared the three estimators with each other by using the best choice of α and β in each scenario; see Supplement C.2. In particular, we arrive at the following tentative conclusion: $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$ seems to be the winner in the case of the real-valued Student distribution for $\alpha = 1$, while $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha, \beta)$ appears to be the most efficient in the case of the non-negative Fréchet and

Pareto distributions, for $\alpha \in \{0.5, 1\}$ and $\beta = 1$.

6.2.2 Estimates of $\mathbf{QES}_{p_n}$

We have also undertaken simulation experiments to evaluate the finite-sample performance of the composite expectile-based estimators $\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)$, $\overline{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)$ and $\widehat{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)$ studied in Theorem 7, with $\widehat{\tau}'_n(p_n)$ being described in (19). They estimate the same conventional expected shortfall QES_{p_n} as the direct quantile-based estimator $\widehat{\text{QES}}^{\star}_{p_n}(\alpha)$ defined in (18). In Supplement C.3, we first examined the accuracy of each estimator for various values of α and β, and then we compared the four estimators with each other. We arrive at the following tentative conclusions:

- In the case of the (real-valued) Student distribution, the best estimator seems to be $\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0)$;
- In the cases of Fréchet and Pareto distributions (both positive), the best estimators seem to be, respectively, $\overline{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0.5, \beta = 1)$ and $\widehat{\text{QES}}^{\star}_{p_n}(\alpha = 1) \equiv \widehat{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 1, \beta = 1)$.

6.2.3 Confidence intervals for $\mathbf{QES}_{p_n}$

By Theorem 7 we have

$$\frac{\sqrt{k}}{\log[k/n(1-p_n)]}\left(\frac{\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)}{\text{QES}_{p_n}} - 1\right) \xrightarrow{d} \mathcal{N}(b_\alpha(\gamma), v_\alpha(\gamma)),$$

where $b_\alpha(\gamma) := b_\alpha$ and $v_\alpha(\gamma) := v_\alpha$ are described in (9) and (10), respectively. Under the bias condition $\lambda_1 = \lambda_2 = 0$ in Theorem 2, the asymptotic bias in (9) reduces to $b_\alpha(\gamma) = 0$.

With this condition, the (symmetric) expectile-based asymptotic confidence interval with confidence level $100\vartheta\%$ has the form $\widetilde{\mathrm{CI}}_\vartheta(k) = \widetilde{\mathrm{XES}}^\star_{\widehat{\tau}'_n(p_n)}(\alpha) \times \mathcal{I}$, where $\mathcal{I}$ stands for the interval

$$\mathcal{I} := \left[1 \pm z_{(1+\vartheta)/2} \log\left(\frac{k}{n(1-p_n)}\right) \sqrt{v_\alpha\left(\overline{\gamma}_{1-k/n}(\alpha)\right)/k}\right],$$

with $z_{(1+\vartheta)/2}$ being the $(1+\vartheta)/2-$quantile of the standard Gaussian distribution. Likewise, the confidence intervals derived from the asymptotic normality of $\overline{\mathrm{XES}}^\star_{\widehat{\tau}'_n(p_n)}(\alpha)$ and $\widehat{\mathrm{XES}}^\star_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)$, in Theorem 7, can be expressed respectively as

$$\overline{\mathrm{CI}}_\vartheta(k) = \overline{\mathrm{XES}}^\star_{\widehat{\tau}'_n(p_n)}(\alpha,\beta) \times \mathcal{I}, \quad \widehat{\mathrm{CI}}_\vartheta(k) = \widehat{\mathrm{XES}}^\star_{\widehat{\tau}'_n(p_n)}(\alpha,\beta) \times \mathcal{I}.$$

Note also that the quantile-based confidence interval, derived from the asymptotic normality of $\widehat{\mathrm{QES}}^\star_{p_n}(\alpha) \equiv \widehat{\mathrm{XES}}^\star_{\widehat{\tau}'_n(p_n)}(\alpha,1)$, is just $\widehat{\mathrm{CI}}_\vartheta(k)$ for $\beta = 1$. In Supplement C.4, we compared the average lengths and the achieved coverages of the three 95% asymptotic confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$, $\overline{\mathrm{CI}}_{0.95}(k)$ and $\widehat{\mathrm{CI}}_{0.95}(k)$. It follows that

