

A quasi-PORT methodology for VaR based on second-order reduced-bias estimation

M. Ivette Gomes

Universidade de Lisboa, FCUL, DEIO and CEAUL

Fernanda Figueiredo

Universidade do Porto, Faculdade de Economia, and CEAUL

Lígia Henriques-Rodrigues

Instituto Politécnico de Tomar, and CEAUL

M. Cristina Miranda

ISCA, Universidade de Aveiro, and CEAUL

Abstract. In this paper, we deal with the estimation, under a semi-parametric framework, of the *Value at Risk* (VaR) at a level p , the size of the loss occurred with a small probability p . Under such a context, the classical VaR estimators are the Weissman-Hill estimators, based on an intermediate number k of top order statistics. But these VaR estimators do not enjoy the adequate linear property of quantiles, contrarily to the PORT VaR-estimators, which depend on an extra *tuning* parameter q , with $0 \leq q < 1$. We shall here consider “quasi-PORT” reduced-bias VaR-estimators, for which such a linear property is obtained approximately. They are based on a partially shifted version of a *minimum-variance reduced-bias* (MVRB) estimator of the extreme value index, the primary parameter in *Statistics of Extremes*. Due to the stability on k of the MVRB extreme value index and associated VaR-estimates, we propose the use of a heuristic stability criterion for the choice of k and q , providing applications of the methodology in the field of finance.

1 Introduction and preliminaries

We shall place ourselves under a semi-parametric framework, to refer the estimation of a positive *extreme value index* (EVI) γ , the primary parameter in *Statistics of Extremes* and the basis for the estimation of the *Value at Risk* (VaR) at a level p , denoted VaR_p , the size of the loss occurred with a small probability p . In other words, we are interested in the estimation of a (high) *quantile*, $\chi_{1-p} := F^{\leftarrow}(1-p)$, of a probability distribution function (d.f.) F , with $F^{\leftarrow}(y) := \inf \{x : F(x) \geq y\}$, the generalized inverse function of F . Let us denote $U(t) := F^{\leftarrow}(1-1/t)$, $t \geq 1$, a reciprocal quantile function such that $\chi_{1-p} \equiv VaR_p = U(1/p)$. We shall

thus consider heavy-tailed parents, quite common in finance, i.e., parents such that, as $t \rightarrow \infty$,

$$U \in RV_\gamma \quad \Longleftrightarrow \quad 1 - F \in RV_{-1/\gamma},$$

where, as usual, the notation RV_α stands for regularly-varying functions with an index of regular variation equal to α , i.e., positive measurable functions $g(\cdot)$ such that for any $x \geq 0$, $g(tx)/g(t) \rightarrow x^\alpha$, as $t \rightarrow \infty$. We are then working in $\mathcal{D}_{\mathcal{M}}(EV_{\gamma>0})$, the domain of attraction for maxima of EV_γ , $\gamma > 0$, with EV_γ denoting the general *extreme value* (EV) d.f., given by

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

with γ the EVI. For heavy-tailed parents and given a sample $\mathbf{X}_n = (X_1, \dots, X_n)$, the classical EVI-estimator is the Hill estimator (Hill, 1975), here denoted $H \equiv H_n(k)$, and given by

$$H_n(k) \equiv H_n(k; \mathbf{X}_n) := \frac{1}{k} \sum_{i=1}^k \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}, \quad k < n, \quad (1.2)$$

the average of the k log-excesses over a high random threshold $X_{n-k:n}$. Consistency of the estimator in (1.2) is achieved if $X_{n-k:n}$ is an *intermediate* order statistic (o.s.), i.e., if

$$k = k_n \rightarrow \infty \quad \text{and} \quad k/n \rightarrow 0, \quad \text{as } n \rightarrow \infty. \quad (1.3)$$

The Hill estimator in (1.2) is scale-invariant but not location invariant, as often desired, and this contrarily to the PORT-Hill estimators, recently introduced in Araújo Santos *et al.* (2006) and further studied in Gomes *et al.* (2008a), with PORT standing for *peaks over a random threshold*. The class of PORT-Hill estimators is based on a sample of excesses over a random threshold $X_{n_q:n}$, $n_q := [nq] + 1$, with $[x]$ denoting, as usual, the integer part of x , i.e., it is based on

$$\mathbf{X}_n^{(q)} := (X_{n:n} - X_{n_q:n}, X_{n-1:n} - X_{n_q:n}, \dots, X_{n_q+1:n} - X_{n_q:n}), \quad n_q = [nq] + 1. \quad (1.4)$$

We need to have $0 < q < 1$, for d.f.'s with an infinite left endpoint $x_F := \inf\{x : F(x) > 0\}$ (*the random threshold is an empirical quantile*). We can also have $q = 0$, provided that the underlying model has a finite left endpoint x_F (*the random threshold is then the minimum*). These new classes of EVI-estimators are the so-called PORT-Hill estimators, denoted $H_n^{(q)}$, and, for $0 \leq q < 1$ and $k < n - n_q$, they are given by

$$H_n^{(q)}(k) := H_n(k; \mathbf{X}_n^{(q)}) = \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n} - X_{n_q:n}}{X_{n-k:n} - X_{n_q:n}}, \quad (1.5)$$

i.e., they have the same functional form of the Hill estimator in (1.2), but with the original sample $\underline{\mathbf{X}}_n = (X_1, \dots, X_n)$ replaced by the sample of excesses $\underline{\mathbf{X}}_n^{(q)}$ in (1.4). These estimators are now invariant for both changes of scale and location in the data, and depend on the *tuning parameter* q , which provides a highly flexible class of EVI-estimators. Provided that we adequately choose the tuning parameter q , the PORT estimators may even compare favorably with the second-order *minimum-variance reduced-bias* (MVRB) EVI-estimators, recently introduced in the literature and briefly discussed in the following.

Indeed, due to the high bias of the Hill estimator, in (1.2), for moderate up to large k , several authors have dealt with bias reduction in the field of extremes, working then in a slightly more strict class than $\mathcal{D}_{\mathcal{M}}(EV_{\gamma>0})$, the class of models $U(\cdot)$ such that

$$U(t) = C t^\gamma (1 + A(t)/\rho + o(t^\rho)), \quad A(t) = \gamma\beta t^\rho, \quad (1.6)$$

as $t \rightarrow \infty$, where $\rho < 0$ and $\beta \neq 0$. This means that the slowly varying function $L(t)$ in $U(t) = t^\gamma L(t)$ is assumed to behave asymptotically as a constant C . Note that to assume (1.6) is equivalent to saying that we can choose $A(t) = \gamma\beta t^\rho$, $\rho < 0$, in the more general second-order condition

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho}. \quad (1.7)$$

The Hill estimator, in (1.2), reveals usually a high asymptotic bias, i.e., as $n \rightarrow \infty$, $\sqrt{k}(H_n(k) - \gamma)$ is asymptotically normal with variance γ^2 and a non-null mean value, equal to $\lambda_A/(1 - \rho)$, whenever $\sqrt{k} A(n/k) \rightarrow \lambda_A \neq 0$, finite, with $A(\cdot)$ the function in (1.7). This non-null asymptotic bias, together with a rate of convergence of the order of $1/\sqrt{k}$, leads to sample paths with a high variance for small k , a high bias for large k , and a very sharp mean squared error (MSE) pattern, as a function of k . A simple class of second-order MVRB EVI-estimators is the one in Caeiro *et al.* (2005), used for a semi-parametric estimation of $\ln VaR_p$ in Gomes and Pestana (2007b). This class, here denoted $\bar{H} \equiv \bar{H}_n(k)$, depends upon the estimation of the second-order parameters (β, ρ) in (1.6). Its functional form is

$$\bar{H}_n(k) \equiv \bar{H}_n(k; \underline{\mathbf{X}}_n) \equiv \bar{H}_{\hat{\beta}, \hat{\rho}}(k) := H_n(k) (1 - \hat{\beta}(n/k)^{\hat{\rho}} / (1 - \hat{\rho})), \quad (1.8)$$

with $H_n(k)$ the Hill estimator in (1.2), and where $(\hat{\beta}, \hat{\rho})$ needs to be an adequate consistent estimator of (β, ρ) . Algorithms for the estimation of (β, ρ) are provided in Gomes and Pestana (2007a,b), with one of them reformulated in Section 2 of this paper.

