A sturdy second order reduced bias' Value at Risk estimator*

M. Ivette Gomes

University of Lisbon, DEIO, CEAUL and FCUL

and

Dinis Pestana

University of Lisbon, DEIO, CEAUL and FCUL

June 16, 2005

Abstract. The main objective of *Statistics of Extremes* lies in the estimation of quantities related to extreme events. In many areas of application, like for instance *Insurance*, *Finance* and *Statistical Quality Control*, a typical requirement is to estimate a high quantile, or equivalently, the *Value at Risk* at a level p (VaR_p), a value, high enough, so that the chance of an exceedance of that value is equal to p, small. In this paper we deal with the semi-parametric estimation of VaR_p, for heavy tails. Since the classical semi-parametric estimators of extreme events' parameters usually exhibit a reasonably high bias for low thresholds, i.e., for large values of k, the number of top order statistics used for the estimation, we shall here deal with bias reduction techniques for heavy tails, trying to improve the performance of the classical high quantile estimators. High quantiles depend strongly on the tail index γ. Recently, new interesting classes of reduced bias' γ-estimators have been introduced in the literature. In those classes, the second order parameters in the bias are estimated at a level k_1 of a larger order than that of the level k at which we compute the tail index estimators,

*Research partially supported by FCT / POCTI / FEDER.

and doing this, it is possible to keep the asymptotic variance of the new estimators equal to the asymptotic variance of the Hill estimator, the maximum likelihood tail index estimator, under a strict Pareto model. The use of one of those classes of γ-estimators in quantile estimation enables us to introduce new classes of high quantiles' estimators. The asymptotic distributional properties of the proposed classes of estimators are derived and the estimators are compared with alternative ones, not only asymptotically, but also for finite samples through Monte Carlo techniques. An application to the log-exchange rates of the Euro against the Sterling Pound is also provided.

AMS 2000 subject classification. Primary 62G32, 62E20; Secondary 65C05.

Keywords and phrases. *Heavy tails; High quantiles; Semi-parametric estimation; Statistics of Extremes.*

1 Introduction and preliminaries

A model F is said to belong to the *domain of attraction for maxima* ($\mathcal{D}_{\mathcal{M}}$) of an *Extreme Value* distribution function (d.f.),

$$EV_\gamma(x) = \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right), \ 1+\gamma x \geq 0 & \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)), \ x \in \mathbb{R} & \text{if} \quad \gamma = 0 \end{cases}, \tag{1.1}$$

if and only if the maximum $M_n = \max(X_1, X_2, \ldots, X_n)$, linearly normalized, of an independent, identically distributed (i.i.d.) sample of size n, converges weakly, as $n \to \infty$, towards a non-degenerate random variable (r.v.), with a d.f. necessarily of the type of EV_γ in (1.1), for some $\gamma \in \mathbb{R}$. We then write $F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma)$.

Let us denote

$$U(t) := F^{\leftarrow}(1 - 1/t) = \inf\{x : F(x) \geq 1 - 1/t\}$$

and let RV_α denote the class of regularly varying functions with index of regular variation equal to α, i.e., non-negative measurable functions $g(\cdot)$ such that, for all $x > 0$, $g(tx)/g(t) \to x^{\alpha}$, as $t \to \infty$.

Heavy-tailed models have revealed to be quite useful in the most diversified areas, like insurance, economics, finance, telecommunications and biostatistics, among others. A model F is said to be heavy-tailed if and only if we have $\gamma > 0$ in (1.1). It is well-known that for $\gamma > 0$,

$$F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma) \iff U \in RV_\gamma \iff 1 - F \in RV_{-1/\gamma}. \tag{1.2}$$

For small values of p, we want to estimate $\chi_p \equiv VaR_p$, a value such that $1 - F(VaR_p) = \mathbb{P}(X > VaR_p) = p$, with VaR standing for *Value at Risk*, a typical parameter in the areas of insurance and finance. More specifically, we want to estimate the parameter,

$$VaR_p = U\left(\frac{1}{p}\right), \quad p = p_n \to 0, \quad n\, p_n \to K, \text{ as } n \to \infty, \ K \in [0,1], \tag{1.3}$$

and we shall assume to be working in Hall's class of models (Hall and Welsh, 1985), where there exist $\gamma > 0$, $\rho < 0$, $C > 0$ and $\beta \neq 0$ such that

$$U(t) = Ct^{\gamma}\left(1 + \frac{\gamma\, \beta\, t^{\rho}}{\rho} + o(t^{\rho})\right), \qquad \text{as } t \to \infty. \tag{1.4}$$

We are then working in a wide subclass of $\mathcal{D}_{\mathcal{M}}(EV_\gamma)$, $\gamma > 0$. Such a class contains most of the heavy-tailed models important in applications, like the *Fréchet*, the *Generalized Pareto* and the *Student*-t_ν, with ν degrees of freedom.

We are going to base inference on the largest k top order statistics (o.s.), and as usual in semi-parametric estimation of parameters of extreme events, we shall assume that k is an *intermediate* sequence of integers in $[1, n]$, i.e.,

$$k = k_n \to \infty, \quad k \in [1, n], \qquad k = o(n) \qquad \text{as } n \to \infty. \tag{1.5}$$

Since, from (1.3) and (1.4), $VaR_p = U(1/p) \sim C\ p^{-\gamma}$, as $p \to 0$, an obvious estimator of VaR_p is $\widehat{VaR}_p = \widehat{C}\ p^{-\widehat{\gamma}}$, with $\widehat{C}$ and $\widehat{\gamma}$ any consistent estimators of C and γ, respectively. Given a sample $(X_1, X_2, \cdots, X_n)$, let us denote

$X_{i:n}$, $1 \leq i \leq n$, the set of associated ascending o.s. Denoting Y a standard Pareto model, i.e, a model such that $F_Y(y) = 1 - 1/y,\ y \geq 1$, the use of the universal uniform transformation enables us to write $X_{n-k:n} \stackrel{d}{=} U(Y_{n-k:n})$. Next, since $Y_{n-k:n} \stackrel{p}{\sim} (n/k)$ for intermediate k and (1.4) holds true, we get $X_{n-k:n} \stackrel{p}{\sim} C\, Y^{\gamma}_{n-k:n} \stackrel{p}{\sim} C\,(n/k)^{\gamma}$, as $n \to \infty$. Consequently, an obvious estimator of C is $\widehat{C} = (k/n)^{\widehat{\gamma}}\, X_{n-k:n}$, and

$$Q^{(p)}_{\widehat{\gamma}}(k) := X_{n-k:n} \left(\frac{k}{np}\right)^{\widehat{\gamma}} \tag{1.6}$$

is the obvious quantile or VaR-estimator at the level p (Weissman, 1978).