- $\widetilde{\mathrm{CI}}_{0.95}(k)$ performs best in the case of the Student distribution, for the selected weight $\alpha = 1$;
- $\widehat{\mathrm{CI}}_{0.95}(k)$ performs quite well in the case of the Fréchet distribution, for the selected weights $\alpha = 1$ and $\beta = 1$;
- $\overline{\mathrm{CI}}_{0.95}(k)$ performs quite well in the case of the Pareto distribution, for the selected weights $\alpha = 1$ and $\beta = 0.5$.

7 Applications

This section applies our expectile-based method to estimate the tail expected shortfall on medical insurance data and financial returns data.

7.1 Medical insurance data

We first illustrate the methodology via the Society of Actuaries group medical insurance large claims data discussed in Beirlant *et al.* [3] and Daouia *et al.* [11], among others. The database contains $n = 75{,}789$ claim amounts exceeding 25,000 USD, collected over the year 1991 from 26 insurers. The scatterplot and histogram of the log-claim amounts, shown in Figure 1(a), clearly exhibit an important right-skewness. Beirlant *et al.* ([3], p.123) have argued that the underlying distribution satisfies the model assumption (2) with a γ estimate around 0.35. A popular measure to assess the magnitude of future unexpected higher claim amounts is the expected shortfall QES_{p_n} defined in (13). Insurance companies typically are interested in an extremely low exceedance probability, say $1 - p_n = 1/100{,}000$, which corresponds to a rare event that occurs on average only once every 100,000 cases.

In this setting of non-negative data with heavy right tail, our experience with simulated data indicates that $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0.5, \beta = 1)$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha = 1)$ provide the best extrapolated pointwise estimates of the extreme value QES_{p_n} in terms of MSE and bias. As such, these are the estimates we adopt here. For the sake of simplicity, they will be denoted by $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}$, respectively.

The evolution of the composite expectile-based estimator $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ as a function of the sample fraction k is represented in Figure 1(b) as rainbow curve, for the selected range of intermediate values of $k = 10, 11, \ldots, 700$. The effect of the *expectHill* estimate $\overline{\gamma}_{1-k/n}(\alpha = 0.5)$ on $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ is highlighted by a colour-scheme, ranging from dark red (low $\overline{\gamma}_{1-k/n}$) to dark violet (high $\overline{\gamma}_{1-k/n}$). This γ estimate seems to mainly vary within the interval $[0.35, 0.36]$, which corresponds to the stable (green) part of the plot. The curve $k \mapsto \overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ exceeds overall the sample maximum $Y_{n,n} = 4.51$ million (indicated by the horizontal pink dashed line). To select a reasonable pointwise estimate, we applied a simple automatic data-driven device that consists first in computing the standard deviations of $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ over a moving window large enough to cover 20% of the possible values of k in the selected range $10 \leqslant k \leqslant 700$. Then the k where the standard deviation is minimal defines the desired sample fraction. The resulting estimate $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)} = 5.99$ million is obtained for the value $k = 208$ in the window $[119, 259]$.

The graph of the pure quantile-based estimator $\widehat{\mathrm{QES}}^{\star}_{p_n}$ against k is superimposed in the same figure as dashed black curve. It is broadly similar to that of $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$, but the latter is smoother and more stable. The pointwise estimate $\widehat{\mathrm{QES}}^{\star}_{p_n} = 6.37$ million is indicated by the minimal standard deviation achieved at $k = 222$ over the window $[119, 259]$. It is more pessimistic (in risk assessment terminology) than $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)} = 5.99$ million, probably due to the instability of the quantile-based plot in dashed black.

