The Bootstrap Methodology and Adaptive Reduced-bias Tail Index and Value-at-Risk Estimation*

M. Ivette Gomes

University of Lisbon, DEIO, CEAUL and FCUL

Sandra Mendonça

CEAUL and University of Madeira, Dept. Mathematics and Engineering

and

Dinis Pestana

University of Lisbon, DEIO, CEAUL and FCUL

June 14, 2009

Abstract. