Improved Reduced Bias Tail Index and Quantile Estimators*

Jan Beirlant
Catholic University of Leuven, Belgium

Fernanda Figueiredo
University of Porto and CEAUL, Portugal

M. Ivette Gomes and Björn Vandewalle
University of Lisbon, Faculty of Science
DEIO and CEAUL, Portugal

Abstract. In this paper, we deal with bias reduction techniques for heavy tails, trying to improve mainly upon the performance of classical high quantile estimators. High quantiles depend strongly on the tail index γ, for which new classes of reduced bias estimators have recently been introduced, where the second order parameters in the bias are estimated at a level k_1 of a larger order than the level k at which the tail index is estimated. Doing this, it was seen that the asymptotic variance of the new estimators could be kept equal to the one of the popular Hill estimator. In a similar way, we now introduce new classes of tail index and associated high quantile estimators, with an asymptotic mean squared error smaller than that of the classical ones for all k. We derive their asymptotic distributional properties and compare them with those of alternative estimators. Next to that, an illustration of the finite sample behavior of the estimators is also provided through a Monte Carlo simulation study and the application to a set of real data in the field of insurance.

AMS 2000 subject classification: Primary 62G32, 62E20; Secondary 65C05.
Keywords and phrases: Extreme value statistics, heavy tails, high quantiles.

*Research partially supported by FCT / POCTI and POCI / FEDER.

1 Introduction.

Denote by $\mathcal{R}_\alpha$ the class of functions that are regularly varying at infinity with real index α, i.e. the class of positive measurable functions g such that

$$\lim_{t\to\infty}\frac{g(tx)}{g(t)} = x^\alpha, \tag{1.1}$$

for all $x > 0$. Then, in the context of Extreme Value Theory, a distribution function (d.f.) F is said to have a heavy right tail whenever there exists a positive γ such that the tail function $\overline{F} := 1 - F \in \mathcal{R}_{-1/\gamma}$. Equivalently, defining the tail quantile function of F as $U(t) := F^{\leftarrow}(1-1/t) = \inf\{x \mid F(x) \geq 1 - 1/t\}$ for $1 < t < \infty$, it can be argued that F is heavy-tailed, i.e. that $\overline{F} \in \mathcal{R}_{-1/\gamma}$ if and only if $U \in \mathcal{R}_\gamma$.

The parameter γ is referred to as the tail index of F and helps to indicate the size and frequency of certain extreme events. In general, the larger the tail index, the heavier the tail. Heavy-tailed models have revealed to be quite useful in areas ranging from insurance, economics and finance until telecommunication and biostatistics. In this paper, we focus on the estimation of high quantiles, in risk management often referred to as the Value-at-Risk. More specifically, given a sequence $X_1, \ldots, X_n$ of independent and identically distributed random variables (r.v.'s) with common d.f. F and a small value p, we are interested in estimating $\chi_p = U(1/p)$.

Here, extreme value methodology typically applies to probabilities p for which $p = O(1/n)$, meaning that in view of the asymptotic theory (with n tending to infinity) we have to assume that p depends on n. Therefore, writing $p = p_n$, we consider the estimation of the quantile $\chi_{p_n} = U(1/p_n)$, where $p_n \to 0$ and $p_n = O(1/n)$, as $n \to \infty$. Denoting by $X_{1,n} \leq \ldots \leq X_{n,n}$ the order statistics (o.s.'s) corresponding to our original sample, one semi-parametric estimator for χ_p can arise quite naturally, under the above framework, as the statistic

$$\widehat{\chi}_{p,k} \equiv \widehat{\chi}_{p,k}(\widehat{\gamma}) := X_{n-k,n}\left(\frac{k+1}{(n+1)p}\right)^{\widehat{\gamma}}, \tag{1.2}$$

introduced by Weissman (1978), where $\widehat{\gamma}$ is an estimator for the tail index γ, typically based on the $k+1$ top o.s.'s $X_{n-k,n} \leq \ldots \leq X_{n,n}$, and denoted by $\widehat{\gamma}_k$.

The classical semi-parametric estimator for the tail index which is commonly plugged into

(1.2), is the popular Hill (1975) estimator, defined as

$$\widehat{\gamma}_{H,k} = \frac{1}{k} \sum_{i=1}^{k} i \left(\ln X_{n-i+1,n} - \ln X_{n-i,n} \right), \tag{1.3}$$

the maximum likelihood tail index estimator under an exponential model (with mean value γ) for the scaled log-spacings $U_{i,k} = i\,(\ln X_{n-i+1,n} - \ln X_{n-i,n})$, for $1 \le i \le k$. Furthermore, as usual for semi-parametric estimators of extreme event parameters, in order to balance certain asymptotic requirements with respect to consistency of the considered estimators, we will also assume that $k = k_n$ is an intermediate sequence such that $k_n \to \infty$ and $k_n = o(n)$, as $n \to \infty$.

In recent literature, it has become a well-established fact that the main source of bias for the semi-parametric quantile estimator (1.2) can be attributed to the bias introduced through the estimation of the tail index γ. This bias, in turn, mostly occurs due to a slow convergence rate in limiting result (1.1) of first order condition $U \in \mathcal{R}_\gamma$. Using the above maximum likelihood point of view, for instance, the assumption that above a certain high threshold $X_{n-k,n}$ the scaled log-spacings $U_{i,k}$, $1 \le i \le k$, behave as data from an exponential distribution, is sometimes over-optimistic and in practice often results in severe bias for values of k that are too large.

Our main aim now is essentially to present new estimators for χ_p in the lines of both Gomes and Figueiredo (2003) and Matthys *et al.* (2004), i.e., based on adequate reduced bias estimation of the tail index γ and a direct accommodation of the bias for high quantiles, respectively, in a wide sub-class of Hall's class of models (Hall and Welsh, 1985). The key feature of these new estimators exists in the fact that the estimation of the second order parameters in the respective bias-terms, described in section 2 of this paper, is performed at a level k_1 of a larger order than the level k at which the parameters γ and χ_p are classically estimated. Doing this, we are able to guarantee a mean squared error smaller than that of the classical estimators for all levels k and both for tail index and quantile estimation. In addition, the asymptotic behavior of the new classes of tail index and of quantile estimators will be derived under appropriate higher order conditions, in sections 3 and 4, respectively. The Monte Carlo simulation study in section 5 will enable us to obtain some of the features of these new estimators for finite samples and an illustration of the behavior of the estimators will be provided, in section 6, for a set of real data in the field of insurance (automobile claims from a European car insurance portfolio).

2 A brief description of second order parameters' estimators

In order to be able to correctly assess the asymptotic non-degenerate behavior of semi-parametric estimators of extreme event parameters as above, we need more than just the first order condition $U \in \mathcal{R}_\gamma$. A convenient refinement can be found in the assumption that there exists a constant $\rho \le 0$ and a function $A(t) \to \infty$, as $t \to \infty$, with constant sign for large values of t, such that

$$\lim_{t\to\infty} \frac{1}{A(t)} \left(\frac{U(tx)}{U(t)} - x^\gamma \right) = x^\gamma \ \frac{x^\rho - 1}{\rho} \tag{2.1}$$

for all $x > 0$. The limiting function in (2.1) is necessarily of the above form and $|A|$ is regularly varying at infinity with index ρ (Geluk and de Haan, 1987). This so-called second order condition (which we will denote by $U \in \mathcal{R}_\gamma^\rho$) can easily be understood to specify the rate of convergence in the limiting result (1.1) of the first order condition $U \in \mathcal{R}_\gamma$. We shall often further assume that we are working in the sub-class of Hall's class of models (Hall and Welsh, 1985), characterized by the existence of constants $\gamma > 0$, $\rho < 0$, $C > 0$, $D \neq 0$ and $\beta \neq 0$ such that

$$U(t) = Ct^\gamma \left(1 + \frac{\gamma\beta}{\rho} t^\rho + Dt^{2\rho} + o\big(t^{2\rho}\big) \right), \tag{2.2}$$

as $t \to \infty$, where, comparatively to Hall's class of models, we merely made explicit a third order term of power 2ρ. Then, or more generally in Hall's class of models, we may consider the parameterization

$$A(t) := \gamma \ \beta \ t^\rho, \quad \beta \neq 0, \ \rho < 0, \tag{2.3}$$

for the regularly varying function $A(\cdot)$ in (2.1), and we are interested in the estimation of the shape second order parameter ρ and the scale second order parameter β in (2.3). We note that most common heavy-tailed distributions, like the Fréchet, the Burr and the Student-t distribution, belong to the class in (2.2), and consequently belong to Hall's class of models.

2.1 Estimation of the shape second order parameter

We consider particular members of the class of estimators for the second order parameter ρ as proposed in Fraga Alves *et al.* (2003). As first argued in Caeiro and Gomes (2004), these estimators are understood to depend on a real tuning parameter τ and the corresponding statistic

$$T_{n,k}^{(\tau)} = \begin{cases} \dfrac{\left(M_{n,k}^{(1)}\right)^{\tau} - \left(M_{n,k}^{(2)}/2\right)^{\tau/2}}{\left(M_{n,k}^{(2)}/2\right)^{\tau/2} - \left(M_{n,k}^{(3)}/6\right)^{\tau/3}} & \text{if} \quad \tau \neq 0 \\ \\ \dfrac{\ln\left(M_{n,k}^{(1)}\right) - \frac{1}{2}\ln\left(M_{n,k}^{(2)}/2\right)}{\frac{1}{2}\ln\left(M_{n,k}^{(2)}/2\right) - \frac{1}{3}\ln\left(M_{n,k}^{(3)}/6\right)} & \text{if} \quad \tau = 0 \end{cases},$$

where for any positive real ξ, we define

$$M_{n,k}^{(\xi)} = \frac{1}{k} \sum_{i=1}^{k} \{\ln X_{n-i+1,n} - X_{n-k,n}\}^{\xi}.$$

The statistics $T_{n,k}^{(\tau)}$ can be seen to converge towards $3(1-\rho)/(3-\rho)$, independently of τ, whenever the second order condition $U \in \mathcal{R}_{\gamma}^{\rho}$ holds, and the intermediate sequence k is such that $\sqrt{k}\; A(n/k) \to \infty$, as $n \to \infty$.

From this, a suitable class of estimators for the second order parameter ρ can then easily be defined as

$$\widehat{\rho}_{\tau,k} = \min\left(0,\; \frac{3\left(T_{n,k}^{(\tau)} - 1\right)}{T_{n,k}^{(\tau)} - 3}\right). \tag{2.4}$$

Under adequate general conditions, these semi-parametric estimators are consistent and asymptotically normal and show highly stable sample paths for a wide range of large k-values, the number of top o.s.'s used. Below, we state without a proof, a particular case of the main theorem in Fraga Alves *et al.* (2003), where a more general result may be found.

Proposition 2.1. *If the second order condition $U \in \mathcal{R}_{\gamma}^{\rho}$ holds, and the intermediate sequence k_1 is such that $\sqrt{k_1}\; A(n/k_1) \to \infty$, as $n \to \infty$, then the statistics $\widehat{\rho}_{\tau} \equiv \widehat{\rho}_{\tau,k_1}$ in* (2.4) *converge in probability towards ρ, as $n \to \infty$, for any $\tau \in \mathbb{R}$. Furthermore, for models in* (2.2)*, the choice*

$$k_1 = \left[n^{0.995}\right] \tag{2.5}$$

enables us to guarantee that for any intermediate sequence k, we have $(\widehat{\rho}_{\tau} - \rho)\ln(n/k) = o_p(1)$, as $n \to \infty$, provided that $|\rho| < 99.5$, which is an almost irrelevant restriction.

Further theoretical and simulated results in Fraga Alves *et al.* (2003), together with the use of these estimators in different reduced bias statistics, has led different authors to advise in practice the drawing of a few sample paths of $\widehat{\rho}_{\tau,k}$ (for some τ-values), electing the value of τ which provides the highest stability for large values of k, by means of any stability criterion, like for instance the ones in Gomes and Figueiredo (2003) or Gomes and Pestana (2004). Indeed, the adequate choice of τ is more crucial than the choice of the level k_1. Next to the use of k_1, in (2.5), in the simulations of section 5 we have essentially restricted the choice of the tuning parameter in (2.4) between $\tau = 0$ and $\tau = 1$, with the advise of considering $\tau = 0$ whenever $\rho \in [-1, 0)$ and $\tau = 1$ for $\rho \in (-\infty, -1)$.

As already mentioned in Proposition 2.1, we will denote generically $\widehat{\rho}_\tau$ any of the estimators in (2.4), computed at the level k_1 in (2.5).

2.2 Estimation of the scale second order parameter

With respect to the estimation of second order parameter β, we consider the class of estimators as proposed in the maximum likelihood set-up of Gomes and Martins (2002). There, relying on the fact that in Hall's class of models the scaled log-spacings $U_{i,k}$ are approximately exponentially distributed with mean value $\gamma\left(1 + \beta\left(\frac{i}{n+1}\right)^{-\rho}\right)$, for $1 \le i \le k$, the joint maximization in γ, β and ρ of the corresponding log-likelihood quite naturally leads to the definition of

$$\widehat{\beta}_{\widehat{\rho},k} = \left(\frac{k+1}{n+1}\right)^{\widehat{\rho}} \frac{d_{\widehat{\rho},k}\ D_{0,k} - D_{\widehat{\rho},k}}{d_{\widehat{\rho},k}\ D_{\widehat{\rho},k} - D_{2\widehat{\rho},k}}, \tag{2.6}$$

as a suitable estimator for the scale second order parameter β, where, for any non-positive real α, $d_{\alpha,k}$ and the statistic $D_{\alpha,k}$ are defined by

$$d_{\alpha,k} = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k+1}\right)^{-\alpha} \quad \text{and} \quad D_{\alpha,k} = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k+1}\right)^{-\alpha} U_{i,k}\ . \tag{2.7}$$

We use the simple notation $\widehat{\beta}_\tau = \widehat{\beta}_{\widehat{\rho}_\tau,k_1}$ to denote the above estimator, computed at the same level k_1 at which we compute the estimator $\widehat{\rho}_\tau$ as defined in (2.4). We now formalize, without proofs, the main distributional result needed and available in the literature, related to the class of estimators in (2.6). For more details, see for instance Gomes and Martins (2002) and Gomes *et al.* (2004b).

Proposition 2.2. *If $U \in \mathcal{R}_\gamma^\rho$, with $A(t)$ specified as in (2.3), and k_1 is an intermediate sequence such that $\sqrt{k_1}\ A(n/k_1) \to \infty$, as $n \to \infty$, then the statistic $\widehat{\beta}_{\widehat{\rho},k_1}$ in (2.6) is consistent for the estimation of β provided that $(\widehat{\rho} - \rho)\ln(n/k_1) = o_p(1)$, as $n \to \infty$.*

3 Second order reduced bias tail index estimation

The classical tail index estimates, like the Hill estimates in (1.3), exhibit, most of the times, a strong bias for moderate values of k and sample paths with very short stability regions around the target value. Theoretically, this is due to the fact that whenever $U \in \mathcal{R}_\gamma^\rho$ and we consider intermediate levels k such that $\sqrt{k}\ A(n/k) \to \lambda$, finite, as $n \to \infty$, we may write (de Haan and Peng, 1998),

$$\sqrt{k}\ \left(\widehat{\gamma}_{H,k} - \gamma\right) \stackrel{d}{=} \gamma P_k + \frac{\sqrt{k}A(n/k)}{1-\rho} + o_p\left(\sqrt{k}A(n/k)\right), \tag{3.1}$$

as $n \to \infty$, where P_k is an asymptotically standard normal r.v. This means that $\sqrt{k}\left(\widehat{\gamma}_{H,k} - \gamma\right)$ converges weakly towards a normal r.v., with mean value $\lambda/(1-\rho)$, possibly non-null.

The problem of reduced bias tail index estimation has been addressed recently by several authors, among whom we mention Peng (1988), Beirlant *et al.* (1999), Feuerverger and Hall (1999) and Gomes *et al.* (2000). All these researchers consider the possibility of dealing with the bias term in an appropriate way, building different new second order reduced bias estimators. For this type of estimators, $\widehat{\gamma}_{R,k}$ say, and under the same conditions as above, with $\widetilde{P}_k$ an asymptotically standard normal r.v. and $\sigma_R > 0$, we may write

$$\sqrt{k}\left(\widehat{\gamma}_{R,k} - \gamma\right) \stackrel{d}{=} \sigma_R \widetilde{P}_k + o_p\left(\sqrt{k}A(n/k)\right), \tag{3.2}$$

as $n \to \infty$. Note that in the above mentioned classes of reduced bias tail index estimators we have $\sigma_R \geq \gamma(1-\rho)/|\rho|$, the minimal asymptotic standard deviation in Drees' class of models (Drees, 1998), i.e. $\sigma_R > \gamma$, the asymptotic standard deviation of the Hill estimator. Under the same conditions as above, $\sqrt{k}\left(\widehat{\gamma}_{R,k} - \gamma\right)$ has then an asymptotic null mean value even for $\lambda \neq 0$. Furthermore, under mild additional conditions (e.g. Gomes *et al.*, 2004a), we may even guarantee the asymptotic normality of these estimators for levels k such that $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$.

Among the "asymptotically unbiased" or second order reduced bias tail index estimators considered in Gomes and Figueiredo (2003) for quantile estimation, we make use here of the one with the smallest asymptotic variance, i.e. the "maximum likelihood" estimator for γ under the same exponential model for the scaled log-spacings as used above for estimating the second order parameter β, and introduced in Gomes and Martins (2002) as

$$M_{\widehat{\rho},k} = D_{0,k} - D_{\widehat{\rho},k}\, \frac{d_{\widehat{\rho},k}\ D_{0,k} - D_{\widehat{\rho},k}}{d_{\widehat{\rho},k}\ D_{\widehat{\rho},k} - D_{2\widehat{\rho},k}}, \tag{3.3}$$

with $d_{\alpha,k}$ and $D_{\alpha,k}$ provided in (2.7). Note that the estimator depends on the estimation of the shape second order parameter ρ and the statistic $D_{0,k}$ corresponds to the previously discussed Hill estimator $\widehat{\gamma}_{H,k}$, in (1.3). An estimator of the type of the one in (3.3), but implicit, was first introduced in Beirlant *et al.* (1999) and Feuerverger and Hall (1999), and has been studied (with a misspecification of ρ at $\rho = -1$) in Gomes and Martins (2004). The estimator in (3.3) attains the minimal asymptotic variance in Drees' class of functionals. Indeed, we may state:

Proposition 3.1 (Gomes and Martins, 2002)**.** *If the second order condition $U \in \mathcal{R}_\gamma^\rho$ holds, with $\rho < 0$, and k is an intermediate sequence such that $\sqrt{k}\ A(n/k) \to \lambda$, finite and non necessarily null, as $n \to \infty$, then*

$$\sqrt{k}\,(M_{\rho,k} - \gamma) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(0,\ \frac{\gamma^2(1-\rho)^2}{\rho^2}\right).$$

Provided that $\widehat{\rho}$ is consistent for the estimation of ρ, the same limiting behavior also holds if we consider $M_{\widehat{\rho},k}$ instead of $M_{\rho,k}$.

Considering the estimators $\widehat{\gamma}_{H,k}$, $\widehat{\beta}_{\widehat{\rho},k}$ and $M_{\widehat{\rho},k}$ as in (1.3), (2.6) and (3.3), respectively, we can write

$$M_{\widehat{\rho},k} = \widehat{\gamma}_{H,k} - \widehat{\beta}_{\widehat{\rho},k}\left(\frac{n}{k}\right)^{\widehat{\rho}} D_{\widehat{\rho},k}.$$

In a spirit similar to the one in Gomes *et al.* (2004b) and Caeiro *et al.* (2004), Gomes *et al.* (2005) have considered the estimator $\widehat{\beta} = \widehat{\beta}_{\widehat{\rho},k_1}$ for the scale second order parameter, i.e. computed at an intermediate higher level k_1, suggesting the consideration of the estimator

$$\overline{M}_{\widehat{\beta},\widehat{\rho},k} = \widehat{\gamma}_{H,k} - \widehat{\beta}_{\widehat{\rho},k_1}\left(\frac{n}{k}\right)^{\widehat{\rho}} D_{\widehat{\rho},k}, \tag{3.4}$$

for a suitable ρ-estimator. This is one of the tail index estimators we shall consider here for quantile estimation.

Apart from the estimator $\overline{M}_{\widehat{\beta},\widehat{\rho},k}$ in (3.4), we shall also consider, now in a spirit similar to the one used in Gomes and Pestana (2004), the computation of $D_{\widehat{\rho},k}$, a consistent estimator of $\gamma/(1-\rho)$, at its estimated optimal level. From Gomes and Martins (2004), we know that if second order condition $U \in \mathcal{R}_\gamma^\rho$ and k is a intermediate sequence, then

$$D_{\alpha,k} \stackrel{d}{=} \frac{\gamma}{1-\alpha} + \frac{\gamma}{\sqrt{(1-2\alpha)\ k}}\ \overline{Z}_k^{(\alpha)} + \frac{A(n/k)}{1-\alpha-\rho}\big(1+o_p(1)\big), \tag{3.5}$$

for any $\alpha \le 0$, where $\overline{Z}_k^{(\alpha)}$ is asymptotically standard normal. Consequently, if $A(t)$ may be specified as in (2.3), the asymptotic optimal level for $D_{\rho,k}$ is provided by

$$k_0 = k_0(n) = \left(\frac{(1-2\ \rho)\ n^{-2\ \rho}}{-2\ \rho\ \beta^2}\right)^{1/(1-2\ \rho)}$$

and with the obvious notation $\widehat{k}_0$, we can define

$$\overline{\overline{M}}_{\widehat{\beta},\widehat{\rho},k} = \widehat{\gamma}_{H,k} - \widehat{\beta}_{\widehat{\rho},k_1}\left(\frac{n}{k}\right)^{\widehat{\rho}} D_{\widehat{\rho},\widehat{k}_0}, \tag{3.6}$$

again for an adequate consistent ρ-estimator. Using the notation $\widetilde{M}$ to denote either $\overline{M}$ or $\overline{\overline{M}}$, we may state:

Proposition 3.2. *If the second order condition $U \in \mathcal{R}_\gamma^\rho$ holds, with $A(t)$ given as in (2.3), and k is an intermediate sequence such that $\sqrt{k}\ A(n/k) \to \lambda$, finite and non necessarily null, as $n \to \infty$, then*

$$\sqrt{k}\left(\widetilde{M}_{\beta,\rho,k} - \gamma\right) \underset{n\to\infty}{\stackrel{d}{\longrightarrow}} Normal\left(0,\gamma^2\right).$$

Provided that $(\widehat{\rho}-\rho)\ln(n/k) = o_p(1)$, as $n \to \infty$, being $\widehat{\beta}$ consistent for the estimation of β, the same limiting behavior holds if we consider $\widetilde{M}_{\widehat{\beta},\widehat{\rho},k}$ instead of $\widetilde{M}_{\beta,\rho,k}$.

