Composite bias-reduced Lp–quantile-based estimators of extreme quantiles and expectiles

Gilles Stupfler, Antoine Usseglio-Carleve

Composite bias-reduced L^p−quantile-based estimators of extreme quantiles and expectiles

Gilles Stupfler[(1)] & Antoine Usseglio-Carleve[(1,2)]

[(1)]Univ Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes, France
[(2)]Toulouse School of Economics, University of Toulouse Capitole, France

Abstract. Quantiles are a fundamental concept in extreme value theory. They can be obtained from a minimization framework using an absolute error loss criterion. The companion notion of expectiles, based on squared rather than absolute error loss minimization, has received substantial attention from the fields of actuarial science, finance and econometrics over the last decade. Quantiles and expectiles can be embedded in a common framework of L^p−quantiles, whose extreme value properties have been explored very recently. Although this generalized notion of quantiles has shown potential for the estimation of extreme quantiles and expectiles, available estimators remain quite difficult to use: they suffer from substantial bias and the question of the choice of the tuning parameter p remains open. In this paper, we work in a context of heavy tails, and we construct composite bias-reduced estimators of extreme quantiles and expectiles based on L^p−quantiles. We provide a discussion of the data-driven choice of p and of the anchor L^p−quantile level in practice. The proposed methodology is compared to existing approaches on simulated data and real data.

Keywords. Bias reduction, expectiles, extrapolation, extremes, heavy tails, L^p−quantiles.

1 Introduction

Carrying out inference about the extremes of a random phenomenon is an important goal in several fields of statistical applications. The first motivating such problem was to determine how high dikes constructed in the Netherlands should be so that areas below sea level can be adequately protected from flooding (de Haan and Ferreira, 2006). Extreme phenomena are also of interest in insurance and finance, as they may negatively affect financial institutions or insurance companies; examples of studies in this direction include the analysis of extreme log-returns of financial time series (Drees, 2003) and the inference about extreme risks linked to large losses in insurance (Rootzén and Tajvidi, 1997). Providing a solution to such problems has typically relied on estimating an extreme quantile of a well-chosen univariate random variable.

Even though quantiles are straightforward to calculate and interpret, they are not devoid of drawbacks: for instance, a quantile only takes into account information on the frequency of a tail event, ignoring information about its magnitude, and does not generally induce a coherent risk measure in the sense of Artzner et al. (1999). This has been the rationale behind the study of alternative extreme value indicators in the recent literature. Among them is the family of expectiles, introduced in Newey and Powell (1987) through an asymmetric least squares minimization problem. Expectiles have the advantage to be a coherent risk measure, contrary to quantiles, making them particularly appealing in actuarial and financial applications (see, among others, Taylor, 2008; Cai and Weng, 2016). In fact, Bellini et al. (2014) have shown that expectiles are the only L^p−quantiles that are also coherent risk measures. Expectiles, however, are not comonotonically additive (Emmer et al., 2015), while quantiles are.

This motivates the use of expectiles as a complement, or alternative, to quantiles for extreme risk assessment. Inference on, and using, extreme expectiles has recently been studied by, among others, Daouia et al. (2018), Daouia et al. (2020), Padoan and Stupfler (2020) and Daouia et al. (2021).

This is done in the context of a heavy-tailed underlying distribution, which is also our context in the present paper: a random variable Y is said to be heavy-tailed when its survival function can be written $\mathbb{P}(Y > y) = y^{-1/\gamma}\ell(y)$, where $\gamma > 0$ and ℓ is a slowly varying function, i.e. such that $\ell(ty)/\ell(t) \to 1$ as $t \to \infty$ for all $y > 0$. The parameter γ is the so-called tail index of Y. In this setup, it is a consequence of the Weissman extrapolation relationship (Weissman, 1978) and of an asymptotic proportionality relationship between extreme expectiles and quantiles that extreme expectiles can be estimated at arbitrarily large levels, using an extrapolation methodology featuring an estimator of the tail index.

Existing expectile estimators have certain weaknesses. In particular, they tend to become unstable as the tail of the underlying distribution gets heavier, because of the inherent non-robustness of expectiles to extreme observations. This general unstability of extreme expectile estimators with respect to large observations was the reason for the investigation of extreme L^p-quantile estimators in Daouia et al. (2019): if $p \geq 1$ and $\alpha \in (0, 1)$, the L^p-quantile $q_\alpha(p)$ of Y is given by

$$q_\alpha(p) \in \underset{t\in\mathbb{R}}{\arg\min}\, \mathbb{E}\left[\eta_\alpha^{(p)}(Y - t) - \eta_\alpha^{(p)}(Y)\right], \tag{1.1}$$

where $\eta_\alpha^{(p)}(y) = \left|\alpha - \mathbb{1}_{\{y<0\}}\right| |y|^p$. Taking $p = 1$ leads to the quantile, see Koenker and Bassett (1978); the case $p = 2$ leads to the expectile. When $\mathbb{E}|Y|^{p-1} < \infty$, problem (1.1) has a unique solution as soon as $p > 1$. For $p \notin \{1, 2\}$, L^p-quantiles are not necessarily coherent nor comonotonically additive, but like expectiles, extreme L^p-quantiles are asymptotically proportional to extreme quantiles (and therefore to extreme expectiles as well). It follows that extreme expectiles and quantiles can in fact be estimated using L^p-quantiles through inverting this asymptotic proportionality relationship. This gives rise to classes of L^p-quantile-based composite estimators of extreme expectiles and extreme quantiles. For expectiles, this can beneficial if p is chosen between 1 and 2, because L^p-quantiles are then more robust than expectiles to extreme values. For quantiles, this may also be interesting because L^p-quantiles take into account both the magnitude and frequency of extreme observations, and so may provide estimators that take a more holistic perspective of the distribution, and especially of its right tail, than traditional extreme quantile estimators which extrapolate a single order statistic. Of course, the use of L^p-quantiles requires making the moment assumption $\mathbb{E}|Y|^{p-1} < \infty$ that strictly speaking is not required for the estimation of extreme quantiles, but in the kind of risk assessment situations we are interested in, a first or indeed second moment will often exist: see *e.g.* recently Chavez-Demoulin et al. (2014), Cai et al. (2015) and Alm (2016), as well as the R package `CASdatasets`. Even if this is not the case, the fact that L^p-quantiles with low p require only fractional moments ensures that the L^p-quantile approach can always be employed for a (possibly small) range of values of $p > 1$.

However, these composite estimators typically suffer from bias that can be substantial in finite-sample situations. This is due to (i) the error made in the use of the asymptotic proportionality relationship, and (ii) the use of the Weissman extrapolation relationship, which strictly speaking will only be exactly valid when the underlying distribution is purely Pareto rather than just heavy-tailed. The contribution of the present paper is to develop and analyze (theoretically and empirically) bias-reduced versions of the L^p-quantile-based composite estimators of extreme expectiles and extreme quantiles introduced in Daouia et al. (2019). Our work is based on an asymptotic expansion of the proportionality relationship between extreme L^p-quantiles and expectiles or quantiles, and on an estimation of the errors terms arising. Because bias also features at the extrapolation step, we propose to use an L^p-quantile-based, bias-corrected estimator of the tail index γ to eliminate this extrapolation-specific bias. This is paired with new selection strategies for the tuning parameters involved in the estimators, in particular the power p, to result in fully data-driven estimators that can be readily used as part of the R package `Expectrem`, available at `https://github.com/AntoineUC/Expectrem`. The bias-reduced estimators shall be compared to their original versions, as well as to bias-reduced versions that do not rely on the connection to L^p-quantiles, introduced by Gomes and Pestana (2007) for extreme quantile estimation and recently by Girard et al. (2020b) for extreme expectile estimation. We will especially show that our proposed method provides a very substantial improvement when the response variable has a finite variance and a second-order parameter close to 0, which is an interesting and difficult case in practice.

The paper is organized as follows. Section 2 introduces our setting, notation, and modeling assumptions. Section 3 discusses the construction of the proposed bias-corrected estimators, first by focusing on the bias terms arising from the use of the asymptotic proportionality relationship, and then by constructing an L^p-quantile-based, asymptotically unbiased estimator of the tail index. Section 4 introduces selection rules for our tuning parameters and examines the finite-sample performance of our estimators on a simulation study, and Section 5 showcases our estimators on two real insurance data sets. Proofs of all our auxiliary and main results are postponed to the Appendices.

2 Setting, notation and model

Let $p \geq 1$ and let Y be a real random variable such that $\mathbb{E}[|Y|^{p-1}] < \infty$ having a **continuous distribution function** $F = F^{(1)}$. Differentiating the cost function (1.1), one finds that the L^p-quantile $q_\alpha(p)$ of level $\alpha \in (0,1)$ of Y is equivalently defined as follows:

$$q_\alpha(p) = \inf\left\{y \in \mathbb{R} \colon \overline{F}^{(p)}(y) \leq 1-\alpha\right\} \text{ with } \overline{F}^{(p)}(y) = \frac{\mathbb{E}\left[|Y-y|^{p-1}\,\mathbb{1}_{\{Y>y\}}\right]}{\mathbb{E}\left[|Y-y|^{p-1}\right]}. \tag{2.1}$$

This is indeed justified because $\overline{F}^{(p)}$ is a decreasing function; see Lemma A.1 in the case $p>1$. In view of (2.1), if independent copies $Y_1, \ldots, Y_n$ of Y are available, an estimator of an L^p-quantile can be obtained by replacing expectations by their empirical counterparts:

$$\widehat{q}_\alpha(p) = \inf\left\{y \in \mathbb{R} \colon \widehat{\overline{F}}_n^{(p)}(y) \leq 1-\alpha\right\} \text{ with } \widehat{\overline{F}}_n^{(p)}(y) = \frac{\sum_{i=1}^n |Y_i-y|^{p-1}\,\mathbb{1}_{\{Y_i>y\}}}{\sum_{i=1}^n |Y_i-y|^{p-1}}. \tag{2.2}$$

The asymptotic properties of $\widehat{q}_\alpha(p)$ are established in Daouia et al. (2019) for high levels $\alpha \to 1$ when Y is heavy-tailed. This corresponds to making the following fundamental assumption on the right tail of Y:

$\mathcal{C}_1(\gamma)$ The survival function $\overline{F}^{(1)}$ of Y is regularly varying at infinity with index $-1/\gamma < 0$, that is:

$$\lim_{t\to\infty} \frac{\overline{F}^{(1)}(ty)}{\overline{F}^{(1)}(t)} = y^{-1/\gamma}, \ \forall y > 0.$$

According to de Haan and Ferreira (2006, Theorem 1.2.1), this is equivalent to assuming that the distribution of Y belongs to the Fréchet maximum domain of attraction, with extreme value index $\gamma > 0$. Under this assumption, one may prove the following asymptotic proportionality relationship between high quantiles and L^p-quantiles: for all $p > 1$ and $0 < \gamma < 1/(p-1)$,

$$\lim_{\alpha\to 1} \frac{\overline{F}^{(1)}(q_\alpha(p))}{1-\alpha} = \lim_{\alpha\to 1} \frac{\overline{F}^{(1)}(q_\alpha(p))}{\overline{F}^{(1)}(q_\alpha(1))} = \frac{\gamma}{B(p, \gamma^{-1}-p+1)} = g_p(\gamma), \tag{2.3}$$

where $B(x,y) = \int_0^1 u^{x-1}(1-u)^{y-1}du$ is the Beta function. See Daouia et al. (2019, Proposition 1). This motivates the following Weissman-type estimator of an extreme L^p-quantile $\widehat{q}^{\star}_{\alpha'_n}(p)$: if $\alpha'_n \to 1$ is such that $n(1-\alpha'_n) \to c < \infty$,

$$\widehat{q}^{\star}_{\alpha'_n}(p) = \left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(p), \tag{2.4}$$

where $\overline{\gamma}$ is a consistent estimator of γ and $\widehat{q}_{\alpha_n}(p)$ is the estimator in (2.2) at a much lower, in-the-sample intermediate level α_n (i.e. such that $n(1-\alpha_n) \to \infty$). This class of estimators is used in Daouia et al. (2019) as vehicles to estimate extreme quantiles and expectiles. A drawback of these extreme L^p-quantile estimators is that their finite-sample bias can be fairly substantial; this can be seen from e.g. Daouia et al. (2019, Proposition 3), whereby the higher-order error terms in the asymptotic proportionality relationship (2.3) linking extreme L^p-quantiles to extreme quantiles can be quite large. In the following section, we first provide expressions for these higher-order error terms, before using them to construct bias-reduced, L^p-quantile-based estimators of extreme quantiles and expectiles.

3 Main results

3.1 Construction of bias-reduced L^p-quantile-based extreme quantile and expectile estimators

The idea behind the construction of L^p-quantile-based extreme quantile and expectile estimators is to use the asymptotic proportionality relationship

$$\lim_{\alpha\to 1}\frac{q_\alpha(p)}{q_\alpha(1)} = [g_p(\gamma)]^{-\gamma} \Leftrightarrow q_\alpha(1) = [g_p(\gamma)]^{\gamma} q_\alpha(p)(1+o(1)) \text{ as } \alpha\to 1, \tag{3.1}$$

in order to connect extreme L^p-quantiles to extreme quantiles (corresponding to L^1-quantiles). It is then possible to connect extreme L^p-quantiles to extreme expectiles by using relationship (3.1) for $p=2$:

$$\lim_{\alpha\to 1}\frac{q_\alpha(p)}{q_\alpha(2)} = \lim_{\alpha\to 1}\frac{q_\alpha(p)}{q_\alpha(1)}\times\frac{q_\alpha(1)}{q_\alpha(2)} = \left[\frac{g_p(\gamma)}{g_2(\gamma)}\right]^{-\gamma} \Leftrightarrow q_\alpha(2) = \left[\frac{g_p(\gamma)}{g_2(\gamma)}\right]^{\gamma} q_\alpha(p)(1+o(1)) \text{ as } \alpha\to 1.$$

These two approximations motivate the following extreme quantile and expectile estimators based on the Weissman-type extreme L^p-quantile estimator $\widehat{q}^{\star}_{\alpha'_n}(p)$ in Equation (2.4):

$$\widetilde{q}^{\star}_{\alpha'_n,p}(1) = \widehat{q}^{\star}_{\alpha'_n}(p)[g_p(\overline{\gamma})]^{\overline{\gamma}} = \left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(p)[g_p(\overline{\gamma})]^{\overline{\gamma}}$$

$$\text{and}\quad \widetilde{q}^{\star}_{\alpha'_n,p}(2) = \widehat{q}^{\star}_{\alpha'_n}(p)\left(\frac{g_p(\overline{\gamma})}{g_2(\overline{\gamma})}\right)^{\overline{\gamma}} = \left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(p)\left(\frac{g_p(\overline{\gamma})}{g_2(\overline{\gamma})}\right)^{\overline{\gamma}}.$$

These estimators were originally proposed in Daouia et al. (2019, Sections 4 and 5). Their construction combines several approximations (the Weissman extrapolation relationship and the asymptotic proportionality between L^p-quantiles, quantiles and expectiles) whose use is the main generator of bias in these estimators. We therefore first carry out a detailed study of the Weissman and proportionality approximations, which will be instrumental in providing a bias-reduced version of these estimators. The crucial condition for doing so is the following second-order refinement of the heavy tail assumption.

$\mathcal{C}_2(\gamma,\rho,A)$ The survival function $\overline{F}^{(1)}$ is second-order regularly varying with index $-1/\gamma<0$, second-order parameter $\rho\le 0$ and auxiliary function A having constant sign and converging to 0 at infinity, i.e.

$$\lim_{t\to\infty}\frac{1}{A\left(1/\overline{F}^{(1)}(t)\right)}\left(\frac{\overline{F}^{(1)}(ty)}{\overline{F}^{(1)}(t)} - y^{-1/\gamma}\right) = y^{-1/\gamma}\frac{y^{\rho/\gamma}-1}{\gamma\rho},\ \forall y>0.$$

Here $(y^x-1)/x$ should be read as $\log(y)$ when $x=0$.

Condition $\mathcal{C}_2(\gamma,\rho,A)$ is equivalent to (see de Haan and Ferreira, 2006, Theorem 2.3.9)

$$\lim_{t\to\infty}\frac{1}{A(t)}\left(\frac{q_{1-1/(ty)}(1)}{q_{1-1/t}(1)} - y^{\gamma}\right) = y^{\gamma}\frac{y^{\rho}-1}{\rho},\ \forall y>0. \tag{3.2}$$

Such a second-order condition allows the control of bias terms in statistical extreme value procedures through the function A, and as such, is the cornerstone for asymptotic normality results in extreme value theory. Beirlant et al. (2004) provide a large number of examples of commonly used continuous distributions satisfying $\mathcal{C}_2(\gamma,\rho,A)$. Note also that any distribution satisfying

$$\overline{F}^{(1)}(y) = y^{-1/\gamma}\left(a + by^{\rho/\gamma} + \mathrm{o}(y^{\rho/\gamma})\right) \text{ as } y\to\infty,$$

where $a,b>0$ and $\rho<0$, will automatically satisfy $\mathcal{C}_2(\gamma,\rho,A)$. This contains in particular the Hall-Weiss class of models, see Hua and Joe (2011).