With Q standing for quantile function, the classical Weissman-Hill Var_p -estimator,

$$Q_{p|H}(k) := X_{n-k+1:n} c_k^{H_n(k)}, \quad c_k \equiv c_{k,n,p} = \frac{k}{np}, \quad (1.9)$$

has been introduced in Weissman (1978). The MVRB VaR-estimator $Q_{p|\bar{H}}$, with $Q_{p|H}$ given in (1.9), was studied in Gomes and Pestana (2007b). However, for any positive real δ , and with

$Q_{p|\bullet}$ denoting either $Q_{p|H}$ or $Q_{p|\overline{H}}$, $Q_{p|\bullet}(k; \delta \underline{\mathbf{X}}_n) = \delta Q_{p|\bullet}(k; \underline{\mathbf{X}}_n)$, as desirable, but contrarily to the linear property for quantiles, $\chi_p(\delta X + s) = \delta \chi_p(X) + s$ for any real s and positive real δ , we no longer have $Q_{p|\bullet}(k; s \underline{\mathbf{1}}_n + \delta \underline{\mathbf{X}}_n) = s + \delta Q_{p|\bullet}(k; \underline{\mathbf{X}}_n)$, with $\underline{\mathbf{1}}_n$ denoting, as usual, a vector with n unit elements. Araújo Santos *et al.* (2006) have developed a class of high quantile estimators based on the sample of excesses over a random threshold $X_{n_q:n}$, provided in (1.4), and, among others, they propose the so-called PORT-Weisman-Hill Var_p -estimators,

$$Q_{p|H}^{(q)}(k) := (X_{n-k:n} - X_{n_q:n}) c_k^{H_n^{(q)}(k)} + X_{n_q:n}, \quad (1.10)$$

where $H_n^{(q)}(k)$ is the Hill estimator of γ , made location/scale invariant by using the transformed sample $\underline{\mathbf{X}}_n^{(q)}$, i.e. $H_n^{(q)}(k)$ is the estimator in (1.5). They consequently obtain exactly the above mentioned linear property for the quantile estimators.

The second-order MVRB EVI-estimators in (1.8) are not location invariant, but they are “approximately” location invariant. If we merely replace, in (1.10), $H_n^{(q)}(k)$ by $\overline{H}_n(k)$ in (1.8), we have practically no improvement comparatively with the MVRB-estimator $Q_{p|\overline{H}}$. With $H_n^{(q)}(k)$ defined in (1.5), we shall consider here the “quasi-PORT” EVI-estimator,

$$\overline{H}_n^{(q)}(k) \equiv \overline{H}_n^{(q)}(k; \hat{\beta}, \hat{\rho}) := H_n^{(q)}(k) \left(1 - \hat{\beta}(n/k)^{\hat{\rho}} / (1 - \hat{\rho}) \right) \quad (1.11)$$

and the associated “quasi-PORT” Var_p -estimator, with the functional form

$$Q_{p|\overline{H}}^{(q)}(k) := (X_{n-k:n} - X_{n_q:n}) c_k^{\overline{H}_n^{(q)}(k)} + X_{n_q:n}. \quad (1.12)$$

In Section 2 of this paper, we briefly discuss the estimation of the second-order parameters β and ρ . In Section 3, we briefly review the main asymptotic properties of the estimators under consideration. In Section 4, we describe the results associated with a Monte-Carlo simulation study of the new VaR-estimators, in (1.12). Finally, in Section 5, due to the stability on k of the MVRB estimates \overline{H} , in (1.8), and $Q_{p|\overline{H}}$, with $Q_{p|H}$ provided in (1.9), as well as the new VaR-estimates in (1.12), we propose the use of a heuristic stability criterion for the choice of k and q , providing applications of the methodology in the field of finance.

2 Estimation of second-order parameters

All reduced-bias EVI-estimators, like the one in (1.8), and associated VaR-estimators, require the estimation of scale and shape second-order parameters, (β, ρ) , in (1.6). Such an estimation will now be briefly discussed.

2.1 Estimation of the shape second-order parameter

For models in (1.6), we shall consider the class of estimators introduced in Fraga Alves *et al.* (2003), possibly parameterized in a tuning real parameter $\tau \in \mathbb{R}$, as suggested in Caeiro and Gomes (2006). Those estimators are based on the statistics

$$T_n^{(\tau)}(k) := \frac{\left(M_n^{(1)}(k)\right)^\tau - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}}, \quad \tau \in \mathbb{R},$$

with the notation $a^{b\tau} = b \ln a$, for $\tau = 0$, and where

$$M_n^{(j)}(k) := \frac{1}{k} \sum_{i=1}^k \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}^j, \quad j = 1, 2, 3.$$

They are given by

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

Under mild restrictions on k , these statistics converge towards ρ , independently of the *tuning* parameter τ . Distributional properties of the estimators in (2.1) can be found in Fraga Alves *et al.* (2003). Consistency is achieved in the class of models in (1.6), for *intermediate* k -values, denoted k_1 , such that apart from condition (1.3), with k replaced by k_1 , we have $\sqrt{k_1} A(n/k_1) \rightarrow \infty$, as $n \rightarrow \infty$. We have here decided for the choice

$$k_1 = \lceil n^{1-\epsilon} \rceil, \quad \epsilon = 0.001. \quad (2.2)$$

Remark 2.1. *With the choice of k_1 in (2.2), and whenever $\sqrt{k_1} A(n/k_1) \rightarrow \infty$, we get $\hat{\rho} - \rho := \hat{\rho}_\tau(k_1) - \rho = o_p(1/\ln n)$, a condition needed, in order not to have any increase in the asymptotic variance of the new bias-corrected Hill estimator in equation (1.8), comparatively with the Hill estimator in (1.2). Note that with the choice of k_1 in (2.2), we get $\sqrt{k_1} A(n/k_1) \rightarrow \infty$ if and only if $\rho > 1/2 - 1/(2\epsilon) = -499.5$, an irrelevant restriction, from a practical point of view.*

Remark 2.2. *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.'s used, for a range of large k -values (see, for instance, the pattern of $\hat{\rho}_0(k)$ in figures 7, 9 and 11).*

Remark 2.3. *It is sensible to advise practitioners not to choose blindly the value of τ in (2.1): sample paths of $\hat{\rho}_\tau(k)$, as functions of k , for a few values of τ , should be drawn, in order to select the value of τ which provides higher stability for large k , by means of any stability criterion, like the one proposed in the Algorithm of Section 2.3.*

2.2 Estimation of a scale second-order parameter

For the estimation of the scale second-order parameter β , in (1.6), we shall here consider

$$\hat{\beta}_{\hat{\rho}}(k) := \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{d_{\hat{\rho}}(k) D_0(k) - D_{\hat{\rho}}(k)}{d_{\hat{\rho}}(k) D_{\hat{\rho}}(k) - D_{2\hat{\rho}}(k)}, \quad k < n, \quad (2.3)$$

dependent on the estimator $\hat{\rho} = \hat{\rho}_{\tau}(k_1)$ suggested in Section 2.1 and where, for any $\alpha \leq 0$,

$$d_{\alpha}(k) := \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{-\alpha} \quad \text{and} \quad D_{\alpha}(k) := \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{-\alpha} U_i,$$

with $U_i := i (\ln X_{n-i+1:n} - \ln X_{n-i:n})$, $1 \leq i \leq k$, the *scaled log-spacings*.

Details on the distributional behaviour of the estimator in (2.3) can be found in Gomes and Martins (2002) and more recently in Gomes *et al.* (2008b) and Caeiro *et al.* (2009). Consistency is achieved for models in (1.6), k values such that (1.3) holds and $\sqrt{k} A(n/k) \rightarrow \infty$, as $n \rightarrow \infty$, and estimators $\hat{\rho}$ of ρ such that $\hat{\rho} - \rho = o_p(1/\ln n)$. Alternative estimators of β can be found in Caeiro and Gomes (2006) and Gomes *et al.* (2009).

2.3 An algorithm for second-order parameters estimation

Based on the algorithms proposed before, we now consider the following simple algorithm:

Algorithm (Second-order estimation):

1. Given a sample (x_1, x_2, \dots, x_n) , plot the observed values of $\hat{\rho}_{\tau}(k)$ in (2.1), for $\tau = 0$ and $\tau = 1$;
2. Consider $\{\hat{\rho}_{\tau}(k)\}_{k \in \mathcal{K}}$, for k in the interval $\mathcal{K} = ([n^{0.995}], [n^{0.999}])$, compute their median, denoted η_{τ} , and compute $I_{\tau} := \sum_{k \in \mathcal{K}} (\hat{\rho}_{\tau}(k) - \eta_{\tau})^2$, for $\tau = 0, 1$. Next choose the *tuning parameter* $\tau = 0$ if $I_0 \leq I_1$; otherwise, choose $\tau = 1$;
3. Work with $\hat{\rho} \equiv \hat{\rho}_{\tau} = \hat{\rho}_{\tau}(k_1)$ and $\hat{\beta} \equiv \hat{\beta}_{\tau} := \hat{\beta}_{\hat{\rho}_{\tau}}(k_1)$, $\hat{\rho}_{\tau}(k)$, k_1 and $\hat{\beta}_{\hat{\rho}}(k)$ given in (2.1), (2.2) and (2.3), respectively.

Remark 2.4. *If there are negative elements in the sample, the sample size n should be replaced by n_0 , the number of positive elements in the sample.*

For models with $|\rho| \leq 1$, the most common in practice and the ones for which bias-reduction is neatly needed, this algorithm leads in almost all situations to the *tuning parameter* $\tau = 0$, the value considered in this paper, both in simulations and in case-studies. For details on this and similar algorithms, see Gomes and Pestana (2007a).