Proof. Since we may write $A(t) = \gamma\ \beta\ t^\rho$, we have that

$$\overline{M}_{\beta,\rho,k} = \widehat{\gamma}_{H,k} - \frac{A(n/k)}{\gamma}\ D_{\rho,k} \quad \text{and} \quad \overline{\overline{M}}_{\beta,\rho,k} = \widehat{\gamma}_{H,k} - \frac{A(n/k)}{\gamma}\ D_{\rho,k_0}.$$

From (3.1) and (3.5), denoting $\widetilde{M}$, either $\overline{M}$ or $\overline{\overline{M}}$,

$$\begin{aligned}\widetilde{M}_{\beta,\rho,k} &\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ \overline{Z}_k^{(1)} + \frac{A(n/k)}{1-\rho} + o_p(A(n/k)) - \frac{A(n/k)}{\gamma} \times \left(\frac{\gamma}{1-\rho} + o_p(1)\right)\\ &\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ \overline{Z}_k^{(1)} + o_p(A(n/k),\end{aligned}$$

and the first part of the proposition follows. In the lines of Gomes *et al.* (2005), now under a second order framework, and with $D'_{\rho,k} = -\sum_{i=1}^{k}(i/(k+1))^{-\rho}\ln(i/(k+1))\ U_i/k$, we may use the delta-method (Casela and Berger, 2002, pages 240-245), and due to the fact that $\widehat{\rho}-\rho = o_p(1/\ln(n/k))$, we may write

$$\widetilde{M}_{\widehat{\beta},\widehat{\rho},k} - \widetilde{M}_{\beta,\rho,k} \overset{p}{\sim} \left(\frac{\widehat{\beta}-\beta}{\beta}\right) A(n/k)\ D_{\rho,k} + (\widehat{\rho}-\rho)\,\frac{A(n/k)}{\gamma}\left\{\ln(n/k)\ D_{\rho,k} + D'_{\rho,k}\right\}.$$

This expression enables us to replace (β,ρ) by the estimators $(\widehat{\beta},\widehat{\rho})$, under the conditions in the proposition, and still guarantee the same asymptotic properties. □

Provided that we adequately choose $\widehat{\beta}$ and $\widehat{\rho}$, the reduced bias tail index estimators $\overline{M}_{\widehat{\beta},\widehat{\rho},k}$ and $\overline{\overline{M}}_{\widehat{\beta},\widehat{\rho},k}$ in (3.4) and (3.6), respectively, have thus a smaller bias than the classical Hill estimator for all k, without any increase in the asymptotic variance, which is kept equal to γ^2.

4 Reduced bias quantile estimation

When considering the classical quantile estimator $\widehat{\chi}_{p,k} \equiv \widehat{\chi}_{p,k}\left(\widehat{\gamma}_{H,k}\right)$, and with the notation $c_n := \frac{k+1}{(n+1)p}$, a sequence that goes to infinity with n, it is well-known that if we choose a level k such that $\sqrt{k}\ A(n/k) \to \lambda \neq 0$ and finite, as $n\to\infty$, then $\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{\chi}_{p,k}}{\chi_p}-1\right)$ is asymptotically normal, with a non-null bias given by $\lambda/(1-\rho)$ and a variance equal to γ^2.

We shall consider two types of reduced bias semi-parametric quantile estimators, to be introduced in subsections 4.1 and 4.2. Under the same conditions as above these estimators have a null bias and an asymptotic variance γ^2.

4.1 Alternative I, based on reduced bias tail index estimation

Gomes and Figueiredo (2003) suggest the use, in (1.2), of the reduced bias tail index estimators in Gomes and Martins (2001, 2002) and Gomes *et al.* (2004a), all of them second order reduced bias estimators with asymptotic variance $\sigma_R^2 > \gamma^2$. As a result, they have that, under the same mild restriction on k, $\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{\chi}_{p,k}}{\chi_p}-1\right)$ is asymptotically normal with a null bias even when $\sqrt{k}\ A(n/k)\to\lambda\neq 0$ and finite, as $n\to\infty$, but at expenses of a larger variance $\sigma_R^2>\gamma^2$.

We now consider the use, in (1.2), of the reduced bias tail index estimators $\overline{M}_{\widehat{\beta},\widehat{\rho},k}$ and $\overline{\overline{M}}_{\widehat{\beta},\widehat{\rho},k}$ in (3.4) and (3.6), respectively. Doing that we are able to guarantee that if we choose a level k such that $\sqrt{k}\ A(n/k) \to \lambda \neq 0$ and finite, as $n \to \infty$, then, under the same mild restrictions as before, $\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{\chi}_{p,k}}{\chi_p} - 1\right)$ is asymptotically normal, with a null bias and a variance equal to γ^2.

4.2 Alternative II, based on direct accommodation of the bias

Alternatively, trying to reduce the bias of the classical quantile estimator by going directly into the second order framework, in Matthys *et al.* (2004) the semi-parametric estimator defined as

$$\widehat{\chi}^*_{p,k} \equiv \widehat{\chi}^*_{p,k}(\widehat{\gamma}) := X_{n-k,n}\left(\frac{k+1}{(n+1)p}\right)^{\widehat{\gamma}} \exp\left(\widehat{a}_{n,k}\ \frac{\left(\frac{k+1}{(n+1)p}\right)^{\widehat{\rho}} - 1}{\widehat{\rho}}\right),$$

was introduced, where $(\widehat{\gamma}, \widehat{\rho}, \widehat{a}_{n,k})$ is a suitable estimator for $(\gamma, \rho, a_{n,k})$ and $a_{n,k} = A\left(\frac{n+1}{k+1}\right)$. There, these authors suggested the use of the maximum likelihood estimators $\widehat{\gamma}_k$, $\widehat{\rho}_k$ and $\widehat{a}_{n,k}$ under a second order exponential regression model for the scaled log-spacings $U_{i,k}$, with mean value $\gamma + a_{n,k}\left(\frac{i}{k+1}\right)^{-\rho}$, for $1 \le i \le k$.

In what follows, we assume to be working in Hall's class of models, we may thus choose $A(t)$ as in (2.3), and as a result, the quantile estimator above can be rewritten as

$$\widehat{\chi}^*_{p,k} = X_{n-k,n}\left(\frac{k+1}{(n+1)p}\right)^{\widehat{\gamma}} \exp\left(\widehat{\gamma}\ \widehat{\beta}\left(\frac{n+1}{k+1}\right)^{\widehat{\rho}}\ \frac{\left(\frac{k+1}{(n+1)p}\right)^{\widehat{\rho}} - 1}{\widehat{\rho}}\right), \tag{4.1}$$

with $\widehat{\gamma}$, $\widehat{\beta}$ and $\widehat{\rho}$ suitable estimators for γ, β and ρ. Also here, $\widehat{\gamma}$ will typically be an estimator based on the $k+1$ top o.s.'s (denoted by $\widehat{\gamma}_k$). Again, just as in subsection 4.1 and for obvious reasons, we consider, in (4.1), the reduced bias tail index estimators $\overline{M}_{\widehat{\beta},\widehat{\rho},k}$ and $\overline{\overline{M}}_{\widehat{\beta},\widehat{\rho},k}$ in (3.4) and (3.6), respectively.

4.3 Asymptotic behavior of the quantile estimators

Almost directly from Gomes and Figueiredo (2003), we may state:

Theorem 4.1. *Suppose that* $U \in \mathcal{R}_\gamma^\rho$ *and* $k = k_n$ *is an intermediate sequence such that* $c_n := \frac{k+1}{(n+1)p} \to \infty$ *and* $\frac{\ln c_n}{\sqrt{k}} \to 0$*, as* $n \to \infty$*. Then, with* $\widehat{\gamma}_k$ *any consistent estimator of the tail index* γ*, as* $n \to \infty$*, we have that*

$$\frac{\sqrt{k}}{\ln c_n}\left(\frac{\widehat{\chi}_{p,k}}{\chi_p} - 1\right) \stackrel{d}{=} \sqrt{k}\,\left(\widehat{\gamma}_k - \gamma\right) + O_p(1/\ln c_n) + o_p\big(\sqrt{k}\ A(n/k)\big). \tag{4.2}$$

If $(\widehat{\gamma}_k, \widehat{\beta}, \widehat{\rho})$ *is any consistent estimator of* (γ, β, ρ) *for which* $(\widehat{\rho} - \rho)\ln(n/k) = o_p(1)$*, as* $n \to \infty$*, we may replace, in* (4.2)*,* $\widehat{\chi}_{p,k}$ *in* (1.2) *by* $\widehat{\chi}^*_{p,k}$ *in* (4.1)*, getting the same distributional result.*

Proof. Since

$$\frac{X_{n-k,n}}{U\left(\frac{n+1}{k+1}\right)} \stackrel{d}{=} 1 + \frac{\gamma\ B_k}{\sqrt{k}} + o_p(A(n/k)),$$

with B_k asymptotically standard normal and $\chi_p = U(1/p) = U\left(\frac{n+1}{k+1}c_n\right)$, we have

$$\frac{\widehat{\chi}_{p,k}}{\chi_p} - 1 \stackrel{d}{=} (\widehat{\gamma}_k - \gamma)\ln c_n + \frac{\gamma\ B_k}{\sqrt{k}} - \frac{c_n^\rho - 1}{\rho}\ A(n/k) + o_p(A(n/k)),$$

and consequently

$$\frac{\widehat{\chi}_{p,k}}{\chi_p} - 1 \stackrel{d}{=} (\widehat{\gamma}_k - \gamma)\ln c_n + \frac{\gamma\ B_k}{\sqrt{k}} + \frac{A(n/k)}{\rho} + o_p(A(n/k)), \tag{4.3}$$

as $n \to \infty$. If on the other hand, we consider $\widehat{\chi}^*_{p,k}$, we get

$$\frac{\widehat{\chi}^*_{p,k}}{\chi_p} - 1 \stackrel{d}{=} (\widehat{\gamma}_k - \gamma)\ln c_n + \frac{\gamma\ B_k}{\sqrt{k}} - \frac{c_n^\rho - 1}{\rho}\ A(n/k) + \widehat{\gamma}_k\ \widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}\ \frac{c_n^{\widehat{\rho}} - 1}{\widehat{\rho}} + o_p(A(n/k)),$$

as $n \to \infty$. Now, $\widehat{\gamma}_k = \gamma(1 + o_p(1))$ and

$$\begin{aligned}\gamma\ \widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}\ \frac{c_n^{\widehat{\rho}} - 1}{\widehat{\rho}} - A(n/k)\ \frac{c_n^\rho - 1}{\rho} \ &\stackrel{p}{\sim}\ \left(\frac{\widehat{\beta} - \beta}{\beta}\right) A(n/k)\ \frac{c_n^\rho - 1}{\rho} \\ &+ (\widehat{\rho} - \rho)\ A(n/k)\left[\ln\left(\frac{n}{k}\right)\ \frac{c_n^\rho - 1}{\rho} + \frac{1}{\rho}\left(c_n^\rho \ln c_n - \frac{c_n^\rho - 1}{\rho}\right)\right].\end{aligned}$$