For heavy tails, the classical tail index estimator, usually the one which is plugged in (1.6), for a semi-parametric quantile estimation, is the Hill estimator $\widehat{\gamma} = \widehat{\gamma}(k) =: H(k)$ (Hill, 1975), with the functional expression,

$$H(k) := \frac{1}{k} \sum_{i=1}^{k} U_i, \tag{1.7}$$

$$U_i := i\,(\ln X_{n-i+1:n} - \ln X_{n-i:n})\,, \quad 1 \leq i \leq k. \tag{1.8}$$

We thus get the so-called classical quantile or VaR estimator, based on the Hill tail index estimator H, with the obvious notation, $Q^{(p)}_{H}(k)$.

We shall here work with ln-VaR instead of VaR, particularly due to the slightly skewed distribution of the VaR-estimators, like is shown in Table 1, where for different sample sizes n, from a Fréchet(γ) parent, with d.f. $F(x) = \exp\left(-x^{-1/\gamma}\right),\ x \geq 0$, $\gamma = 0.25$, we present the skewness and kurtosis of simulated samples of size $N = 1000$ from $\left\{Q^{(p)}_{H}(k_0)\right\}$ and $\left\{\ln Q^{(p)}_{H}(k_0)\right\}$, $p = 1/n$, $k_0 = 2\ n^{2/3}$. The skewness and kurtosis sample coefficients considered are the common ones, the sample counterparts of $\beta_1 = \mu_3/\mu_2^{3/2}$ and $\beta_2 = \mu_4/\mu_2^2 - 3$, which are null for a normal model, being $\mu_k = \mathbb{E}\,(X - \mathbb{E}X)^k$, $k \geq 1$, the k-th central moment of the underlying X.

We shall thus consider the classical ln-VaR estimator,

$$\ln Q^{(p)}_{H}(k) := \ln X_{n-k:n} + H(k)\ \ln\left(\frac{k}{np}\right), \tag{1.9}$$

	Skewness		Kurtosis	
n	Q_H	$\ln Q_H$	Q_H	$\ln Q_H$
100	0.8167	0.2689	1.5508	0.2478
200	0.8347	0.3739	1.3558	0.1987
500	0.5547	0.1994	0.5289	0.0866
1000	0.3512	0.0718	0.1119	-0.0353

Table 1: Skewness and kurtosis of simulated samples of size $N = 1000$ from $\left\{Q_H^{(p)}(k_0)\right\}$ and $\left\{\ln Q_H^{(p)}(k_0)\right\}$, $p = 1/n$, $k_0 = 2n^{2/3}$, in a Fréchet model with $\gamma = 0.25$.

with $H(k)$ the Hill estimator in (1.7).

In order to derive the asymptotic non-degenerate behaviour of semi-parametric estimators of extreme events' parameters, we need more than the first order condition in (1.2). As usual, we shall here assume that

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \tag{1.10}$$

for all $x > 0$, where A is a function of constant sign near infinity (positive or negative), and $\rho \le 0$ is the shape second order parameter. The limit function in (1.10) is necessarily of this given form, and $|A| \in RV_\rho$ (Geluk and de Haan, 1987). We shall assume throughout the paper that $\rho < 0$, and that we are working with models in (1.4), for which (1.10) holds true, with

$$A(t) = \gamma\ \beta\ t^\rho, \quad \gamma > 0,\ \beta \ne 0,\ \rho < 0. \tag{1.11}$$

Under the second order framework in (1.10), and for intermediate k, i.e., whenever (1.5) holds, we may guarantee the asymptotic normality of the Hill estimator $H(k)$, for an adequate k. Indeed, we may write (de Haan and Peng, 1998),

$$H(k) \ \stackrel{d}{=}\ \gamma + \frac{\gamma\ P_k}{\sqrt{k}} + \frac{A(n/k)}{1-\rho}(1 + o_p(1)), \tag{1.12}$$

$P_k = \sqrt{k}\left(\sum_{i=1}^k E_i/k - 1\right)$, with $\{E_i\}$ standard exponential i.i.d. r.v.'s. Consequently, if we choose a level k such that $\sqrt{k}\ A(n/k) \to \lambda \ne 0$, finite, as $n \to \infty$, $\sqrt{k}\,(H(k) - \gamma)$ is asymptotically normal, with variance equal to γ^2 and a non-null bias given by $\lambda/(1-\rho)$. Most of the times, this type of estimates exhibit a strong bias for moderate k and sample paths with very short stability regions around the target value γ. This has recently led researchers to consider the possibility of dealing with the bias term in an appropriate way, building new estimators, $\widehat{\gamma}_R(k)$ say, the so-called second order reduced bias' estimators discussed by Peng (1998), Beirlant *et al.* (1999), Feuerverger and Hall (1999), Gomes *et al.* (2000), among others. Then, for k intermediate, i.e., such that (1.5) holds, and under the second order framework in (1.10), we may write, with P_k^R an asymptotically standard normal r.v.,

$$\widehat{\gamma}_R(k) \overset{d}{=} \gamma + \frac{\sigma_R \, P_k^R}{\sqrt{k}} + o_p(A(n/k)),$$

where $\sigma_R > 0$, being $A(\cdot)$ again the function in (1.10). Consequently, the sequence of r.v.'s, $\sqrt{k}\,(\widehat{\gamma}_R(k) - \gamma)$ is asymptotically normal with variance equal to σ_R^2 and a null mean value even when $\sqrt{k}\ A(n/k) \to \lambda \neq 0$, finite, as $n \to \infty$.

Gomes and Figueiredo (2003) suggest the use, in (1.6), of reduced bias' tail index estimators, like the ones in Gomes and Martins (2001, 2002) and Gomes *et al.* (2004), being then able to reduce also the dominant component of the classical quantile estimator's asymptotic bias. Recently, Caeiro *et al.* (2004), consider the tail index estimator

$$\overline{H}_{\widehat{\beta},\,\widehat{\rho}}(k) := H(k)\left(1 - \frac{\widehat{\beta}}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right), \tag{1.13}$$

where $\left(\widehat{\beta},\ \widehat{\rho}\right)$ is an adequate consistent estimator of $(\beta,\ \rho)$, the vector of the second order parameters in (1.11). Notice that the dominant component of the bias of Hill's estimator in (1.12), $A(n/k)/(1-\rho) = \gamma\ \beta(n/k)^\rho/(1-\rho)$, is thus estimated through $H(k)\ \widehat{\beta}\ (n/k)^{\widehat{\rho}}/(1-\widehat{\rho})$ and directly removed from the Hill estimator $H(k)$ in (1.7). This is the tail index estimator we intend to use here for the ln-VaR estimation, i.e., we shall consider, as a possible alternative to the classical estimator in (1.9), the estimator,

$$\ln Q_{\overline{H}}^{(p)}(k) := \ln X_{n-k:n} + \overline{H}(k)\ \ln\left(\frac{k}{np}\right), \tag{1.14}$$

with $\overline{H}$ given in (1.13).

Mathys and Beirlant (2003) try also to reduce the bias of the classical quantile estimators, going directly into the second order framework. Denoting $a_n = k/(n\ p)$, which converges towards infinity, due to the validity of (1.3), we get

$$\begin{aligned} \frac{\chi_p}{X_{n-k:n}} &= \frac{U(1/p)}{U(Y_{n-k:n})} \stackrel{d}{=} \frac{U(Y_{n-k:n}a_n(1+o_p(1)))}{U(Y_{n-k:n})} \\ &\stackrel{p}{\sim} a_n^\gamma\left(1 + A(n/k)\ \left(a_n^\rho - 1\right)/\rho\right). \end{aligned} \tag{1.15}$$

With the parameterization $A(t) = \gamma\ \beta\ t^\rho$, (1.15) suggested to the above mentioned authors, the consideration of the ln-Var estimator

$$\ln \overline{Q}_{\widehat{\gamma}}^{(p)}(k) := \ln X_{n-k:n} + \widehat{\gamma} \ln\left(\frac{k}{np}\right) + \widehat{\gamma}\ \widehat{\beta}\ \left(\frac{n}{k}\right)^{\widehat{\rho}}\ \frac{(k/(np))^{\widehat{\rho}} - 1}{\widehat{\rho}},$$

with $(\widehat{\gamma},\ \widehat{\beta},\ \widehat{\rho})$ a suitable estimator of $(\gamma,\ \beta,\ \rho)$. In this paper, together with the class of estimators in (1.14), we shall also be interested in the new class of estimators

$$\ln \overline{Q}_{\overline{H}}^{(p)}(k) := \ln X_{n-k:n} + \overline{H}(k)\left(\ln\left(\frac{k}{np}\right) + \widehat{\beta}\ \left(\frac{n}{k}\right)^{\widehat{\rho}}\ \frac{(k/(np))^{\widehat{\rho}} - 1}{\widehat{\rho}}\right), \tag{1.16}$$

with $\overline{H}$ the tail index estimator in (1.13), and $\left(\widehat{\beta},\ \widehat{\rho}\right)$ a vector of second order parameters' estimators, to be specified later on, in section 2 of this paper. Our aim is essentially to consider the new classes of estimators for ln-VaR_p, in (1.14) and (1.16), proving, in section 3 of this paper, their consistency and asymptotic normality under appropriate conditions. We shall thus be working in the lines of both Gomes and Figueiredo (2003) and Matthys and Beirlant

(2003): we are going to base our quantile estimation on an adequate reduced bias' tail index estimator, the reduced bias tail index estimator $\overline{H}$ in (1.13), in a spirit similar to the one in Gomes and Figueiredo (2003), possibly together with an adequate direct accomodation of bias of high quantiles, like in Matthys and Beirlant (2003). Section 4 is devoted to a Monte Carlo simulation, that enables the derivation of the distributional properties of the class of estimators in (1.16), comparatively to the classical estimator in (1.9) and the reduced bias' estimators in (1.14), for finite values of n. Finally, in section 5, we illustrate the performance of the new ln-VaR estimators through the analysis of the log-exchange rates of the Euro against the Sterling Pound.

2 Estimation of second order parameters and reduced bias' tail index estimators

The reduced bias' tail index estimators in (1.13) require the estimation of the second order parameters β and ρ in (1.11). Such an estimation will now be briefly discussed.

2.1 Estimation of the shape second order parameter ρ

We shall consider here again particular members of the class of estimators of the second order parameter ρ proposed in Fraga Alves *et al.* (2003). Such a class of estimators may be parameterized in a tuning real parameter $\tau \in \mathbb{R}$. These ρ-estimators depend on the statistics

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

which converge towards $3(1-\rho)/(3-\rho)$, independently of the *tuning* parameter τ, whenever the second order condition (1.10) holds true and k is such that (1.5) holds and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$. The ρ-estimators herewith considered have the functional expression,

$$\widehat{\rho}_\tau(k) \equiv \widehat{\rho}_n^{(\tau)}(k) := \min\left(0,\ \frac{3(T_n^{(\tau)}(k)-1)}{T_n^{(\tau)}(k)-3}\right). \tag{2.1}$$

We shall formalize, without proofs, the distributional properties of the estimators in (2.1), needed in this paper. Proofs may be found in Fraga Alves *et al.* (2003).

Proposition 2.1 (Fraga Alves *et al.*, 2003)**.** *If the second order condition* (1.10) *holds, with* $\rho \le 0$*, k is a sequence of intermediate integers, i.e.,* (1.5) *holds, and*

$$\lim_{n\to\infty} \sqrt{k}\ A(n/k) = \infty,$$

then $\widehat{\rho}_n^{(\tau)}(k)$ *in* (2.1) *converges in probability towards* ρ*, as* $n \to \infty$*. Moreover,* $\widehat{\rho}_n^{(\tau)}(k) - \rho = O_p\left(1/\left(\sqrt{k}\ A(n/k)\right)\right)$.