3 A brief note on the normal asymptotic non-degenerate behaviour of the estimators

The asymptotic normality, as well as full information on a possibly non-null asymptotic bias, of the estimators $\hat{\rho}_\tau(k)$ and $\hat{\beta}_\rho(k)$ in (2.1) and (2.3), respectively, as well as of reduced-bias estimators, like $\overline{H}_n(k)$ in (1.8), is easier to derive if we slightly restrict the class of models in (1.6). In this paper, similarly to what has been done in Gomes *et al.* (2007), and for convenience of exposition, we consider a third-order framework where we merely make explicit a third order term in (1.6), assuming that

$$U(t) = Ct^\gamma(1 + \gamma\beta t^\rho + \beta' t^{2\rho} + o(t^{2\rho})), \quad (3.1)$$

as $t \rightarrow \infty$, with $C, \gamma > 0$, $\beta, \beta' \neq 0$, $\rho < 0$.

Remark 3.1. Note that to assume (3.1) is equivalent to saying that the more general third-order condition

$$\lim_{t \rightarrow \infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} - \frac{x^\rho - 1}{\rho}}{B(t)} = \frac{x^{\rho+\rho'} - 1}{\rho + \rho'} \quad (3.2)$$

holds with $\rho = \rho' < 0$ and that we can choose, in (3.2), $A(t) = \alpha t^\rho =: \gamma \beta t^\rho$, $B(t) = \beta' t^\rho = \beta' A(t)/(\beta\gamma)$, $\beta, \beta' \neq 0$, with β and β' “scale” second- and third-order parameters, respectively.

Remark 3.2. Note also that several common heavy-tailed models belong to the class in (3.1). Among them we mention:

- the Fréchet model, with d.f. $\Phi_{1/\gamma}(x) = \exp(-x^{-1/\gamma})$, $x \geq 0$, $\gamma > 0$, for which $\rho' = \rho = -1$, $\beta = 0.5$ and $\beta' = 5/6$;
- the EV model, with d.f. $EV_\gamma(x)$ in (1.1), for $\gamma = 1/2$ ($\rho = \rho' = -0.5$) and for $\gamma = 1$ or $\gamma \geq 2$ ($\rho = \rho' = -1$). The parameter β is equal to 1 for $\gamma = 1/2$, 3/2 for $\gamma = 1$ and 1/2 for $\gamma \geq 2$. The parameter β' is $-1/4$ for $\gamma = 1/2$, $-1/12$ for $\gamma = 1$, $-11/12$ for $\gamma = 2$ and $\gamma(3\gamma - 5)/24$ for $\gamma > 2$;
- the Generalized Pareto (GP) model, with d.f. $GP_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \geq 0$, $\gamma > 0$, for which $\rho' = \rho = -\gamma$ and $\beta = \beta' = 1$;
- the Burr model, with d.f. $B_{\gamma,\rho}(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, $\gamma > 0$, $\rho' = \rho < 0$ and, as for the GP model, $\beta = \beta' = 1$;
- the Student’s t_ν -model with ν degrees of freedom, with a probability density function

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

for which $\gamma = 1/\nu$ and $\rho' = \bar{\rho} = -2/\nu$. For an explicit expression of β and β' as a function of ν , see Caeiro and Gomes (2008).

Given an estimator $S_n(k)$ of a parameter of extreme events ξ , based on the k top o.s.'s in the available sample, and under an adequate second-order framework, let us say the more strict second-order framework in (1.6), there exist a regularly varying function $C \in RV_{c(\rho)}$, $c(\rho) < 0$, a sequence of standard normal r.v.'s P_k^S , an asymptotic variance, $\sigma_S^2 > 0$, an asymptotic bias $b_S \in \mathbb{R}$ and a rate of convergence $\sqrt{r_k} \rightarrow 0$, as $k \rightarrow \infty$, such that, as $n \rightarrow \infty$, the asymptotic distributional representation

$$S_n(k) \stackrel{d}{=} \xi + \sigma_S \sqrt{r_k} P_k^S + b_S C(n/k)(1 + o_p(1)) \quad (3.3)$$

holds. Consequently, $S_n(k)$ is consistent for the estimation of ξ provided that k is intermediate, i.e., (1.3) holds. If we further assume that $\sqrt{k} C(n/k) \rightarrow \lambda_C$, finite,

$$\frac{S_n(k) - \xi}{\sqrt{r_k}} \stackrel{d}{\underset{n \rightarrow \infty}{\rightarrow}} \text{Normal}(\lambda_C b_S, \sigma_S^2). \quad (3.4)$$

Whenever we are dealing with EVI-estimators, the rate of convergence is usually of the order of $1/\sqrt{k}$, i.e., $r_k = 1/k$. Then, for the associated semi-parametric quantile estimators, $r_k = (\ln c_k)^2/k$, with c_k defined in (1.9). For the classical EVI-estimation, through an estimator like the Hill estimator, in (1.2), and under the second-order framework in (1.7), $C \equiv A \in RV_\rho$, i.e., $c(\rho) = \rho$, whereas for reduced-bias EVI-estimation, through an estimator like the one in (1.8), and under the more restrict third-order framework in (3.1), $C \equiv A^2 \in RV_{2\rho}$, i.e., $c(\rho) = 2\rho$. A similar but more technical comment applies to the PORT-estimation based on a classical EVI-estimator or on a reduced-bias EVI-estimator. For details on the asymptotic behaviour of PORT-estimators, see Araújo Santos *et al.* (2006). For MVRB EVI-estimation through the EVI-estimator in (1.8), and under the third-order framework in (3.1), see Caeiro *et al.* (2009). A link between EVI- and associated VaR-estimators can be found, for instance, in Beirlant *et al.* (2008), Theorem 4.1. Indeed, let us assume that (1.6) holds, $k = k_n$ is an intermediate sequence such that $c_k := k/np \rightarrow \infty$, $\ln c_k/\sqrt{k} \rightarrow 0$ and $\sqrt{k}A(n/k) \rightarrow \lambda_A$, finite, as $n \rightarrow \infty$. Let $\hat{\gamma}_n(k)$ be any consistent estimator of the tail index γ , such that

$$\sqrt{k}(\hat{\gamma}_n(k) - \gamma) \stackrel{d}{\underset{n \rightarrow \infty}{\rightarrow}} \text{Normal}(\lambda_A b_{\hat{\gamma}}, \sigma_{\hat{\gamma}, \rho}^2), \quad (3.5)$$

and let us consider $Q_{p|\hat{\gamma}}(k)$, with $Q_{p|H}(k)$ provided in (1.9). Then,

$$\frac{\sqrt{k}}{\ln c_k} \left(\frac{Q_{p|\hat{\gamma}}(k)}{\text{VaR}_p} - 1 \right) \stackrel{d}{\underset{n \rightarrow \infty}{\rightarrow}} \text{Normal}(\lambda_A b_{\hat{\gamma}}, \sigma_{\hat{\gamma}, \rho}^2)$$

even if we work with reduced-bias tail index estimators like the ones in (1.8), provided that $(\hat{\beta}, \hat{\rho})$ is consistent for the estimation of (β, ρ) and $(\hat{\rho} - \rho) \ln(n/k) = o_p(1)$, as $n \rightarrow \infty$. Similar

remarks apply to PORT and quasi-POR estimation. In (3.5), we have $b_H = 1/(1 - \rho)$ whereas $b_{\overline{H}} = 0$. The asymptotic variance in (3.5) is equal to γ^2 both for the Hill and the corrected-Hill estimators, in (1.2) and (1.8), respectively. Consequently, the \overline{H} -estimators outperform the H -estimators for all k . A similar remark applies to quantile estimation.

4 Finite sample behaviour: a Monte-Carlo simulation

In this section, for $p = 1/n$ and $q = 0, 0.1$ and 0.25 , we are interested in the finite-sample behaviour of the VaR-estimators, $Q_{p|\overline{H}}^{(q)}(k)$, in (1.12), comparatively with the classical Weissman-Hill VaR-estimator, $Q_{p|H}(k)$, in (1.9), the associated PORT-Weissman-Hill VaR-estimators, $Q_{p|H}^{(q)}(k)$, in (1.10), and the MVRB VaR-estimator $Q_{p|\overline{H}}(k)$, with $Q_{p|H}(k)$ given in (1.9). The overall simulation is based on a multi-sample simulation with size 5000×20 , i.e., 20 replicates with 5000 runs each. For details on multi-sample simulation refer to Gomes and Oliveira (2001). The patterns of mean values (E) and root mean squared errors (RMSE) are based on the first replicate. We have considered the following underlying parents, already mentioned in Remark 3.2:

- A. the Burr model, with $\gamma = 0.25$ and $\rho = -0.5$;
- B. the Student's t_ν -model with $\nu = 4$ degrees of freedom ($\gamma = 0.25$ and $\rho = -0.5$);
- C. the general EV model, with $\gamma = 0.5$ ($\rho = -0.5$).

4.1 Mean values and mean squared errors patterns

We shall consider the following normalized VaR_p -estimators, $Q_{p|H}(k)/VaR_p$, $Q_{p|\overline{H}}(k)/VaR_p$, $Q_{p|H}^{(q)}(k)/VaR_p$ and $Q_{p|\overline{H}}^{(q)}(k)/VaR_p$. For the sake of simplicity, we denote these quotients by Q_H , $Q_{\overline{H}}$, $Q_{H|q}$ and $Q_{\overline{H}|q}$, respectively. In Figures 1, 2 and 3, for the models in A., B. and C., respectively, we show the simulated patterns of mean value, $E(Q_\bullet)$, and root mean squared error, $RMSE(Q_\bullet)$, of these normalized estimators.