Since $(\widehat{\rho} - \rho)\ln(n/k) = o_p(1)$ and furthermore $(c_n^\rho - 1)/\rho \to -1/\rho$ and $c_n^\rho \ln c_n \to 0$, as $n \to \infty$, we have

$$\frac{\widehat{\chi}^*_{p,k}}{\chi_p} - 1 \stackrel{d}{=} (\widehat{\gamma}_k - \gamma)\ln c_n + \frac{\gamma\ B_k}{\sqrt{k}} + o_p(A(n/k)), \tag{4.4}$$

as $n \to \infty$. □

If we look at the asymptotic distributional representations in (4.2) we notice immediately that the main contribution in terms of bias is provided by a possible bias of $\widehat{\gamma}_k$ and it thus seems obvious that, both in (1.2) and (4.1), it is better to use a reduced bias tail index estimator rather than a classical one. Remark however that although the main contribution comes from the first term on the right of (4.2), we still expect to get slightly different results when using $\widehat{\chi}^*_{p,k}$ instead of $\widehat{\chi}_{p,k}$ in practice, due to the difference in the remainder terms (see equations (4.3) and (4.4)).

4.4 Some general comments

The results obtained in subsection 4.3 as well as the ones obtained in the Monte Carlo simulation in subsection 5.2 lead us to strongly advise the use of the quantile estimator in (4.1), with the tail index estimator $\overline{M}_k \equiv \overline{M}_{\widehat{\beta},\widehat{\rho},k}$ in (3.4), or the quantile estimator in (1.2), with the tail index estimator $\overline{\overline{M}}_k \equiv \overline{\overline{M}}_{\widehat{\beta},\widehat{\rho},k}$ in (3.6). Anyway, the estimator in (4.1), with the tail index estimator $M_k \equiv M_{\widehat{\rho},k}$ in (3.3), does also exhibit an interesting mean value stability around the true target value, particularly for all the simulated models with $\rho \neq -1$.

5 Simulated behavior of the tail index and quantile estimators

We have here implemented multi-sample Monte Carlo simulation experiments of size 5000×10 both for the tail index and quantile estimators. For any details on multi-sample simulation refer to Gomes and Oliveira (2001). Let us generically denote $T_{n,k}$ any statistic or r.v. dependent on k, the number of top o.s.'s to be used in an inferential procedure related to any parameter of extreme or even rare events. The optimal sample fraction for $T_{n,k}$ is denoted $k_0^T(n)/n$, with $k_0^T(n) := \arg\min_k MSE\,[T_{n,k}]$ and we shall use the common notation $T_{n0} := T_{n,k_0^T(n)}$.

5.1 Distributional properties of the tail index estimators

As mentioned before, we shall use the notation $\widehat{\beta}_j = \widehat{\beta}_{\widehat{\rho}_j,k_1}$, $\widehat{\rho}_j = \widehat{\rho}_{j,k_1}$, $j = 0,1$, with $\widehat{\rho}_{\tau,k}$, k_1 and $\widehat{\beta}_{\widehat{\rho},k}$ given in (2.4), (2.5) and (2.6), respectively. The estimators $\left(\widehat{\beta}_j,\ \widehat{\rho}_j\right)$ of $(\beta,\ \rho)$ have been incorporated in the tail index estimators under study, leading to $M_j \equiv M_{j,k} \equiv M_{\widehat{\rho}_j,k}$, $\overline{M}_j \equiv \overline{M}_{j,k} \equiv \overline{M}_{\widehat{\beta}_j,\widehat{\rho}_j,k}$ and $\overline{\overline{M}}_j \equiv \overline{\overline{M}}_{j,k} \equiv \overline{\overline{M}}_{\widehat{\beta}_j,\widehat{\rho}_j,k}$, $j = 0,1$. The simulations show that the tail index estimators $\overline{M}_j$ and $\overline{\overline{M}}_j$, with j equal to either 0 or 1, according as $|\rho| \le 1$ or $|\rho| > 1$, seem to work quite well, as illustrated in the sequel. Indeed, for all models simulated, the use of either $\overline{M}_1$ or $\overline{\overline{M}}_1$ always enables us to achieve a better performance than the one we get with the Hill estimator H. In a "blind" way, we might advise such a choice. But $\overline{M}_0$ and $\overline{\overline{M}}_0$ provide much better results than $\overline{M}_1$ and $\overline{\overline{M}}_1$, respectively, whenever $|\rho|$, unknown, is smaller than or equal to 1.

In Figures 1, 2 and 3, for samples of size $n = 1000$ from Fréchet(γ), $\gamma = 0.25$, and Burr(γ, ρ), $(\gamma, \rho) = (0.25, -0.5)$ and $(0.25, -2)$, parents, respectively, we show the simulated patterns of the scaled mean values, $E[\bullet]/\gamma$, and mean squared errors, $MSE[\bullet]/\gamma^2$, of $\overline{M}_{j,k}$ and $\overline{\overline{M}}_{j,k}$ in (3.4) and (3.6), respectively, together with the ones of the Hill estimator $\widehat{\gamma}_{H,k}$ in (1.3), denoted H, and the $M_{j,k}$ estimator in (3.3), $j = 0$ or 1 according as $|\rho| \le 1$ or $|\rho| > 1$. The Fréchet(γ) d.f. is $F(x) = \exp\{-x^{-1/\gamma}\}$, $x \ge 0$, $\gamma > 0$, and the Burr(γ, ρ) d.f. has the functional form, $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \ge 0$, $\gamma > 0$, $\rho < 0$. The mean values and mean squared errors of the estimators are based on the first replicate, with 5000 runs.

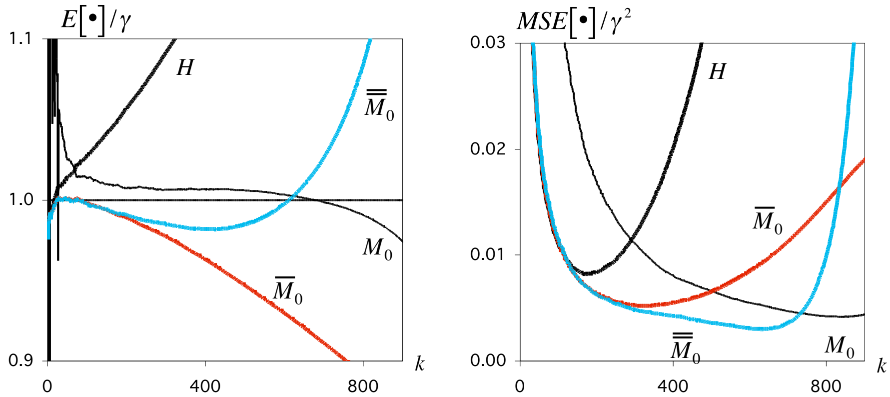

Figure 1: Simulated scaled mean values (*left*) and mean squared errors (*right*) of M_0, $\overline{M}_0$, $\overline{\overline{M}}_0$ and the Hill estimator H, for samples of size $n = 1000$ from a Fréchet parent with $\gamma = 0.25$ ($\rho = -1$)

We may draw the following specific comments:

- For a Fréchet model (with $\rho = -1$) (Figure 1) both the bias and the mean squared error of $\overline{M}_0$ and $\overline{\overline{M}}_0$ are smaller than the corresponding ones of the H-estimator, for all k. Both statistics are close to each other and close to the target value γ for small values of

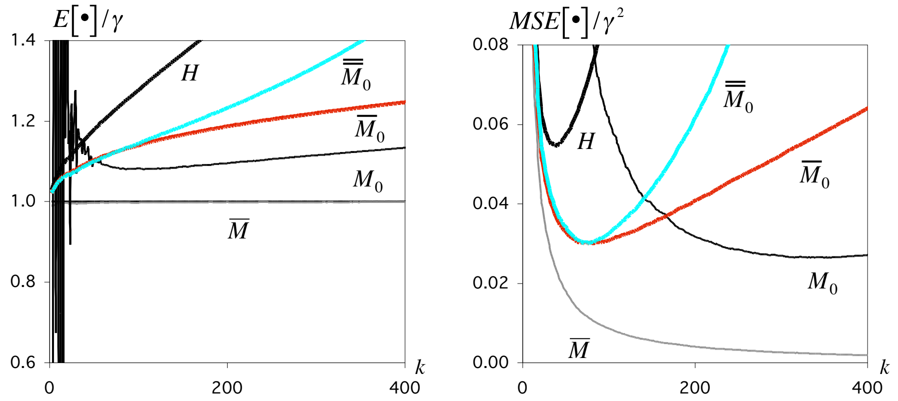

Figure 2: Simulated scaled mean values (*left*) and mean squared errors (*right*) of M_0, $\overline{M}_0$, $\overline{\overline{M}}_0$, the Hill estimator H and the r.v. $\overline{M} \equiv \overline{M}_{\beta,\rho}$, for samples of size $n = 1000$ from a Burr parent with $(\gamma, \rho) = (0.25, -0.5)$

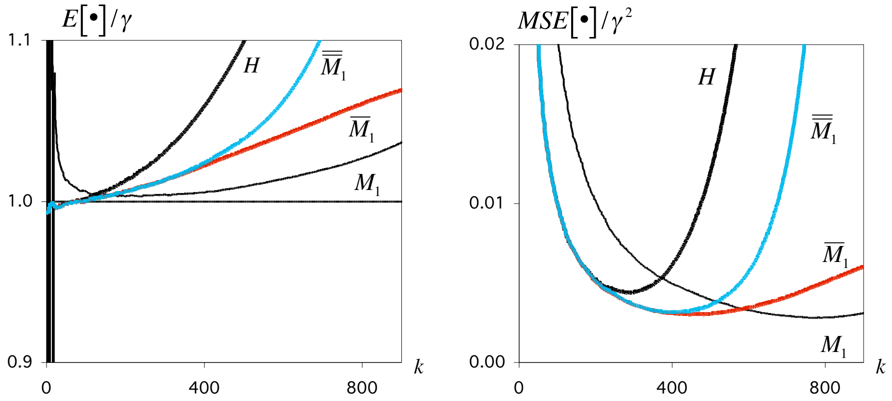

Figure 3: Simulated scaled mean values (*left*) and mean squared errors (*right*) of M_0, $\overline{M}_0$, $\overline{\overline{M}}_0$ and the Hill estimator H, for samples of size $n = 1000$ from a Burr parent with $(\gamma, \rho) = (0.25, -2)$

k. Next they diverge from each other, and it may be even sensible to investigate whether this feature would enable us to develop an adaptive statistical choice of the threshold. The statistic $\overline{\overline{M}}_0$ attains the smallest minimum mean squared error, among the ones considered. This type of behavior is more generally true for all simulated parents with $\rho = -1$.

- For values of $\rho > -1$ (Figure 2), and apart from the M_0-statistic, which exhibits a high volatility for small k, all the other statistics considered are positively biased for all k. The $\overline{M}_0$ and $\overline{\overline{M}}_0$ statistics are much better than the H-statistic, both regarding bias and mean squared error. In this region of ρ-values, further improvement in the second order parameters' estimation is still welcome. Indeed, and for comparison, we have also pictured in Figure 2 the analogue behavior of the r.v. $\overline{M} \equiv \overline{M}_{\beta,\rho}$: the improvement is obvious, with $\overline{M}$ behaving like an unbiased estimator of γ.

- For $\rho < -1$ (Figure 3), we need to use $\widehat{\rho}_1$ (instead of $\widehat{\rho}_0$). The associated $\overline{M}_1$ and $\overline{\overline{M}}_1$ statistics perform better than the H-statistic, but not a long way from it. Indeed, the Hill estimator already exhibits a quite interesting performance for this region of ρ-values.

5.2 Distributional properties of the quantile estimators

5.2.1 Mean values and mean squared error patterns

In Figures from 4 until 8, again on the basis of the first 5000 runs from a multi-sample experiment of size 5000×10, we show, for $p = 1/n$ and $j = 0$ or 1, the simulated patterns of mean value, $E[\bullet]$, and root mean squared error, $RMSE[\bullet]$, of some of the following normalized χ_p-estimators: $\widehat{\chi}_{p,k}(H)/\chi_p$, $\widehat{\chi}_{p,k}(M_j)/\chi_p$, $\widehat{\chi}_{p,k}(\overline{M}_j)/\chi_p$, $\widehat{\chi}_{p,k}(\overline{\overline{M}}_j)/\chi_p$, $\widehat{\chi}^*_{p,k}(M_j)/\chi_p$, $\widehat{\chi}^*_{p,k}(\overline{M}_j)/\chi_p$ and $\widehat{\chi}^*_{p,k}(\overline{\overline{M}}_j)/\chi_p$, based thus on the Hill estimator, H, and the "maximum likelihood" reduced bias estimators M_j, $\overline{M}_j$ and $\overline{\overline{M}}_j$, $j = 0, 1$, in (3.3), (3.4) and (3.6), respectively. For sake of simplicity, we shall denote these quotients by Q_H, Q_{M_j}, $Q_{\overline{M}_j}$, $Q_{\overline{\overline{M}}_j}$, $Q^*_{M_j}$, $Q^*_{\overline{M}_j}$ and $Q^*_{\overline{\overline{M}}_j}$, respectively, with $j = 0, 1$. The models underlying the simulated data, in Figures from 4 until 7, are the Fréchet model with $\gamma = 0.25$ and the Burr models with $(\gamma, \rho) = (0.25, -0.5), (0.25, -1)$ and $(0.25, -2)$, respectively, where we have pictured only the case $\tau = 0$ in (2.4). Figure 9 is the only one where we represent the scaled quantile estimators associated to $\tau = 1$ in (2.4) for the Burr model with $(\gamma,\ \rho) = (0.25,\ -2)$. In all figures, for any of the tail index estimators M, $\overline{M}_j$ and $\overline{\overline{M}}_j$, $j = 0$ or 1, we place usually only one of the scaled quantile estimators (either $Q_\bullet$ or $Q^*_\bullet$, the one providing the smaller minimum mean squared error).

Remark 5.1. *Note that, similarly to what has happened before with the tail index estimation, the computation of both second order parameters' estimators, at the high value k_1 in (2.5),*

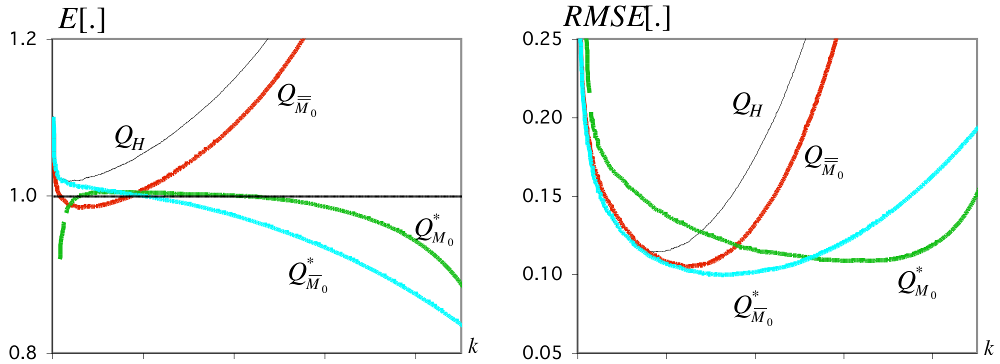

Figure 4: Underlying *Fréchet* parent with $\gamma = 0.25$ $(\rho = -1)$.

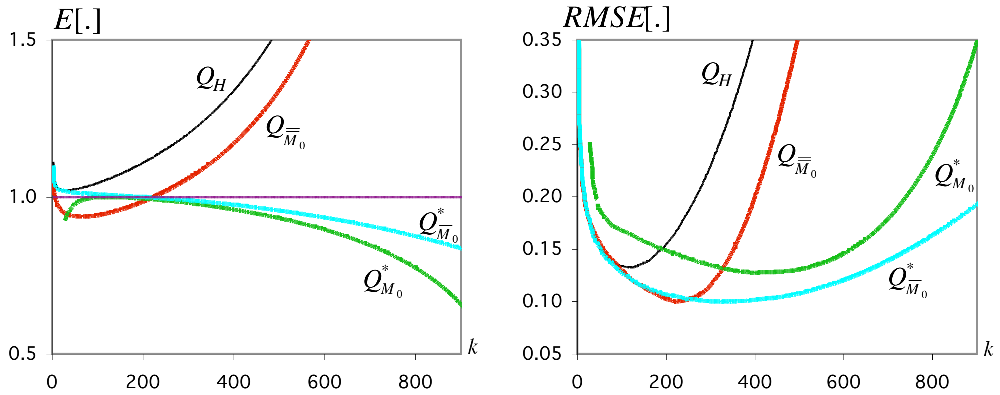

Figure 5: Underlying *Burr* parent with $\gamma = 0.25$ and $\rho = -1$.

enables us to work with high quantiles' estimators, with a mean squared error smaller than the mean squared error of the classical estimator $\widehat{\chi}_{p,k}(\widehat{\gamma}_{H,k})$*, with* $\widehat{\chi}_p$ *and* $\widehat{\gamma}_{H,k}$ *given in* (1.2) *and* (1.3)*, respectively, for most values of the threshold* k*. Those high quantile estimators are provided by the use in either* (1.2) *or* (4.1)*, of the tail index estimator* $\overline{M}$ *or* $\overline{\overline{M}}$ *in* (3.4) *and* (3.6)*, respectively.*

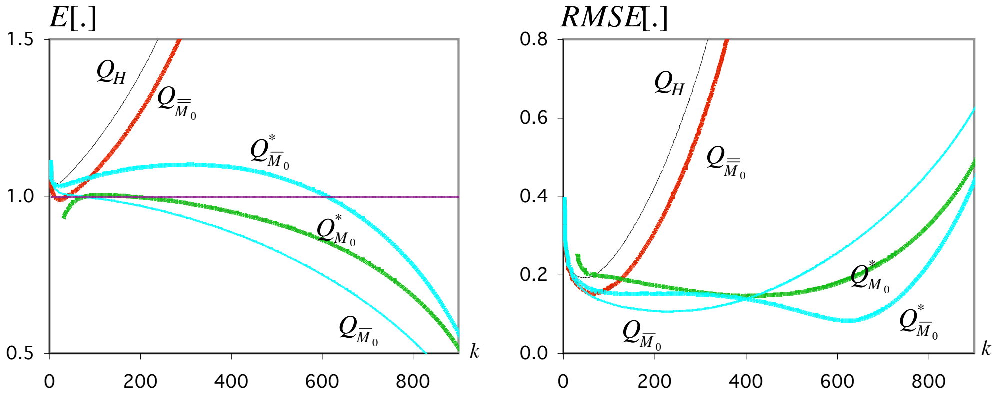

Figure 6: Underlying *Burr* parent with $\gamma = 0.25$ and $\rho = -0.5$.

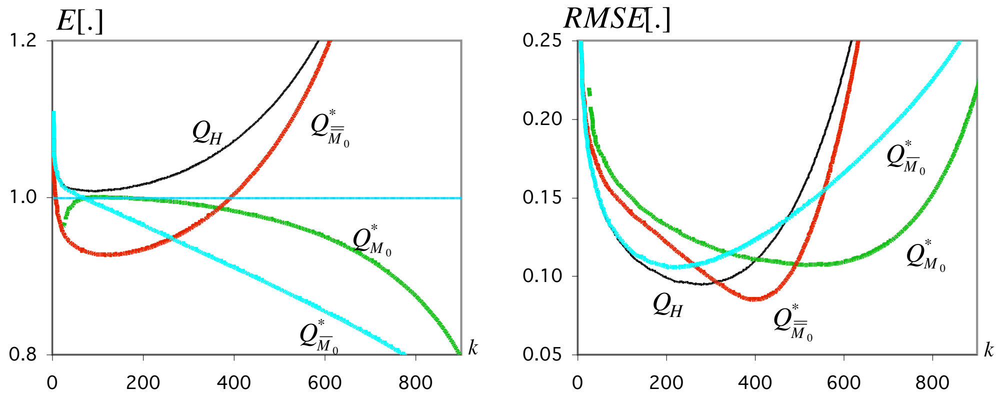

Figure 7: Underlying *Burr* parent with $\gamma = 0.25$ and $\rho = -2$ (estimators related to $\tau = 0$ in (2.4)).

5.2.2 Behavior of quantile estimators at optimal levels

We shall also present, for $n = 200,\ 500,\ 1000,\ 2000$ and 5000, the simulated mean values and relative efficiencies of $Q_\bullet$ and $Q^*_\bullet$, at their optimal levels, for the different second order reduced bias estimators under study. The search of the minimum mean squared error has been performed over the region of k-values between 1 and $[0.95 \times n]$. The root mean squared error ($RMSE$) of Q_H is also provided so that it is possible to recover the $RMSE$ of any other quantile estimator. For a certain Q_T, the $REFF^T_0$ indicator is given by

$$REFF^T_0 := \sqrt{MSE\left[Q_H(k^H_0)\right] / MSE\left[Q_T(k^T_0)\right]} =: \frac{RMSE^H_0}{RMSE^T_0}.$$

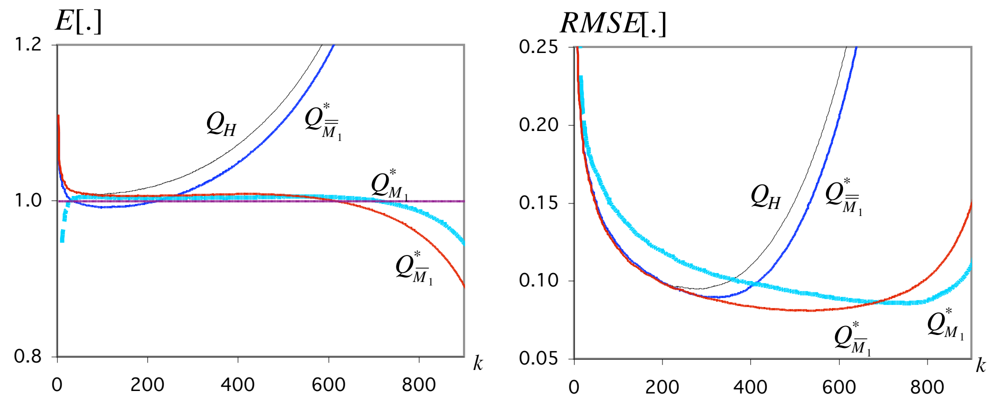

Figure 8: Underlying *Burr* parent with $\gamma = 0.25$ and $\rho = -2$ (estimators related to $\tau = 1$ in (2.4)).