The key point for an accurate quantification of bias is to note that by definition of L^p-quantiles (see Equation (2.1)),

$$\mathbb{E}\left[\left(\frac{Y}{q_\alpha(p)}-1\right)^{p-1}\mathbb{1}_{\{Y>q_\alpha(p)\}}\right] = (1-\alpha)\mathbb{E}\left[\left|\frac{Y}{q_\alpha(p)}-1\right|^{p-1}\right].$$

Moreover, according to Equation (A.9) in the supplementary material of Daouia et al. (2019), the left-hand side of the above equation can be expanded as

$$\begin{aligned}&\frac{1}{\overline{F}^{(1)}(q_\alpha(p))}\mathbb{E}\left[\left(\frac{Y}{q_\alpha(p)}-1\right)^{p-1}\mathbb{1}_{\{Y>q_\alpha(p)\}}\right]\\ &= [g_p(\gamma)]^{-1}\left(1 + A\left(1/\overline{F}^{(1)}(q_\alpha(p))\right)[g_p(\gamma)]^{1+\rho}K(p,\gamma,\rho)(1+o(1))\right)\end{aligned}$$

as $\alpha\to1$, where

$$K(p,\gamma,\rho) = \begin{cases}\dfrac{[g_p(\gamma)]^{-\rho}}{\gamma^2\rho}\left[(1-\rho)B(p,(1-\rho)\gamma^{-1}-p+1) - B(p,\gamma^{-1}-p+1)\right] & \text{if } \rho<0,\\[2ex] \dfrac{p-1}{\gamma^2}\displaystyle\int_1^\infty (u-1)^{p-2}u^{-1/\gamma}\log(u)du & \text{if } \rho=0.\end{cases} \tag{3.3}$$

Therefore

$$\frac{\overline{F}^{(1)}(q_\alpha(p))}{1-\alpha} = g_p(\gamma)(1+r(p,\alpha)) \tag{3.4}$$

$$\text{with } 1+r(p,\alpha) = \mathbb{E}\left[\left|\frac{Y}{q_\alpha(p)}-1\right|^{p-1}\right]\left(1 + A\left(1/\overline{F}^{(1)}(q_\alpha(p))\right)K(p,\gamma,\rho)[g_p(\gamma)]^{1+\rho}(1+o(1))\right)^{-1}$$

as $\alpha\to1$. An asymptotic inversion lemma, such as Lemma 2 in the supplementary material of Daouia et al. (2019), combined with (3.2), then suggests that

$$q_\alpha(1) = [g_p(\gamma)]^\gamma(1+r(p,\alpha))^\gamma\left(1 + \frac{[g_p(\gamma)]^{-\rho}(1+r(p,\alpha))^{-\rho}-1}{\rho}A\left((1-\alpha)^{-1}\right)(1+o(1))\right)^{-1}q_\alpha(p)$$

as $\alpha\to1$. This quantifies the bias in the asymptotic proportionality relationship (3.1). Besides, the bias in the Weissman extrapolation formula for extreme quantiles is well-known (see for instance Chapter 4 in de Haan and Ferreira, 2006):

$$q_{\alpha_n'}(1) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\gamma}q_{\alpha_n}(1)\left(1 + \frac{\left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1))\right).$$

Combining these two bias quantifications results in the following asymptotic expansion linking an extreme quantile to an intermediate L^p-quantile:

$$\begin{aligned}q_{\alpha_n'}(1) &= \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\gamma}q_{\alpha_n}(p)[g_p(\gamma)]^\gamma\frac{(1+r(p,\alpha_n))^\gamma}{1+\frac{[g_p(\gamma)]^{-\rho}(1+r(p,\alpha_n))^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1))}\\ &\times\left(1 + \frac{\left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1))\right).\end{aligned}$$

This identity is the basis for the construction of a bias-reduced version of $\widetilde{q}^{\star}_{\alpha'_n,p}(1)$ through a plug-in of estimators of all the unknown quantities on the right-hand side. In particular, this requires the estimation of the quantity $A((1-\alpha_n)^{-1})$. We assume throughout that the function A can be expressed as $A(t) = b\gamma t^{\rho}$; this amounts to assuming that the underlying distribution belongs to the Hall-Welsh class in the sense of Gomes and Pestana (2007), which is reasonable for modeling purposes. This reduces the estimation of $A((1-\alpha_n)^{-1})$ to the estimation of b and ρ; consistent estimators of b and ρ are available from various sources, such as the R package `evt0`. Assuming then that consistent estimators $\overline{b}$, $\overline{\rho}$ and $\overline{\gamma}$ of b, ρ and γ, have been chosen, we construct a bias-reduced version of the extreme quantile estimator $\widetilde{q}^{\star}_{\alpha'_n,p}(1)$ as

$$\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p}(1) = \left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(p)[g_p(\overline{\gamma})]^{\overline{\gamma}} \frac{(1+\overline{r}(p,\alpha_n))^{\overline{\gamma}}}{1+\frac{[g_p(\overline{\gamma})]^{-\overline{\rho}}(1+\overline{r}(p,\alpha_n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\alpha_n)^{-\overline{\rho}}} \times\left(1+\frac{\left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\alpha_n)^{-\overline{\rho}}\right),$$

where

$$1+\overline{r}(p,\alpha_n) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i}{\widehat{q}_{\alpha_n}(p)}-1\right|^{p-1}\left(1+\overline{b}\overline{\gamma}\left[\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))\right]^{-\overline{\rho}}K(p,\overline{\gamma},\overline{\rho})[g_p(\overline{\gamma})]^{1+\overline{\rho}}\right)^{-1}.$$

The estimation of extreme expectiles from L^p-quantiles uses two approximations, first at the level α'_n to connect $q_{\alpha'_n}(2)$ to $q_{\alpha'_n}(1)$, and then to connect an extreme quantile to an extreme L^p-quantile. Compared to the previous construction, reducing the bias of the extreme expectile estimator $\widetilde{q}^{\star}_{\alpha'_n,p}(2)$ then involves an extra bias correction at the level α'_n. This results in the following bias-reduced estimator:

$$\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p}(2) = \left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(p)\left(\frac{g_p(\overline{\gamma})}{g_2(\overline{\gamma})}\right)^{\overline{\gamma}}\left(\frac{1+\overline{r}(p,\alpha_n)}{1+\overline{r}(2,\alpha'_n)}\right)^{\overline{\gamma}} \times \frac{1+\frac{[g_2(\overline{\gamma})]^{-\overline{\rho}}(1+\overline{r}(2,\alpha'_n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\alpha'_n)^{-\overline{\rho}}}{1+\frac{[g_p(\overline{\gamma})]^{-\overline{\rho}}(1+\overline{r}(p,\alpha_n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\alpha_n)^{-\overline{\rho}}} \times\left(1+\frac{\left(\frac{1-\alpha'_n}{1-\alpha_n}\right)^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\alpha_n)^{-\overline{\rho}}\right)$$

where

$$1+\overline{r}(2,\alpha'_n) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i}{\widehat{q}^{\star}_{\alpha'_n}(2)}-1\right|\left(1+\overline{b}\overline{\gamma}(\overline{\gamma}^{-1}-1)^{-\overline{\rho}}(1-\alpha'_n)^{-\overline{\rho}}K(2,\overline{\gamma},\overline{\rho})[g_2(\overline{\gamma})]^{1+\overline{\rho}}\right)^{-1}.$$

It is important to note that the estimators $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p}(1)$ and $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p}(2)$ are not straightforward plug-in estimators of the bias terms obtained in Daouia et al. (2019). In particular, using a linearization of the bias term as in Daouia et al. (2019) would have a substantial negative impact on finite-sample performance.

A crucial component of the estimators $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p}(1)$ and $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p}(2)$ is the tail index estimator $\overline{\gamma}$. We discuss the construction of a purely L^p-quantile-based estimator of the tail index in the next section.

3.2 An L^p-quantile based methodology for bias-reduced tail index estimation

A very popular tail index estimator is the Hill estimator of Hill (1975),

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

Like the Hill estimator, available tail index estimators (see de Haan and Ferreira, 2006, Chapter 3) are based on quantiles. We suggest here a construction of a bias-reduced tail index estimator based on L^p-quantiles, which will complement our construction of bias-reduced, L^p-quantile-based extreme quantile and expectile estimators.

For that purpose, we recall (2.3) and note that the function g_p is strictly decreasing for all $p > 1$ (see Lemma A.2 in the Appendix). A natural idea may then be to estimate the tail index γ by inverting the asymptotic relationship (2.3): in other words, to consider an intermediate sequence (α_n) and define a tail index estimator $\widehat{\gamma}_{\alpha_n}^{(p)}$ as

$$\widehat{\gamma}_{\alpha_n}^{(p)} = \inf\left\{\gamma > 0 \colon g_p(\gamma) \leq \frac{\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))}{1-\alpha_n}\right\}.$$

Since L^p-quantiles are equivariant by increasing affine transformations (see Bellini et al. (2014)), it is noteworthy that the new estimator $\widehat{\gamma}_{\alpha_n}^{(p)}$ is both shift- and scale-invariant, in contrast to the Hill estimator which is not shift-invariant. The particular cases $p \in \{2, 3, 4\}$ yield closed formulas for $\widehat{\gamma}_{\alpha_n}^{(p)}$. For instance, $p = 2$ leads to $g_p(\gamma) = \gamma^{-1} - 1$ and therefore to the expectile-based estimator

$$\widehat{\gamma}_{\alpha_n}^{(2)} = \left(1 + \frac{\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(2))}{1-\alpha_n}\right)^{-1}.$$

This estimator was introduced by Girard et al. (2020a) in a conditional setting. The asymptotic normality of $\widehat{\gamma}_{\alpha_n}^{(2)}$ is established therein under the condition $\gamma < 1/2$. The objective of this section is to deal with the case $p > 1$ in its full generality, to establish a general asymptotic normality result for $\widehat{\gamma}_{\alpha_n}^{(p)}$, and in particular to derive its asymptotic bias. This will be a prerequisite for the construction of a bias-reduced version of $\widehat{\gamma}_{\alpha_n}^{(p)}$.

In the sequel, we denote by $IB(t, x, y) = \int_0^t u^{x-1}(1-u)^{y-1}du$ the incomplete Beta function (for $0 < t < 1$) and by Ψ the digamma function i.e. the log-derivative of Euler's Gamma function. Our first theoretical result provides the joint asymptotic distribution of the empirical intermediate quantiles and L^p-quantiles, for $p > 1$. Denote by $\wedge$ and $\vee$ the minimum and maximum operators, i.e. $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$.

Proposition 1. *Assume $\overline{F}^{(1)}$ satisfies $\mathcal{C}_2(\gamma, \rho, A)$ with $\gamma < 1/[2(p-1)]$ and $\mathbb{E}(|\min(Y, 0)|^{2(p-1)}) < \infty$. Let (α_n) and (β_n) be two intermediate sequences such that $\sqrt{n(1-\alpha_n)}A\left((1-\alpha_n)^{-1}\right) = O(1)$ and $(1-\beta_n)/(1-\alpha_n) \to \theta > 0$ as $n \to \infty$. Then*

$$\sqrt{n(1-\alpha_n)}\left(\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)} - 1, \frac{\widehat{q}_{\beta_n}(1)}{q_{\beta_n}(1)} - 1\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \gamma^2\mathbf{\Lambda}\right),$$

where $\mathbf{\Lambda}$ is a symmetric 2×2 matrix having entries

$$\begin{cases} \Lambda_{1,1} &= \dfrac{B\left(2p-1, \gamma^{-1} - 2p + 2\right)}{B\left(p, \gamma^{-1} - p + 1\right)}, \\ \Lambda_{1,2} &= (p-1)\gamma(g_p(\gamma)/\theta)IB\left((g_p(\gamma)/\theta)^{-\gamma} \wedge 1, \gamma^{-1} - p + 1, p - 1\right) + \left((g_p(\gamma)/\theta)^{\gamma} \vee 1 - 1\right)^{p-1}, \\ \Lambda_{2,2} &= 1/\theta. \end{cases}$$

This result extends Theorem 3 of Daouia et al. (2020) which is dedicated to the case $p = 2$. It also extends, in the independent and identically distributed case specifically, Theorem 1 of Daouia et al. (2019) which focuses on the marginal asymptotic distribution of $\widehat{q}_{\alpha_n}(p)$ only.

Using Proposition 1, the asymptotic normality of $\widehat{\gamma}_{\alpha_n}^{(p)}$ can then be established, under further conditions which allows one to evaluate the error in approximation (2.3). Since L^p-quantiles are calculated using the whole of the underlying distribution, it should be expected that these bias conditions will involve both distribution tails. We therefore assume that $\overline{F}^{(1)}_{-}$, the survival function of $-Y$, either satisfies $y^a\overline{F}^{(1)}_{-}(y) \to 0$ as $y \to +\infty$ for all $a > 0$ (in particular, it can be light-tailed or short-tailed, and this condition is especially automatically satisfied when $Y \geq 0$ with probability 1) or satisfies a first-order condition of the form $\mathcal{C}_1(\gamma_\ell)$. In the case where $\overline{F}^{(1)}_{-}$ satisfies $y^a\overline{F}^{(1)}_{-}(y) \to 0$ as $y \to +\infty$ for all $a > 0$, our subsequent conditions and results may be read with the convention $\gamma_\ell = 0$.

With this further assumption, we are now ready to state our next result.

Theorem 1. *Assume $\overline{F}^{(1)}$ satisfies $\mathcal{C}_2(\gamma, \rho, A)$ and $\overline{F}^{(1)}_{-}$ either satisfies $y^a\overline{F}^{(1)}_{-}(y) \to 0$ as $y \to +\infty$ for all $a > 0$, or $\mathcal{C}_1(\gamma_\ell)$. Let (α_n) be an intermediate sequence such that, as $n \to \infty$,*

i) $\sqrt{n(1-\alpha_n)}A\left((1-\alpha_n)^{-1}\right) \to \lambda_1 \in \mathbb{R}$,

ii) $\sqrt{n(1-\alpha_n)}\left(\dfrac{\mathbb{E}\left[Y\mathbb{1}_{\{0<Y<q_{\alpha_n}(1)\}}\right]}{q_{\alpha_n}(1)} \vee (1-\alpha_n)\right) \to \lambda_2 \in \mathbb{R}$,

iii) $\sqrt{n(1-\alpha_n)}\left(\dfrac{\mathbb{E}\left[-Y\mathbb{1}_{\{0<-Y<q_{\alpha_n}(1)\}}\right]}{q_{\alpha_n}(1)} \vee F^{(1)}(-q_{\alpha_n}(1))\right) \to \lambda_3 \in \mathbb{R}$.

Then, for all $1 < p < (\gamma \vee \gamma_\ell)^{-1}/2 + 1$,

$$\sqrt{n(1-\alpha_n)}\left(\widehat{\gamma}^{(p)}_{\alpha_n} - \gamma\right) \xrightarrow{d} \mathcal{N}(b_p(\gamma, \gamma_\ell, \rho), v_p(\gamma)),$$

where

$$\begin{cases} b_p(\gamma,\gamma_\ell,\rho) &= \dfrac{-\gamma}{1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)}\left[\lambda_1 b_p^{(1)}(\gamma,\rho)+\lambda_2 b_p^{(2)}(\gamma)+\lambda_3 b_p^{(3)}(\gamma,\gamma_\ell)\right], \\ v_p(\gamma) &= \dfrac{\gamma B\left(p,\gamma^{-1}-p+1\right)}{\left(1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)\right)^2}\left[\gamma\dfrac{B(2p-1,\gamma^{-1}-2p+2)}{B(p,\gamma^{-1}-p+1)^2}-1\right] \end{cases}$$

and

$$\begin{cases} b_p^{(1)}(\gamma,\rho) &= g_p(\gamma)K(p,\gamma,\rho), \\ b_p^{(2)}(\gamma) &= (p-1)[g_p(\gamma)]^{\gamma}\mathbb{1}_{\{\gamma\leq 1\}}+\left[(p-1)g_p(\gamma)B(p-1,1-\gamma^{-1})-(1-g_p(\gamma))\right]\mathbb{1}_{\{\gamma>1\}}, \\ b_p^{(3)}(\gamma,\gamma_\ell) &= -(p-1)\left[[g_p(\gamma)]^{\gamma}\mathbb{1}_{\{\gamma_\ell\leq 1\}}+[g_p(\gamma)]^{\gamma/\gamma_\ell}B(\gamma_\ell^{-1}-p+1,1-\gamma_\ell^{-1})\mathbb{1}_{\{\gamma_\ell>1\}}\right] \end{cases}$$

where $K(p, \gamma, \rho)$ is given in Equation (3.3)*.*

In Figure 1, the asymptotic variance $v_p(\gamma)$ is compared to the asymptotic variance $v_H(\gamma) = \gamma^2$ of the Hill estimator, for various values of p and $\gamma \in (0, 1)$. It can be seen therein that the asymptotic variance of $\widehat{\gamma}^{(p)}_{\alpha_n}$ is smaller than $v_H(\gamma)$ when γ is small, for all considered values of p. On the contrary, $v_p(\gamma)$ increases faster than $v_H(\gamma)$ as γ increases. It also appears that $p = 1.4$ seems to be a good compromise when γ is less than $\simeq 2/3$, which is in line with what Daouia et al. (2019) found regarding extreme L^p-quantile estimation (see Figure 2 therein). This includes the important case $\gamma < 1/2$ of a finite variance, which is widespread in insurance and finance. Note that Theorem 1 involves the condition $\gamma < (p-1)^{-1}/2$, which is less restrictive than $\gamma < 1/2$ when $p \in (1, 2)$. As such, the estimator $\widehat{\gamma}^{(p)}_{\alpha_n}$ is applicable in a larger range of situations compared to the estimator $\widehat{\gamma}^{(2)}_{\alpha_n}$ studied in Girard et al. (2020a). Note also that, when $|\rho|$ is larger than 1 and for nonnegative random variables, the constraint on α_n is condition ii) in Theorem 1, which restricts α_n to be such that $1 - \alpha_n = \mathrm{O}(n^{-1/3})$, and therefore the rate of convergence of our estimator is capped at $n^{-1/3}$. This stands in contrast of the Hill estimator, whose optimal convergence rate is $n^{\rho/(1-2\rho)}$. On the contrary, when $|\rho|$ is smaller than 1, the estimator $\widehat{\gamma}^{(p)}_{\alpha_n}$ attains this optimal convergence rate of $n^{\rho/(1-2\rho)}$. The value of the estimator will thus intuitively reside in the cases when $|\rho|$ is small, which are known to be difficult cases in extreme

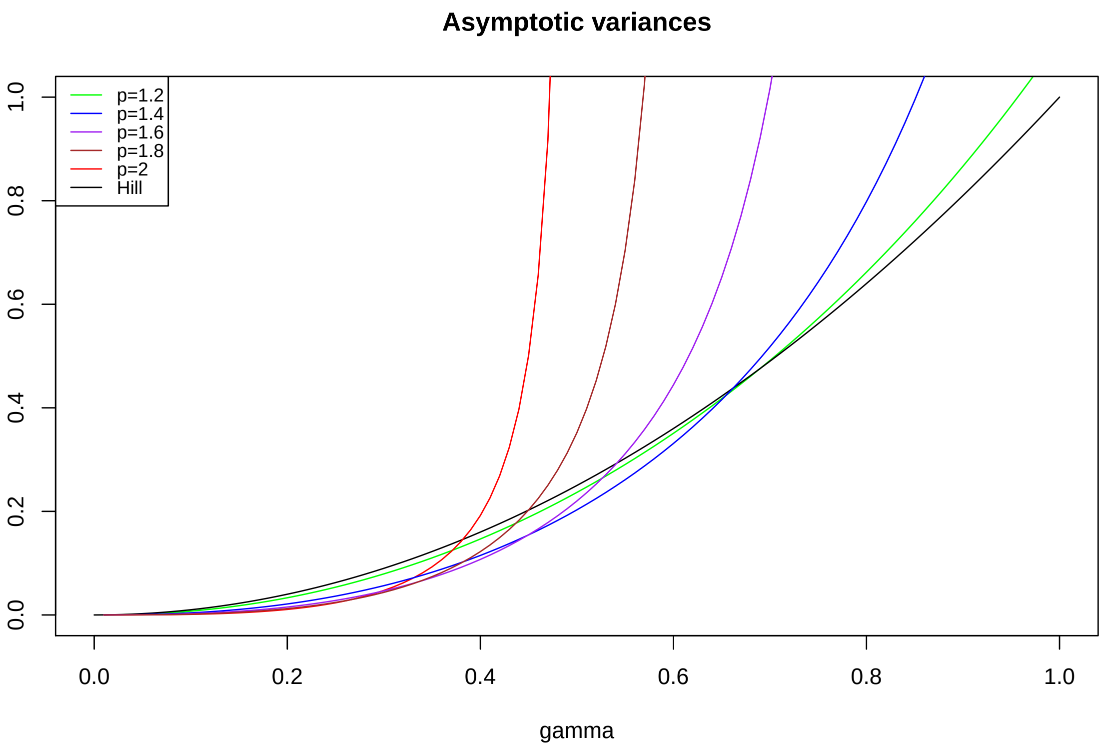

Figure 1: Asymptotic variance $v_p(\gamma)$ for $p = 1.2$ (green), 1.4 (blue), 1.6 (violet), 1.8 (brown) and 2 (red), as a function of $\gamma \in (0, 1)$. The black curve represents $v_H(\gamma)$, the asymptotic variance of the Hill estimator.

value analysis where the Hill estimator tends to perform poorly. A detailed discussion on the choice of p in practice and the performance of our estimator is proposed in Section 4.

The asymptotic bias component involves three parts, all of which tend to make substantial contributions to the total bias of the estimator. The quantity $b_p^{(1)}(\gamma, \rho)$ is proportional to the auxiliary function A, while the quantities $b_p^{(2)}(\gamma)$ and $b_p^{(3)}(\gamma, \gamma_\ell)$ are specific to the use of L^p-quantiles. We now correct the bias term in Theorem 1, in order to obtain the most accurate tail index estimator possible from $\widehat{\gamma}_{\alpha_n}^{(p)}$. One possibility for doing so would be to directly estimate the three bias components, to divide the estimators by $\sqrt{n(1-\alpha_n)}$, and to subtract the obtained quantities from $\widehat{\gamma}_{\alpha_n}^{(p)}$. This does not tend to yield good results in practice, not least because $b_p^{(3)}(\gamma, \gamma_\ell)$ is somewhat difficult to estimate, as it requires the estimation of the left and right tail indices. We suggest here a simpler and more efficient methodology whose justification goes back to the construction of the estimator $\widehat{\gamma}_{\alpha_n}^{(p)}$. According to Equation (3.4),

$$g_p(\gamma) = \frac{\overline{F}^{(1)}(q_\alpha(p))}{1-\alpha} \frac{1 + A\left(1/\overline{F}^{(1)}(q_\alpha(p))\right) K(p,\gamma,\rho)[g_p(\gamma)]^{1+\rho}(1+o(1))}{\mathbb{E}\left[\left|\frac{Y}{q_\alpha(p)} - 1\right|^{p-1}\right]}$$

as $\alpha \to 1$. This suggests to refine the estimator $\widehat{\gamma}_{\alpha_n}^{(p)}$ by considering instead

$$\widetilde{\gamma}_{\alpha_n}^{(p)} = \inf\left\{\gamma > 0\colon g_p(\gamma) \le \frac{\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))}{1-\alpha_n} \frac{1 + \bar{b}\bar{\gamma}\left[\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))\right]^{-\bar{\rho}} K(p,\bar{\gamma},\bar{\rho})[g_p(\bar{\gamma})]^{1+\bar{\rho}}}{n^{-1}\sum_{i=1}^n \left|\frac{Y_i}{\widehat{q}_{\alpha_n}(p)} - 1\right|^{p-1}}\right\}.$$

The asymptotic normality of this estimator can be established as a straightforward corollary of Theorem 1.

Corollary 1. *Assume that the conditions of Theorem 1 hold. If moreover $A(t) = b\gamma t^{\rho}$ and $\overline{\gamma}$, $\overline{\rho}$ and $\overline{b}$ are consistent estimators of γ, ρ and b such that $(\overline{\rho}-\rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$, then*

$$\sqrt{n(1-\alpha_n)}\left(\widetilde{\gamma}_{\alpha_n}^{(p)} - \gamma\right) \xrightarrow{d} \mathcal{N}(0, v_p(\gamma)).$$

3.3 Final class of estimators

Here we define our final versions of the estimators as implemented in our simulation study and real data analysis. To calculate the estimator $\widetilde{\gamma}_{\alpha_n}^{(p)}$, we use the estimators $\overline{b}$ and $\overline{\rho}$ of Gomes and Martins (2002) and Fraga Alves et al. (2003), implemented in the R package `evt0` and particularly the function `mop`. We also have to use an estimator $\overline{\gamma}$, which in this case we take to be a bias-reduced version of the Hill estimator $\widehat{\gamma}_{\alpha_n}^{(H)}$, introduced by Caeiro et al. (2005):

$$\widetilde{\gamma}_{\alpha_n}^{(H)} = \widehat{\gamma}_{\alpha_n}^{(H)}\left(1 - \frac{\overline{b}}{1-\overline{\rho}}(1-\alpha_n)^{-\overline{\rho}}\right).$$

This results in the following, final version of the estimator $\widetilde{\gamma}_{\alpha_n}^{(p)}$:

$$\widetilde{\gamma}_{\alpha_n}^{(p)} = \inf\left\{\gamma > 0\colon g_p(\gamma) \leq \frac{\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))}{1-\alpha_n} \frac{1 + \overline{b}\widetilde{\gamma}_{\alpha_n}^{(H)}\left[\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))\right]^{-\overline{\rho}} K\left(p, \widetilde{\gamma}_{\alpha_n}^{(H)}, \overline{\rho}\right)\left[g_p\left(\widetilde{\gamma}_{\alpha_n}^{(H)}\right)\right]^{1+\overline{\rho}}}{n^{-1}\sum_{i=1}^n \left|\frac{Y_i}{\widehat{q}_{\alpha_n}(p)} - 1\right|^{p-1}}\right\}.$$

Note that this estimator is computed in the R function `lpindex` in the package `Expectrem`. This is a novel estimator which cannot be deduced in a direct way from the earlier work in Daouia et al. (2019).

To define our class of estimators of an extreme quantile $q_{\alpha_n'}(1)$, we plug in our general expression of $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1)$ the estimator $\widetilde{\gamma}_{\alpha_n}^{(p)}$ in place of $\overline{\gamma}$ and we keep our estimators $\overline{b}$ and $\overline{\rho}$ from the R function `mop`. This results in the bias-reduced, L^p-quantile-based extreme quantile estimator

$$\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\widetilde{\gamma}_{\alpha_n}^{(p)}} \widehat{q}_{\alpha_n}(p)\left[g_p\left(\widetilde{\gamma}_{\alpha_n}^{(p)}\right)\right]^{\widetilde{\gamma}_{\alpha_n}^{(p)}} \frac{(1+\overline{r}(p,\alpha_n))^{\widetilde{\gamma}_{\alpha_n}^{(p)}}}{1 + \frac{[g_p(\widetilde{\gamma}_{\alpha_n}^{(p)})]^{-\overline{\rho}}(1+\overline{r}(p,\alpha_n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\widetilde{\gamma}_{\alpha_n}^{(p)}(1-\alpha_n)^{-\overline{\rho}}}$$
$$\times\left(1 + \frac{\left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\rho}} - 1}{\overline{\rho}}\overline{b}\widetilde{\gamma}_{\alpha_n}^{(p)}(1-\alpha_n)^{-\overline{\rho}}\right).$$

This quantile estimator is also available in the R package `Expectrem`, using the function `extQuantlp`. A similar procedure applied to $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(2)$ yields the bias-reduced L^p-quantile-based extreme expectile estimator

$$\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(2) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\widetilde{\gamma}_{\alpha_n}^{(p)}} \widehat{q}_{\alpha_n}(p)\left(\frac{g_p\left(\widetilde{\gamma}_{\alpha_n}^{(p)}\right)}{g_2\left(\widetilde{\gamma}_{\alpha_n}^{(p)}\right)}\right)^{\widetilde{\gamma}_{\alpha_n}^{(p)}} \left(\frac{1+\overline{r}(p,\alpha_n)}{1+\overline{r}(2,\alpha_n')}\right)^{\widetilde{\gamma}_{\alpha_n}^{(p)}}$$
$$\times \frac{1 + \frac{\left[g_2\left(\widetilde{\gamma}_{\alpha_n}^{(p)}\right)\right]^{-\overline{\rho}}(1+\overline{r}(2,\alpha_n'))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\widetilde{\gamma}_{\alpha_n}^{(p)}(1-\alpha_n')^{-\overline{\rho}}}{1 + \frac{\left[g_p\left(\widetilde{\gamma}_{\alpha_n}^{(p)}\right)\right]^{-\overline{\rho}}(1+\overline{r}(p,\alpha_n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\widetilde{\gamma}_{\alpha_n}^{(p)}(1-\alpha_n)^{-\overline{\rho}}} \times \left(1 + \frac{\left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\rho}} - 1}{\overline{\rho}}\overline{b}\widetilde{\gamma}_{\alpha_n}^{(p)}(1-\alpha_n)^{-\overline{\rho}}\right).$$

This estimator can be computed using the function `extExpectlp` in the R package `Expectrem`.

Our final theoretical result shows that these estimators are asymptotically Gaussian under the conditions of Corollary 1.

Theorem 2. *Assume that the conditions of Corollary 1 hold. Assume also that $\rho < 0$, and that $(1-\alpha_n')/(1-\alpha_n) \to 0$ with $\sqrt{n(1-\alpha_n)}/\log[(1-\alpha_n)/(1-\alpha_n')] \to \infty$.*

(i) Then

$$\frac{\sqrt{n(1-\alpha_n)}}{\log[(1-\alpha_n)/(1-\alpha_n')]}\left(\frac{\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p}(1)}{q_{\alpha_n'}(1)}-1\right)\xrightarrow{d}\mathcal{N}(0,v_p(\gamma)).$$

(ii) If moreover $\gamma<1$ *and* $\mathbb{E}(|\min(Y,0)|)<\infty$ *then*

$$\frac{\sqrt{n(1-\alpha_n)}}{\log[(1-\alpha_n)/(1-\alpha_n')]}\left(\frac{\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p}(2)}{q_{\alpha_n'}(2)}-1\right)\xrightarrow{d}\mathcal{N}(0,v_p(\gamma)).$$

In the next section we examine the finite-sample behavior of the estimators $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p}(1)$ and $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p}(2)$ on simulated data, and we discuss the choice of their tuning parameters (the power p and the anchor intermediate level α_n).

4 Simulation study

4.1 Experimental design

We consider the following distributions:

- A Burr distribution with parameters $\gamma>0$ and $\rho<0$, whose survival function is defined by

$$\overline{F}^{(1)}(y)=\left(1+y^{-\rho/\gamma}\right)^{1/\rho},\ y>0.$$

  The second-order condition $\mathcal{C}_2(\gamma,\rho,A)$ holds with $A(y)$ proportional to y^{ρ}.

- The Fréchet distribution with parameter $\gamma>0$, whose survival function is defined by

$$\overline{F}^{(1)}(y)=1-e^{-y^{-1/\gamma}},\ y>0.$$

  The second-order condition $\mathcal{C}_2(\gamma,-1,A)$ holds with $A(y)$ proportional to y^{-1}.

- The Inverse-Gamma distribution with survival function defined by

$$\overline{F}^{(1)}(y)=\int_y^{\infty}\frac{1}{\Gamma(1/\gamma)}t^{-1/\gamma-1}e^{-1/t}dt,\ y>0.$$

  The second-order condition $\mathcal{C}_2(\gamma,-\gamma,A)$ holds with $A(y)$ proportional to $y^{-\gamma}$.

- The absolute value of a Student distribution with $1/\gamma>0$ degrees of freedom, whose survival function is defined by

$$\overline{F}^{(1)}(y)=2\sqrt{\frac{\gamma}{\pi}}\frac{\Gamma\left(\frac{\gamma^{-1}+1}{2}\right)}{\Gamma\left(\frac{\gamma^{-1}}{2}\right)}\int_y^{\infty}\left(1+\gamma t^2\right)^{-\frac{\gamma^{-1}+1}{2}}dt,\ y>0.$$

  The second-order condition $\mathcal{C}_2(\gamma,-2\gamma,A)$ holds with $A(y)$ proportional to $y^{-2\gamma}$.

- The Generalized Pareto Distribution GPD(γ) with shape parameter $\gamma>0$ and unit scale, whose survival function is defined by

$$\overline{F}^{(1)}(y)=(1+\gamma y)^{-1/\gamma},\ y>0.$$

  The second-order condition $\mathcal{C}_2(\gamma,-\gamma,A)$ holds with $A(y)$ proportional to $y^{-\gamma}$.

For each distribution, $N=500$ samples of size $n=1{,}000$ are simulated. Our goal is to compare the finite-sample performance of the proposed estimators $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p}(1)$ and $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p}(2)$ with:

- The classical Weissman-type extrapolated estimators of extreme quantiles (Weissman, 1978) and expectiles (Daouia et al., 2018). These are respectively

$$\widehat{q}^{\star}_{\alpha_n'}(1) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(1) \text{ and } \widehat{q}^{\star}_{\alpha_n'}(2) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(2), \text{ with } \overline{\gamma} = \widetilde{\gamma}^{(H)}_{\alpha_n}.$$

- The composite L^p-quantile-based extreme quantile and expectile estimators originally suggested in Daouia et al. (2019) (and thus not featuring bias reduction). These are $\widetilde{q}^{\star}_{\alpha_n',p}(1)$ and $\widetilde{q}^{\star}_{\alpha_n',p}(2)$, see Section 3.1. We take again $\overline{\gamma} = \widetilde{\gamma}^{(H)}_{\alpha_n}$, and we choose $p = 1.4$ as recommended below Figure 2 in Daouia et al. (2019).

- The bias-reduced extreme quantile estimator of Gomes and Pestana (2007), namely

$$\widehat{q}^{\star,\mathrm{RB}}_{\alpha_n'}(1) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(1) \left(1 + \frac{\left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\rho}} - 1}{\overline{\rho}} \overline{b}\overline{\gamma}(1-\alpha_n)^{-\overline{\rho}}\right) \text{ with } \overline{\gamma} = \widetilde{\gamma}^{(H)}_{\alpha_n}.$$

- The bias-reduced extreme expectile estimator of Girard et al. (2020b). This is

$$\widehat{q}^{\star,\mathrm{RB}}_{\alpha_n'}(2) = \left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\gamma}} \widehat{q}_{\alpha_n}(2) \left(1 + \frac{\left(\frac{1-\alpha_n'}{1-\alpha_n}\right)^{-\overline{\rho}} - 1}{\overline{\rho}} \overline{b}\overline{\gamma}(1-\alpha_n)^{-\overline{\rho}}\right) (1+\overline{B}_n)(1+\overline{B}_n').$$

  Here $\overline{B}_n$ and $\overline{B}_n'$ are bias correction terms defined as

$$1+\overline{B}_n = (1+\overline{R}(\alpha_n))^{\overline{\gamma}} \times \left(1 + \frac{\frac{\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}}{\left(1+\overline{R}(\alpha_n)\right)^{\overline{\rho}}} - 1}{\overline{\rho}} \overline{b}\overline{\gamma}\,(1-\alpha_n)^{-\overline{\rho}}\right)^{-1}$$

$$\text{and } 1+\overline{B}_n' = (1+\overline{R}^{\star}(\alpha_n'))^{-\overline{\gamma}} \times \left(1 + \frac{\frac{\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}}{(1+\overline{R}^{\star}(\alpha_n'))^{\overline{\rho}}} - 1}{\overline{\rho}} \overline{b}\overline{\gamma}(1-\alpha_n')^{-\overline{\rho}}\right)$$

  where the quantities $\overline{R}(\alpha_n)$ and $\overline{R}^{\star}(\alpha_n')$ are defined as

$$1+\overline{R}(\alpha_n) = \left(1 - \frac{n^{-1}\sum_{i=1}^n Y_i}{\widehat{q}_{\alpha_n}(2)}\right) \frac{1}{2\alpha_n - 1} \times \left(1 + \frac{\overline{b}[\widehat{\overline{F}}^{(1)}_n(\widehat{q}_{\alpha_n}(2))]^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}\right)^{-1}$$

$$\text{and } 1+\overline{R}^{\star}(\alpha_n') = \left(1 - \frac{n^{-1}\sum_{i=1}^n Y_i}{\widehat{q}^{\star}_{\alpha_n'}(2)}\right) \frac{1}{2\alpha_n' - 1} \times \left(1 + \frac{\overline{b}\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}(1-\alpha_n')^{-\overline{\rho}}\right)^{-1}.$$

  We take $\overline{\gamma} = \widetilde{\gamma}^{(H)}_{\alpha_n}$.

For all the considered estimators and throughout our finite-sample experiments, as in Section 3.3, the estimators $\overline{b}$ and $\overline{\rho}$ are those of Gomes and Martins (2002) and Fraga Alves et al. (2003), implemented in the function `mop` of the R package `evt0`.

Each of the considered estimators requires to select an intermediate sequence (α_n), and our proposed estimators require a choice of p. We propose next a technique based on Asymptotic Mean Squared Error minimization.

4.2 Choices of p and α_n

The choice of the sequence α_n is crucial because of its consequences on the tail index estimator featuring in the extrapolation procedure: taking α_n too close to 1 translates into a large variance, while taking α_n too far from 1 translates into a large bias. This choice therefore leads to solving a trade-off between the bias and variance of the tail index estimator to be used. One possibility is to balance bias and variance by calculating an Asymptotic Mean Squared Error. For the Hill estimator specifically, such a selection rule is discussed in Section 3.2 in de Haan and Ferreira (2006): the idea is to note that under the second-order condition $\mathcal{C}_2(\gamma, \rho, A)$ and the assumption $\sqrt{n(1-\alpha_n)}A((1-\alpha_n)^{-1}) \to \lambda \in \mathbb{R}$,

$$\sqrt{n(1-\alpha_n)}(\widehat{\gamma}_{\alpha_n}^{(H)} - \gamma) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \gamma^2\right).$$

Minimizing the Asymptotic Mean Squared Error is therefore tantamount to minimizing the quantity

$$\left[\frac{1}{1-\rho} A\left((1-\alpha_n)^{-1}\right)\right]^2 + \frac{\gamma^2}{n(1-\alpha_n)}$$

with respect to α_n. With $A(t) = b\gamma t^\rho$, this yields an optimal value $\alpha_n^{(H)}$ defined by

$$1 - \alpha_n^{(H)} = \left(\frac{(1-\rho)^2}{-2\rho b^2}\right)^{\frac{1}{1-2\rho}} n^{-\frac{1}{1-2\rho}}.$$

For the L^p-quantile-based tail index estimator, the discussion is more involved because, following Theorem 1, there are three sources of bias involved:

- The first one is proportional to $A(n/k_n)$ and represented by the quantity $\lambda_1 \mathfrak{b}_p^{(1)}(\gamma, \rho)$, where

$$\mathfrak{b}_p^{(1)}(\gamma, \rho) = \frac{-\gamma}{1 + \frac{1}{\gamma}\left(\Psi\left(\gamma^{-1} - p + 1\right) - \Psi\left(\gamma^{-1} + 1\right)\right)} b_p^{(1)}(\gamma, \rho).$$

- The second one is proportional to $\frac{\mathbb{E}\left[Y \mathbb{1}_{\{0<Y<q_{\alpha_n}(1)\}}\right]}{q_{\alpha_n}(1)} \vee (1-\alpha_n)$ and represented by the quantity $\lambda_2 \mathfrak{b}_p^{(2)}(\gamma)$, where

$$\mathfrak{b}_p^{(2)}(\gamma) = \frac{-\gamma}{1 + \frac{1}{\gamma}\left(\Psi\left(\gamma^{-1} - p + 1\right) - \Psi\left(\gamma^{-1} + 1\right)\right)} b_p^{(2)}(\gamma).$$

- The third one is proportional to $\frac{\mathbb{E}\left[-Y \mathbb{1}_{\{0<-Y<q_{\alpha_n}(1)\}}\right]}{q_{\alpha_n}(1)} \vee F^{(1)}(-q_{\alpha_n}(1))$ and represented by the quantity $\lambda_3 \mathfrak{b}_p^{(3)}(\gamma, \gamma_\ell)$, where

$$\mathfrak{b}_p^{(3)}(\gamma, \gamma_\ell) = \frac{-\gamma}{1 + \frac{1}{\gamma}\left(\Psi\left(\gamma^{-1} - p + 1\right) - \Psi\left(\gamma^{-1} + 1\right)\right)} b_p^{(3)}(\gamma, \gamma_\ell).$$

Intuitively, the second and third sources of bias should be easiest to correct, because they involve quantities that can be estimated at the rate $\sqrt{n(1-\alpha_n)}$ or more. In practice, this means that the trade-off to be solved when using the L^p-quantile-based tail index estimator will essentially involve this first source of bias. Hence the idea of minimizing the *Partial Asymptotic Mean Squared Error*

$$\text{PAMSE}_p(\alpha_n) = \left[\mathfrak{b}_p^{(1)}(\gamma, \rho) A\left((1-\alpha_n)^{-1}\right)\right]^2 + \frac{v_p(\gamma)}{n(1-\alpha_n)}.$$

When $A(t) = b\gamma t^\rho$, minimizing this quantity with respect to $1-\alpha_n$ yields an optimal $\alpha_n^{(p)}$ satisfying

$$1 - \alpha_n^{(p)} = \left(\frac{v_p(\gamma)}{-2\rho b^2 \gamma^2 [\mathfrak{b}_p^{(1)}(\gamma, \rho)]^2}\right)^{\frac{1}{1-2\rho}} n^{-\frac{1}{1-2\rho}}.$$

Estimators $\alpha_n^{\star,(H)}$ and $\alpha_n^{\star,(p)}$ of $\alpha_n^{(H)}$ and $\alpha_n^{(p)}$ are then readily obtained by plugging in estimators of γ, b, ρ: for the estimation of γ in these data-driven choices of α_n, we use the bias-reduced Hill estimator $\widetilde{\gamma}_{1-50/n}^{(H)}$ (here $1-50/n = 0.95$ for our sample size $n = 1{,}000$).

An additional quantity to be chosen when using the tail index estimator $\widetilde{\gamma}_{\alpha_n}^{(p)}$ is the tuning parameter p. To this end we note that the optimal value of $\mathrm{PAMSE}_p(\alpha_n)$ at $\alpha_n = \alpha_n^{(p)}$ is

$$\mathrm{PAMSE}_p(\alpha_n^{(p)}) = \left([v_p(\gamma)]^{-\rho}\mathfrak{b}_p^{(1)}(\gamma,\rho)\right)^{\frac{2}{1-2\rho}}(b\gamma)^{\frac{2}{1-2\rho}}(1-2\rho)(-2\rho)^{\frac{2\rho}{1-2\rho}}n^{\frac{2\rho}{1-2\rho}}.$$

It is therefore reasonable to choose p which minimizes this quantity and to plug in estimates of γ and ρ, leading us to consider the optimal value

$$p^\star = \underset{p>1}{\arg\min}\,[v_p(\overline{\gamma})]^{-\overline{\rho}}\mathfrak{b}_p^{(1)}(\overline{\gamma},\overline{\rho}).$$

Here once again we set $\overline{\gamma} = \widetilde{\gamma}_{1-50/n}^{(H)}$. This choice of p is represented in Figure 2 as a function of γ and ρ. It is clearly seen that the optimal value of p decreases as γ increases, suggesting that more robust L^p-quantile-based estimates should be used when the tail gets heavier. Interestingly it also seems that the selected value of p should be lower as ρ gets away from 0, i.e. when the underlying extremes get closer to the extremes of the Pareto distribution. The selected value of p appears to be a convex function of γ and ρ.

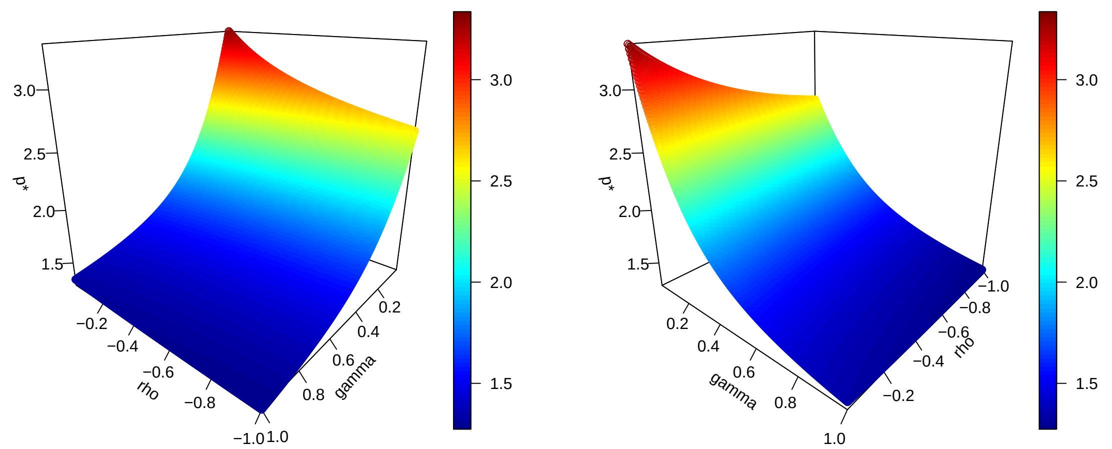

Figure 2: Optimal value of p minimizing $\mathrm{PAMSE}_p(\alpha_n^{(p)})$ as a function of ρ and γ.

4.3 Results

We display in Figures 3 and 4 boxplots of the estimators $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1)$ and $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(2)$, with $\alpha_n = \alpha_n^{(p)}$ and the data-driven choice $p = p^\star$ (some information about the chosen $p^\star$ is reported in Table 1), alongside boxplots of their competitors with $\alpha_n = \alpha_n^{(H)}$. In Figure 5, we also provide a comparison between the tail index estimators $\widetilde{\gamma}_{\alpha_n}^{(p)}$ and $\widetilde{\gamma}_{\alpha_n}^{(H)}$.

The proposed estimators seem to perform well overall. They are less biased than the alternative estimators tested here, particularly when $\gamma \leq 0.5$ and in the difficult cases of the Inverse Gamma, Student and GPD distributions, where the second-order parameter ρ is close to 0. The improvement is particularly obvious with respect to the Weissman estimators, and there is also a visible improvement with respect to the bias-reduced versions of these Weissman estimators. The tail index estimator $\widetilde{\gamma}_{\alpha_n}^{(p)}$ appears to be less biased than $\widetilde{\gamma}_{\alpha_n}^{(H)}$ for $\gamma \in \{0.1, 0.3\}$, and comparable when $\gamma \in \{0.5, 0.7\}$.

Model	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0.7$
(A) Burr, $\rho = -1$	2.33 (0.072)	1.85 (0.086)	1.58 (0.077)	1.42 (0.064)
(B) Fréchet	2.23 (0.086)	1.80 (0.090)	1.54 (0.079)	1.40 (0.066)
(C) Burr, $\rho = -0.5$	2.31 (0.049)	1.79 (0.075)	1.52 (0.065)	1.38 (0.053)
(D) Inverse Gamma	2.15 (0.061)	1.75 (0.077)	1.54 (0.075)	1.40 (0.063)
(E) Student	2.07 (0.062)	1.81 (0.079)	1.59 (0.073)	1.43 (0.065)
(F) GPD	1.86 (0.068)	1.68 (0.070)	1.52 (0.065)	1.41 (0.058)

Table 1: Average chosen $p^\star$ using the data-driven selection rule in 500 simulated datasets of size 1,000 (standard deviations reported between brackets). In each case, the average selected $p^\star$ is written in green if the observed MSE of the bias-reduced L^p-quantile-based tail index estimator $\widetilde{\gamma}_{\alpha_n}^{(p)}$ (for $\alpha_n = \alpha_n^{(p)}$ and $p = p^\star$) is lower than the observed MSE of the bias-reduced Hill estimator $\widetilde{\gamma}_{\alpha_n}^{(H)}$ (for $\alpha_n = \alpha_n^{(H)}$), in blue if this is not the case but if it is lower than the observed MSE of the Hill estimator $\widehat{\gamma}_{\alpha_n}^{(H)}$ (for $\alpha_n = \alpha_n^{(H)}$), and in red otherwise.

5 Real data analysis

5.1 Tornado losses data set

We consider a data set on monetary losses consecutive to tornadoes in the USA in 2018. The data, available at `https://www.spc.noaa.gov/wcm/#data`[1], contains in particular the variables `loss` (total financial losses in USD), `len` and `wid` (respectively length and width of the area affected by the tornado). We propose to consider the loss per square yard

$$Y = \frac{\texttt{loss}}{\texttt{len} \times \texttt{wid}}.$$

This leads to a data set $Y_1, \ldots, Y_n$ of $n = 578$ nonnegative losses per unit of surface. A histogram of the recorded log-values is proposed in the top left panel of Figure 6. The heavy-tailed behavior of the data is assessed through a Generalized Pareto quantile-quantile plot in the top right panel of Figure 6, inspired by the discussion in pp. 90–91 of de Haan and Ferreira (2006): it can be seen in the Figure that the excesses $Y_{n-i+1,n} - Y_{n-k,n}$, for $1 \leq i \leq k-1$ and $k = 100$, are approximately linearly related to quantiles of a heavy-tailed Generalized Pareto distribution with tail index around 0.8, obtained via maximum likelihood fitting with the `gpd` function in the R package `evir`. This is a characteristic of a heavy tail.

We represent in the bottom panels of Figure 6 the bias-reduced Hill estimator $\widetilde{\gamma}_{1-k/n}^{(H)}$ and the bias-reduced L^p-quantile-based estimator $\widetilde{\gamma}_{1-k/n}^{(p)}$ as a function of k, with a selected $p = p^\star = 1.29$ (left panel). The estimated value of γ lies around 1, so that the existence of expectiles of Y is unclear. We therefore focus on quantile estimation and we represent the four quantile estimators $\widehat{q}_{\alpha_n'}^{\star}(1)$, $\widehat{q}_{\alpha_n'}^{\star,\mathrm{RB}}(1)$, $\widetilde{q}_{\alpha_n',p}^{\star}(1)$ and $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1)$ for $\alpha_n' = 0.995 \approx 1 - 3/n$. In the bias-reduced versions of our composite estimators, we take again $p = p^\star = 1.29$. Our composite quantile estimators (non-bias-reduced and bias-reduced) largely agree in this context of a large value of γ (in line with the results obtained in the simulation study of Section 4), and they seem a bit more stable than the classical Weissman estimators. The bias-reduced estimated values of the quantile (at their respective optimal levels chosen as in Section 4.2) are $\widehat{q}_{0.995}^{\star,\mathrm{RB}}(1) = 73.49$ and $\widetilde{q}_{0.995,p^\star}^{\star,\mathrm{RB}}(1) = 53.97$. In this example, our proposed bias-reduced composite version therefore translates into a substantially lower risk estimate than the bias-reduced Weissman approach.

5.2 Medical claims data set

We next study the SOA Group Medical Insurance Large Claims Database, available for instance from the R package `ReIns`. This data set, made of $n = 75{,}789$ claims exceeding \$25,000, has been considered

[1] See the file `https://www.spc.noaa.gov/wcm//data/2018_torn.csv`

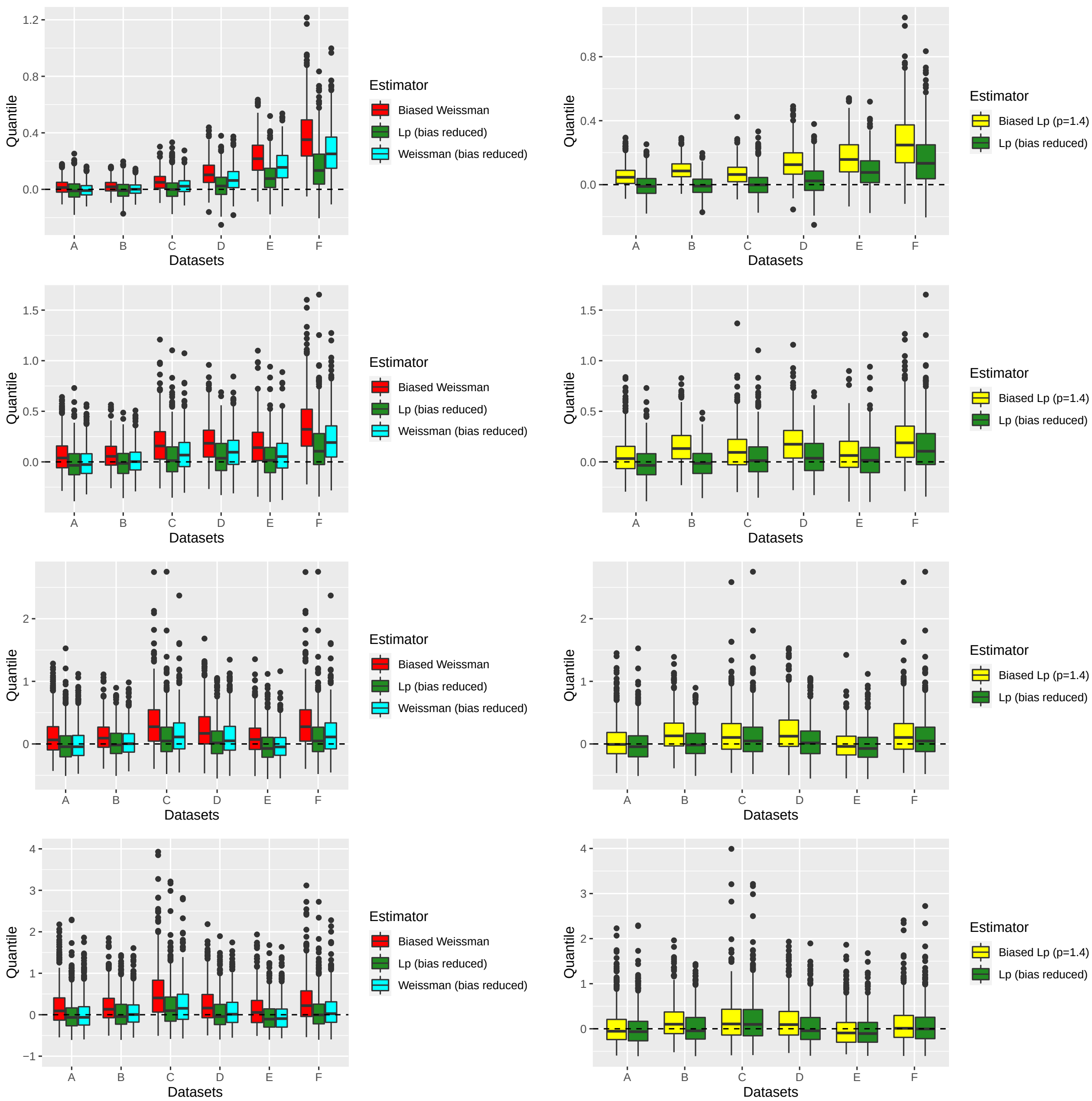

Figure 3: Extreme quantile estimation at level $\alpha'_n = 1 - 1/n = 0.999$ for $n = 1{,}000$. Left panels: boxplots of $\widehat{q}^{\star}_{\alpha'_n}(1)/q_{\alpha'_n}(1) - 1$ (red boxes), $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p^\star}(1)/q_{\alpha'_n}(1) - 1$ (green boxes) and $\widehat{q}^{\star,\mathrm{RB}}_{\alpha'_n}(1)/q_{\alpha'_n}(1) - 1$ (blue boxes). Right panels: boxplots of $\widetilde{q}^{\star}_{\alpha'_n,1.4}(1)/q_{\alpha'_n}(1) - 1$ (yellow boxes) and $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha'_n,p^\star}(1)/q_{\alpha'_n}(1) - 1$ (green boxes). From top to bottom: $\gamma = 0.1, 0.3, 0.5$ and 0.7. Model A is the Burr distribution with $\rho = -1$, model B is the Fréchet distribution, model C is the Burr distribution with $\rho = -0.5$, model D is the Inverse Gamma distribution, model E is the Student distribution and model F is the Generalized Pareto distribution.

a number of times in the extreme value literature, including in the monograph by Beirlant et al. (2004) where the heavy-tailed character of the data is discussed, and in Daouia et al. (2018) in the context of extreme expectile estimation. An overview of the data is provided through a histogram in the top left panel of Figure 7. The literature has found that the heavy tail assumption is reasonable for this data set, with an estimated tail index of around 1/3 (see for instance Beirlant et al., 2004, p.123); this is essentially what we find using the estimators $\widetilde{\gamma}^{(H)}_{1-k/n}$ and $\widetilde{\gamma}^{(p)}_{1-k/n}$ (with $p = p^\star = 1.95$), see the top right panel in Figure 7. We therefore propose here to estimate both extreme quantiles and expectiles, at the level $1 - 1/100{,}000 = 0.99999 > 1 - 1/n$. The bottom two panels of Figure 7

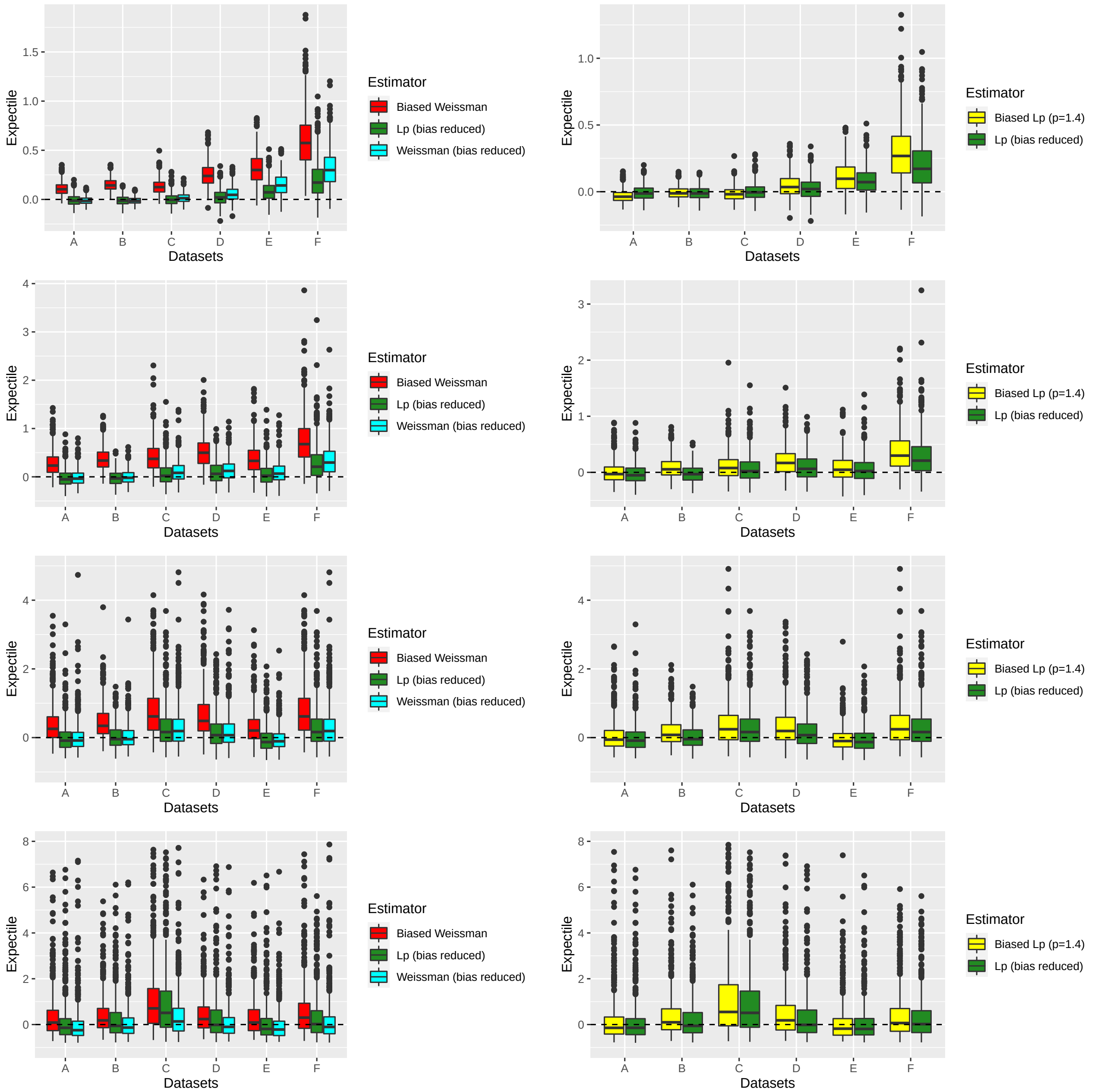

Figure 4: Extreme expectile estimation at level $\alpha_n' = 1 - 1/n = 0.999$ for $n = 1{,}000$. Left panels: boxplots of $\widehat{q}^{\star}_{\alpha_n'}(2)/q_{\alpha_n'}(2) - 1$ (red boxes), $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p^\star}(2)/q_{\alpha_n'}(2) - 1$ (green boxes) and $\widehat{q}^{\star,\mathrm{RB}}_{\alpha_n'}(2)/q_{\alpha_n'}(2) - 1$ (blue boxes). Right panels: boxplots of $\widetilde{q}^{\star}_{\alpha_n',1.4}(2)/q_{\alpha_n'}(2) - 1$ (yellow boxes) and $\widetilde{q}^{\star,\mathrm{RB}}_{\alpha_n',p^\star}(2)/q_{\alpha_n'}(2) - 1$ (green boxes). From top to bottom: $\gamma = 0.1, 0.3, 0.5$ and 0.7. Model A is the Burr distribution with $\rho = -1$, model B is the Fréchet distribution, model C is the Burr distribution with $\rho = -0.5$, model D is the Inverse Gamma distribution, model E is the Student distribution and model F is the Generalized Pareto distribution.

provide graphical representations of the different considered estimators. In this situation of a tail index around $1/3$, the bias correction in the composite estimators has a substantial effect. The bias-corrected composite estimator (again with a chosen $p = p^\star = 1.95$) and its extrapolated competitor essentially agree both on the quantile and expectile estimation problem, although the composite versions seem to display slightly more stable sample paths as a function of k. In detail, using the selection rules of α_n and p introduced in Section 4.2, we find estimated extreme quantiles (resp. expectiles) as $\widehat{q}^{\star,\mathrm{RB}}_{0.99999}(1) =$ 3,544,379 (resp. $\widehat{q}^{\star,\mathrm{RB}}_{0.99999}(2) = 2{,}856{,}904$) and $\widetilde{q}^{\star,\mathrm{RB}}_{0.99999,p^\star}(1) = 3{,}888{,}743$ (resp. $\widetilde{q}^{\star,\mathrm{RB}}_{0.99999,p^\star}(2) = 3{,}142{,}720$). The composite estimators here yield a slightly more conservative assessment of risk than their quantile

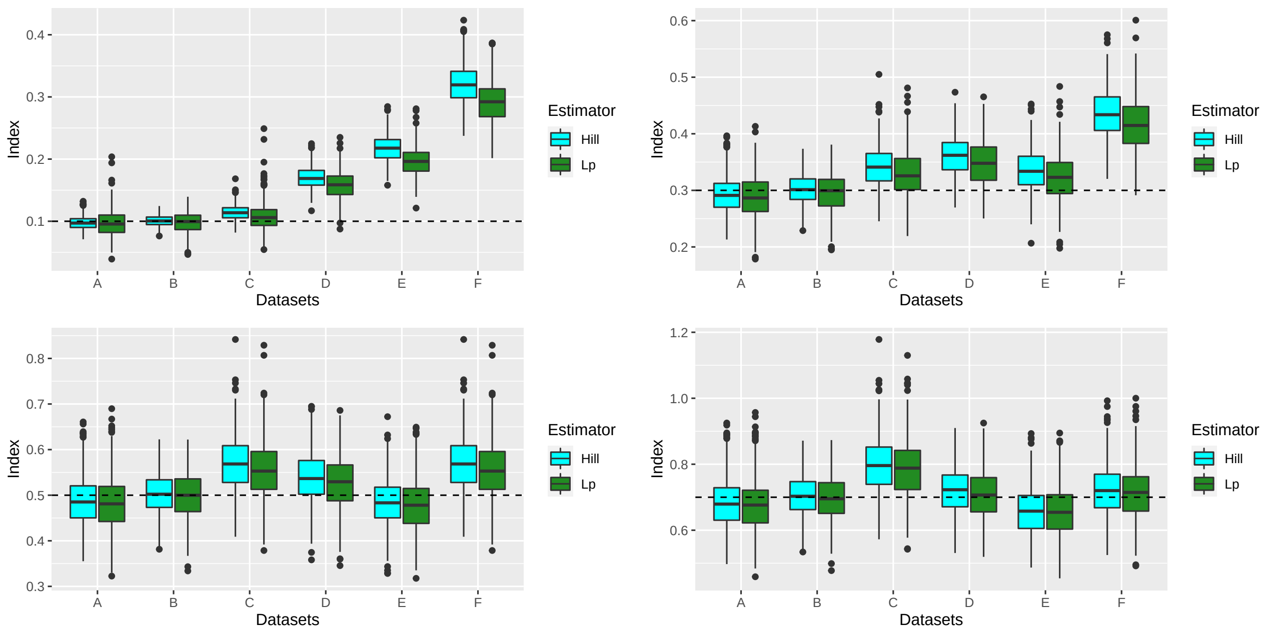

Figure 5: Boxplots of the bias-reduced tail index estimators $\widetilde{\gamma}_{\alpha_n}^{(H)}$ (blue) and $\widetilde{\gamma}_{\alpha_n}^{(p)}$ (green) for each simulated dataset. From left to right and then top to bottom: $\gamma = 0.1$, 0.3, 0.5 and 0.7. Model A is the Burr distribution with $\rho = -1$, model B is the Fréchet distribution, model C is the Burr distribution with $\rho = -0.5$, model D is the Inverse Gamma distribution, model E is the Student distribution and model F is the Generalized Pareto distribution.

or expectile-based counterparts.

Acknowledgements

This research is supported by the French National Research Agency under the grants ANR-17-EURE-0010 and ANR-19-CE40-0013 (ExtremReg project). G. Stupfler also acknowledges support from an AXA Research Fund Award on "Mitigating risk in the wake of the COVID-19 pandemic".

References

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions (10th printing)*. U.S. National Bureau of Standards.

Alm, J. (2016). Signs of dependence and heavy tails in non-life insurance data. *Scandinavian Actuarial Journal*, 2016(10):859–875.

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

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

Bellini, F., Klar, B., Müller, A., and Gianin, E. R. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54:41–48.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular Variation*. Cambridge University Press.

Caeiro, F., Gomes, M. I., and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat*, 3(2):113–136.

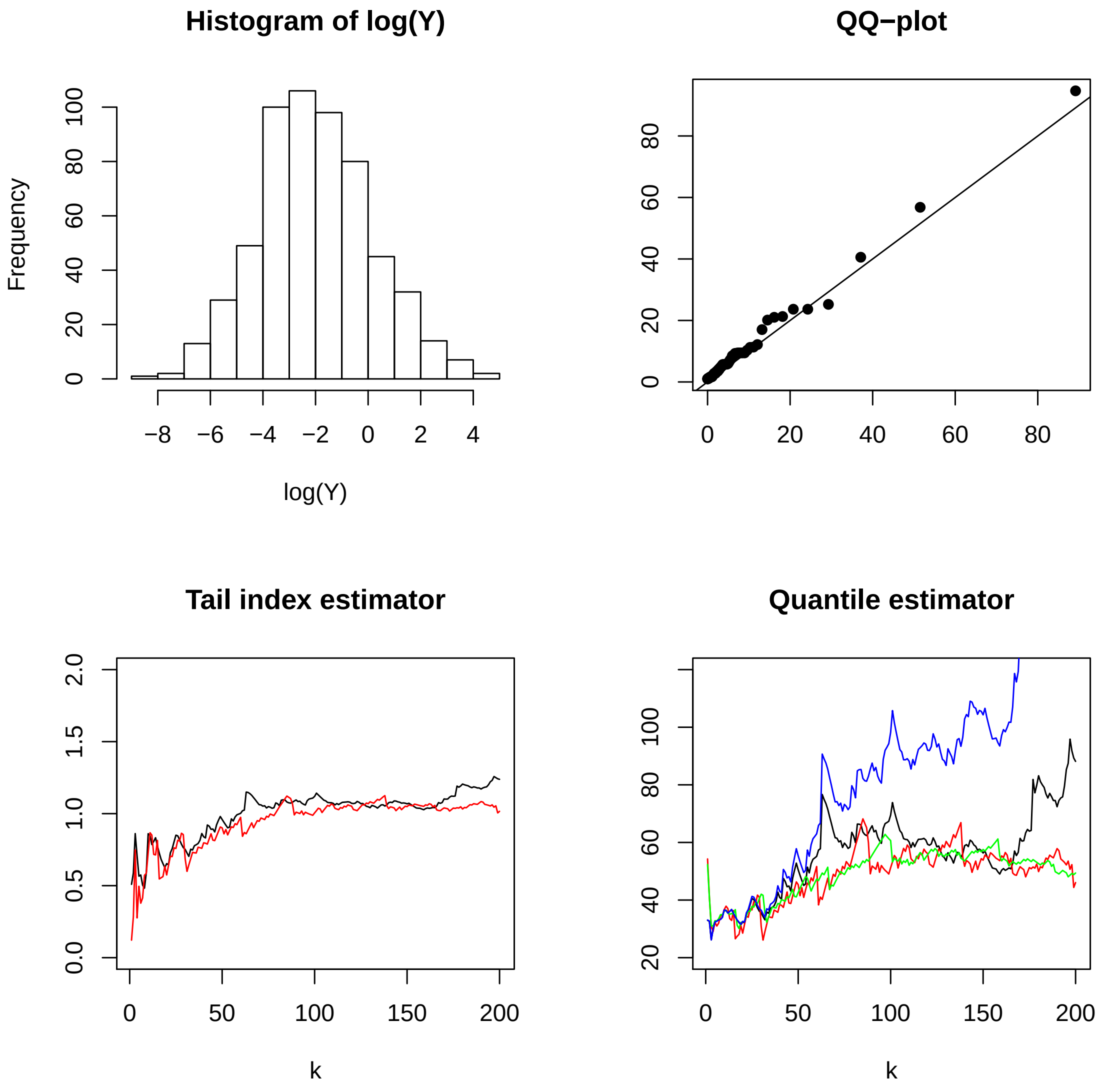

Figure 6: Tornado losses data analysis. Top left: Histogram of the log-claims. Top right: Generalized Pareto quantile-quantile plot of the $Y_{n-i+1,n} - Y_{n-k,n}$, for $1 \le i \le k-1$, with $k = 100$ (scale and shape parameter estimates obtained via the `gpd` function in the R package `evir`; the shape parameter estimate is around 0.8). Bottom left: Tail index estimators $\widetilde{\gamma}^{(H)}_{1-k/n}$ (black) and $\widetilde{\gamma}^{(p^\star)}_{1-k/n}$ (red) as functions of k. Bottom right: 99.5 % quantile estimators of losses (in dollars) per square yard due to tornadoes, estimated with $\widehat{q}^{\star}_{0.995}(1)$ (blue), $\widetilde{q}^{\star,\mathrm{RB}}_{0.995}(1)$ (black), $\widetilde{q}^{\star}_{0.995,1.4}(1)$ (green) and $\widetilde{q}^{\star,\mathrm{RB}}_{0.995,p^\star}(1)$ (red) as functions of k, where the anchor level α_n is set as $1 - k/n$. Throughout, $p^\star = 1.29$.

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

Cai, J. and Weng, C. (2016). Optimal reinsurance with expectile. *Scandinavian Actuarial Journal*, 2016(7):624–645.

Chavez-Demoulin, V., Embrechts, P., and Sardy, S. (2014). Extreme-quantile tracking for financial time series. *Journal of Econometrics*, 181(1):44–52.

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(2):263–292.

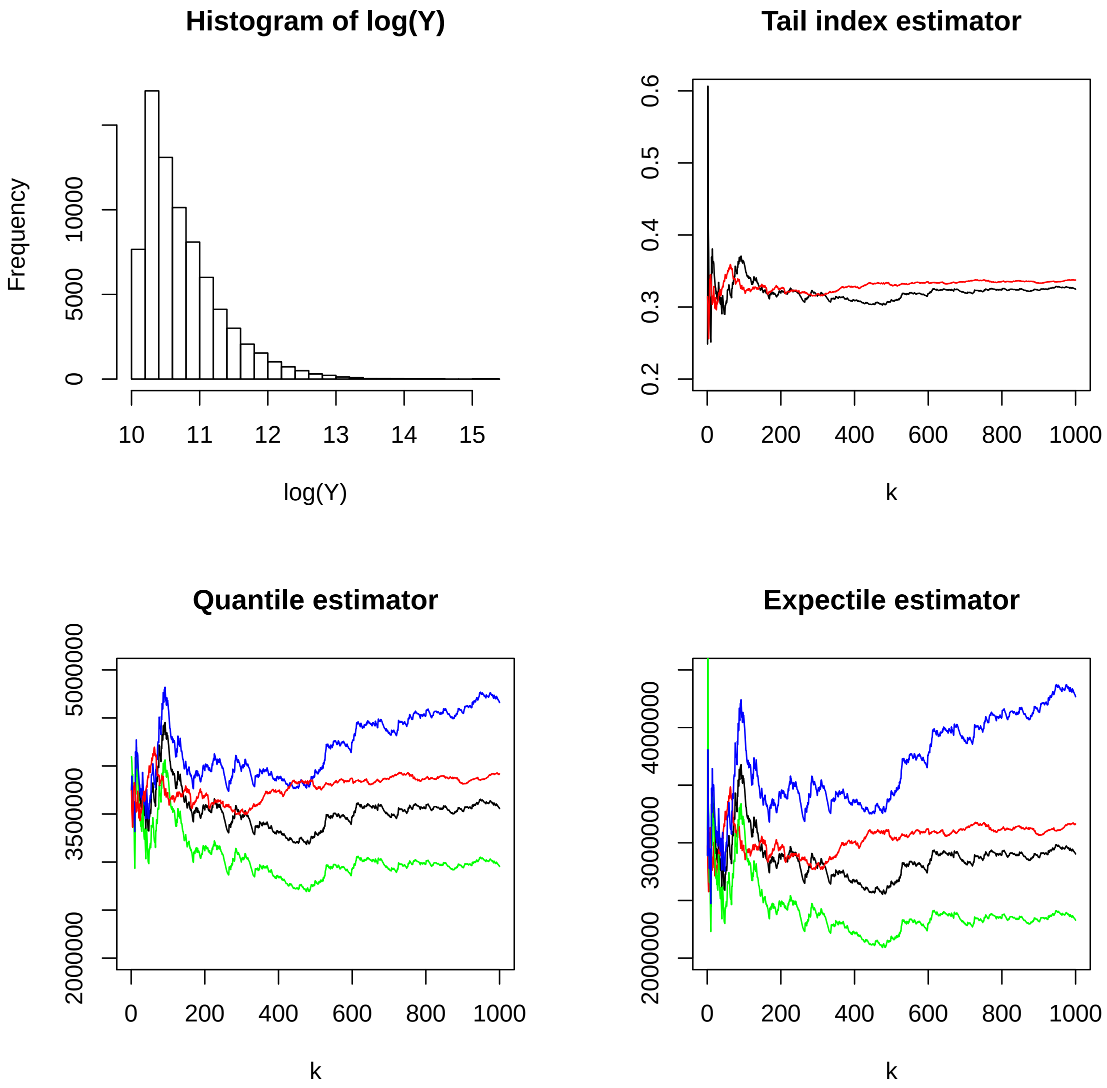

Figure 7: SOA Group Medical Insurance Large Claims database analysis. Top left: Histogram of the log-claims. Top right: Tail index estimators $\widetilde{\gamma}^{(H)}_{1-k/n}$ (black) and $\widetilde{\gamma}^{(p^\star)}_{1-k/n}$ (red) as functions of k. Bottom left: Extreme quantile estimators $\widehat{q}^{\star}_{0.99999}(1)$ (blue), $\widetilde{q}^{\star,\mathrm{RB}}_{0.99999}(1)$ (black), $\widetilde{q}^{\star}_{0.99999,1.4}(1)$ (green) and $\widetilde{q}^{\star,\mathrm{RB}}_{0.99999,p^\star}(1)$ (red) as functions of k, where the anchor level α_n is set as $1-k/n$. Bottom right: Extreme expectile estimators $\widehat{q}^{\star}_{0.99999}(2)$ (blue), $\widetilde{q}^{\star,\mathrm{RB}}_{0.99999}(2)$ (black), $\widetilde{q}^{\star}_{0.99999,1.4}(2)$ (green) and $\widetilde{q}^{\star,\mathrm{RB}}_{0.99999,p^\star}(2)$ (red) as functions of k, where the anchor level α_n is set as $1-k/n$. Throughout, $p^\star = 1.95$.

Daouia, A., Girard, S., and Stupfler, G. (2019). Extreme M-quantiles as risk measures: from L^1 to L^p optimization. *Bernoulli*, 25(1):264–309.

Daouia, A., Girard, S., and Stupfler, G. (2020). Tail expectile process and risk assessment. *Bernoulli*, 26(1):531–556.

Daouia, A., Girard, S., and Stupfler, G. (2021). ExpectHill estimation, extreme risk and heavy tails. *Journal of Econometrics*, 221(1):97–117.

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

Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, 9(1):617–657.

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

Fraga Alves, M. I., Gomes, M. I., and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60(2):193–214.

Girard, S., Stupfler, G., and Usseglio-Carleve, A. (2020a). Nonparametric extreme conditional expectile estimation, *Scandinavian Journal of Statistics*, to appear. `https://hal.archives-ouvertes.fr/hal-02114255`.

Girard, S., Stupfler, G., and Usseglio-Carleve, A. (2020b). On automatic bias reduction for extreme expectile estimation. `https://hal.archives-ouvertes.fr/hal-03086048`.

Gomes, M. I. and Martins, M. J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes*, 5(1):5–31.

Gomes, M. I. and Pestana, D. (2007). A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, 102(477):280–292.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.

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

Jones, M. (1994). Expectiles and M-quantiles are quantiles. *Statistics & Probability Letters*, 20:149–153.

Koenker, R. and Bassett, G. J. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Newey, W. and Powell, J. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55(4):819–847.

Padoan, S. and Stupfler, G. (2020). Joint inference on extreme expectiles for multivariate heavy-tailed distributions. `https://arxiv.org/abs/2007.08944`.

Rootzén, H. and Tajvidi, N. (1997). Extreme value statistics and wind storm losses: a case study. *Scandinavian Actuarial Journal*, 1997(1):70–94.

Stupfler, G. (2019). On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails. *Extremes*, 22(4):749–769.

Taylor, J. W. (2008). Estimating Value at Risk and Expected Shortfall Using Expectiles. *Journal of Financial Econometrics*, 6(2):231–252.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815.

Supplementary Material to "Composite bias-reduced L^p−quantile-based estimators of extreme quantiles and expectiles"

Gilles Stupfler and Antoine Usseglio-Carleve

Appendix A Preliminary results

The first result states that the function $\overline{F}^{(p)}$ defined in Equation (2.1) is continuous and strictly decreasing on the range of values of Y, and is therefore in particular a survival function, as soon as $\overline{F}^{(1)}$ is continuous (which we assume throughout). A general result for M-quantiles has been proved in Jones (1994) under conditions whose validity for L^p−quantiles is not completely clear.

Lemma A.1. *Denote the lower and upper endpoints of Y by $y_* = \inf\left\{y \in \mathbb{R} \colon \overline{F}^{(1)}(y) < 1\right\}$ and $y^* = \sup\left\{y \in \mathbb{R} \colon \overline{F}^{(1)}(y) > 0\right\}$. Then for all $p > 1$ such that $\mathbb{E}\left[|Y|^{p-1}\right] < +\infty$, $\overline{F}^{(p)} : (y_*, y^*) \to (0,1)$ is continuous and strictly decreasing.*

Proof. For all $y \in (y_*, y^*)$, the integration by parts formula easily leads to

$$\frac{1}{\overline{F}^{(p)}(y)} = 1 + \frac{\mathbb{E}\left[(y-Y)^{p-1}\mathbb{1}_{\{Y \le y\}}\right]}{\mathbb{E}\left[(Y-y)^{p-1}\mathbb{1}_{\{Y > y\}}\right]} = 1 + \frac{\int_0^{+\infty} t^{p-2} F^{(1)}(y-t)dt}{\int_0^{+\infty} t^{p-2}\overline{F}^{(1)}(y+t)dt}.$$

Then for all $h \in (0, y^* - y)$, $\overline{F}^{(p)}(y+h)^{-1} - \overline{F}^{(p)}(y)^{-1} =$

$$\frac{\int_{(0,+\infty)^2} s^{p-2} t^{p-2}\left\{F^{(1)}(y+h-t)\overline{F}^{(1)}(y+s) - F^{(1)}(y-t)\overline{F}^{(1)}(y+h+s)\right\} ds\, dt}{\int_0^{+\infty} t^{p-2}\overline{F}^{(1)}(y+h+t)dt \int_0^{+\infty} t^{p-2}\overline{F}^{(1)}(y+t)dt}.$$

Notice that

$$\begin{aligned} & F^{(1)}(y+h-t)\overline{F}^{(1)}(y+s) - F^{(1)}(y-t)\overline{F}^{(1)}(y+h+s) \\ = \; & \overline{F}^{(1)}(y+s)\left\{F^{(1)}(y+h-t) - F^{(1)}(y-t)\right\} + F^{(1)}(y-t)\left\{\overline{F}^{(1)}(y+s) - \overline{F}^{(1)}(y+h+s)\right\} \ge 0. \end{aligned}$$

Let us suppose now that there exist $y \in (y_*, y^*)$ and $h \in (0, y^* - y)$ such that $\overline{F}^{(p)}(y+h) = \overline{F}^{(p)}(y)$. In that case, by continuity of $\overline{F}^{(1)}$, we necessarily have

$$\overline{F}^{(1)}(y+s)\left\{F^{(1)}(y+h-t) - F^{(1)}(y-t)\right\} = 0 \text{ for all } s,t > 0.$$

Since $y \in (y_*, y^*)$, $\overline{F}^{(1)}(y+s) > 0$ for s small enough, hence

$$F^{(1)}(y+h-t) = F^{(1)}(y-t) \text{ for all } t > 0.$$

The contradiction $F^{(1)}(y) = 0$ is then readily obtained by iterating the above relationship for $t = h$:

$$F^{(1)}(y) = F^{(1)}(y+h-h) = F^{(1)}(y-h) = \cdots = F^{(1)}(y-Nh) \to 0 \;\text{ as } N \to \infty.$$

Hence $\overline{F}^{(p)}(y+h)^{-1} - \overline{F}^{(p)}(y)^{-1} > 0$, proving that $\overline{F}^{(p)}$ is strictly decreasing on (y_*, y^*). Continuity of this mapping can be shown by proving the continuity of both

$$y \mapsto \mathbb{E}\left[(y-Y)^{p-1}\mathbb{1}_{\{Y \le y\}}\right] \text{ and } y \mapsto \mathbb{E}\left[(Y-y)^{p-1}\mathbb{1}_{\{Y > y\}}\right].$$

This directly follows from the dominated convergence theorem. □

Our second lemma establishes that the function g_p defined in (2.3) is indeed monotonic.

Lemma A.2. *For all $p > 1$, $g_p(\cdot)$ is strictly decreasing on $(0, 1/(p-1))$ and has derivative*

$$g_p'(\gamma) = \frac{1 - \gamma^{-1}\left(\Psi\left(\gamma^{-1}+1\right) - \Psi\left(\gamma^{-1}-p+1\right)\right)}{B\left(p, \gamma^{-1}-p+1\right)}, \text{ for all } \gamma \in (0, 1/(p-1)).$$

Proof. The expression of $g_p'(\cdot)$ follows from direct calculations. Since $p > 1$, we have $\Psi\left(\gamma^{-1}+1\right) - \Psi\left(\gamma^{-1}-p+1\right) > \Psi\left(\gamma^{-1}+1\right) - \Psi\left(\gamma^{-1}\right) = \gamma$, see Abramowitz and Stegun (1972, Chapter 6). Thus $g_p'(\gamma) < 0$ and g_p is strictly decreasing as required. □

For all $y \in \mathbb{R}$, let us denote by

$$\varphi^{(k)}(y) = \mathbb{E}\left[|Y-y|^k \mathbb{1}_{\{Y>y\}}\right] \text{ and } m^{(k)}(y) = \mathbb{E}\left[|Y-y|^k\right]$$

the right-tail and total moments of $|Y-y|$. The associated empirical estimators are

$$\widehat{\varphi}_n^{(k)}(y) = \frac{1}{n}\sum_{i=1}^{n} |Y_i - y|^k \mathbb{1}_{\{Y_i>y\}} \text{ and } \widehat{m}_n^{(k)}(y) = \frac{1}{n}\sum_{i=1}^{n} |Y_i - y|^k.$$

Note that, for all $p \geq 1$, $\overline{F}^{(p)}(y) = \varphi^{(p-1)}(y)/m^{(p-1)}(y)$ and in particular $\overline{F}^{(1)}(y) = \varphi^{(0)}(y)$. The next lemma states some properties of $\varphi^{(k)}(y)$, $m^{(k)}(y)$ and $\overline{F}^{(p)}(y)$ as $y \to \infty$.

Lemma A.3. *Assume $\overline{F}^{(1)}$ satisfies $\mathcal{C}_1(\gamma)$.*

i) For all $k \in [0, 1/\gamma)$,

$$\varphi^{(k)}(y) = \frac{B\left(k+1, \gamma^{-1}-k\right)}{\gamma} y^k \overline{F}^{(1)}(y)(1+o(1)) \text{ as } y \to \infty.$$

ii) If $\mathbb{E}(|\min(Y,0)|^k) < \infty$ and $k \in [0, 1/\gamma)$, then

$$m^{(k)}(y) = y^k(1+o(1)) \text{ as } y \to \infty.$$

iii) If $\mathbb{E}(|\min(Y,0)|^{p-1}) < \infty$ and $\gamma < 1/(p-1)$, then

$$\overline{F}^{(p)}(y) = \frac{B\left(p, \gamma^{-1}-p+1\right)}{\gamma} \overline{F}^{(1)}(y)(1+o(1)) \text{ as } y \to \infty,$$

and the function $\overline{F}^{(p)}$ is regularly varying:

$$\forall y > 0, \lim_{t\to\infty} \frac{\overline{F}^{(p)}(ty)}{\overline{F}^{(p)}(t)} = y^{-1/\gamma} \text{ or equivalently } \lim_{t\to\infty} \frac{q_{1-1/(ty)}(p)}{q_{1-1/t}(p)} = y^{\gamma}.$$

iv) Assume further that $\overline{F}^{(1)}$ satisfies $\mathcal{C}_2(\gamma, \rho, A)$. Let $u_n \to \infty$ and $\varepsilon_n \to 0$ be sequences such that $A\left(1/\overline{F}^{(1)}(u_n)\right) = O(\varepsilon_n)$. If moreover $\mathbb{E}(|\min(Y,0)|^{p-1}) < \infty$ and $\gamma < 1/(p-1)$, then

$$\frac{\overline{F}^{(p)}\left(u_n(1+\varepsilon_n)\right)}{\overline{F}^{(p)}\left(u_n\right)} = 1 - \frac{\varepsilon_n}{\gamma}(1+o(1)).$$

Proof. i) The case $k=0$ is straightforward, since $\varphi^{(0)}(y)=\overline{F}^{(1)}(y)$. If $k>0$, remark that

$$\varphi^{(k)}(y)=y^k\mathbb{E}\left[\left(\frac{Y}{y}-1\right)^k\mathbb{1}_{\{Y>y\}}\right],$$

and apply Lemma 1(i) of Daouia et al. (2019) with $H(x)=(x-1)^k$ and $b=1$. ii) follows from noting that, since $\mathbb{E}(|Y|^k)<\infty$, the dominated convergence theorem yields $\mathbb{E}\left[|Y/y-1|^k\right]\to 1$ as $y\to\infty$, from which ii) follows easily. iii) follows from i) and ii). iv) is intuitively suggested by iii), but its proof is more involved. We start by using Lemma 1(ii) of Daouia et al. (2019) with $H(x)=(x-1)^{p-1}$ and $b=1$ to get

$$\varphi^{(p-1)}(y)=(p-1)y^{p-1}\overline{F}^{(1)}(y)\int_1^\infty (x-1)^{p-2}x^{-1/\gamma}\left[1+A(1/\overline{F}^{(1)}(y))\frac{x^{\rho/\gamma}-1}{\gamma\rho}(1+o(1))\right]dx.$$

Note that $\overline{F}^{(1)}(u_n(1+\varepsilon_n))=\overline{F}^{(1)}(u_n)(1+\mathrm{o}(1))$ by the regular variation property of $\overline{F}^{(1)}$, which is known to be true locally uniformly (see Theorem B.1.4 p.363 in de Haan and Ferreira, 2006). Apply this idea again to the function A to get

$$\frac{\varphi^{(p-1)}(u_n(1+\varepsilon_n))}{\varphi^{(p-1)}(u_n)}=(1+\varepsilon_n)^{p-1}\frac{\overline{F}^{(1)}(u_n(1+\varepsilon_n))}{\overline{F}^{(1)}(u_n)}\left(1+o(A(1/\overline{F}^{(1)}(u_n)))\right).$$

Now, by local uniformity of condition $\mathcal{C}_2(\gamma,\rho,A)$ (see e.g. Lemma 2 in Stupfler, 2019) combined with the assumption $A(1/\overline{F}^{(1)}(u_n))=O(\varepsilon_n)$, we get

$$\frac{\varphi^{(p-1)}(u_n(1+\varepsilon_n))}{\varphi^{(p-1)}(u_n)}=1+\left(p-1-\frac{1}{\gamma}\right)\varepsilon_n(1+o(1)). \tag{A.1}$$

We focus then on $m^{(p-1)}(u_n(1+\varepsilon_n))/m^{(p-1)}(u_n)$. Clearly $m^{(p-1)}=\varphi^{(p-1)}+\psi_1^{(p-1)}+\psi_2^{(p-1)}$, with

$$\psi_1^{(p-1)}(y)=\mathbb{E}\left[(y-Y)^{p-1}\mathbb{1}_{\{Y\le y/2\}}\right] \text{ and } \psi_2^{(p-1)}(y)=\mathbb{E}\left[(y-Y)^{p-1}\mathbb{1}_{\{y/2<Y\le y\}}\right].$$

Notice also that, as $y\to\infty$, $\psi_1^{(p-1)}(y)=y^{p-1}(1+o(1))$ by the dominated convergence theorem, and $\psi_2^{(p-1)}(y)=o(y^{p-1})$. Recalling i) and using (A.1) yields

$$\begin{aligned}\frac{m^{(p-1)}(u_n(1+\varepsilon_n))}{m^{(p-1)}(u_n)}-1 &= \left(\frac{\psi_1^{(p-1)}(u_n(1+\varepsilon_n))}{\psi_1^{(p-1)}(u_n)}-1\right)(1+o(1))+o\left(\left|\frac{\psi_2^{(p-1)}(u_n(1+\varepsilon_n))}{\psi_2^{(p-1)}(u_n)}-1\right|\right)\\ &+ o(\varepsilon_n).\end{aligned} \tag{A.2}$$

We thus focus on $\psi_1^{(p-1)}(u_n(1+\varepsilon_n))/\psi_1^{(p-1)}(u_n)$ and $\psi_2^{(p-1)}(u_n(1+\varepsilon_n))/\psi_2^{(p-1)}(u_n)$. To control the first of these two terms, we start by writing, for n large enough,

$$\begin{aligned}\psi_1^{(p-1)}(u_n(1+\varepsilon_n))-\psi_1^{(p-1)}(u_n) &= \mathbb{E}\left[\left\{(u_n(1+\varepsilon_n)-Y)^{p-1}-(u_n-Y)^{p-1}\right\}\mathbb{1}_{\{Y\le u_n(1+\varepsilon_n)/2\}}\right]\\ &+ \mathbb{E}\left[(u_n-Y)^{p-1}\left\{\mathbb{1}_{\{Y\le u_n(1+\varepsilon_n)/2\}}-\mathbb{1}_{\{Y\le u_n/2\}}\right\}\right].\end{aligned} \tag{A.3}$$

To control the first term on the right-hand side of (A.3), we use a Taylor expansion with remainder in integral form in order to write, on the event $\{Y\le u_n(1+\varepsilon_n)/2\}$,

$$\begin{aligned}(u_n(1+\varepsilon_n)-Y)^{p-1}-(u_n-Y)^{p-1} &= (p-1)u_n\varepsilon_n(u_n-Y)^{p-2}\\ &+ (p-1)(p-2)\int_{u_n-Y}^{u_n(1+\varepsilon_n)-Y}(u_n(1+\varepsilon_n)-Y-t)t^{p-3}\,dt.\end{aligned}$$

To control the integral on the right-hand side, use the change of variables $t=(u_n-Y)z$ to get

$$\int_{u_n-Y}^{u_n(1+\varepsilon_n)-Y}(u_n(1+\varepsilon_n)-Y-t)t^{p-3}\,dt=(u_n-Y)^{p-1}\int_1^{1+u_n\varepsilon_n/(u_n-Y)}\left(1+\frac{u_n\varepsilon_n}{u_n-Y}-z\right)z^{p-3}\,dz.$$

For n large enough, $\{Y \le u_n(1+\varepsilon_n)/2\} \subset \{Y \le 3u_n/4\}$, on which $u_n - Y \ge u_n/4$. On this event, it follows that $|u_n\varepsilon_n/(u_n-Y)| \le 4|\varepsilon_n|$ and thus the segment $[1, 1+u_n\varepsilon_n/(u_n-Y)]$ is contained in the interval $[1-4|\varepsilon_n|, 1+4|\varepsilon_n|] \subset [1/2, 2]$ for n large enough. On this last interval, the function $z \mapsto z^{p-3}$ is bounded from above by a positive constant. Writing

$$\left| \int_1^{1+u_n\varepsilon_n/(u_n-Y)} \left(1 + \frac{u_n\varepsilon_n}{u_n-Y} - z\right) z^{p-3}\, dz \right| \le \left|\frac{u_n\varepsilon_n}{u_n-Y}\right| \int_{1-|u_n\varepsilon_n/(u_n-Y)|}^{1+|u_n\varepsilon_n/(u_n-Y)|} z^{p-3}\, dz$$

then entails

$$\left| \int_{u_n-Y}^{u_n(1+\varepsilon_n)-Y} (u_n(1+\varepsilon_n) - Y - t) t^{p-3}\, dt \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} \right| \le C(u_n\varepsilon_n)^2 (u_n-Y)^{p-3} \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}}$$

for n large enough, where C is some positive constant. Consequently

$$\begin{aligned} & \mathbb{E}\left[\left| (u_n(1+\varepsilon_n) - Y)^{p-1} - (u_n-Y)^{p-1} - (p-1)u_n\varepsilon_n(u_n-Y)^{p-2} \right| \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} \right] \\ \le\ & C(u_n\varepsilon_n)^2 \mathbb{E}\left[(u_n-Y)^{p-3} \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} \right]. \end{aligned} \tag{A.4}$$

Remark that for any $\alpha > 0$, $\mathbb{E}\left[(u_n-Y)^{p-\alpha-1}\mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}}\right] = u_n^{p-\alpha-1}(1+o(1))$. This is a consequence of the dominated convergence theorem and the bounds

$$(1 - Y/u_n)^{p-\alpha-1} \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} \le (1/4)^{p-\alpha-1},$$

valid for n large enough when $p - \alpha - 1 \le 0$, and

$$(1 - Y/u_n)^{p-\alpha-1} \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} \le (1+|Y|)^{p-\alpha-1} \le 2^{p-\alpha-1}\mathbb{1}_{\{|Y| \le 1\}} + 2^{p-\alpha-1}|Y|^{p-\alpha-1}\mathbb{1}_{\{|Y|>1\}}$$

when $p - \alpha - 1 > 0$. Using (A.4) then entails

$$\mathbb{E}\left[\left\{ (u_n(1+\varepsilon_n) - Y)^{p-1} - (u_n-Y)^{p-1} \right\} \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} \right] = (p-1)u_n^{p-1}\varepsilon_n(1+o(1)). \tag{A.5}$$

To control the second term on the right-hand side of (A.3), write

$$\begin{aligned} & \left| \mathbb{E}\left[(u_n-Y)^{p-1} \left\{ \mathbb{1}_{\{Y \le u_n(1+\varepsilon_n)/2\}} - \mathbb{1}_{\{Y \le u_n/2\}} \right\} \right] \right| \\ \le\ & \left(\frac{u_n}{2}(1+|\varepsilon_n|) \right)^{p-1} \left[\overline{F}^{(1)}(u_n(1-|\varepsilon_n|)/2) - \overline{F}^{(1)}(u_n(1+|\varepsilon_n|)/2) \right] \\ \le\ & u_n^{p-1}\overline{F}^{(1)}(u_n/2) \left[\frac{\overline{F}^{(1)}(u_n(1-|\varepsilon_n|)/2)}{\overline{F}^{(1)}(u_n/2)} - \frac{\overline{F}^{(1)}(u_n(1+|\varepsilon_n|)/2)}{\overline{F}^{(1)}(u_n/2)} \right] = o(u_n^{p-1}\varepsilon_n) \end{aligned} \tag{A.6}$$

by condition $\mathcal{C}_2(\gamma, \rho, A)$ and its local uniformity. Combining (A.3), (A.5), (A.6) and the asymptotic equivalent $\psi_1^{(p-1)}(y) = y^{p-1}(1+o(1))$ (as $y \to \infty$) results in

$$\frac{\psi_1^{(p-1)}(u_n(1+\varepsilon_n))}{\psi_1^{(p-1)}(u_n)} - 1 = (p-1)\varepsilon_n(1+o(1)). \tag{A.7}$$

We turn to the control of $\psi_2^{(p-1)}(u_n(1+\varepsilon_n))/\psi_2^{(p-1)}(u_n)$. For this we write

$$\psi_2^{(p-1)}(y) = (y/2)^{p-1}\left[\overline{F}^{(1)}(y/2) - \overline{F}^{(1)}(y)\right] + \mathbb{E}\left[\{(y-Y)^{p-1} - (y/2)^{p-1}\}\mathbb{1}_{\{y/2 < Y \le y\}}\right].$$

Using condition $\mathcal{C}_2(\gamma, \rho, A)$ gives, as $y \to \infty$,

$$\overline{F}^{(1)}(y/2) - \overline{F}^{(1)}(y) = \overline{F}^{(1)}(y)\left(2^{1/\gamma} - 1 + A(1/\overline{F}^{(1)}(y))\left[2^{1/\gamma}\frac{2^{-\rho/\gamma}-1}{\gamma\rho} + o(1)\right]\right).$$

Meanwhile, an integration by parts yields

$$\mathbb{E}\left[\{(y-Y)^{p-1} - (y/2)^{p-1}\}\mathbb{1}_{\{y/2 < Y \le y\}}\right] = -(p-1)y^{p-1}\overline{F}^{(1)}(y)\int_{1/2}^1 (1-z)^{p-2}\left[\frac{\overline{F}^{(1)}(yz)}{\overline{F}^{(1)}(y)} - 1\right] dz.$$

Using the local uniformity of condition $\mathcal{C}_2(\gamma,\rho,A)$, we find, as $y\to\infty$,

$$\mathbb{E}\left[\{(y-Y)^{p-1}-(y/2)^{p-1}\}\mathbb{1}_{\{y/2<Y\le y\}}\right]=-(p-1)y^{p-1}\overline{F}^{(1)}(y)\left(\int_{1/2}^{1}(1-z)^{p-2}\left[z^{-1/\gamma}-1\right]dz\right.$$
$$\left.+A(1/\overline{F}^{(1)}(y))\int_{1/2}^{1}(1-z)^{p-2}z^{-1/\gamma}\frac{z^{\rho/\gamma}-1}{\gamma\rho}dz(1+o(1))\right).$$

It then follows from the regular variation properties of $\overline{F}^{(1)}$ and A that

$$\frac{\psi_2^{(p-1)}(u_n(1+\varepsilon_n))}{\psi_2^{(p-1)}(u_n)}-1=O(\varepsilon_n). \tag{A.8}$$

Conclude, by combining (A.2), (A.7) and (A.8), that

$$\frac{m^{(p-1)}(u_n(1+\varepsilon_n))}{m^{(p-1)}(u_n)}-1=\left(\frac{\psi_1^{(p-1)}(u_n(1+\varepsilon_n))}{\psi_1^{(p-1)}(u_n)}-1\right)(1+o(1))+o(\varepsilon_n)=(p-1)\varepsilon_n(1+o(1)). \tag{A.9}$$

Writing

$$\frac{\overline{F}^{(p)}\left(u_n(1+\varepsilon_n)\right)}{\overline{F}^{(p)}\left(u_n\right)}=\frac{\varphi^{(p-1)}\left(u_n(1+\varepsilon_n)\right)}{\varphi^{(p-1)}\left(u_n\right)}\times\frac{m^{(p-1)}\left(u_n\right)}{m^{(p-1)}\left(u_n(1+\varepsilon_n)\right)}$$

and combining (A.1) and (A.9) completes the proof. □

In the following two lemmas, some (joint) asymptotic normality results are established for the estimators $\widehat{\varphi}_n^{(k)}(y_n)$ and $\widehat{\overline{F}}_n^{(k)}(y_n)$ where $y_n\to\infty$ and $n\to\infty$.

Lemma A.4. *Assume $\overline{F}^{(1)}$ satisfies $\mathcal{C}_2(\gamma,\rho,A)$ with $\gamma<1/[2(p-1)]$ and $\mathbb{E}(|\min(Y,0)|^{2(p-1)})<\infty$. Let (y_n) and (y_n') be two sequences such that $y_n\to\infty$, $n\overline{F}^{(1)}(y_n)\to\infty$ and $y_n'/y_n\to\lambda>0$ as $n\to\infty$. Then*

$$\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\widehat{\varphi}_n^{(p-1)}(y_n)}{\varphi^{(p-1)}(y_n)}-1,\frac{\widehat{\varphi}_n^{(0)}(y_n')}{\varphi^{(0)}(y_n')}-1\right)\xrightarrow{d}\mathcal{N}\left(\mathbf{0},\mathbf{\Sigma}\right), \tag{A.10}$$

where $\mathbf{\Sigma}$ is a symmetric matrix having entries

$$\begin{cases}\Sigma_{1,1} &= \gamma\dfrac{B\left(2p-1,\gamma^{-1}-2p+2\right)}{B\left(p,\gamma^{-1}-p+1\right)^2},\\[2ex] \Sigma_{1,2} &= \gamma\dfrac{(p-1)\lambda^{1/\gamma}IB\left((\lambda\vee 1)^{-1},\gamma^{-1}-p+1,p-1\right)+(\lambda\vee 1-1)^{p-1}}{B\left(p,\gamma^{-1}-p+1\right)},\\[2ex] \Sigma_{2,2} &= \lambda^{1/\gamma}.\end{cases} \tag{A.11}$$

Proof. Let $\boldsymbol{\beta}=(\beta_1,\beta_2)\in\mathbb{R}^2$ and focus on the asymptotic distribution of

$$Z_n=\sqrt{n\overline{F}^{(1)}(y_n)}\left\{\beta_1\left(\frac{\widehat{\varphi}_n^{(p-1)}(y_n)}{\varphi^{(p-1)}(y_n)}-1\right)+\beta_2\left(\frac{\widehat{\varphi}_n^{(0)}(y_n')}{\varphi^{(0)}(y_n')}-1\right)\right\}.$$

We clearly have $\mathbb{E}[Z_n]=0$ and $\mathbb{V}\mathrm{ar}[Z_n]=\overline{F}^{(1)}(y_n)\boldsymbol{\beta}^\top\boldsymbol{S}^{(n)}\boldsymbol{\beta}$, where $\boldsymbol{S}^{(n)}$ is the symmetric matrix having entries

$$\begin{cases}S_{1,1}^{(n)} &= \dfrac{\mathbb{V}\mathrm{ar}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>y_n\}}\right]}{\varphi^{(p-1)}(y_n)^2},\\[2ex] S_{1,2}^{(n)} &= \dfrac{\mathrm{cov}\left(|Y-y_n|^{p-1}\mathbb{1}_{\{Y>y_n\}},\mathbb{1}_{\{Y>y_n'\}}\right)}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')},\\[2ex] S_{2,2}^{(n)} &= \dfrac{\mathbb{V}\mathrm{ar}\left[\mathbb{1}_{\{Y>y_n'\}}\right]}{\varphi^{(0)}(y_n')^2}.\end{cases}$$

Let us first focus on $S_{1,1}^{(n)}$. Lemma A.3i) yields

$$S_{1,1}^{(n)} = \frac{\varphi^{(2p-2)}(y_n)}{\varphi^{(p-1)}(y_n)^2} - 1 = \gamma \frac{B\left(2p-1, \gamma^{-1}-2p+2\right)}{B\left(p, \gamma^{-1}-p+1\right)^2} \overline{F}^{(1)}(y_n)^{-1}(1+o(1)).$$

The calculation of the covariance term $S_{1,2}^{(n)}$ relies on Lemma A.3i) and the decomposition

$$S_{1,2}^{(n)} = \frac{\mathbb{E}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>by_n\}}\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} + \frac{\mathbb{E}\left[|Y-y_n|^{p-1}\left(\mathbb{1}_{\{Y>y_n\vee y_n'\}} - \mathbb{1}_{\{Y>by_n\}}\right)\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} - 1$$

where $b = \lambda \vee 1$. The first term above is evaluated by using Lemma A.3i) to write

$$\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>by_n\}}\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} = \frac{\mathbb{E}\left[\left\{\left|\frac{Y}{y_n}-1\right|^{p-1} - (b-1)^{p-1}\right\}\mathbb{1}_{\{Y>by_n\}}\right] + (b-1)^{p-1}\,\overline{F}^{(1)}(by_n)}{\frac{B(p,\gamma^{-1}-p+1)}{\gamma}\overline{F}^{(1)}(y_n)\overline{F}^{(1)}(y_n')(1+o(1))}.$$

Applying Daouia et al. (2019, Lemma 1(i)) with $H(x) = (x-1)^{p-1}$ and carrying out straightforward calculations based on the change of variable $t = 1/x$ and the regular variation property of $\overline{F}^{(1)}$, we obtain

$$\begin{aligned}
&\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>by_n\}}\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} \\
&= \gamma \frac{\overline{F}^{(1)}(y_n)\int_b^\infty (p-1)(x-1)^{p-2}x^{-1/\gamma}dx(1+o(1)) + (b-1)^{p-1}\,\overline{F}^{(1)}(by_n)}{B\left(p,\gamma^{-1}-p+1\right)\overline{F}^{(1)}(y_n)\overline{F}^{(1)}(y_n')(1+o(1))} \\
&= \left[(p-1)\gamma\lambda^{1/\gamma}\frac{IB\left(b^{-1},\gamma^{-1}-p+1,p-1\right)}{B\left(p,\gamma^{-1}-p+1\right)} + \gamma\frac{(b-1)^{p-1}}{B\left(p,\gamma^{-1}-p+1\right)}\right]\overline{F}^{(1)}(y_n)^{-1}(1+o(1)).
\end{aligned}$$

Besides, by Lemma A.3i) and the regular variation property of $\overline{F}^{(1)}$ again,

$$\left|\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\left(\mathbb{1}_{\{Y>y_n\vee y_n'\}} - \mathbb{1}_{\{Y>by_n\}}\right)\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')}\right| = O\left(\overline{F}^{(1)}(y_n)^{-1}\left|\frac{\overline{F}^{(1)}(y_n\vee y_n')}{\overline{F}^{(1)}(y_n)} - \frac{\overline{F}^{(1)}(by_n)}{\overline{F}^{(1)}(y_n)}\right|\right).$$

Noting that $(y_n \vee y_n')/y_n \to b$ and using local uniformity of the regular variation property of $\overline{F}^{(1)}$ (Bingham et al., 1989, Theorem 1.2.1), we find

$$\left|\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\left(\mathbb{1}_{\{Y>y_n\vee y_n'\}} - \mathbb{1}_{\{Y>by_n\}}\right)\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')}\right| = o\left(\overline{F}^{(1)}(y_n)^{-1}\right).$$

Conclude that

$$S_{1,2}^{(n)} = \left[(p-1)\gamma\lambda^{1/\gamma}\frac{IB\left(b^{-1},\gamma^{-1}-p+1,p-1\right)}{B\left(p,\gamma^{-1}-p+1\right)} + \gamma\frac{(b-1)^{p-1}}{B\left(p,\gamma^{-1}-p+1\right)}\right]\overline{F}^{(1)}(y_n)^{-1}(1+o(1)).$$

Finally, using regular variation again,

$$S_{2,2}^{(n)} = \frac{\overline{F}^{(1)}(y_n')}{\overline{F}^{(1)}(y_n')^2} - 1 = \lambda^{1/\gamma}\overline{F}^{(1)}(y_n)^{-1}(1+o(1)),$$

and therefore, $\mathbb{V}\mathrm{ar}[Z_n] \to \boldsymbol{\beta}^\top\boldsymbol{\Sigma}\boldsymbol{\beta}$ as $n\to\infty$, where $\boldsymbol{\Sigma}$ is given in (A.11). To prove the asymptotic normality of Z_n, write $Z_n = \sum_{i=1}^n Z_{i,n}$, where

$$Z_{i,n} = \frac{\sqrt{n\overline{F}^{(1)}(y_n)}}{n}\left\{\beta_1\left(\frac{|Y_i-y_n|^{p-1}\mathbb{1}_{\{Y_i>y_n\}}}{\varphi^{(p-1)}(y_n)} - 1\right) + \beta_2\left(\frac{\mathbb{1}_{\{Y_i>y_n'\}}}{\varphi^{(0)}(y_n')} - 1\right)\right\}.$$

Since $Z_{1,n}, \dots, Z_{n,n}$ are independent and identically distributed centered random variables, according to the Lyapunov central limit theorem, a sufficient condition for asymptotic normality (A.10) is the existence of $\delta > 0$ such that $n\mathbb{E}\left[|Z_{1,n}|^{2+\delta}\right] \to 0$ as $n \to \infty$. To this end, note that, if $T_1, \dots, T_q$ have finite $(2+\delta)-$moments, then

$$\mathbb{E}\left[\left|\sum_{j=1}^{q}(T_j - \mathbb{E}[T_j])\right|^{2+\delta}\right]^{1/(2+\delta)} \leq 2q \max_{1\leq j\leq q} \mathbb{E}\left[|T_j|^{2+\delta}\right]^{1/(2+\delta)},$$

by the triangle inequality. Therefore, letting $\delta > 0$ be so small that $\delta < \gamma^{-1}/(p-1) - 2$, one has

$$\mathbb{E}\left[|Z_{1,n}|^{2+\delta}\right] \leq 4^{2+\delta}\overline{F}^{(1)}(y_n)^{1+\delta/2}n^{-1-\delta/2} \max\left\{\beta_1^{2+\delta}\frac{\varphi^{(p-1)(2+\delta)}(y_n)}{\varphi^{(p-1)}(y_n)^{2+\delta}}, \beta_2^{2+\delta}\frac{\overline{F}^{(1)}(y_n')}{\overline{F}^{(1)}(y_n')^{2+\delta}}\right\}.$$

Lemma A.3i) entails $n\mathbb{E}\left[|Z_{1,n}|^{2+\delta}\right] = O\left(\left(n\overline{F}^{(1)}(y_n)\right)^{-\delta/2}\right) = o(1)$ and the result is proved. □

Lemma A.5. *Under the conditions of Lemma A.4 and with* $\mathbf{\Sigma}$ *given in* (A.11),

$$\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\widehat{\overline{F}}_n^{(p)}(y_n)}{\overline{F}^{(p)}(y_n)} - 1, \frac{\widehat{\overline{F}}_n^{(1)}(y_n')}{\overline{F}^{(1)}(y_n')} - 1\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}). \tag{A.12}$$

Proof. Let $\boldsymbol{\beta} = (\beta_1, \beta_2) \in \mathbb{R}^2$, and consider

$$\begin{aligned} Z_n &= \sqrt{n\overline{F}^{(1)}(y_n)}\beta_1\left(\frac{\widehat{\overline{F}}_n^{(p)}(y_n)}{\overline{F}^{(p)}(y_n)} - 1\right) + \sqrt{n\overline{F}^{(1)}(y_n)}\beta_2\left(\frac{\widehat{\overline{F}}_n^{(1)}(y_n')}{\overline{F}^{(1)}(y_n')} - 1\right) \\ &= \sqrt{n\overline{F}^{(1)}(y_n)}\left\{\beta_1\left(\frac{\widehat{\varphi}_n^{(p-1)}(y_n)}{\varphi^{(p-1)}(y_n)} - 1\right)\frac{m^{(p-1)}(y_n)}{\widehat{m}_n^{(p-1)}(y_n)} + \beta_1\left(\frac{m^{(p-1)}(y_n)}{\widehat{m}_n^{(p-1)}(y_n)} - 1\right) + \beta_2\left(\frac{\widehat{\varphi}_n^{(0)}(y_n')}{\varphi^{(0)}(y_n')} - 1\right)\right\}. \end{aligned}$$

Straightforward calculations yield $\mathbb{E}\left[\widehat{m}_n^{(p-1)}(y_n)/m^{(p-1)}(y_n)\right] = 1$, and $\mathbb{V}\mathrm{ar}\left[\widehat{m}_n^{(p-1)}(y_n)/m^{(p-1)}(y_n)\right] = O(1/n)$ by Lemma A.3ii). Thus

$$\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{m^{(p-1)}(y_n)}{\widehat{m}_n^{(p-1)}(y_n)} - 1\right) = O_{\mathbb{P}}\left(\overline{F}^{(1)}(y_n)\right) = o_{\mathbb{P}}(1),$$

and the result follows from Lemma A.4. □

The final auxiliary result is a corrected version of Proposition 2 in Daouia et al. (2019).