Remark 2.1. *Under adequate general conditions, and for an appropriate tuning parameter* τ*, the* ρ*-estimators in* (2.1) *show highly stable sample paths as functions of* k*, the number of top o.s. used, for a wide range of large* k*-values.*

Remark 2.2. *The theoretical and simulated results in Fraga Alves et al.* (2003)*, together with the use of these estimators in different reduced bias' statistics, has led us to advise in practice the consideration of the level*

$$k_1 = \min(n-1, [2n/\ln\ln n]) \tag{2.2}$$

(*not chosen in any optimal way, but under the conditions of Proposition 2.1*)*, and of the tuning parameters* $\tau = 0$ *for the region* $\rho \in [-1, 0)$ *and* $\tau = 1$ *for the region* $\rho \in (-\infty, -1)$*. In the simulations of section* 4 *we have indeed done* *this. Anyway, we again advise practitioners not to choose blindly the value of* τ *in* (2.1)*. It is sensible to draw a few sample paths of* $\widehat{\rho}_\tau(k)$*, as functions of* k*, electing the value of* τ *which provides higher stability for large* k*, by means of any stability criterion.*

2.2 Estimation of the scale second order parameter β

For the estimation of β we shall here consider the estimator developed in Gomes and Martins (2002), with the functional expression,

$$\widehat{\beta}_{\widehat{\rho}}(k) := \left(\frac{k}{n}\right)^{\widehat{\rho}} \frac{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\left(\frac{1}{k}\sum_{i=1}^{k}U_i\right)-\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\ U_i\right)}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\ U_i\right)-\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-2\widehat{\rho}}\ U_i\right)}, \tag{2.3}$$

with $U_i,\ 1 \le i \le k$, the scaled log-spacings in (1.8).

We shall denote generically $\widehat{\rho}$ any of the estimators in (2.1), computed at the level k_1 in (2.2).

Remark 2.3 (Caeiro *et al.*, 2004)**.** *When we consider the level* k_1 *in* (2.2)*, together with any of the* ρ*-estimators in this section, computed at the level* k_1*,* $\{\widehat{\rho}-\rho\}$ *is of the order of* $1/\left(\sqrt{k_1}A(n/k_1)\right) = O\left((\ln_2 n)^{(1-2\rho)/2}/\sqrt{n}\right)$*. Also, when we consider* $\widehat{\beta} \equiv \widehat{\beta}_{\widehat{\rho}}(k_1)$*, with* $\widehat{\rho}$ *any of the estimator in* (2.1)*, computed also at the same level* k_1 *in* (2.2)*,* $\widehat{\beta}-\beta$ *is of the order of* $\ln(n/k_1)/\left(\sqrt{k_1}A(n/k_1)\right) = O\left(\ln_3 n\ (\ln_2 n)^{(1-2\rho)/2}/\sqrt{n}\right)$*.*

2.3 The asymptotic behaviour of the reduced bias' tail index estimators

We now state the following:

Proposition 2.2 (Caeiro *et al.*, 2004). *If the second order condition* (1.10) *holds, if* $k = k_n$ *is a sequence of intermediate positive integers, i.e.,* (1.5) *holds, and if* $\sqrt{k}\ A(n/k) \to \lambda$*, finite and non necessarily null, as* $n \to \infty$*, then*

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

This same limiting behaviour holds true if we replace $\overline{H}_{\beta,\ \rho}$ *by* $\overline{H}_{\widehat{\beta},\ \widehat{\rho}}$*, provided that* $\widehat{\rho} - \rho = o_p(1)$ *for every* k*-value on which we base the tail index estimation, and we choose* $\widehat{\beta} := \widehat{\beta}_{\widehat{\rho}}(k_1)$*, with* k_1 *and* $\widehat{\beta}_{\widehat{\rho}}(k)$ *given in* (2.2) *and* (2.3)*, respectively. More specifically, and with* W_k *an asymptotic standard normal r.v., we may write*

$$\overline{H}_{\widehat{\beta},\ \widehat{\rho}}(k) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\, W_k + o_p(A(n/k)). \tag{2.4}$$

Remark 2.4. *Notice that, contrarily to what happens in Drees' class of functionals* (*Drees,* 1998)*, where the minimal asymptotic variance of a reduced bias tail index estimator is given by* $(\gamma(1-\rho)/\rho)^2$*, we have been here able to obtain a reduced bias tail index estimator with an asymptotic variance equal to* γ^2*, the asymptotic variance of Hill's estimator, the maximum likelihood estimator of* γ *for a strict Pareto model.*

3 The asymptotic behaviour of reduced bias ln-VaR estimators

As previously mentioned, for intermediate k, i.e., whenever (1.5) holds true, and with $a_n := k/(np_n)$, which, given the conditions in (1.3), goes to infinity as $n \to \infty$, we are here dealing with semi-parametric ln-VaR estimators, of the type

$$\ln Q^{(p)}_{\widehat{\gamma}}(k) := \ln X_{n-k:n} + \widehat{\gamma}(k) \ln a_n, \tag{3.1}$$

or

$$\ln \overline{Q}^{(p)}_{\widehat{\gamma}}(k) := \ln X_{n-k:n} + \widehat{\gamma}(k) \left(\ln a_n + \widehat{\beta} \left(\frac{n}{k}\right)^{\widehat{\rho}} \ \frac{a_n^{\widehat{\rho}} - 1}{\widehat{\rho}} \right), \tag{3.2}$$

where $\widehat{\gamma} \equiv \widehat{\gamma}(k)$ is any semi-parametric estimator of the tail index γ. Details on semi-parametric estimation of extremely high quantiles for a general tail index $\gamma \in \mathbb{R}$ may be found in de Haan and Rootzén (1993) and more recently in Ferreira et al. (2002). Matthys and Beirlant (2003), Gomes and Figueiredo (2003) and Mathys *et al.* (2004) deal with reduced bias quantile estimation for heavy tails.

We may state the following result:

Theorem 3.1. *Under the second order framework in* (1.10)*, for intermediate k, i.e., k such that* (1.5) *holds, in Hall's class of models in* (1.4)*, whenever*

$$\ln (n\ p_n) = o\left(\sqrt{k}\right), \tag{3.3}$$

and if $\sqrt{k}\ A(n/k) \to \lambda$*, finite, possibly non-null,*

$$\frac{\sqrt{k}}{\ln a_n} \left(\ln Q^{(p)}_{H}(k) - \ln \chi_p\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(\frac{\lambda}{1-\rho},\ \gamma^2\right), \tag{3.4}$$

whereas

$$\frac{\sqrt{k}}{\ln a_n} \left(\ln Q^{(p)}_{\overline{H}}(k) - \ln \chi_p\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(0, \gamma^2\right). \tag{3.5}$$

The limiting results in (3.4) *and* (3.5) *hold also true if we replace* $\ln Q^{(p)}_{\bullet}(k)$ *in* (3.1) *by* $\ln \overline{Q}^{(p)}_{\bullet}(k)$ *in* (3.2)*.*