Among the estimators considered, and for every n, the one providing the smallest mean squared error, or equivalently, the highest $REFF$ is underlined and in **bold**. Tables 1, 2, 3 and 4 are related to underlying Fréchet and Burr parents with $\rho = -1$, $\rho = -0.5$ and $\rho = -2$, respectively.

In summary we may draw the following conclusions:

1. The new asymptotically unbiased quantile estimators have in general reasonably stable sample paths, which make less troublesome the choice of the optimal level k.
2. The reduced bias quantile estimator $\widehat{\chi}^*_{p,k}(\overline{M}_0)$ is the one we elect for models with $\rho = -1$, but $\widehat{\chi}_{p,k}(\overline{\overline{M}}_0)$ exhibits also a quite interesting performance.
3. A similar comment applies to models with $\rho > -1$. In this region of ρ-values, $\widehat{\chi}_{p,k}(\overline{M}_0)$ is also worth considering as a valuable alternative to the aforementioned estimators.
4. Finally, for models with $\rho < -1$, the choice of $\widehat{\chi}^*_{p,k}(\overline{\overline{M}}_0)$ seems to be sensible as well as the choices $\widehat{\chi}^*_{p,k}(\overline{M}_1)$ and $\widehat{\chi}^*_{p,k}(\overline{\overline{M}}_1)$
5. The choice $\widehat{\chi}_{p,k}(\overline{M}_1)$ works also nicely for models with $\rho = -1$.

Table 1: Simulated mean values (E_0), root mean squared error of Q_H and relative efficiency measures ($REFF_0$) at optimal levels, together with corresponding 95% confidence intervals, for a Fréchet parent with $\gamma = .25$.

n	200	500	1000	2000	5000
E_0 **(and 95% confidence intervals)**					
Q_H	1.0553 ± 0.0041	1.0523 ± 0.0025	1.0479 ± 0.0022	1.0397 ± 0.0015	1.0364 ± 0.0016
Q_{M_0}	0.9096 ± 0.0055	0.9302 ± 0.0047	0.9460 ± 0.0034	0.9537 ± 0.0025	0.9634 ± 0.0014
$Q_{\overline{M}_0}$	0.9450 ± 0.0030	0.9620 ± 0.0014	0.9681 ± 0.0025	0.9755 ± 0.0012	0.9812 ± 0.0015
$Q_{\overline{\overline{M}}_0}$	1.0126 ± 0.0037	1.0239 ± 0.0036	1.0243 ± 0.0019	1.0242 ± 0.0021	1.0213 ± 0.0017
$Q^*_{M_0}$	0.9592 ± 0.0032	0.9735 ± 0.0014	0.9806 ± 0.0016	0.9861 ± 0.0017	0.9910 ± 0.0013
$Q^*_{\overline{M}_0}$	0.9638 ± 0.0025	0.9773 ± 0.0021	0.9841 ± 0.0014	0.9872 ± 0.0012	0.9920 ± 0.0010
$Q^*_{\overline{\overline{M}}_0}$	1.0270 ± 0.0034	1.0311 ± 0.0015	1.0328 ± 0.0022	1.0297 ± 0.0019	1.0260 ± 0.0013
Q_{M_1}	0.9461 ± 0.0030	0.9718 ± 0.0020	0.9885 ± 0.0017	0.9964 ± 0.0015	0.9999 ± 0.0013
$Q_{\overline{M}_1}$	1.0021 ± 0.0018	1.0331 ± 0.0024	1.0456 ± 0.0059	1.0217 ± 0.0117	0.9983 ± 0.0011
$Q_{\overline{\overline{M}}_1}$	1.0440 ± 0.0032	1.0439 ± 0.0023	1.0444 ± 0.0017	1.0379 ± 0.0029	1.0346 ± 0.0009
$Q^*_{M_1}$	0.9831 ± 0.0030	1.0155 ± 0.0054	1.0376 ± 0.0074	1.0256 ± 0.0099	1.0076 ± 0.0023
$Q^*_{\overline{M}_1}$	1.0341 ± 0.0062	1.0523 ± 0.0025	1.0430 ± 0.0116	1.0324 ± 0.0149	1.0049 ± 0.0027
$Q^*_{\overline{\overline{M}}_1}$	1.0437 ± 0.0040	1.0466 ± 0.0028	1.0445 ± 0.0023	1.0377 ± 0.0016	1.0356 ± 0.0008
$RMSE_0\,[Q_H]$	0.0264 ± 0.0007	0.0181 ± 0.0002	0.0134 ± 0.0002	0.00967 ± 0.0001	0.0063 ± 0.0001
$REFF_0$ **(and 95% confidence intervals)**					
Q_{M_0}	0.7624 ± 0.0093	0.7971 ± 0.0059	0.8293 ± 0.0058	0.8615 ± 0.0045	0.9136 ± 0.0085
$Q_{\overline{M}_0}$	1.1243 ± 0.0063	1.0905 ± 0.0047	1.0884 ± 0.0033	1.0857 ± 0.0030	1.0932 ± 0.0045
$Q_{\overline{\overline{M}}_0}$	0.9724 ± 0.0075	0.9836 ± 0.0053	1.0128 ± 0.0072	1.0485 ± 0.0063	1.1225 ± 0.0081
$Q^*_{M_0}$	0.9638 ± 0.0076	1.0074 ± 0.0051	1.0557 ± 0.0071	1.1133 ± 0.0063	1.2216 ± 0.0089
$Q^*_{\overline{M}_0}$	1.1110 ± 0.0079	1.1160 ± 0.0069	**1.1444** ± 0.0076	**1.1853** ± 0.0075	1.2770 ± 0.0117
$Q^*_{\overline{\overline{M}}_0}$	1.0787 ± 0.0050	1.0557 ± 0.0017	1.0501 ± 0.0022	1.0453 ± 0.0031	1.0512 ± 0.0014
Q_{M_1}	0.9879 ± 0.0090	1.0512 ± 0.0057	1.1072 ± 0.0087	1.1801 ± 0.0063	**1.3158** ± 0.0168
$Q_{\overline{M}_1}$	**1.1992** ± 0.0058	**1.1398** ± 0.0058	1.0718 ± 0.0045	1.0528 ± 0.0080	1.1632 ± 0.0209
$Q_{\overline{\overline{M}}_1}$	1.0646 ± 0.0040	1.0312 ± 0.0019	1.0202 ± 0.0026	1.0136 ± 0.0016	1.0119 ± 0.0004
$Q^*_{M_1}$	1.1113 ± 0.0056	1.0892 ± 0.0100	1.0479 ± 0.0147	1.0696 ± 0.0345	1.2725 ± 0.0730
$Q^*_{\overline{M}_1}$	1.1647 ± 0.0048	1.0654 ± 0.0040	1.0438 ± 0.0043	1.0487 ± 0.0270	1.2627 ± 0.0713
$Q^*_{\overline{\overline{M}}_1}$	1.0353 ± 0.0014	1.0160 ± 0.0003	1.0092 ± 0.0018	1.0069 ± 0.0004	1.0054 ± 0.0003

Table 2: Simulated mean values (E_0), root mean squared error of Q_H and relative efficiency measures ($REFF_0$) at optimal levels, together with corresponding 95% confidence intervals, for a Burr parent with $(\gamma, \rho) = (0.25, -0.5)$.

n	200	500	1000	2000	5000
E_0 **(and 95% confidence intervals)**					
Q_H	1.0843 ± 0.0067	1.0799 ± 0.0040	1.0805 ± 0.0065	1.0725 ± 0.0049	1.0700 ± 0.0034
Q_{M_0}	0.7874 ± 0.0079	0.8343 ± 0.0083	0.8594 ± 0.0049	0.8817 ± 0.0048	0.9037 ± 0.0030
$Q_{\overline{M}_0}$	0.9211 ± 0.0032	0.9413 ± 0.0022	0.9605 ± 0.0020	0.9759 ± 0.0014	0.9912 ± 0.0007
$Q_{\overline{\overline{M}}_0}$	0.9974 ± 0.0093	1.0204 ± 0.0062	1.0338 ± 0.0051	1.0376 ± 0.0040	1.0405 ± 0.0034
$Q^*_{M_0}$	0.9119 ± 0.0052	0.928 ± 0.0052	0.9452 ± 0.0031	0.9605 ± 0.0020	0.9776 ± 0.0013
$Q^*_{\overline{M}_0}$	0.9615 ± 0.0021	0.9821 ± 0.0016	0.9888 ± 0.0011	0.9936 ± 0.0008	0.9975 ± 0.0005
$Q^*_{\overline{\overline{M}}_0}$	1.0432 ± 0.0030	1.0443 ± 0.0033	1.0498 ± 0.0046	1.0457 ± 0.0038	1.0491 ± 0.0046
Q_{M_1}	0.9535 ± 0.0047	0.9834 ± 0.0040	0.9890 ± 0.0022	0.9908 ± 0.0013	1.1055 ± 0.0006
$Q_{\overline{M}_1}$	0.9843 ± 0.0016	0.9896 ± 0.0019	0.9925 ± 0.0012	1.0571 ± 0.0022	1.0698 ± 0.0038
$Q_{\overline{\overline{M}}_1}$	1.0806 ± 0.0074	1.0798 ± 0.0053	1.0797 ± 0.0064	1.0718 ± 0.0046	1.0699 ± 0.0038
$Q^*_{M_1}$	0.9729 ± 0.0056	0.9810 ± 0.0033	1.0642 ± 0.0017	1.0711 ± 0.0047	1.0679 ± 0.0036
$Q^*_{\overline{M}_1}$	0.9743 ± 0.0018	1.1179 ± 0.0070	1.0802 ± 0.0062	1.0723 ± 0.0045	1.0700 ± 0.0038
$Q^*_{\overline{\overline{M}}_1}$	1.0814 ± 0.0068	1.0787 ± 0.0048	1.0780 ± 0.0056	1.0719 ± 0.0049	1.0697 ± 0.0035
$RMSE_0\,[Q_H]$	0.0647 ± 0.0020	0.0477 ± 0.0009	0.0383 ± 0.0007	0.0303 ± 0.0004	0.0224 ± 0.0004
$REFF_0$ **(and 95% confidence intervals)**					
Q_{M_0}	0.7988 ± 0.0119	0.8409 ± 0.0103	0.8773 ± 0.0071	0.9120 ± 0.0082	0.9698 ± 0.0082
$Q_{\overline{M}_0}$	1.4638 ± 0.0253	1.6197 ± 0.0177	1.8397 ± 0.0200	2.1339 ± 0.0251	2.7468 ± 0.0327
$Q_{\overline{\overline{M}}_0}$	1.3062 ± 0.0141	1.2784 ± 0.0080	1.2604 ± 0.0077	1.2332 ± 0.0038	1.2149 ± 0.0061
$Q^*_{M_0}$	1.0853 ± 0.0162	1.1990 ± 0.0111	1.335 ± 0.0131	1.4995 ± 0.0154	1.8375 ± 0.0227
$Q^*_{\overline{M}_0}$	**1.7520** ± 0.0290	**2.0468** ± 0.0171	**2.3391** ± 0.0174	**2.6593** ± 0.0268	**3.2509** ± 0.0333
$Q^*_{\overline{\overline{M}}_0}$	1.1454 ± 0.0052	1.1242 ± 0.0029	1.1141 ± 0.0018	1.1050 ± 0.0019	1.0983 ± 0.0013
Q_{M_1}	1.3074 ± 0.0252	1.5311 ± 0.0152	1.7473 ± 0.0154	2.0543 ± 0.0211	1.2275 ± 0.0117
$Q_{\overline{M}_1}$	1.4805 ± 0.0259	1.6386 ± 0.0163	1.8414 ± 0.0208	1.6868 ± 0.0276	1.0026 ± 0.0004
$Q_{\overline{\overline{M}}_1}$	1.0169 ± 0.0015	1.0086 ± 0.0005	1.0056 ± 0.0005	1.0035 ± 0.0004	1.0021 ± 0.0004
$Q^*_{M_1}$	1.2407 ± 0.0211	1.4687 ± 0.0152	1.4287 ± 0.0157	1.0223 ± 0.0023	1.0259 ± 0.0052
$Q^*_{\overline{M}_1}$	1.2898 ± 0.0232	1.0989 ± 0.0251	1.0040 ± 0.0008	1.0024 ± 0.0006	1.0018 ± 0.0003
$Q^*_{\overline{\overline{M}}_1}$	1.0107 ± 0.0004	1.0054 ± 0.0000	1.0033 ± 0.0003	1.0022 ± 0.0000	1.0013 ± 0.0004