Our experience with simulated data also indicates that reasonably good asymptotic 95% confidence intervals for QES_{p_n}, in terms of average lengths and achieved coverages, are provided by $\widehat{\mathrm{CI}}_{0.95}(k)$, constructed via $\widehat{\mathrm{QES}}^{\star}_{p_n}$, and $\overline{\mathrm{CI}}_{0.95}(k)$ constructed on $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 1, \beta = 0.5)$. The two confidence intervals $\overline{\mathrm{CI}}_{0.95}(k)$ and $\widehat{\mathrm{CI}}_{0.95}(k)$ are superimposed in Figure 1(b) as well, respectively, in dotted blue and solid grey lines. Though $\overline{\mathrm{CI}}_{0.95}(k)$ gives slightly more pessimistic confidence bounds than $\widehat{\mathrm{CI}}_{0.95}(k)$, both confidence intervals point towards similar conclusions. In particular, the stable parts of their lower boundaries (around $k \in [100, 500]$) remain quite conservative as they are very close to the maximum recorded claim amount.

We finally comment on the estimator $\widehat{\tau}_n'(p_n)$ of the extreme expectile level $\tau_n'(p_n)$ which ensures that $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}_n'(p_n)}$ is an asymptotically normal estimator for both $\mathrm{XES}_{\tau_n'(p_n)}$ and QES_{p_n}. The graph of $\widehat{\tau}_n'(p_n)$ against k is displayed in Figure 1(c) as rainbow curve, and the corresponding optimal pointwise estimate is indicated by the horizontal dashed black line. As is to be expected from (19), since our estimate of γ is less than $1/2$, this selected optimal level $\widehat{\tau}_n'(p_n) = 0.9999944$ is higher than the pre-specified relative frequency $p_n = 0.99999$ indicated by the horizontal dashed pink line.

7.2 Financial returns data

In this section, we apply our method to estimate the ES for three large US financial institutions. We consider the same investment banks as in the study of Cai *et al.* [7], namely Goldman Sachs, Morgan Stanley and T. Rowe Price. All of these banks had a market capitalization greater than 5 billion USD at the end of June 2007. The dataset consists of the negative log-returns (Y_i) on their equity prices at a daily frequency during 10 years from July 3rd, 2000, to June 30th, 2010. The choice of the frequency of data and time horizon follows the same setup as in Cai *et al.* [7] and Daouia *et al.* [10]. This results in the sample size $n = 2513$. We use our composite expectile-based method to estimate the standard quantile-based expected shortfall QES_{p_n}, or equivalently the expectile-based expected shortfall $\mathrm{XES}_{\tau_n'(p_n)}$, with an extreme relative frequency $p_n = 1 - \frac{1}{n}$ that corresponds to a once-per-decade rare event.

In this setting of real-valued profit-loss distributions, our experience with simulated data