Proof. We may write

$$\ln X_{n-k:n} \overset{d}{=} \ln U(n/k) + \frac{\gamma\ B_k}{\sqrt{k}} + o_p(A(n/k)),$$

with B_k asymptotically standard normal. Since

$$\ln \chi_p = \ln U(1/p) = \ln U(n a_n/k),$$

we have

$$\ln Q^{(p)}_{\widehat{\gamma}}(k) - \ln \chi_p \overset{d}{=} -(\ln U(na_n/k) - \ln U(n/k)) + \frac{\gamma\, B_k}{\sqrt{k}} + \widehat{\gamma}(k)\ \ln a_n + o_p(A(n/k))$$

$$\overset{d}{=} (\widehat{\gamma}(k)-\gamma)\ln a_n + \frac{\gamma\, B_k}{\sqrt{k}} - \frac{a_n^\rho - 1}{\rho}\ A(n/k)(1+o(1)) + o_p(A(n/k)).$$

Consequently, since $a_n^\rho = o(1)$,

$$\ln Q^{(p)}_{\widehat{\gamma}}(k) - \ln \chi_p \overset{d}{=} (\widehat{\gamma}(k)-\gamma)\ln a_n + \frac{\gamma\, B_k}{\sqrt{k}} + \frac{A(n/k)}{\rho} + o_p(A(n/k)).$$

The dominant term is thus of the order of $\left\{\ln a_n/\sqrt{k}\right\}$, that must converge towards zero, and this is true due to (3.3). The results in (3.4) and (3.5) follow thus from (1.12) and (2.4), respectively.

Similarly,

$$\ln \overline{Q}^{(p)}_{\widehat{\gamma}}(k) - \ln \chi_p \overset{d}{=} \frac{\gamma\, B_k}{\sqrt{k}} + (\widehat{\gamma}(k)-\gamma)\ln a_n + \widehat{\gamma}(k)\ \widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}\ \frac{a_n^{\widehat{\rho}}-1}{\widehat{\rho}} - \frac{a_n^\rho - 1}{\rho}\ A(n/k)(1+o(1)).$$

Now, $\widehat{\gamma}(k) = \gamma(1+o_p(1))$ and the use of the delta-method (Casela and Berger, 2002, pages 240-245), enables us to write

$$\gamma\ \widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}\ \frac{a_n^{\widehat{\rho}}-1}{\widehat{\rho}} - A(n/k)\ \frac{a_n^\rho-1}{\rho} \overset{p}{\sim} \left(\frac{\widehat{\beta}-\beta}{\beta}\right) A(n/k)\ \frac{a_n^\rho-1}{\rho} + (\widehat{\rho}-\rho)\, A(n/k)\left[\ln\left(\frac{n}{k}\right)\ \left(\frac{a_n^\rho-1}{\rho}\right) + \frac{1}{\rho}\left(a_n^\rho \ln a_n - \frac{a_n^\rho - 1}{\rho}\right)\right].$$

Since $(\widehat{\rho}-\rho)\ln(n/k) = o_p(1)$ (see Remark 2.3), $(a_n^\rho - 1)/\rho \to -1/\rho$ and $a_n^\rho \ln a_n \to 0$, as $n\to\infty$, we have

$$\ln \overline{Q}^{(p)}_{\widehat{\gamma}}(k) - \ln \chi_p \overset{d}{=} (\widehat{\gamma}(k)-\gamma)\ln a_n + \frac{\gamma\, B_k}{\sqrt{k}} + o_p(A(n/k)),$$

and the final part of the theorem follows as well. □

Remark 3.1. *Since* (3.4) *holds true for* $\left\{\ln \overline{Q}_{H}^{(p)}(k)\right\}$, *the use of the Hill estimator in a reduced bias quantile estimator like the one proposed by Matthys and Beirlant* (2003), *does not enable the reduction of the dominant component of the bias of the classical log-quantile estimator* $\left\{\ln Q_{H}^{(p)}\right\}$ *in* (1.9). *Consequently, we shall no longer consider* $\left\{\ln \overline{Q}_{H}^{(p)}\right\}$.

4 Simulated behaviour of the ln-VaR estimators

For the estimation of the second order parameter ρ, and as mentioned before, we have here used $\left(\widehat{\beta}_j,\ \widehat{\rho}_j\right)$, $j = 0$ or 1 according as $|\rho| \le 1$ or $|\rho| > 1$, respectively. We use the notation $\widehat{\rho}_j \equiv \widehat{\rho}_j(k_1)$, $\widehat{\beta}_j \equiv \widehat{\beta}_{\widehat{\rho}_j}(k_1)$, $\overline{H}_j(k) = \overline{H}_{\widehat{\beta}_j,\ \widehat{\rho}_j}(k)$, with $\overline{H}_{\widehat{\beta}_j,\ \widehat{\rho}_j}(k)$, $\widehat{\rho}_j(k)$, k_1 and $\widehat{\beta}_{\widehat{\rho}}(k)$ given in (1.13), (2.1), (2.2) and (2.3), respectively. In Figures 1, 2 and 3 we show, for $p = 1/(2n)$ and on the basis of $N = 5000$ runs, the simulated patterns of mean value, $E[\cdot]$, and root mean squared error, $RMSE[\cdot]$, of $\{\ln Q_{H}^{(p)}(k)\}$, $\left\{\ln Q_{\overline{H}_j}^{(p)}(k)\right\}$ and $\left\{\ln \overline{Q}_{\overline{H}_j}^{(p)}(k)\right\}$ in (1.9), (1.14) and (1.16), respectively. These estimators will be denoted $\{\ln Q_H\}$, $\left\{\ln Q_{\overline{H}_j}\right\}$ and $\left\{\ln \overline{Q}_{\overline{H}_j}\right\}$, respectively. Figure 1 is related to the *Fréchet* model with $\gamma = 0.25$ $(\rho = -1,\ \beta = 0.5)$. The models underlying the simulated data, in Figures 2 and 3, are the *Generalized Pareto* (GP) models with $\rho = -0.5$ and -2 $(\gamma = -\rho)$, respectively. The GP(γ) d.f. is given by $F(x) = 1 - (1 + \gamma\, x)^{-1/\gamma}$, $x \ge -1/\gamma, \quad \gamma > 0$ $(\rho = -\gamma,\ \beta = 1)$.

The results obtained lead us to strongly advise the use of the log-quantile estimator in (1.16) for models with $\rho = -1$ or close to it. The estimator in (1.14) does exhibit an interesting performance for all the simulated models with $\rho \ne -1$, being preferable to the estimator in (1.16).