Table 3: Simulated mean values (E_0), root mean squared error of Q_H and relative efficiency measures ($REFF_0$) at optimal levels, together with corresponding 95% confidence intervals, for a Burr parent with $(\gamma, \rho) = (0.25, -1)$.

n	200	500	1000	2000	5000
			E_0 (and 95% confidence intervals)		
Q_H	1.0673 ± 0.0072	1.0566 ± 0.0037	1.0544 ± 0.0016	1.0478 ± 0.0032	1.0410 ± 0.0016
Q_{M_0}	0.8623 ± 0.0065	0.8887 ± 0.0053	0.9039 ± 0.0026	0.9199 ± 0.0039	0.9370 ± 0.0023
$Q_{\overline{M}_0}$	0.9236 ± 0.0032	0.9336 ± 0.0030	0.9412 ± 0.0020	0.9465 ± 0.0022	0.9548 ± 0.0016
$Q_{\overline{\overline{M}}_0}$	0.9940 ± 0.0033	1.0023 ± 0.0029	1.0021 ± 0.0021	1.0070 ± 0.0027	1.0075 ± 0.0020
$Q^*_{M_0}$	0.9383 ± 0.0063	0.9475 ± 0.0029	0.9539 ± 0.0024	0.9601 ± 0.0017	0.9661 ± 0.0018
$Q^*_{\overline{M}_0}$	0.9461 ± 0.0027	0.9523 ± 0.0026	0.9559 ± 0.0014	0.9593 ± 0.0011	0.9665 ± 0.0019
$Q^*_{\overline{\overline{M}}_0}$	1.0017 ± 0.0027	1.0086 ± 0.0044	1.0130 ± 0.0036	1.0137 ± 0.0020	1.0122 ± 0.0021
Q_{M_1}	0.9412 ± 0.0035	0.9614 ± 0.0023	0.9760 ± 0.0020	0.9893 ± 0.0014	0.9962 ± 0.0008
$Q_{\overline{M}_1}$	0.9877 ± 0.0025	0.9947 ± 0.0012	0.9964 ± 0.0011	0.9974 ± 0.0006	0.9983 ± 0.0006
$Q_{\overline{\overline{M}}_1}$	1.0557 ± 0.0032	1.0514 ± 0.0027	1.0494 ± 0.0038	1.0469 ± 0.0016	1.0406 ± 0.0014
$Q^*_{M_1}$	0.9800 ± 0.0019	0.9868 ± 0.0018	0.9919 ± 0.0017	0.9942 ± 0.0007	0.9977 ± 0.0004
$Q^*_{\overline{M}_1}$	0.9971 ± 0.0011	0.9948 ± 0.0016	0.9961 ± 0.0014	0.9970 ± 0.0010	1.0337 ± 0.0004
$Q^*_{\overline{\overline{M}}_1}$	1.0607 ± 0.0042	1.0513 ± 0.0035	1.0501 ± 0.0030	1.0460 ± 0.0029	1.0394 ± 0.0018
$RMSE_0\,[Q_H]$	0.0354 ± 0.0005	0.0243 ± 0.0004	0.0180 ± 0.0003	0.0133 ± 0.0001	0.0087 ± 0.0001
			$REFF_0$ (and 95% confidence intervals)		
Q_{M_0}	0.7523 ± 0.0055	0.7550 ± 0.0053	0.7565 ± 0.0045	0.7627 ± 0.0057	0.7643 ± 0.0059
$Q_{\overline{M}_0}$	1.1234 ± 0.0084	1.0996 ± 0.0077	1.0908 ± 0.0074	1.0799 ± 0.0073	1.0647 ± 0.0078
$Q_{\overline{\overline{M}}_0}$	1.2949 ± 0.0129	1.3192 ± 0.0044	1.3399 ± 0.0074	1.3663 ± 0.0072	1.3981 ± 0.0085
$Q^*_{M_0}$	0.9930 ± 0.0067	1.0238 ± 0.0051	1.0556 ± 0.0079	1.0878 ± 0.0078	1.1274 ± 0.0081
$Q^*_{\overline{M}_0}$	1.3270 ± 0.0112	1.3371 ± 0.0072	1.3416 ± 0.0107	1.3425 ± 0.0093	1.3323 ± 0.0098
$Q^*_{\overline{\overline{M}}_0}$	1.1924 ± 0.0064	1.1918 ± 0.0047	1.2020 ± 0.0075	1.2175 ± 0.0068	1.2368 ± 0.0045
Q_{M_1}	1.1437 ± 0.0107	1.2774 ± 0.0057	1.4247 ± 0.0105	1.6053 ± 0.0105	1.8833 ± 0.0153
$Q_{\overline{M}_1}$	**1.5085** ± 0.0196	**1.6404** ± 0.0128	**1.7567** ± 0.0154	**1.8914** ± 0.0092	**2.1133** ± 0.0153
$Q_{\overline{\overline{M}}_1}$	1.0535 ± 0.0030	1.0343 ± 0.0020	1.0286 ± 0.0016	1.0237 ± 0.0011	1.0187 ± 0.0008
$Q^*_{M_1}$	1.2383 ± 0.0092	1.3886 ± 0.0101	1.5254 ± 0.0124	1.6952 ± 0.0105	2.0010 ± 0.0162
$Q^*_{\overline{M}_1}$	1.4090 ± 0.0180	1.5390 ± 0.0122	1.6554 ± 0.0140	1.8010 ± 0.0096	1.6131 ± 0.0109
$Q^*_{\overline{\overline{M}}_1}$	1.0345 ± 0.0010	1.0234 ± 0.0008	1.0195 ± 0.0000	1.0160 ± 0.0005	1.0131 ± 0.0003

Table 4: Simulated mean values (E_0), root mean squared error of Q_H and relative efficiency measures ($REFF_0$) at optimal levels, together with corresponding 95% confidence intervals, for a Burr parent with $(\gamma, \rho) = (0.25, -2)$.

n	200	500	1000	2000	5000
E_0 **(and 95% confidence intervals)**					
Q_H	1.0447 ± 0.0026	1.0380 ± 0.0030	1.0333 ± 0.0019	1.0276 ± 0.0018	1.0219 ± 0.0013
Q_{M_0}	0.9104 ± 0.0035	0.9309 ± 0.0037	0.9476 ± 0.0038	0.9567 ± 0.0018	0.9672 ± 0.0020
$Q_{\overline{M}_0}$	0.9374 ± 0.0023	0.9459 ± 0.0041	0.9551 ± 0.0027	0.9609 ± 0.0016	0.9676 ± 0.0013
$Q_{\overline{\overline{M}}_0}$	0.9862 ± 0.0025	0.9956 ± 0.0019	0.9971 ± 0.0012	1.0002 ± 0.0012	1.0010 ± 0.0010
$Q^*_{M_0}$	0.9573 ± 0.0016	0.9620 ± 0.0032	0.9668 ± 0.0019	0.9723 ± 0.0016	0.9772 ± 0.0010
$Q^*_{\overline{M}_0}$	0.9462 ± 0.0022	0.9535 ± 0.0022	0.9597 ± 0.0017	0.9650 ± 0.0019	0.9709 ± 0.0017
$Q^*_{\overline{\overline{M}}_0}$	0.9987 ± 0.0036	1.0020 ± 0.0017	1.0026 ± 0.0022	1.0045 ± 0.0017	1.0037 ± 0.0008
Q_{M_1}	0.9451 ± 0.0031	0.9601 ± 0.0019	0.9677 ± 0.0020	0.9736 ± 0.0019	0.9795 ± 0.0008
$Q_{\overline{M}_1}$	0.9850 ± 0.0013	0.9994 ± 0.0014	1.0074 ± 0.0013	1.0128 ± 0.0007	1.0171 ± 0.0006
$Q_{\overline{\overline{M}}_1}$	1.0191 ± 0.0037	1.0214 ± 0.0028	1.0227 ± 0.0020	1.0195 ± 0.0018	1.0158 ± 0.0011
$Q^*_{M_1}$	0.9792 ± 0.0025	0.9901 ± 0.0017	0.9961 ± 0.0019	0.9992 ± 0.0019	0.9992 ± 0.0016
$Q^*_{\overline{M}_1}$	1.0172 ± 0.0023	1.0261 ± 0.0016	1.0257 ± 0.0013	1.0233 ± 0.0015	1.0193 ± 0.0009
$Q^*_{\overline{\overline{M}}_1}$	1.0213 ± 0.0031	1.0225 ± 0.0015	1.0214 ± 0.0010	1.0204 ± 0.0012	1.0158 ± 0.0013
$RMSE_0\,[Q_H]$	0.0207 ± 0.0004	0.0130 ± 0.0002	0.0090 ± 0.0002	0.0061 ± 0.0001	0.0036 ± 0.0000
$REFF_0$ **(and 95% confidence intervals)**					
Q_{M_0}	0.7127 ± 0.0055	0.7086 ± 0.0046	0.7109 ± 0.0048	0.7111 ± 0.0046	0.7129 ± 0.0054
$Q_{\overline{M}_0}$	0.8781 ± 0.0053	0.8326 ± 0.0059	0.8133 ± 0.0047	0.7967 ± 0.0046	0.7776 ± 0.0069
$Q_{\overline{\overline{M}}_0}$	1.0724 ± 0.0087	1.0692 ± 0.0073	1.0876 ± 0.0115	1.1163 ± 0.0058	1.1670 ± 0.0074
$Q^*_{M_0}$	0.8963 ± 0.0047	0.8903 ± 0.0041	0.8915 ± 0.0052	0.8911 ± 0.0055	0.8908 ± 0.0073
$Q^*_{\overline{M}_0}$	0.9887 ± 0.0044	0.9301 ± 0.0063	0.9010 ± 0.0054	0.8724 ± 0.0036	0.8434 ± 0.0081
$Q^*_{\overline{\overline{M}}_0}$	1.1109 ± 0.0070	1.1030 ± 0.0075	1.1102 ± 0.0080	1.1247 ± 0.0044	1.1512 ± 0.0069
Q_{M_1}	0.9366 ± 0.0042	0.9529 ± 0.0066	0.9729 ± 0.0068	0.9950 ± 0.0048	1.0251 ± 0.0102
$Q_{\overline{M}_1}$	**1.1621** ± 0.0061	**1.1534** ± 0.0084	**1.1600** ± 0.0130	**1.1711** ± 0.0067	1.1629 ± 0.0138
$Q_{\overline{\overline{M}}_1}$	1.0968 ± 0.0039	1.0704 ± 0.0044	1.0605 ± 0.0041	1.0564 ± 0.0037	1.0468 ± 0.0103
$Q^*_{M_1}$	1.0549 ± 0.0033	1.0775 ± 0.0086	1.1063 ± 0.0126	1.1494 ± 0.0073	**1.2182** ± 0.0202
$Q^*_{\overline{M}_1}$	1.1600 ± 0.0080	1.1237 ± 0.0056	1.1103 ± 0.0092	1.1069 ± 0.0041	1.0950 ± 0.0103
$Q^*_{\overline{\overline{M}}_1}$	1.0725 ± 0.0023	1.0519 ± 0.0023	1.0438 ± 0.0036	1.0416 ± 0.0022	1.0336 ± 0.0093