These parents were chosen just to illustrate the fact that:

- the quasi-POR VaR-estimators can be unable to improve the performance of $Q_{\overline{H}}$, as happens also with the PORT-Weissman-Hill estimators when compared with the Weissman-Hill estimator Q_H (see Figure 1, associated with the above mentioned Burr model);
- the PORT-Weissman-Hill estimators can outperform the MVRB-estimator, $Q_{\overline{H}}$ (see Figure 2, associated with a Student t_4 underlying parent);

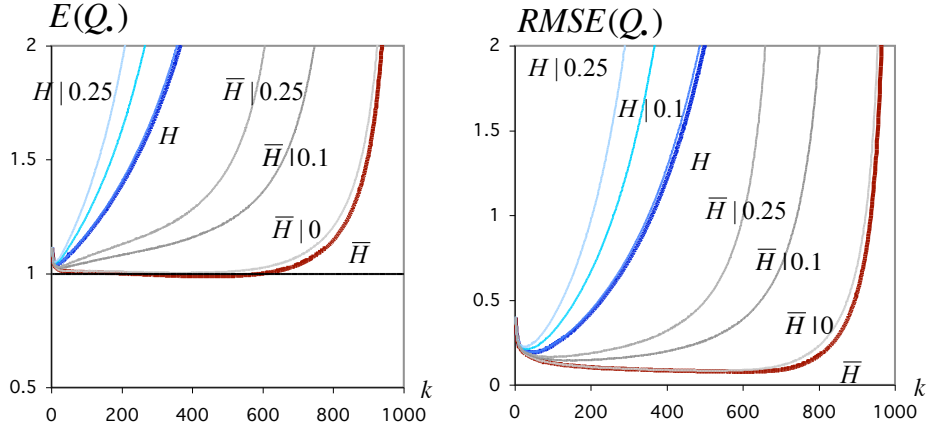


Figure 1: Underlying Burr parent with $(\gamma, \rho) = (0.25, -0.5)$.

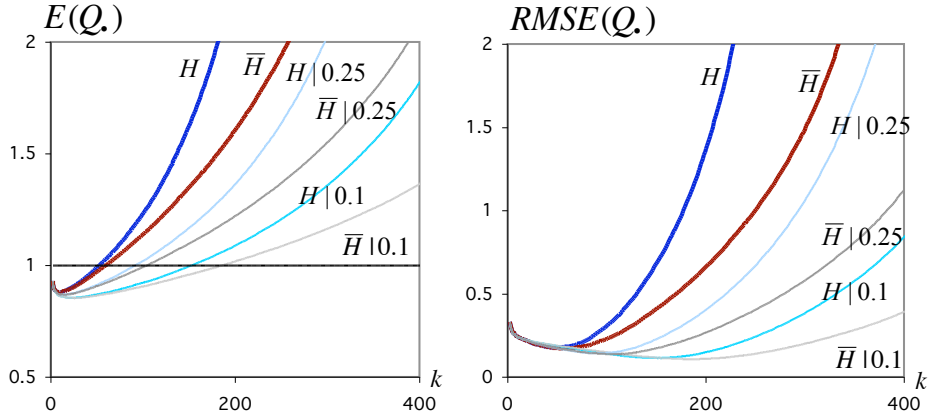


Figure 2: Underlying Student t_4 parent ($\gamma = 0.25, \rho = -0.5$).

- we can often find a value of q that provides the best estimator of $VarR_p$ through the use of the new class of estimators $Q_{p|\bar{H}}^{(q)}(k)$, in (1.12) (like the value $q = 0.1$ in Figure 2 and the value $q = 0.25$ in Figure 3).

4.2 Relative efficiency and bias indicators at optimal levels

Given a sample $\underline{X}_n = (X_1, \dots, X_n)$, let us again denote $S(k) = S_n(k)$ any statistic dependent on k , the number of top o.s.'s used in an inferential procedure related with a parameter of extreme events ξ . The optimal sample fraction for $S(k)$ is denoted $k_0^S(n)/n$, with $k_0^S(n) := \arg \min_k MSE(S_n(k))$, the so-called optimal level for the estimation of the parameter ξ .

We shall now present, for $n = 200, 500, 1000, 2000$ and 5000 , and with \bullet denoting H or

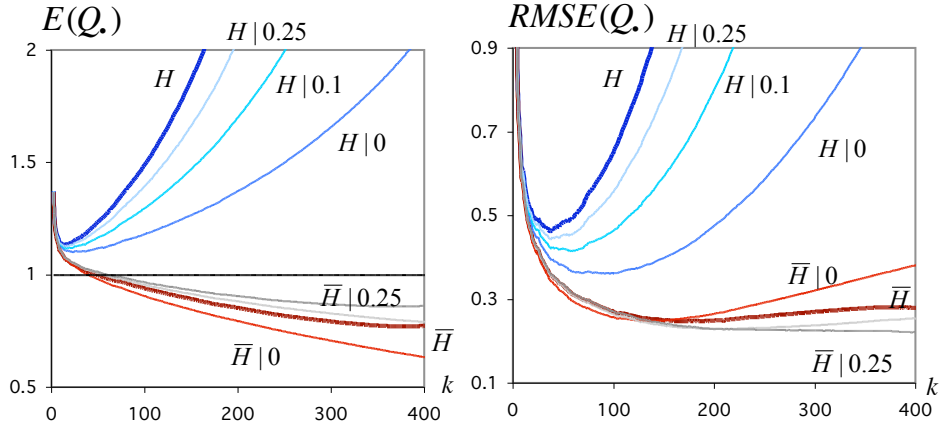


Figure 3: Underlying Extreme Value parent with $\gamma = 0.5$ ($\rho = -0.5$).

\bar{H} or $H|q$ or $\bar{H}|q$, the simulated optimal sample fraction (k_0^\bullet/n), mean values (E_0^\bullet) and relative efficiencies ($REFF_0^\bullet$) of Q_\bullet , at their optimal levels. The search of the minimum MSE has been performed over the region of k -values between 1 and $[0.95 \times n]$. The MSE of $Q_H(k_0^H)$ is also provided so that it is possible to recover the MSE of any other quantile estimator. For a certain Q_\bullet , the $REFF_0^\bullet$ indicator is given by

$$REFF_0^\bullet := \sqrt{\frac{MSE\{Q_H(k_0^H)\}}{MSE\{Q_\bullet(k_0^\bullet)\}}} =: \frac{RMSE_0^H}{RMSE_0^\bullet}.$$

Among the estimators considered, and for all n , the one providing the smallest squared bias and smallest MSE, or equivalently, the highest $REFF$ is underlined and in **bold**. Tables 1, 2 and 3 are related with the underlying parents in A., B., and C., respectively.

For an easier visualization, we present, in Figure 4, the $REFF$ -indicators of the new VaR-estimators, in (1.12), as well as of the PORT-Weissman-Hill VaR-estimators, at optimal levels, comparatively with the classical VaR-estimator, in (1.9), also at its optimal level. Figure 5 is equivalent to Figure 4, but with the simulated mean value of $Q_{\bullet 0} = Q_\bullet(k_0^\bullet)$.

Regarding the $REFF$ -indicators, we would like to draw the following comments:

- For models like the Burr, with a left endpoint equal to zero, we cannot achieve any improvement with the shifted estimators.
- For a model like the Student t_ν , here illustrated for $\nu = 4$, the quasi-PORT VaR-estimators have the best performance for all n , if $q = 0.1$. However, the PORT-Weissman-Hill VaR-estimators associated with $q = 0.1$ exhibit also a high efficiency.
- For an underlying EV model, we reach a clear improvement in the estimation of a high quantile, whenever we consider the quasi-PORT estimators, in (1.12). Note however that