Remark 4.1. *Note that, similarly to what has happened before with the tail index estimation* (*Gomes et al.*, 2004), *and for most of the models in Hall's*

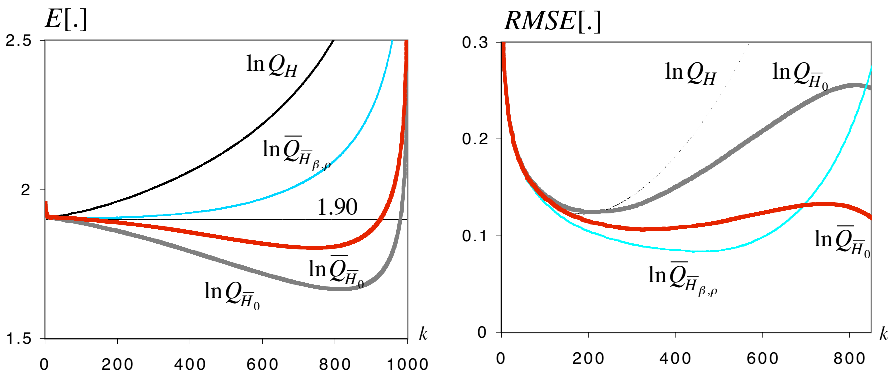

Figure 1: Underlying *Fréchet* parent with $\gamma = 1$ $(\rho = -1)$.

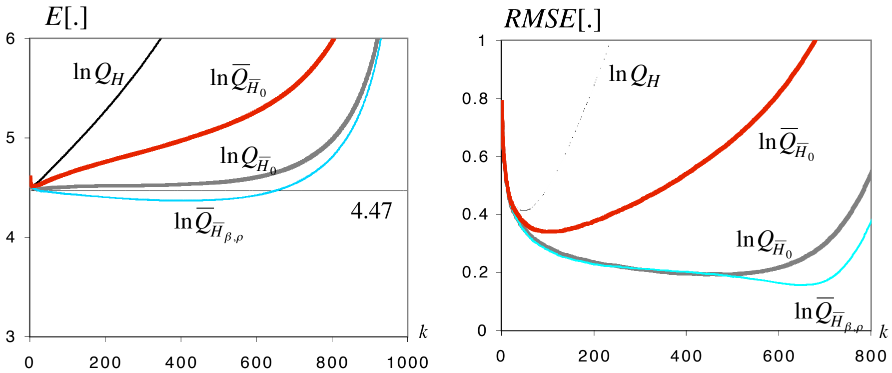

Figure 2: Underlying *GP* parent with $\gamma = 0.5$.

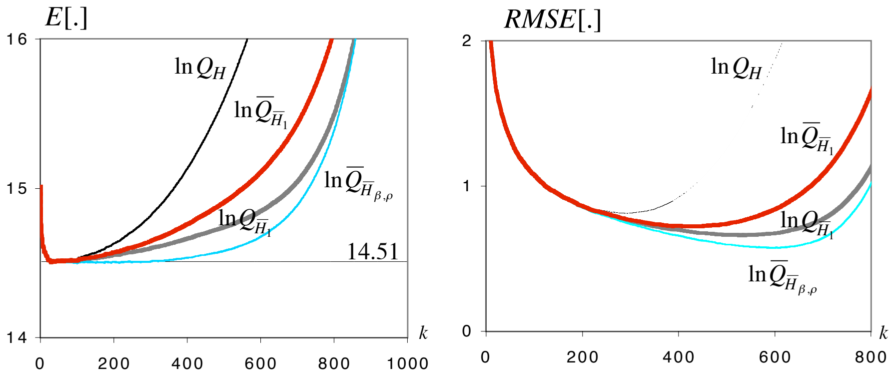

Figure 3: Underlying *GP* parent with $\gamma = 2$.

class in (1.4)*, the computation of both second order parameters' estimators, at the high value* k_1 *in* (2.2)*, enables us to work with high log-quantiles' estimators with a mean squared error smaller than the mean squared error of the classical estimator in* (1.9)*, for all* k*. Those high log-quantile estimators are provided by the use in either* (3.1) *or* (3.2)*, of the tail index estimator* $\overline{H}$ *in* (1.13)*.*

In Tables 2, 3 and 4 we present, for sample sizes $n = 200$, 500, 1000, 2000 and 5000, and $p = 1/(2n)$, the mean values and mean squared errors of the different estimators under study, at their optimal levels, for *Fréchet*, *Generalized Pareto* and *Student*'s underlying parents, respectively. The Student-t_ν has a probability density function given by

$$f_{t_\nu}(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi(\nu-2)}\;\Gamma(\nu/2))}\left[1+\frac{t^2}{\nu-2}\right]^{-(\nu+1)/2},\ t \in \mathbb{R} \quad (\nu > 0).$$

For this model, $\gamma = 1/\nu$ and $\rho = -2/\nu$. Moreover, β is equal to $2\pi^2/3$, 3 and $10\sqrt{3}/9$ for $\nu = 1$, 2 and 4, respectively. We also present the same characteristics for the r.v. $\left\{\ln \overline{Q}_{\overline{H}_{\beta,\ \rho}}\right\}$. If we generically denote $\left\{\ln \widetilde{Q}^{(p)}_{\bullet}(k)\right\}$ any of the estimators (or r.v.'s) under study, we have here simulated the distributional properties of $\left\{\ln \widetilde{Q}^{(p)}_{\bullet}(\widetilde{k}^{\bullet}_0)\right\}$, $\widetilde{k}^{\bullet}_0 := \arg\min_k MSE\left[\ln \widetilde{Q}^{(p)}_{\bullet}(k)\right]$. The simulation experiment is a multi-sample simulation of size 5000×10. For any details on multi-sample simulation, refer to Gomes and Oliveira (2001). Among the estimators and for each value of n, we underline the entries related to the smallest bias and smallest mean squared error. Information on the standard errors associated to the presented characteristics have not been shown, but are available from the authors.

We have also simulated a measure of relative efficiency ($REFF$) of the different ln-VaR-estimators. More precisely, we have simulated

$$REFF_{\widetilde{Q}_{\overline{H}}} := \sqrt{\frac{MSE\left[\ln Q^{(p)}_{H}(k^{H}_0)\right]}{MSE\left[\ln \widetilde{Q}^{(p)}_{\overline{H}}\left(\widetilde{k}^{\overline{H}}_0\right)\right]}}, \quad \text{for} \quad \widetilde{Q} = Q \text{ and } \overline{Q},$$

with $k_0^H := \arg\min_k MSE\left[\ln Q_{\underline{H}}^{(p)}(k)\right]$.