6 A case-study

We shall first consider an illustration of the performance of the above mentioned estimators, through the analysis of the $n = 371$ automobile claim amounts exceeding 1,200.000 Euro over the period 1988-2001 and gathered from several European insurance companies co-operating with the same re-insurer (Secura Belgian Re). This data set was studied both in Beirlant *et al.* (2004) and Vandewalle and Beirlant (2005) as an example to excess-of-loss reinsurance rating and heavy-tailed distributions in car insurance.

In Figure 9, we present the sample path of the $\widehat{\rho}_\tau$ estimates in (2.4) (*left*), as function of k, for $\tau = 0$ and $\tau = 1$, together with the sample paths of the β-estimators in (2.6), also for $\tau = 0$ and $\tau = 1$ (*right*).

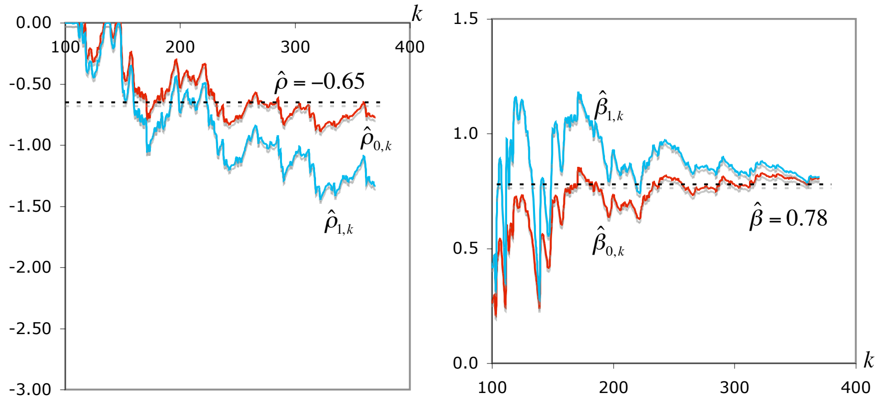

Figure 9: Estimates of the shape second order parameter ρ (*left*) and of the scale second order parameter β (*right*) for the Secura Belgian Re data.

Note that the sample paths of the ρ-estimates associated to $\tau = 0$ and $\tau = 1$ lead us to choose, on the basis of any stability criterion for large k, the estimate associated to $\tau = 0$. From previous experience with this type of estimates, we conclude that the underlying ρ-value is larger than or equal to -1, and the consideration of $\tau = 0$ is then advisable. The estimate of ρ is in this case $\widehat{\rho}_0 = -0.65$, obtained at the level $k_1 = 360$. The associated β-estimator is $\widehat{\beta}_0 = 0.78$.

In Figure 10, we present, at the left, the estimates of the tail index γ, provided by the Hill estimator, $\widehat{\gamma}_{H,k}$ in (1.3), denoted H, the $\overline{M}$-estimator in (3.4) and the $\overline{\overline{M}}$ estimator in (3.6). At the right we present the corresponding quantile estimators associated to $p = 0.001$, i.e., the quantile estimates Q_H, $Q^*_{\overline{M}}$ and $Q^*_{\overline{\overline{M}}}$.

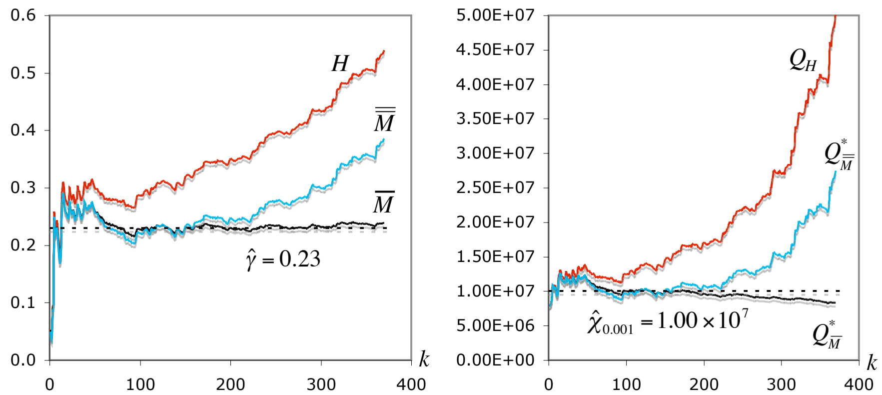

Figure 10: Estimates of the tail index γ (*left*) and of the quantile χ_p, associated to $p = 0.001$ (*right*) for the Secura Belgian Re data.

Regarding the tail index estimation, note that whereas the Hill estimator is unbiased for the estimation of the tail index γ when the underlying model is a strict Pareto model, it exhibits a relevant bias when we have only Pareto-like tails, as happens here, and may be seen from Figure 10 (*left*). The other estimators, which are "asymptotically unbiased" reveal a smaller bias, and enable us to take a decision upon the estimate of γ to be used, with the help of any heuristic stability criterion, like the "largest run" suggested in Gomes and Figueiredo (2003). For the Hill estimator, as we know how to estimate the second order parameters β and ρ, we have simple techniques to estimate the optimal sample fraction. Indeed, we get $\widehat{k}_0^H = \left((1-\widehat{\rho})^2 n^{-2\widehat{\rho}} / \left(-2\, \widehat{\rho}\, \widehat{\beta}^2\right)\right)^{1/(1-2\widehat{\rho})} = 58$. Unfortunately, we do not have yet the possibility of adaptively estimate the optimal sample fraction associated to the second order reduced bias estimates. The estimate pictured, $\widehat{\gamma} = 0.23$, is the median of the $\overline{M}(k)$ estimates for k between $\widehat{k}_0^H$ and $4 \times \widehat{k}_0^H$. A similar technique led us to the quantile estimate $\widehat{\chi}_{0.001} = 10,009.158$, as pictured in Figure 10.

References

[1] Beirlant, J., Dierckx, G., Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177-200.

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

[3] Caeiro, F. and Gomes, M. I. (2004). *A new class of estimators of the "scale" second order parameter.* Notas e Comunicações CEAUL 20/04. Submitted.

[4] Caeiro, F., Gomes, M. I. and Pestana, D. D. (2004). Direct reduction of bias of the classical Hill estimator. *RevStat* **3** (2), 111-136.

[5] Drees, H. (1998). A general class of estimators of the extreme value index. *J. Statist. Planning and Inference* **98**, 95-112.

[6] Ferreira, A., de Haan, L. and Peng, L. (2003). On optimising the estimation of high quantiles of a probability distribution. *Statistics* **37**(5), 401-434.

[7] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.

[8] 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, 194-213.

[9] Geluk, J. and L. de Haan (1987). *Regular Variation, Extensions and Tauberian Theorems.* CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, Netherlands.

[10] Gomes, M. I. and Figueiredo, F. (2003). Bias reduction in risk modelling: semi-parametric quantile estimation. To appear at *Test*.

[11] Gomes, M. I., Caeiro, F. and Figueiredo, F. (2004a). Bias reduction of a tail index estimator through an external estimation of the second order parameter. *Statistics* **38**(6), 497-510.

[12] Gomes, M. I., de Haan, L. and Rodrigues, L. (2004b). *Tail index estimation through accommodation of bias in the weighted log-excesses.* Notas e Comunicações C.E.A.U.L. 14/2004. Submitted.

[13] Gomes, M. I. and M. J. Martins (2001). Alternatives to Hill's estimator — asymptotic versus finite sample behaviour. *J. Statist. Planning and Inference* **93**, 161-180.

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

[15] Gomes, M. I. and M. J. Martins (2004). Bias reduction and explicit estimation of the tail index. *J. Statist. Planning and Inference* **124**, 361-378.

[16] Gomes, M. I., Martins, M. J. and Neves, M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events – the Jackknife methodology. *Extremes* **3**:3, 207-229.

[17] Gomes, M. I., Martins, M. J. and Neves, M. (2005). *Revisiting the second order reduced bias "maximum likelihood" tail index estimators.* Notas e Comunicações CEAUL 10/2005. Submitted.

[18] Gomes, M. I. and Pestana, D. (2004). A simple second order reduced bias' tail index estimator. To appear at *J. Statist. Comp. and Simul.*.

[19] Gomes, M. I. and Pestana, D. (2005). *A sturdy second order reduced bias' Value at Risk estimator.* Notas e Comunicações 05/2005. Submitted.

[20] Haan, L. de and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica* **52**, 60-70.

[21] Haan, L. de and H. Rootzén (1993). On the estimation of high quantiles. *J. Statist. Planning and Inference* **35**, 1-13.

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

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

[24] Matthys, G. and Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica* **13**, 853-880.

[25] Matthys, G., Delafosse, M., Guillou, A. and Beirlant, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions. *Insurance: Mathematics and Economics* **34**, 517-537.

[26] Peng, L.(1998). Asymptotically unbiased estimator for the extreme-value index. *Statistics and Probability Letters* **38**(2), 107-115.

[27] Vandewalle, B. and Beirlant, J. (2005). On univariate extreme value statistics and the estimation of reinsurance premiums. Accepted at *Insurance Mathematics and Economics.*

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