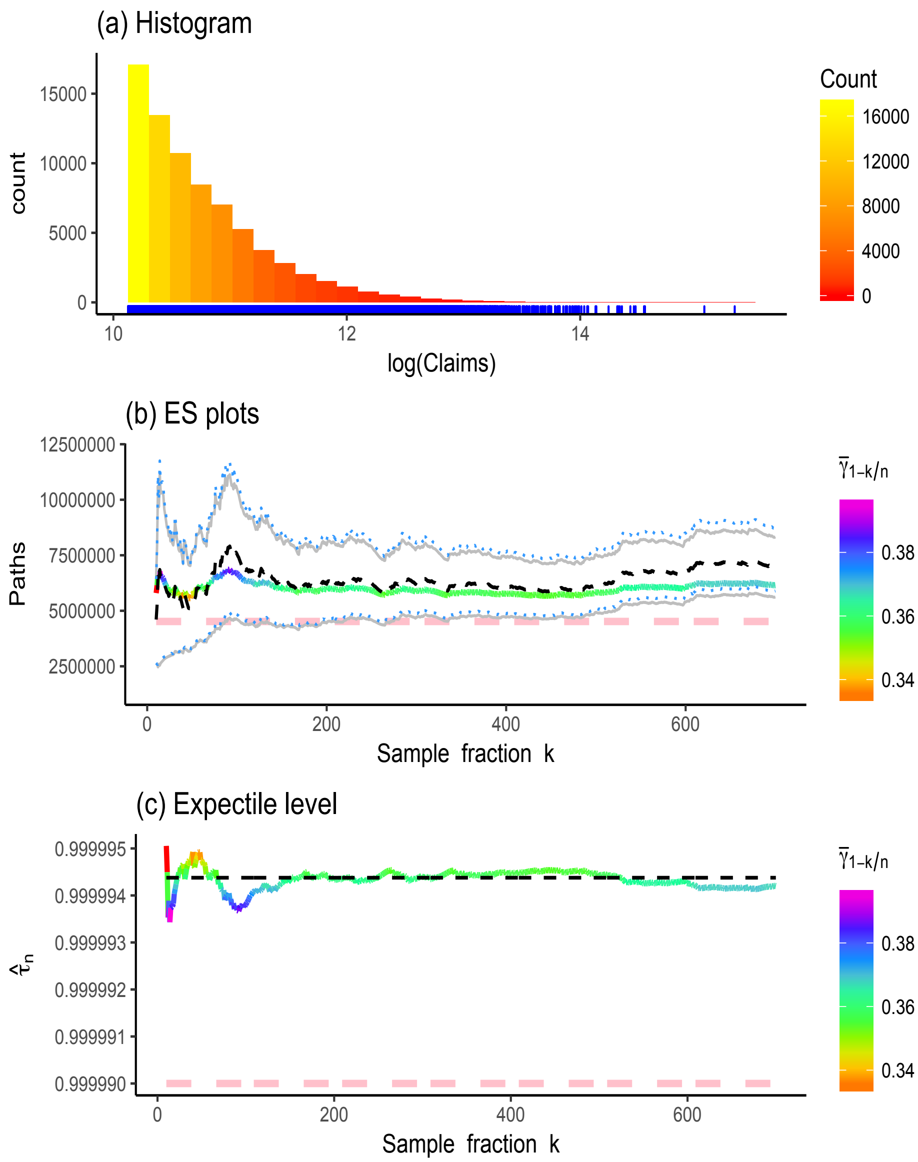

Figure 1: *(a) Scatterplot and histogram of the log-claim amounts. (b) The ES plots* $k \mapsto \overline{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0.5, \beta = 1)$ *as rainbow curve, and* $k \mapsto \widehat{QES}^{\star}_{p_n}(\alpha = 1)$ *in dashed black, along with the constant sample maximum* $Y_{n,n}$ *in horizontal dashed pink. The confidence intervals* $\overline{CI}_{0.95}(k)$ *in dotted blue lines and* $\widehat{CI}_{0.95}(k)$ *in solid grey lines. (c) The plot of* $k \mapsto \widehat{\tau}'_n(p_n)$ *as rainbow curve, along with the selected optimal pointwise estimate in horizontal dashed black line, and the constant tail probability* p_n *in horizontal dashed pink.*

indicates that the composite estimator $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)$ provides the best QES_{p_n} estimates in terms of MSE and bias for the special weight $\alpha = 0$, while it provides reasonably good asymptotic 95% confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ for the different weight $\alpha = 1$. In the estimation, we employ the intermediate sequence $\tau_n = 1 - k/n$ as before, for the selected range of values $k = 1, \ldots, 150$. For our comparison purposes, we use as a benchmark the direct quantile-based estimator $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha = 1) \equiv \widehat{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 1, \beta = 1)$ of El Methni *et al.* [15], as well as the corresponding asymptotic 95% confidence interval $\widehat{\mathrm{CI}}_{0.95}(k)$. We will denote in the sequel the rival estimates $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0)$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha = 1)$ simply as $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}$.

For each bank, we superimpose in Figure 2 the plots of the two estimates $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}$ against k, as rainbow and dashed black curves respectively, along with the competing confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ in dotted blue lines and $\widehat{\mathrm{CI}}_{0.95}(k)$ in solid grey lines. The effect of the *expectHill* estimate $\overline{\gamma}_{1-k/n}(\alpha = 0) \equiv \widetilde{\gamma}_{1-k/n}$ on the estimate $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ is highlighted by a colour-scheme, ranging from dark red (low $\widetilde{\gamma}_{1-k/n}$) to dark violet (high $\widetilde{\gamma}_{1-k/n}$).