Proposition 2. *Assume* $\overline{F}^{(1)}$ *satisfies condition* $\mathcal{C}_2(\gamma, \rho, A)$ *and* $\overline{F}_-^{(1)}$ *either satisfies* $y^a\overline{F}_-^{(1)}(y) \to 0$ *as* $y \to +\infty$ *for all* $a > 0$ *(in which case we set* $\gamma_\ell = 0$*), or* $\mathcal{C}_1(\gamma_\ell)$*. Then, for all* $1 < p < (\gamma \vee \gamma_\ell)^{-1} + 1$,

$$\frac{\overline{F}^{(1)}(q_\tau(p))}{1-\tau} = \frac{\gamma}{B(p, \gamma^{-1} - p + 1)}(1 + R(\tau, p))$$

where

$$\begin{aligned} R(\tau, p) &= (1 - g_p(\gamma))(1-\tau)(1 + \mathrm{o}(1)) - g_p(\gamma)K(p, \gamma, \rho)A((1-\tau)^{-1})(1 + \mathrm{o}(1)) \\ &\quad - (p-1)\left(g_p(\gamma)^{\min(\gamma,1)}R_r(q_\tau(1), p, \gamma) - g_p(\gamma)^{\gamma/\max(\gamma_\ell,1)}R_\ell(q_\tau(1), p, \gamma_\ell)\right) \end{aligned}$$

as $\tau \uparrow 1$, with $K(p,\gamma,\rho)$ as in Theorem 1 and

$$R_r(q,p,\gamma) = \begin{cases} \dfrac{\mathbb{E}(X\mathbb{1}_{\{0<X<q\}})}{q}(1+\mathrm{o}(1)) & \text{if } \gamma \le 1, \\ \overline{F}^{(1)}(q)B(p-1,1-\gamma^{-1})(1+\mathrm{o}(1)) & \text{if } \gamma > 1, \end{cases}$$

$$\text{and } R_\ell(q,p,\gamma_\ell) = \begin{cases} -\dfrac{\mathbb{E}(X\mathbb{1}_{\{-q<X<0\}})}{q}(1+\mathrm{o}(1)) & \text{if } \gamma_\ell \le 1, \\ F^{(1)}(-q)B(\gamma_\ell^{-1}-p+1,1-\gamma_\ell^{-1})(1+\mathrm{o}(1)) & \text{if } \gamma_\ell > 1. \end{cases}$$

The difference between Proposition 2 and Proposition 2 in Daouia et al. (2019) is the remainder term of order $(1-\tau)$, which is incorrect in the latter.

Proof. Retrace the steps of the proof of Proposition 2 in Daouia et al. (2019) and note that (with the notation therein) Equation (A.15) is incorrect because it does not take term $I_1(q;p)$ into account. Keeping the notation of this paper, using the asymptotic equivalent of this term provided right after Equation (A.10) and accounting properly for it in (A.15) yields the desired result after straightforward computations. □

Appendix B Proofs of main results

B.1 Proof of Proposition 1

Let us introduce $(z_1,z_p)\in\mathbb{R}^2$, $\sigma_n = 1/\sqrt{n(1-\alpha_n)}$, and, for all $p>1$, focus on the probability:

$$\begin{aligned} \Phi_n(z_1,z_p) &= \mathbb{P}\left(\left\{\sigma_n^{-1}\left(\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)}-1\right)\le z_p\right\}\bigcap\left\{\sigma_n^{-1}\left(\frac{\widehat{q}_{\beta_n}(1)}{q_{\beta_n}(1)}-1\right)\le z_1\right\}\right) \\ &= \mathbb{P}\left(\{\widehat{q}_{\alpha_n}(p)\le q_{\alpha_n}(p)(1+z_p\sigma_n)\}\bigcap\{\widehat{q}_{\beta_n}(1)\le q_{\beta_n}(1)(1+z_1\sigma_n)\}\right). \end{aligned}$$

Remarking that $1-\alpha_n = \overline{F}^{(p)}(q_{\alpha_n}(p))$ by continuity of $\overline{F}^{(p)}$ (see Lemma A.1) and $1-\beta_n = \overline{F}^{(1)}(q_{\beta_n}(1))$ by continuity of $\overline{F}^{(1)}$, and using the fact that, for all y, α and k, $\widehat{q}_n^{(k)}(\alpha)\le y \Leftrightarrow \widehat{\overline{F}}_n^{(k)}(y)\le 1-\alpha$, it follows that $\Phi_n(z_1,z_p)$ can be rewritten as

$$\mathbb{P}\left(\left\{\widehat{\overline{F}}_n^{(p)}(q_{\alpha_n}(p)(1+z_p\sigma_n))\le\overline{F}^{(p)}(q_{\alpha_n}(p))\right\}\bigcap\left\{\widehat{\overline{F}}_n^{(1)}(q_{\beta_n}(1)(1+z_1\sigma_n))\le\overline{F}^{(1)}(q_{\beta_n}(1))\right\}\right).$$

Equivalently, letting $y_n = q_{\alpha_n}(p)(1+z_p\sigma_n)$ and $y_n' = q_{\beta_n}(1)(1+z_1\sigma_n)$, one has

$$\begin{aligned} \Phi_n(z_1,z_p) = \mathbb{P}\Bigg(&\left\{\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\widehat{\overline{F}}_n^{(p)}(y_n)}{\overline{F}^{(p)}(y_n)}-1\right)\le\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\overline{F}^{(p)}(q_{\alpha_n}(p))}{\overline{F}^{(p)}(q_{\alpha_n}(p)(1+z_p\sigma_n))}-1\right)\right\} \\ &\bigcap\left\{\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\widehat{\overline{F}}_n^{(1)}(y_n')}{\overline{F}^{(1)}(y_n')}-1\right)\le\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\overline{F}^{(1)}(q_{\beta_n}(1))}{\overline{F}^{(1)}(q_{\beta_n}(1)(1+z_1\sigma_n))}-1\right)\right\}\Bigg). \end{aligned}$$

Note now that a combination of the local uniformity of condition $\mathcal{C}_2(\gamma,\rho,A)$ (see e.g. Lemma 2 in Stupfler, 2019) with assumption $A((1-\alpha_n)^{-1}) = O(1/\sqrt{n(1-\alpha_n)}) = O(\sigma_n)$ yields:

$$\overline{F}^{(1)}(q_{\beta_n}(1)(1+z_1\sigma_n)) = \overline{F}^{(1)}(q_{\beta_n}(1))\left(1-\frac{z_1\sigma_n}{\gamma}(1+o(1))\right).$$

Besides, applying Lemma A.3iv) gives

$$\overline{F}^{(p)}(q_{\alpha_n}(p)(1+z_p\sigma_n)) = \overline{F}^{(p)}(q_{\alpha_n}(p))\left(1-\frac{z_p\sigma_n}{\gamma}(1+o(1))\right).$$

Here the asymptotic proportionality between $q_{\alpha_n}(p)$ and $q_{\alpha_n}(1)$ was used, together with the regular variation property of A, so as to satisfy the assumptions of Lemma A.3iv) with $u_n = q_{\alpha_n}(p)$ and $\varepsilon_n = z_p\sigma_n$. Therefore:

$$\begin{cases} \dfrac{\overline{F}^{(p)}\left(q_{\alpha_n}(p)\right)}{\overline{F}^{(p)}\left(q_{\alpha_n}(p)(1+z_p\sigma_n)\right)} - 1 &= \dfrac{z_p}{\gamma}\sigma_n(1+o(1)), \\ \dfrac{\overline{F}^{(1)}\left(q_{\beta_n}(1)\right)}{\overline{F}^{(1)}\left(q_{\beta_n}(1)(1+z_1\sigma_n)\right)} - 1 &= \dfrac{z_1}{\gamma}\sigma_n(1+o(1)). \end{cases}$$

Since $\sqrt{n\overline{F}^{(1)}(y_n)} = \sigma_n^{-1}\sqrt{g_p(\gamma)}(1+o(1))$ from (2.3) and local uniformity of the regular variation condition, it follows that

$$\Phi_n(z_1,z_p) = \mathbb{P}\left(\left\{\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\widehat{\overline{F}}_n^{(p)}(y_n)}{\overline{F}^{(p)}(y_n)} - 1\right) \le \frac{\sqrt{g_p(\gamma)}}{\gamma}z_p(1+o(1))\right\}\right.$$
$$\left.\bigcap\left\{\sqrt{n\overline{F}^{(1)}(y_n)}\left(\frac{\widehat{\overline{F}}_n^{(1)}(y_n')}{\overline{F}^{(1)}(y_n')} - 1\right) \le \frac{\sqrt{g_p(\gamma)}}{\gamma}z_1(1+o(1))\right\}\right).$$

Another direct consequence of (2.3), using the regular variation property of $\overline{F}^{(1)}$, is $q_{\alpha_n}(1) = g_p(\gamma)^\gamma q_{\alpha_n}(p)(1+o(1))$, hence $y_n' = \theta^{-\gamma}q_{\alpha_n}(1)(1+o(1)) = \theta^{-\gamma}g_p(\gamma)^\gamma y_n(1+o(1))$. It only remains to apply Lemma A.5 with $\lambda = \theta^{-\gamma}g_p(\gamma)^\gamma$ to conclude the proof.

B.2 Proof of Theorem 1

Let again $\sigma_n = 1/\sqrt{n(1-\alpha_n)}$ and for any $z\in\mathbb{R}$, focus on the probability

$$\Phi_n(z) = \mathbb{P}\left(\sigma_n^{-1}\left(\frac{\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))}{1-\alpha_n} - g_p(\gamma)\right) \le z\right) = \mathbb{P}\left(\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p)) \le (1-\alpha_n)\left(g_p(\gamma)+z\sigma_n\right)\right).$$

Equivalently, $\Phi_n(z) = \mathbb{P}\left(\widehat{q}_{\alpha_n}(p) \ge \widehat{q}_{\beta_n}(1)\right)$, where $\beta_n = 1-(1-\alpha_n)\left(g_p(\gamma)+z\sigma_n\right)$ and therefore,

$$\Phi_n(z) = \mathbb{P}\left(\sigma_n^{-1}\left(\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)} - 1\right) \ge \sigma_n^{-1}\left(\frac{\widehat{q}_{\beta_n}(1)}{q_{\beta_n}(1)} - 1\right)\frac{q_{\beta_n}(1)}{q_{\alpha_n}(p)} + \sigma_n^{-1}\left(\frac{q_{\beta_n}(1)}{q_{\alpha_n}(p)} - 1\right)\right).$$

From de Haan and Ferreira (2006, Theorem 2.3.9), condition $\mathcal{C}_2(\gamma,\rho,A)$ entails

$$\frac{q_{\beta_n}(1)}{q_{\alpha_n}(1)} = g_p(\gamma)^{-\gamma}\left[1 + \frac{g_p(\gamma)^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1)) - \frac{\gamma z\sigma_n}{g_p(\gamma)}(1+o(1))\right]. \tag{B.1}$$

In addition, combining Proposition 2 with Lemma 1(i) in Daouia et al. (2020) yields

$$\frac{q_{\alpha_n}(p)}{q_{\alpha_n}(1)} = g_p(\gamma)^{-\gamma}\left[1 - \gamma R(\alpha_n,p)(1+o(1)) + \frac{g_p(\gamma)^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1))\right], \tag{B.2}$$

with

$$\begin{aligned} R(\alpha_n,p) &= [1-g_p(\gamma)](1-\alpha_n)(1+o(1)) - g_p(\gamma)K(p,\gamma,\rho)A\left((1-\alpha_n)^{-1}\right)(1+o(1)) \\ &- (p-1)g_p(\gamma)^{\gamma\wedge 1}\left(\frac{\mathbb{E}\left[Y\mathbb{1}_{\{0<Y<q_{\alpha_n}(1)\}}\right]}{q_{\alpha_n}(1)}\mathbb{1}_{\{\gamma\le 1\}} + (1-\alpha_n)B\left(p-1,1-\gamma^{-1}\right)\mathbb{1}_{\{\gamma>1\}}\right) \\ &+ (p-1)g_p(\gamma)^{\frac{\gamma}{\gamma_\ell\vee 1}}\frac{\mathbb{E}\left[-Y\mathbb{1}_{\{0<-Y<q_{\alpha_n}(1)\}}\right]}{q_{\alpha_n}(1)}\mathbb{1}_{\{\gamma_\ell\le 1\}} \\ &+ (p-1)g_p(\gamma)^{\frac{\gamma}{\gamma_\ell\vee 1}}F^{(1)}\left(-q_{\alpha_n}(1)\right)B\left(\gamma_\ell^{-1}-p+1,1-\gamma_\ell^{-1}\right)\mathbb{1}_{\{\gamma_\ell>1\}}, \end{aligned}$$

and where $K(p,\gamma,\rho)$ is given in (3.3). Combining (B.1) and (B.2) yields

$$\sigma_n^{-1}\left(\frac{q_{\beta_n}(1)}{q_{\alpha_n}(p)}-1\right) = -z\frac{\gamma}{g_p(\gamma)} + o(1) + \gamma\sigma_n^{-1}R(\alpha_n,p)(1+o(1)).$$

In view of conditions *i*), *ii*) and *iii*), one has

$$\sigma_n^{-1}R(\alpha_n,p) \to \lambda = -\left[\lambda_1 b_p^{(1)}(\gamma,\rho) + \lambda_2 b_p^{(2)}(\gamma) + \lambda_3 b_p^{(3)}(\gamma,\gamma_\ell)\right] \text{ as } n\to\infty.$$

Then,

$$\Phi_n(z) = \mathbb{P}\left(\frac{g_p(\gamma)}{\gamma}\left[-\sigma_n^{-1}\left(\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)}-1\right) + \sigma_n^{-1}\left(\frac{\widehat{q}_{\beta_n}(1)}{q_{\beta_n}(1)}-1\right)(1+o(1))\right] + \lambda g_p(\gamma) + o(1) \le z\right),$$

and the asymptotic distribution of $\sigma_n^{-1}\left(\frac{\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))}{1-\alpha_n} - g_p(\gamma)\right)$ is the same as that of

$$\frac{g_p(\gamma)}{\gamma}(-1,1)\,\sigma_n^{-1}\begin{pmatrix}\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)}-1\\ \frac{\widehat{q}_{\beta_n}(1)}{q_{\beta_n}(1)}-1\end{pmatrix} + \lambda g_p(\gamma).$$

Finally, $\widehat{\gamma}_{\alpha_n}^{(p)}$ being obtained by inverting $g_p(\cdot)$ at $\widehat{\overline{F}}_n^{(1)}(\widehat{q}_{\alpha_n}(p))/(1-\alpha_n)$, its asymptotic distribution is the limit in distribution of

$$\frac{g_p(\gamma)}{\gamma g_p'(\gamma)}(-1,1)\,\sigma_n^{-1}\begin{pmatrix}\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)}-1\\ \frac{\widehat{q}_{\beta_n}(1)}{q_{\beta_n}(1)}-1\end{pmatrix} + \frac{\lambda g_p(\gamma)}{g_p'(\gamma)}$$

by the delta-method. Apply Lemma A.2 and Proposition 1 with $\theta = g_p(\gamma)$ to conclude the proof.

B.3 Proof of Theorem 2

We prove the result on $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1)$; the proof of the result about $\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(2)$ is identical. Using our assumptions on $\overline{b}$ and $\overline{\rho}$ and the $\sqrt{n(1-\alpha_n)}$−consistency of $\widetilde{\gamma}_{\alpha_n}^{(p)}$, we find

$$\begin{aligned}
&\log\left(\frac{\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1)}{q_{\alpha_n'}(1)}\right)\\
&= (\widetilde{\gamma}_{\alpha_n}^{(p)}-\gamma)\log\left(\frac{1-\alpha_n}{1-\alpha_n'}\right) + \log\left(\frac{\widehat{q}_{\alpha_n}(p)}{q_{\alpha_n}(p)}\right) + \log\left(\frac{[g_p(\widetilde{\gamma}_{\alpha_n}^{(p)})]^{\widetilde{\gamma}_{\alpha_n}^{(p)}}}{[g_p(\gamma)]^{\gamma}}\right) + \log\left([g_p(\gamma)]^{\gamma}\frac{q_{\alpha_n}(p)}{q_{\alpha_n}(1)}\right)\\
&- \log\left(\left[\frac{1-\alpha_n'}{1-\alpha_n}\right]^{\gamma}\frac{q_{\alpha_n'}(1)}{q_{\alpha_n}(1)}\right) + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\alpha_n)}).
\end{aligned}$$

Combine then Theorem 1 in Daouia et al. (2019) (for independent data), the $\sqrt{n(1-\alpha_n)}$−consistency of $\widetilde{\gamma}_{\alpha_n}^{(p)}$, Proposition 2, asymptotic inversion of $\overline{F}^{(1)}$ (see e.g. Daouia et al., 2020, Lemma 1(i)) and Theorem 2.3.9 in de Haan and Ferreira (2006) to get

$$\log\left(\frac{\widetilde{q}_{\alpha_n',p}^{\star,\mathrm{RB}}(1)}{q_{\alpha_n'}(1)}\right) = (\widetilde{\gamma}_{\alpha_n}^{(p)}-\gamma)\log\left(\frac{1-\alpha_n}{1-\alpha_n'}\right) + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\alpha_n)}).$$

The proof is complete.