Table 1: Simulated optimal sample fractions (k_0^*/n), mean values (E_0^*), MSE of Q_{H_0} 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
k_0^*/n (and 95% confidence intervals)					
Q_H	0.0903 \pm 0.0039	0.0614 \pm 0.0028	0.0458 \pm 0.0020	0.0328 \pm 0.0012	0.0222 \pm 0.0009
$Q_{H 0}$	0.0848 \pm 0.0041	0.0597 \pm 0.0022	0.0440 \pm 0.0019	0.0321 \pm 0.0010	0.0218 \pm 0.0008
$Q_{H 0.1}$	0.0685 \pm 0.0037	0.0451 \pm 0.0026	0.0331 \pm 0.0014	0.0240 \pm 0.0010	0.0146 \pm 0.0006
$Q_{H 0.25}$	0.0568 \pm 0.0033	0.0367 \pm 0.0022	0.0266 \pm 0.0012	0.0189 \pm 0.0008	0.0121 \pm 0.0005
$Q_{\bar{H}}$	0.8720 \pm 0.0021	0.7632 \pm 0.0025	0.5950 \pm 0.0066	0.4106 \pm 0.0099	0.1203 \pm 0.0051
$Q_{\bar{H} 0}$	0.8088 \pm 0.0044	0.6863 \pm 0.0061	0.5139 \pm 0.0071	0.3592 \pm 0.0100	0.1039 \pm 0.0050
$Q_{\bar{H} 0.1}$	0.4345 \pm 0.0174	0.3035 \pm 0.0091	0.1485 \pm 0.0068	0.0745 \pm 0.0031	0.0340 \pm 0.0012
$Q_{\bar{H} 0.25}$	0.3505 \pm 0.0134	0.2079 \pm 0.0092	0.0943 \pm 0.0048	0.0495 \pm 0.0022	0.0244 \pm 0.0013
E_0^* (and 95% confidence intervals)					
Q_H	1.0854 \pm 0.0034	1.0818 \pm 0.0033	1.0791 \pm 0.0035	1.0737 \pm 0.0026	1.0695 \pm 0.0023
$Q_{H 0}$	1.0856 \pm 0.0036	1.0830 \pm 0.0027	1.0785 \pm 0.0035	1.0740 \pm 0.0023	1.0697 \pm 0.0025
$Q_{H 0.1}$	1.0888 \pm 0.0035	1.0847 \pm 0.0049	1.0839 \pm 0.0033	1.0824 \pm 0.0034	1.0751 \pm 0.0029
$Q_{H 0.25}$	1.0904 \pm 0.0035	1.0864 \pm 0.0045	1.0860 \pm 0.0032	1.0845 \pm 0.0027	1.0823 \pm 0.0029
$Q_{\bar{H}}$	0.9660 \pm 0.0038	0.9873 \pm 0.0018	1.0012 \pm 0.0013	1.0390 \pm 0.0007	1.0596 \pm 0.0011
$Q_{\bar{H} 0}$	0.9566 \pm 0.0043	0.9855 \pm 0.0025	1.0113 \pm 0.0009	1.0481 \pm 0.0007	1.0590 \pm 0.0015
$Q_{\bar{H} 0.1}$	0.9676 \pm 0.0010	1.0469 \pm 0.0016	1.0686 \pm 0.0023	1.0728 \pm 0.0022	1.0706 \pm 0.0019
$Q_{\bar{H} 0.25}$	0.9823 \pm 0.0014	1.0614 \pm 0.0023	1.0739 \pm 0.0028	1.0771 \pm 0.0025	1.0774 \pm 0.0028
$MSE_0^{Q_H}$	0.0648 \pm 0.0013	0.0478 \pm 0.0006	0.0381 \pm 0.0004	0.0303 \pm 0.0002	0.0223 \pm 0.0002
$REFF_0^*$ (and 95% confidence intervals)					
$Q_{H 0}$	0.9846 \pm 0.0009	0.9885 \pm 0.0006	0.9908 \pm 0.0005	0.9926 \pm 0.0002	0.9943 \pm 0.0002
$Q_{H 0.1}$	0.9241 \pm 0.0032	0.9143 \pm 0.0027	0.9035 \pm 0.0027	0.8907 \pm 0.0038	0.8758 \pm 0.0033
$Q_{H 0.25}$	0.8769 \pm 0.0041	0.8635 \pm 0.0045	0.8486 \pm 0.0040	0.8325 \pm 0.0034	0.8107 \pm 0.0031
$Q_{\bar{H}}$	1.8538 \pm 0.0219	2.2510 \pm 0.0187	2.4571 \pm 0.0103	2.1555 \pm 0.0103	1.5554 \pm 0.0092
$Q_{\bar{H} 0}$	1.7285 \pm 0.0174	2.1311 \pm 0.0155	2.2926 \pm 0.0139	1.9577 \pm 0.0081	1.5124 \pm 0.0091
$Q_{\bar{H} 0.1}$	1.5795 \pm 0.0145	1.5271 \pm 0.0103	1.3281 \pm 0.0095	1.1922 \pm 0.0055	1.0883 \pm 0.0038
$Q_{\bar{H} 0.25}$	1.4566 \pm 0.0153	1.3178 \pm 0.0090	1.1593 \pm 0.0062	1.0553 \pm 0.0026	0.9689 \pm 0.0012

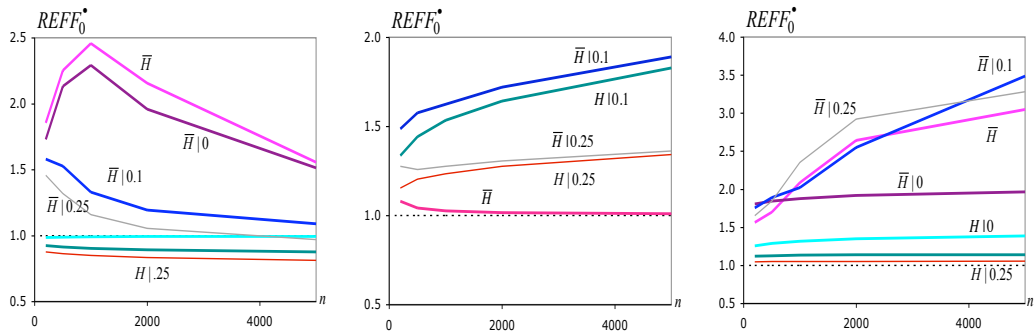


Figure 4: REFF-indicators for a Burr model with $\gamma = 0.25$ and $\rho = -0.5$ (left), a Student t_4 model (center) and an EV model, with $\gamma = 0.5$ (right).

Table 2: Simulated optimal sample fractions (k_0^\bullet/n), mean values (E_0^\bullet), MSE of Q_{H_0} and relative efficiency measures ($REFF_0^\bullet$) at optimal levels, together with corresponding 95% confidence intervals, for a Student t_4 parent ($\gamma = 0.25, \rho = -0.5$).

n	200	500	1000	2000	5000
k_0^\bullet/n (and 95% confidence intervals)					
Q_H	0.0808 \pm 0.0037	0.0610 \pm 0.0017	0.0501 \pm 0.0011	0.0404 \pm 0.0006	0.0297 \pm 0.0005
$Q_{H 0.1}$	0.1983 \pm 0.0034	0.1684 \pm 0.0017	0.1512 \pm 0.0016	0.1377 \pm 0.0010	0.1242 \pm 0.0006
$Q_{H 0.25}$	0.1328 \pm 0.0040	0.1064 \pm 0.0016	0.0895 \pm 0.0015	0.0760 \pm 0.0010	0.0624 \pm 0.0004
$Q_{\bar{H}}$	0.1560 \pm 0.0083	0.0785 \pm 0.0027	0.0568 \pm 0.0014	0.0429 \pm 0.0009	0.0314 \pm 0.0005
$Q_{\bar{H} 0.1}$	0.2843 \pm 0.0032	0.2234 \pm 0.0038	0.1825 \pm 0.0021	0.1574 \pm 0.0010	0.1367 \pm 0.0006
$Q_{\bar{H} 0.25}$	0.2003 \pm 0.0042	0.1311 \pm 0.0031	0.1001 \pm 0.0018	0.0824 \pm 0.0011	0.0650 \pm 0.0007
E_0^\bullet (and 95% confidence intervals)					
Q_H	0.9630 \pm 0.0081	0.9791 \pm 0.0052	0.9933 \pm 0.0040	1.0014 \pm 0.0028	1.0027 \pm 0.0028
$Q_{H 0.1}$	0.9888 \pm 0.0044	0.9956 \pm 0.0024	0.9990 \pm 0.0027	1.0006 \pm 0.0018	1.0000 \pm 0.0012
$Q_{H 0.25}$	0.9789 \pm 0.0065	0.9921 \pm 0.0030	0.9977 \pm 0.0035	1.0012 \pm 0.0029	1.0037 \pm 0.0015
$Q_{\bar{H}}$	0.9661 \pm 0.0047	0.9879 \pm 0.0050	0.9974 \pm 0.0046	1.0014 \pm 0.0036	1.0075 \pm 0.0030
$Q_{\bar{H} 0.1}$	0.9644 \pm 0.0039	0.9993 \pm 0.0034	1.0010 \pm 0.0024	1.0007 \pm 0.0015	1.0007 \pm 0.0009
$Q_{\bar{H} 0.25}$	0.9842 \pm 0.0045	0.9980 \pm 0.0043	0.9980 \pm 0.0031	1.0027 \pm 0.0022	1.0031 \pm 0.0019
$MSE_0^{Q_H}$	0.0649 \pm 0.0008	0.0433 \pm 0.0005	0.0317 \pm 0.0003	0.0232 \pm 0.0003	0.0152 \pm 0.0001
$REFF_0^\bullet$ (and 95% confidence intervals)					
$Q_{H 0.1}$	1.3351 \pm 0.0080	1.4418 \pm 0.0084	1.5339 \pm 0.0068	1.6430 \pm 0.0097	1.8276 \pm 0.0134
$Q_{H 0.25}$	1.1536 \pm 0.0061	1.2034 \pm 0.0058	1.2346 \pm 0.0066	1.2756 \pm 0.0064	1.3415 \pm 0.0062
$Q_{\bar{H}}$	1.0792 \pm 0.0065	1.0412 \pm 0.0035	1.0256 \pm 0.0021	1.0166 \pm 0.0015	1.0092 \pm 0.0014
$Q_{\bar{H} 0.1}$	1.4853 \pm 0.0070	1.5756 \pm 0.0118	1.6243 \pm 0.0058	1.7196 \pm 0.0115	1.8894 \pm 0.0136
$Q_{\bar{H} 0.25}$	1.2752 \pm 0.0090	1.2572 \pm 0.0067	1.2760 \pm 0.0060	1.3062 \pm 0.0065	1.3626 \pm 0.0066

even the PORT-estimators based on the Hill estimator provide REFF-indicators higher than one for all n , with the highest indicator associated with $q = 0$ (a shift induced by the minimum of the available sample). The pattern of the REFF-indicators is not so clear-cut when we consider the quasi-PORT VaR-estimators. Anyway, there is, for all n , a value of q providing the highest efficiency: $q = 0.1$ for $n \leq 500$ and $n = 5000$, and $q = 0.25$ for $n = 1000$ and 2000 .

It is also clear from Figures 4 and 5 that there is not a full agreement between REFF and BIAS indicators, but the discrepancies are moderate. Regarding bias at optimal levels, we can draw the following comments:

- For the simulated Burr model the MVRB VaR-estimators exhibit the smallest bias for all n , but not a long way from the quasi-PORT VaR-estimator associated with $q = 0$, as expected.
- For the Student model, it is almost impossible to draw a clear-cut conclusion, but the quasi-PORT VaR-estimator based on $q = 0.1$ has an interesting bias pattern.

Table 3: Simulated optimal sample fractions (k_0^*/n), mean values (E_0^*), MSE of Q_{H_0} and relative efficiency measures ($REFF_0^*$) at optimal levels, together with corresponding 95% confidence intervals, for a $EV_{0.5}$ parent ($\rho = -\gamma = -0.5$).

n	200	500	1000	2000	5000
k_0^*/n (and 95% confidence intervals)					
$Q_{H 0}$	0.1425 \pm 0.0058	0.1047 \pm 0.0046	0.0859 \pm 0.0038	0.0656 \pm 0.0026	0.0466 \pm 0.0016
$Q_{H 0.1}$	0.1050 \pm 0.0058	0.0743 \pm 0.0040	0.0566 \pm 0.0017	0.0429 \pm 0.0016	0.0281 \pm 0.0012
$Q_{H 0.25}$	0.0828 \pm 0.0054	0.0596 \pm 0.0031	0.0452 \pm 0.0022	0.0340 \pm 0.0011	0.0220 \pm 0.0008
$Q_{\bar{H}}$	0.1795 \pm 0.0071	0.1712 \pm 0.0038	0.5089 \pm 0.0011	0.4646 \pm 0.0015	0.3484 \pm 0.0038
$Q_{\bar{H} 0}$	0.2385 \pm 0.0057	0.1756 \pm 0.0038	0.1302 \pm 0.0024	0.1009 \pm 0.0020	0.0718 \pm 0.0019
$Q_{\bar{H} 0.1}$	0.2653 \pm 0.0085	0.2443 \pm 0.0060	0.2330 \pm 0.0495	0.6638 \pm 0.0019	0.6252 \pm 0.0015
$Q_{\bar{H} 0.25}$	0.2485 \pm 0.0091	0.2841 \pm 0.0286	0.5073 \pm 0.0045	0.4844 \pm 0.0021	0.3880 \pm 0.0026
E_0^* (and 95% confidence intervals)					
Q_H	1.2450 \pm 0.0100	1.2110 \pm 0.0088	1.1889 \pm 0.0077	1.1746 \pm 0.0054	1.1570 \pm 0.0055
$Q_{H 0}$	1.2031 \pm 0.0075	1.1710 \pm 0.0062	1.1571 \pm 0.0062	1.1380 \pm 0.0047	1.1181 \pm 0.0042
$Q_{H 0.1}$	1.2205 \pm 0.0093	1.1931 \pm 0.0075	1.1759 \pm 0.0040	1.1626 \pm 0.0056	1.1407 \pm 0.0045
$Q_{H 0.25}$	1.2260 \pm 0.0102	1.2009 \pm 0.0079	1.1844 \pm 0.0067	1.1707 \pm 0.0035	1.1477 \pm 0.0042
$Q_{\bar{H}}$	0.7997 \pm 0.0108	0.8418 \pm 0.0040	0.9148 \pm 0.0052	0.9724 \pm 0.0023	0.9966 \pm 0.0022
$Q_{\bar{H} 0}$	0.7993 \pm 0.0042	0.8392 \pm 0.0037	0.8669 \pm 0.0027	0.8859 \pm 0.0025	0.9082 \pm 0.0024
$Q_{\bar{H} 0.1}$	0.8174 \pm 0.0059	0.8472 \pm 0.0034	0.8777 \pm 0.0042	0.9559 \pm 0.0024	0.9837 \pm 0.0015
$Q_{\bar{H} 0.25}$	0.8335 \pm 0.0063	0.8528 \pm 0.0069	0.9338 \pm 0.0037	0.9857 \pm 0.0021	1.0036 \pm 0.0018
$MSE_0^{Q_H}$	0.4414 \pm 0.0119	0.2909 \pm 0.0059	0.2185 \pm 0.0035	0.1665 \pm 0.0022	0.1161 \pm 0.0019
$REFF_0^*$ (and 95% confidence intervals)					
$Q_{H 0}$	1.2555 \pm 0.0119	1.2887 \pm 0.0119	1.3190 \pm 0.0093	1.3489 \pm 0.0075	1.3869 \pm 0.0068
$Q_{H 0.1}$	1.1213 \pm 0.0084	1.1275 \pm 0.0064	1.1353 \pm 0.0055	1.1381 \pm 0.0027	1.1398 \pm 0.0041
$Q_{H 0.25}$	1.0480 \pm 0.0047	1.0505 \pm 0.0034	1.0523 \pm 0.0034	1.0527 \pm 0.0019	1.0545 \pm 0.0021
$Q_{\bar{H}}$	1.5639 \pm 0.0183	1.7016 \pm 0.0217	2.0847 \pm 0.0204	2.6439 \pm 0.0204	3.0474 \pm 0.0223
$Q_{\bar{H} 0}$	1.8069 \pm 0.0196	1.8423 \pm 0.0213	1.8781 \pm 0.0160	1.9165 \pm 0.0145	1.9639 \pm 0.0167
$Q_{\bar{H} 0.1}$	1.7560 \pm 0.0190	1.8908 \pm 0.0220	2.0224 \pm 0.0170	2.5494 \pm 0.0236	3.4872 \pm 0.0362
$Q_{\bar{H} 0.25}$	1.6532 \pm 0.0184	1.8430 \pm 0.0230	2.3528 \pm 0.0286	2.9227 \pm 0.0260	3.2823 \pm 0.0297

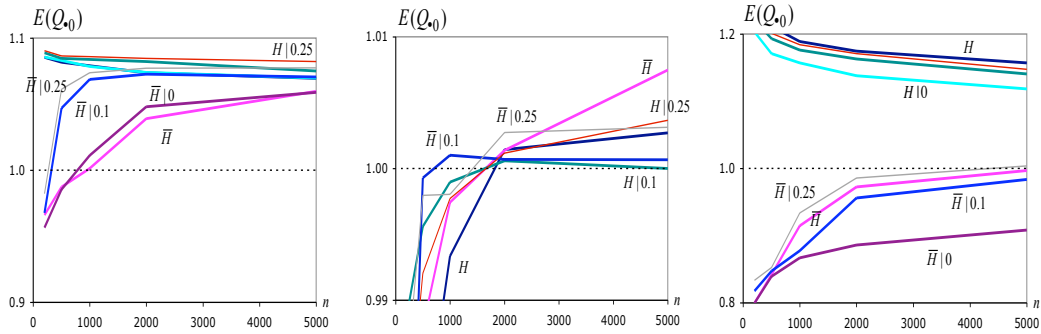


Figure 5: BIAS-indicators for a Burr model with $\gamma = 0.25$ and $\rho = -0.5$ (left), a Student t_4 model (center) and an EV model, with $\gamma = 0.5$ (right).

- For a EV model with $\gamma = 0.5$ and for all n , the smallest bias is achieved by the quasi-PORT quantile estimator based on the shifted version of \bar{H} , for $q = 0.25$. Note also the over-

estimation achieved by the quantile estimators based on the Hill, as a counterpart of an under-estimation achieved by the quantile estimators based on the MVRB EVI-estimator, for small n .

In summary we may draw the following conclusions:

1. If the underlying model has a finite left endpoint at zero, the PORT or quasi-PORT estimators can never beat the original estimators regarding efficiency.
2. For parents with an infinite left endpoint, like the Student parents, or a left endpoint different from zero, like the EV parents, the best performance regarding efficiency is attained by the new estimators for an adequate value of q . Such a q depends on the underlying model and on the sample size n . A similar comment applies to bias reduction.

5 A heuristic choice of tuning parameters and case-studies in the field of finance

5.1 An algorithm for the heuristic choice of k and q

With the notation $X_{0:n} = 0$, and with \bar{H}_n and $\bar{H}_n^{(q)}$, given in (1.8) and (1.11), respectively, we can consider that $\bar{H}_n \equiv \bar{H}_n^{(q)}$ for $q = -1/n$ ($n_q = 0$). Our interest lies now on the estimation of VaR_p through $Q_{p|\bar{H}}^{(q)}$, in (1.12). After Step 3., in the Algorithm provided in Section 2.3, we propose now the following adaptive heuristic estimation of VaR_p .

Algorithm (cont.) (Adaptive estimation of VaR_p).

4. For $q = -1/n, 0(0.05)0.25$ compute the observed values of $Q_{p|\bar{H}}^{(q)}(k)$, $k = 1, 2, \dots, n - [nq] - 1$;
5. Consider the smallest number of decimal places that enables variation in the observed values of $Q_{p|\bar{H}}^{(q)}(k)$, as functions of k (in this case zero decimal places for all data sets considered). Choose q in the following way: for each q consider as possible estimates of VaR_p the values $Q_{p|\bar{H}}^{(q)}(k)$, $k_{min}^{(q)} \leq k \leq k_{max}^{(q)}$, to which is associated the largest run, with a size $l_q = k_{max}^{(q)} - k_{min}^{(q)} + 1$. Choose $q_0 := \arg \max_q l_q$;
6. Consider all those estimates, $Q_{p|\bar{H}}^{(q_0)}(k)$, $k_{min}^{(q_0)} \leq k \leq k_{max}^{(q_0)}$, now with an extra decimal place. Count the frequencies associated to these estimates and obtain the mode of these values, considering them with an extra decimal figure. Let us denote \mathcal{K}^* the set of k -values corresponding to those estimates. Take k_0 as the maximum of \mathcal{K}^* (in order to minimize the variance).

5.2 Applications to data in the field of finance

We have considered the performance of the non-adaptive and adaptive VaR-estimators studied in this paper, when applied to the analysis of the log-returns associated with two of the four sets of finance data considered in Gomes and Pestana (2007b). Those sets of data, collected over the same period, i.e. from January 4, 1999 through November 17, 2005, were the daily closing values of the Dow Jones Industrial Average In (DJI) and Microsoft Corp. (MSFT). Additionally, we have considered over the same period the Euro-GB Pound (EGBP) daily exchange rates, already used in Gomes *et al.* (2008c). All these samples have a size $n = 1762$. The Value at Risk (VaR), defined as a large quantile of negative log-returns, i.e., of $L_i = -\ln(S_{i+1}/S_i)$, $1 \leq i \leq n - 1$, with S_i , $1 \leq i \leq n$, a sample of consecutive close prices, is a common risk measure for large losses. For details about VaR see, for instance, Holton (2003). Here, since we are interested in the analysis of the risk of holding short positions, we have dealt with the positive log-returns, i.e., with $P_i = \ln(S_{i+1}/S_i) = -L_i$, $1 \leq i \leq n - 1$. Although there is some increasing trend in the volatility, stationarity and weak dependence are assumed, under the same considerations as in Drees (2003).

For all data sets we present essentially two figures. In the first one, we picture a box-and-whiskers' plot (*left*) and a histogram (*right*) of the available data. It is clear from all the graphs that all sets of data have heavy left and right tails, and we have thus eliminated the estimators associated with $q = 0$, due to their inconsistency. In the second one, we present at the left the sample path of the $\hat{\rho}_\tau(k)$ estimates in (2.1), as function of k , for $\tau = 0$ and $\tau = 1$, together with the sample paths of the β -estimators in (2.3), also for $\tau = 0$ and $\tau = 1$. At the right, we present, for $p = 1/(2n)$, the estimates of the VaR_p , provided by the $Q_{p|H}$, $Q_{p|\bar{H}}$ and $Q_{p|\bar{H}}^{(q)}$, $q = 0.1, 0.2$, with $Q_{p|H}$ and $Q_{p|\bar{H}}^{(q)}$ given in (1.9) and (1.12), respectively. 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 and for all data sets, the estimate associated with $\tau = 0$.

5.2.1 DJI data

From Figure 6, we immediately see that the underlying model has heavy left and right tails. The number of positive elements in the available sample of log-returns is $n_0 = 867$. Step 3. of the Algorithm here presented led us to the ρ -estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [n_0^{0.999}] = 861$. The associated β -estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 1.03$ (see Figure 7, left). The methodology is quite resistant to different choices of k_1 .

Regarding the VaR-estimation, note that whereas the Weissman-Hill estimator $Q_{p|H}(k)$, in (1.9), is unbiased for the estimation of the tail index VaR_p when the underlying model is a strict Pareto model, it exhibits a relevant bias when we have only Pareto-like tails, as happens

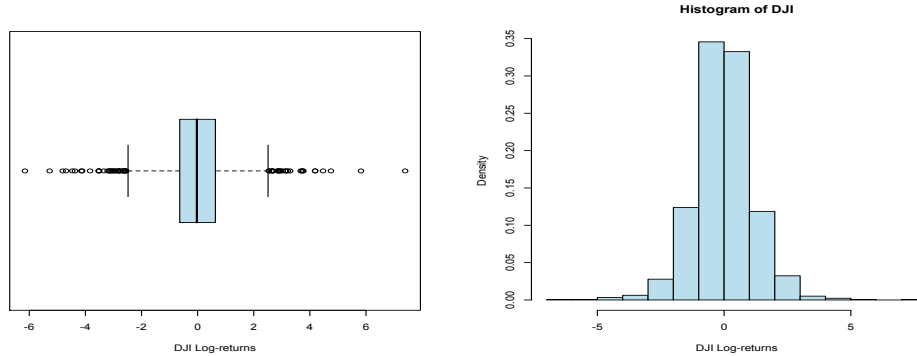


Figure 6: Box-and-whiskers (*right*) and Histogram (*left*) associated with the DJI data.

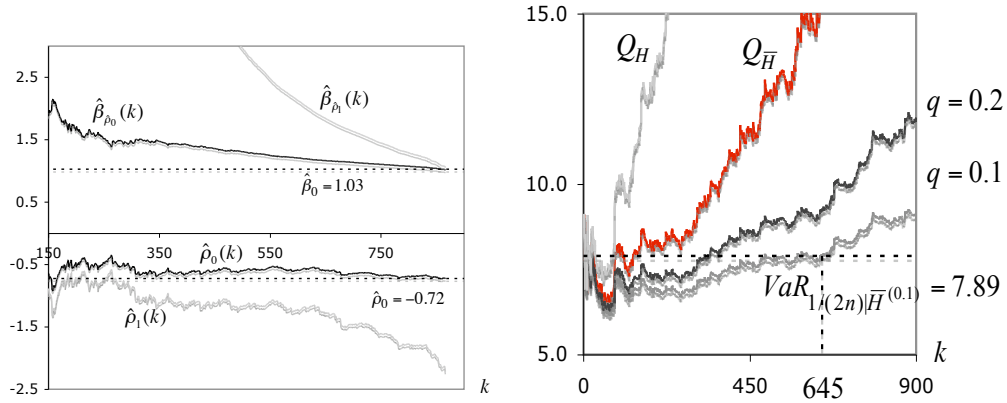


Figure 7: Estimates of the shape second-order parameter ρ and of the scale second-order parameter β (*left*) and quantile estimates (*right*), for the DJI data.

here, and may be seen from Figure 7 (*right*). The quasi-PORT estimators, $Q_{p|\bar{H}}^{(q)}$, in (1.12), based on shifted *MVRB* EVI-estimators, which are “asymptotically unbiased”, have a smaller bias, exhibit more stable sample paths as function of k , and enable us to take a decision upon the estimate of γ and VaR to be used, with the help of any heuristic stability criterion, like the “largest run” suggested in Gomes *et al.* (2004), and written algorithmically in Section 5.1. In this case, the largest run, in Step 5. of the Algorithm, is equal to 327 and was attained by the estimate 8. We have then been led to the choice $q = 0.1$. Next, in Step 6. of the Algorithm, we were led to the choice of an estimate 7.9 (with an associated frequency equal to 75). We finally came to the choice $k = 645$ and to the final estimate $VaR_{1/(2n)|\bar{H}}^{(0.1)} := Q_{1/(2n)|\bar{H}}^{(0.1)}(645) = 7.89$.

5.2.2 MSFT data

Figure 8 and Figure 9 are similar to Figure 6 and Figure 7, respectively, now for the MSFT data.

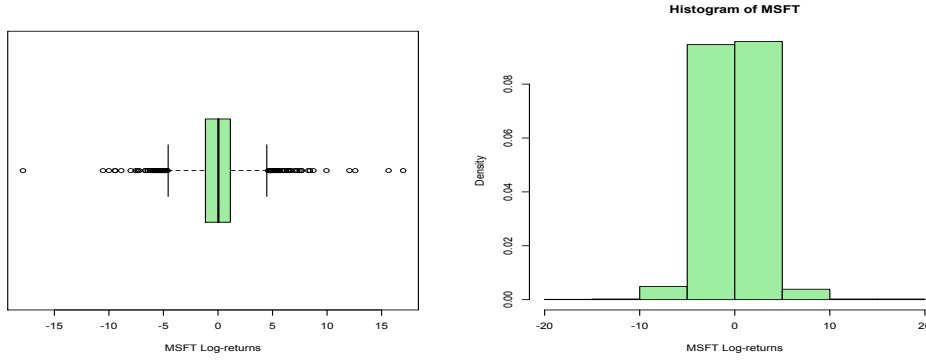


Figure 8: Box-and-whiskers (*right*) and Histogram (*left*) associated with the MSFT data.

As can be inferred from Figure 8, the tails are again heavy. The number of positive elements in the available sample of log-returns is now $n_0 = 882$. Step 3. of the Algorithm here presented led us to the ρ -estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [n_0^{0.999}] = 876$. The associated β -estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 1.02$ (see Figure 9, left).

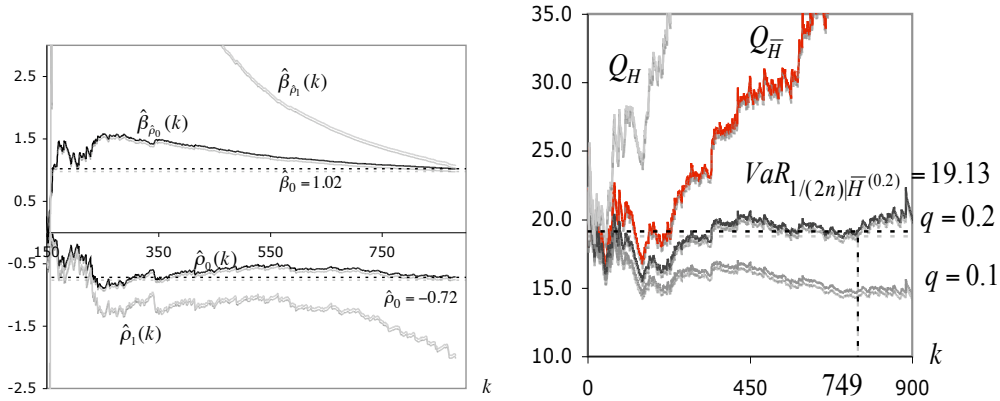


Figure 9: Estimates of the shape second-order parameter ρ and of the scale second-order parameter β (*left*) and quantile estimates (*right*), for the MSFT data.

In this case, the largest run, in Step 5. of the Algorithm, is equal to 113 and was attained by the estimate 19. We have then been led to the choice $q_0 = 0.2$. Next, in Step 6. of the Algorithm, we were led to the choice of an estimate 19.1 (with an associated frequency equal to 24). We finally came to the choice $k = 749$ and to the final estimate $VaR_{1/(2n)|\bar{H}}^{(0.2)} :=$

$Q_{1/(2n)|\bar{H}}^{(0.2)}(749) = 19.13$, the values pictured in Figure 9, right.

5.2.3 EGBP data

Figure 10 and Figure 11 are again similar to Figure 6 and Figure 7, respectively, now for EGBP data, and similar conclusions can be drawn.

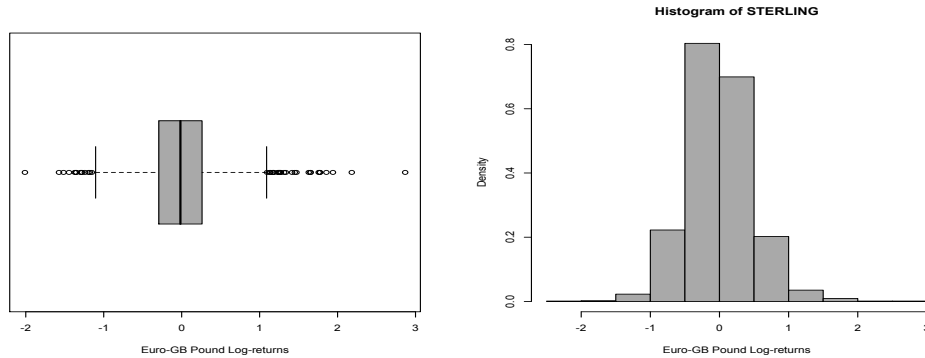


Figure 10: Box-and-whiskers (*right*) and Histogram (*left*) associated with the EGBP data.

The number of positive elements in the available sample of log-returns is now $n_0 = 835$. Step 3. of the Algorithm here presented led us to the ρ -estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.67$, obtained at the level $k_1 = [n_0^{0.999}] = 829$. The associated β -estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 1.03$ (see Figure 11, left).

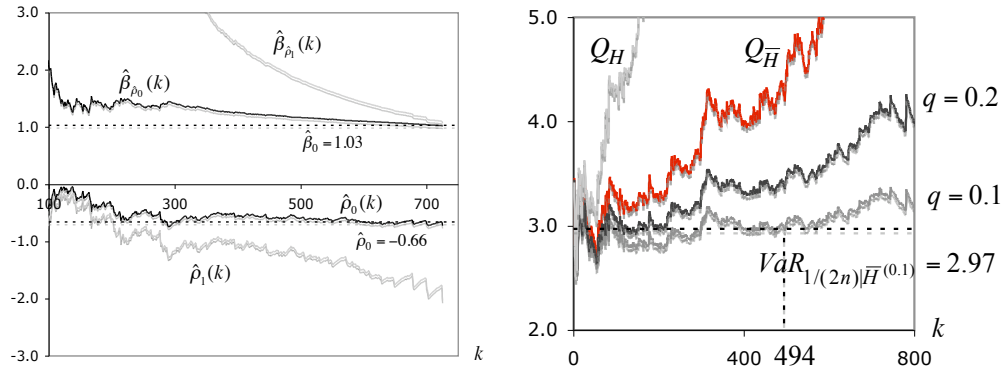


Figure 11: Estimates of the shape second-order parameter ρ and of the scale second-order parameter β (*left*) and quantile estimates (*right*), for the EGBP data.

In this case, the largest run, in Step 5. of the Algorithm, is equal to 886 and was attained by the estimate 3. We have then been led to the choice $q_0 = 0.1$. Next, in Step 6. of the Algorithm, we were led to the choice of an estimate 3.0 (with an associated frequency equal

to 226). We finally came to the choice $k = 494$ and to the final estimate $VaR_{1/(2n)|\bar{H}}^{(0.1)} := Q_{1/(2n)|\bar{H}}^{(0.1)}(494) = 2.97$, the values now pictured in Figure 11, right.

References

- [1] Araújo Santos, P., Fraga Alves, M.I. and Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation. *Revstat* **4**:3, 227-247.
- [2] Beirlant, J., Figueiredo, F., Gomes, M.I. and Vandewalle, B. (2008). Improved reduced-bias tail index and quantile estimators. *J. Statist. Plann. and Inference* **138**:6, 1851-1870.
- [3] Caeiro, F. and Gomes, M.I. (2006). A new class of estimators of a “scale” second order parameter. *Extremes* **9**:3-4, 193-211, 2007.
- [4] Caeiro, F. and Gomes, M.I. (2008). Minimum-variance reduced-bias tail index and high quantile estimation. *Revstat* **6**:1, 1-20.
- [5] Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Commun. in Statist. – Theory and Methods* **38**:7, 1019-1040.
- [6] Caeiro, F., Gomes, M.I. and Pestana, D.D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 111-136.
- [7] Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9**:4, 617-657.
- [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] Gomes, M.I., Caeiro, F. and Figueiredo, F. (2004). Bias reduction of a tail index estimator through an external estimation of the second-order parameter. *Statistics* **38**:6, 497-510.
- [10] Gomes, M.I., Fraga Alves, M.I. and Araújo Santos, P. (2008a). PORT Hill and Moment Estimators for Heavy-Tailed Models. *Commun. Statist. – Simul. and Comput.* **37**, 1281-1306.
- [11] Gomes, M.I., de Haan, L. and Henriques-Rodrigues, L. (2008b). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *J. Royal Statistical Society* **B70**, Issue 1, 31-52.
- [12] Gomes, M.I., Hall, A. and Miranda, C. (2008c). Subsampling techniques and the Jackknife methodology in the estimation of the extremal index. *J. Comput. Statist. and Data Analysis* **52**:4, 2022-2041.
- [13] Gomes, M.I., Henriques-Rodrigues, L., Pereira, H. and Pestana, D. (2009a). Tail index and second order parameters’ semi-parametric estimation based on the log-excesses. *J. Statist. Comput. and Simul.* DOI: 10.1080/00949650902755178.
- [14] Gomes, M.I. and Martins, M.J. (2002). “Asymptotically unbiased” estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5**:1, 5-31.

- [15] Gomes, M.I., Martins, M.J. and Neves, M.M. (2007). Improving second order reduced bias extreme value index estimation. *Revstat* **5**:2, 177-207.
- [16] Gomes, M.I. and Oliveira, O. (2001). The bootstrap methodology in Statistics of Extremes: choice of the optimal sample fraction. *Extremes* **4**:4, 331-358, 2002.
- [17] Gomes, M.I. and Pestana, D.D. (2007a). A simple second order reduced-bias' tail index estimator. *J. Statist. Comp. and Simul.* **77**:6, 487-504.
- [18] Gomes, M.I. and Pestana, D.D. (2007b). A sturdy reduced-bias extreme quantile (VaR) estimator. *J. American Statistical Association*, Vol. **102**, no. 477, 280–292.
- [19] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.
- [20] Holton, G.A. (2003). *Value-at-Risk: Theory and Practice*. Academic Press Advanced Finance Series.
- [21] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. American Statistical Association*, Vol. **73**, 812-815.