We have already provided some Monte Carlo evidence that the composite expectile-based estimates $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ are efficient and accurate relative to the pure quantile-based estimates $\widehat{\mathrm{QES}}^{\star}_{p_n}$ and confidence intervals $\widehat{\mathrm{CI}}_{0.95}(k)$, respectively. Their superiority in terms of plots' stability and confidence intervals' length can clearly be visualized in Figure 2 for the three banks. The final ES levels based on minimizing the standard deviations of the estimates, computed over a moving window covering 20% of the possible values of k, are reported in Table 1, along with the asymptotic 95% confidence intervals of the ES. Based on the reliable $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ estimates (in the second column), the ES levels for Goldman Sachs and T. Rowe Price seem to be very close (around -30% to -34%), whereas the ES level for Morgan Stanley is almost twice higher (around -60%). The $\widehat{\mathrm{QES}}^{\star}_{p_n}$ estimates (in the fourth column) point also towards similar pessimistic results. The lower confidence bands (in third and fifth columns) are themselves quite conservative since they are almost equal to the maximum losses (in the last column) for the three banks.

The theory for our ES estimator $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and for the estimator $\widehat{\mathrm{QES}}^{\star}_{p_n}$ of El Methni *et al.* [15] is derived for independent and identically distributed random variables $Y_1, \ldots, Y_n$. For this application to financial returns, the potential serial dependence may then affect the estimation results. Similarly to our extreme value analysis under mixing conditions in Daouia *et al.* [9], our convergence results may work under serial dependence with enlarged asymptotic variances. A practical solution already employed by Cai *et al.* [7] to reduce substantially the potential serial dependence in this particular dataset is by using weekly loss returns in the same sample period (*i.e.* sums of the daily loss returns during each week). This results in a sample of size $n = 522$. The plots of the two estimates and the asymptotic 95% confidence intervals, against $k \in [1, 80]$, are superimposed in Figure 3 for the three banks, along with the new sample maxima. The final pointwise results are reported in Table 2. By comparing the obtained estimates for the daily and weekly losses, it may be seen that the results are qualitatively robust to the change from daily to weekly data. In particular, the $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ levels for Goldman Sachs and T. Rowe Price are still almost equal, while the estimated level for Morgan Stanley remains almost twice higher. Quantitatively, these ES estimates are much more conservative: around -40% to -43% for Goldman Sachs and T. Rowe Price, and around -87% for Morgan Stanley.