In Figure 4 we present the patterns of the simulated $REFF$'s of the two estimators under study, $Q_{\underline{H}}^{(p)}$ and $\overline{Q}_{\underline{H}}^{(p)}$ in (1.14) and (1.16), respectively (denoted Q and $\overline{Q}$, for sake of simplicity), again for $p = 1/(2n)$, for different sample sizes ($n = 500, 1000, 2000, 5000$) and for the different simulated models with $\rho = -2, -1$ and -0.5.

Table 2: Mean values/mean squared errors of $\ln VaR$ estimators at optimal levels, for a Fréchet underlying model with $\gamma = 0.25$.

n	$\ln \chi_p$	$\ln Q_H$	$\ln Q_{\overline{H}_0}$	$\ln \overline{Q}_{\overline{H}_0}$	$\ln \overline{Q}_{\overline{H}_{\beta,\ \rho}}$
			Fréchet$(\gamma = 0.25)$		
200	1.4976	1.5627 / 0.0309	1.4329 / 0.0361	1.4423 / 0.0253	1.5373 / 0.0173
500	1.7268	1.7846 / 0.0208	1.6765 / 0.0228	1.6852 / 0.0165	1.7599 / 0.0104
1000	1.9002	1.9546 / 0.0152	1.8548 / 0.0156	1.8646 / 0.0114	1.9284 / 0.0069
2000	2.0735	2.1186 / 0.0110	2.0430 / 0.0103	2.0609 / 0.0074	2.0973 / 0.0046
5000	2.3025	2.3423 / 0.0071	2.2825 / 0.0064	2.2973 / 0.0048	2.3209 / 0.0026

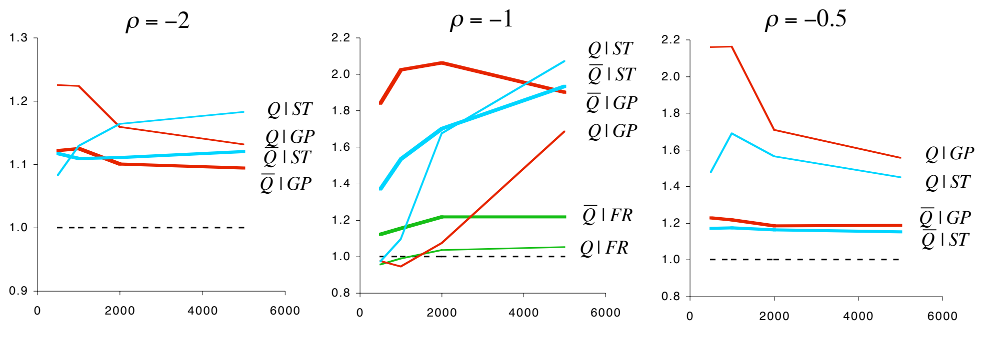

Figure 4: Simulated relative efficiencies for different models with $\rho = -2,\ -1$ and -0.5.

Table 3: Mean values/mean squared errors of ln-VaR estimators at optimal levels, for Generalized Pareto (GP) underlying models.

n	$\ln \chi_p$	$\ln Q_H$	$\ln Q_{\overline{H}_0}$	$\ln \overline{Q}_{\overline{H}_0}$	$\ln \overline{Q}_{\overline{H}_{\beta,\ \rho}}$
GP($\gamma = 0.5$)					
200	3.6376	3.8389 / 0.2963	3.6272 / 0.0896	3.8417 / 0.1917	3.5967 / 0.0809
500	4.1149	4.3055 / 0.2178	4.1225 / 0.0467	4.3033 / 0.1440	4.3033 / 0.0399
1000	4.4710	4.6563 / 0.1737	4.5422 / 0.0371	4.6556 / 0.1169	4.6556 / 0.0238
2000	4.8242	4.9872 / 0.1378	4.9586 / 0.0472	4.9954 / 0.0977	4.9954 / 0.0142
5000	5.2883	5.4480 / 0.1017	5.4166 / 0.0421	5.4371 / 0.0719	5.4371 / 0.0072
GP($\gamma = 1$)					
200	5.9890	6.3025 / 0.6496	5.5818 / 0.6413	5.9676 / 0.2512	6.1173 / 0.2571
500	6.9068	7.1711 / 0.4435	6.5550 / 0.4665	6.8915 / 0.1306	7.0045 / 0.1483
1000	7.6004	7.8518 / 0.3265	7.2757 / 0.3656	7.5919 / 0.0796	7.6941 / 0.0983
2000	8.2938	8.5170 / 0.2414	7.9013 / 0.2091	8.2920 / 0.0567	8.3752 / 0.0644
5000	9.2102	9.3965 / 0.1568	9.2037 / 0.0551	9.2137 / 0.0434	9.2761 / 0.0364
GP($\gamma = 2$)					
200	11.2898	11.6861 / 1.5481	11.5482 / 1.0659	11.6617 / 1.2533	11.5360 / 0.9022
500	13.1224	13.4670 / 0.9613	13.3947 / 0.6400	13.4509 / 0.7631	13.3068 / 0.5109
1000	14.5087	14.7979 / 0.6592	14.7687 / 0.4405	14.7965 / 0.5209	14.6714 / 0.3295
2000	15.8950	16.1460 / 0.4479	16.1036 / 0.3336	16.1153 / 0.3698	16.0284 / 0.2100
5000	17.7275	17.9172 / 0.2612	17.8545 / 0.2039	17.8897 / 0.2179	17.8275 / 0.1135

Table 4: Mean values/mean squared errors of ln-VaR estimators at optimal levels, for Student $ST(\nu)$ underlying models ($\gamma = 1/\nu,\ \rho = -2/\nu$).

n	$\ln \chi_p$	$\ln Q_H$	$\ln Q_{\overline{H}_0}$	$\ln \overline{Q}_{\overline{H}_0}$	$\ln \overline{Q}_{\overline{H}_{\beta,\ \rho}}$
$ST(4)$					
200	1.7223	1.8294 / 0.1155	1.6108 / 0.0932	1.8141 / 0.0816	1.6193 / 0.0640
500	1.9703	2.0813 / 0.0825	1.9065 / 0.0379	2.0669 / 0.0599	1.8839 / 0.0418
1000	2.1529	2.2542 / 0.0656	2.1852 / 0.0230	2.2455 / 0.0474	2.0729 / 0.0297
2000	2.3326	2.4301 / 0.0518	2.3956 / 0.0212	2.4288 / 0.0382	2.2668 / 0.0208
5000	2.5679	2.6594 / 0.0385	2.6371 / 0.0183	2.6549 / 0.0289	2.5146 / 0.0130
$ST(2)$					
200	2.6454	2.8088 / 0.2448	2.4457 / 0.2937	2.5811 / 0.1862	2.5579 / 0.1079
500	3.1059	3.2579 / 0.1701	2.9119 / 0.1800	3.1368 / 0.0906	3.0172 / 0.0646
1000	3.4532	3.5914 / 0.1284	3.2203 / 0.1079	3.4769 / 0.0544	3.3773 / 0.0432
2000	3.8001	3.9329 / 0.0948	3.7501 / 0.0340	3.8198 / 0.0327	3.7363 / 0.0289
5000	4.2585	4.3669 / 0.0629	4.2551 / 0.0147	4.2707 / 0.0168	4.2069 / 0.0163
$ST(1)$					
200	4.8467	5.0861 / 0.6196	4.6267 / 1.2938	4.8667 / 0.4908	4.9295 / 0.3643
500	5.7630	5.9601 / 0.3862	5.7094 / 0.3296	5.8616 / 0.3092	5.8215 / 0.1947
1000	6.4561	6.6368 / 0.2732	6.5087 / 0.2143	6.5733 / 0.2219	6.5065 / 0.1211
2000	7.1493	7.3000 / 0.1875	7.2424 / 0.1384	7.2787 / 0.1518	7.1974 / 0.0752
5000	8.0663	8.1879 / 0.1119	8.1575 / 0.0800	8.1689 / 0.0891	8.1073 / 0.0396

5 An illustration

We shall here consider the performance of the above mentioned estimators in the analysis of Euro-UK Pound daily exchange rates from January 4, 1999 till December 14, 2004. This data has been collected by the European System of Central Banks, and was obtained from *http://www.bportugal.pt/rates/cambtx/*.

In Figure 5, and working with the $n_0 = 725$ positive log-returns, we picture the sample paths of the estimators of the second order parameters ρ and β under discussion.

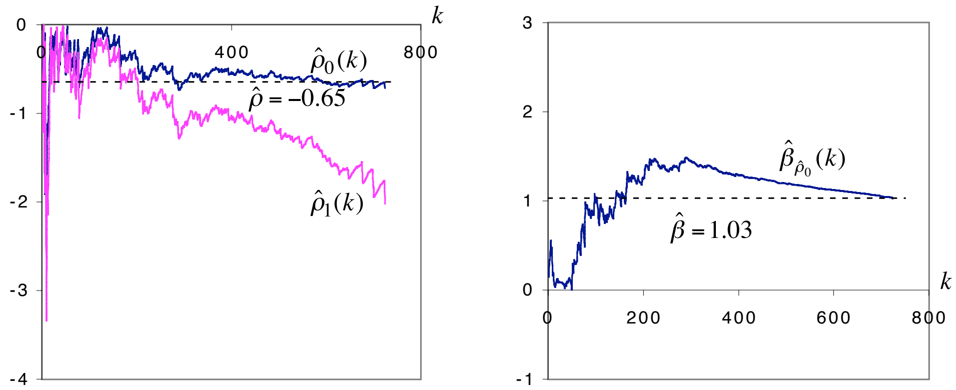

Figure 5: Estimates of the second order parameter ρ, through $\widehat{\rho}_\tau(k)$ in (2.1), $\tau = 0$ and 1 (*left*) and of the second order parameter β, through $\widehat{\beta}_{\widehat{\rho}_0}(k)$ in (2.3) (*right*), for the Daily Log-Returns on the Euro-UK Pound.

Remark 5.1. *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 values of k, the estimate associated to $\tau = 0$. From the experience we have with this class of estimates, this means that $|\rho| \leq 1$ and the tuning parameter $\tau = 0$ is then advisable. We have got $\widehat{\rho}_0 = \widehat{\rho}_0(725) = -0.65$. The use of $\widehat{\beta}_{\widehat{\rho}_0}(k)$ in* (2.3)*, computed at the level k_1 in* (2.2)*, leads then us to the estimate $\widehat{\beta}_0 = 1.03$.*

The sample paths of the classical Hill estimator H in (1.7), the second order reduced bias' tail index estimator $\overline{H}$ in (1.13) and of the ln-Var-estimators in (1.9), (1.14) and (1.16), for $p = 0.001$, are pictured in Figure 6.

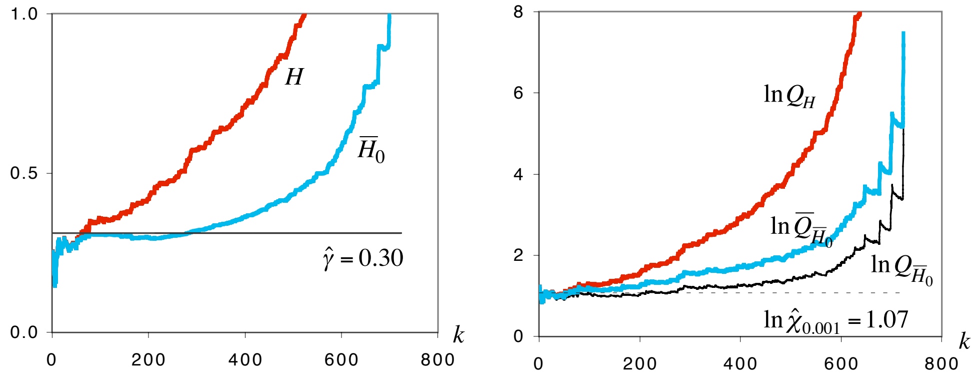

Figure 6: Estimates provided through the Hill and the $\overline{H}$ estimators in (1.7) and (1.13) (*left*) and the ln-VaR_p estimators in (1.9), (1.14) and (1.16), for the Daily Log-Returns on the Euro-UK Pound and $p = 0.001$.

For $p = 0.001$, any stability criterion for moderate values of k leads us to the choice of the estimator $\left\{\ln Q_{\overline{H}}\right\}$ and to the estimate 1.07 for $\ln \chi_{0.001}$.

References

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

[2] Caeiro, F. Gomes, M. I. and Pestana, D. (2004). *Direct reduction of bias of the classical Hill estimator.* Notas e Comunicações CEAUL 18/04. Submitted.

[3] Casela, G. and Berger, R. L. (2002). *Statistical Inference*, 2nd edition. Duxbury, USA.

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

[5] 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.

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

[7] 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.

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

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

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

[12] 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.

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

[14] Gomes, M. I. and O. Oliveira (2001). The bootstrap methodology in Statistical Extremes — choice of the optimal sample fraction. *Extremes* **4**:4, 331-358.

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

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

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

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

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

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

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

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