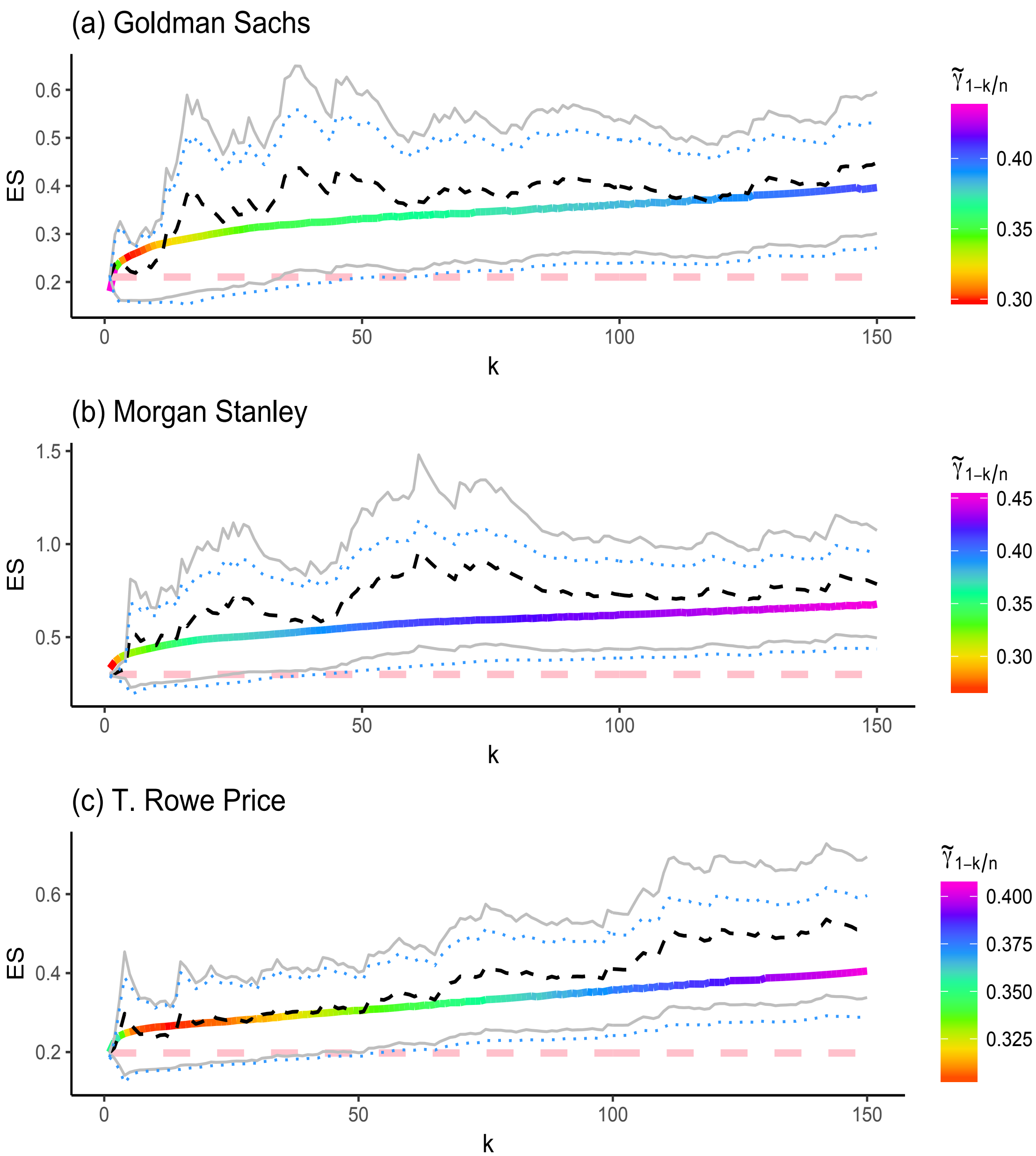

Figure 2: *Results based on daily loss returns of the three investment banks: (a) Goldman Sachs, (b) Morgan Stanley, and (c) T. Rowe Price, with* $n = 2513$ *and* $p_n = 1 - 1/n$*. The estimates* $\widetilde{XES}^{\star}_{\hat{\tau}'_n(p_n)}(\alpha = 0)$ *as rainbow curve and* $\widehat{QES}^{\star}_{p_n}(\alpha = 1)$ *as dashed black curve, along with the asymptotic* 95% *confidence intervals* $\widetilde{CI}_{0.95}(k)$ *in dotted blue lines and* $\widehat{CI}_{0.95}(k)$ *in solid grey lines. The sample maximum* $Y_{n,n}$ *indicated in horizontal dashed pink line.*

Bank	$\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$	$\widetilde{\mathrm{CI}}_{0.95}$	$\widehat{\mathrm{QES}}^{\star}_{p_n}$	$\widehat{\mathrm{CI}}_{0.95}$	$Y_{n,n}$
Goldman Sachs	0.345	(0.210, 0.506)	0.393	(0.235, 0.544)	0.210
Morgan Stanley	0.598	(0.376, 0.785)	0.601	(0.316, 0.984)	0.299
T. Rowe Price	0.308	(0.171, 0.411)	0.301	(0.177, 0.437)	0.197

Table 1: *ES levels of the three investment banks, with the* 95% *confidence intervals and the sample maxima. Results based on daily loss returns, with* $n = 2513$ *and* $p_n = 1 - \frac{1}{n}$.

Bank	$\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$	$\widetilde{\mathrm{CI}}_{0.95}$	$\widehat{\mathrm{QES}}^{\star}_{p_n}$	$\widehat{\mathrm{CI}}_{0.95}$	$Y_{n,n}$
Goldman Sachs	0.436	(0.194, 0.620)	0.495	(0.226, 0.680)	0.365
Morgan Stanley	0.874	(0.384, 1.305)	0.883	(0.366, 1.478)	0.904
T. Rowe Price	0.401	(0.213, 0.511)	0.407	(0.216, 0.548)	0.305

Table 2: *Results based on weekly loss returns, with* $n = 522$ *and* $p_n = 1 - \frac{1}{n}$.

Supplementary Material

The supplement to this article contains simulation results along with the proofs of all our theoretical results.

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., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent Measures of Risk, *Math. Finance*, **9**, 203–228.

[3] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley.

[4] Bellini, F., Klar, B., Müller, A. and Gianina, E.R. (2014). Generalized quantiles as risk measures, *Insurance Math. Econom.*, **54**, 41–48.

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

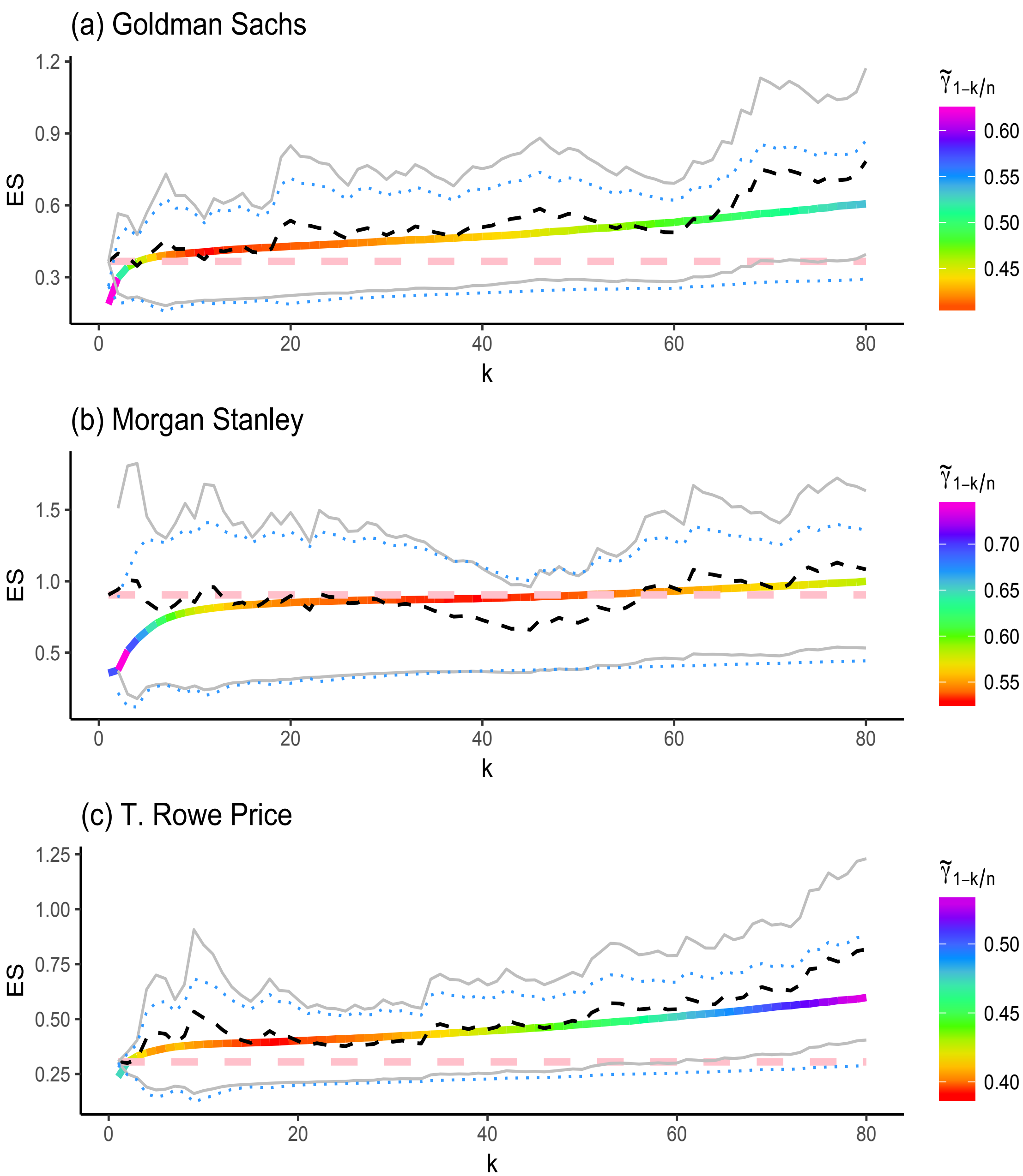

Figure 3: *Results based on weekly loss returns of the three investment banks, with* $n = 522$ *and* $p_n = 1 - 1/n$.

[6] Breckling, J. and Chambers, R. (1988). M-quantiles, *Biometrika*, **75**, 761–772.

[7] 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, *J. R. Stat. Soc. Ser. B*, **77**, 417–442.

[8] Daníelsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O. and Shin, H.S. (2001). An Academic Response to Basel II. Special paper no. 130, Financial Markets Group, London School of Economics.

[9] Daouia, A., Girard, S. and Stupfler, G. (2017). Extreme M-quantiles as risk measures: From L^1 to L^p optimization, *Bernoulli* (to appear). Available at `https://hal.inria.fr/hal-01585215`.

[10] Daouia, A., Girard, S. and Stupfler, G. (2018a). Estimation of tail risk based on extreme expectiles, *J. R. Stat. Soc. Ser. B*, **80**, 263–292.

[11] Daouia, A., Girard, S. and Stupfler, G. (2018b). Tail expectile process and risk assessment, preprint. Available at `https://hal.archives-ouvertes.fr/hal-01744505`.

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

[13] Drees, H., de Haan, L. and Resnick, S. (2000). How to make a Hill plot, *Ann. Statist.*, **28**, 254–274.

[14] Ehm, W., Gneiting, T., Jordan, A. and Krüger, F. (2016). Of quantiles and expectiles: consistent scoring functions, Choquet representations, and forecast rankings, *J. R. Stat. Soc. Ser. B*, **78**, 505–562.

[15] El Methni, J., Gardes, L. and Girard, S. (2014). Nonparametric estimation of extreme risks from conditional heavy-tailed distributions, *Scand. J. Stat.*, **41**, 988–1012.

[16] El Methni, J. and Stupfler, G. (2017). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statist. 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., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer.

[19] Hall, P. and Welsh, A.W. (1985). Adaptive estimates of parameters of regular variation, *Ann. Statist.*, **13**, 331–341.

[20] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Ann. Statist.*, **3**, 1163–1174.

[21] Holzmann, H. and Klar, B. (2016). Expectile asymptotics, *Electron. J. Stat.*, **10**, 2355–2371.

[22] Koenker, R. and Bassett, G.S. (1978). Regression quantiles, *Econometrica*, **46**, 33–50.

[23] Krätschmer, V. and Zähle, H. (2017). Statistical inference for expectile-based risk measures, *Scand. J. Stat.*, **44**, 425–454.

[24] Kuan, C.-M., Yeh, J.-H. and Hsu, Y.-C. (2009). Assessing value at risk with CARE, the Conditional Autoregressive Expectile models, *J. Econometrics*, **2**, 261–270.

[25] Newey, W.K. and Powell, J.L. (1987). Asymmetric least squares estimation and testing, *Econometrica*, **55**, 819–847.

[26] Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer.

[27] Resnick, S. and Stărică, C. (1997). Smoothing the Hill estimator, *Adv. in Appl. Probab.*, **29**, 271–293.

[28] Rockafellar, R.T. and Uryasev, S. (2000). Optimization of Conditional Value-at-Risk, *Journal of Risk*, **2**, 21–42.

[29] Rockafellar, R.T. and Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking and Finance*, **26**, 1443–1471.

[30] Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression, *Comput. Statist. Data Anal.*, **56**, 755–767.

[31] Taylor, J. (2008). Estimating value at risk and expected shortfall using expectiles, *Journal of Financial Econometrics*, **6**, 231–252.

[32] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *J. Amer. Statist. Assoc.*, **73**, 812–815.

[33] Ziegel, J.F. (2016). Coherence and elicitability, *Math. Finance*, **26**, 901–918.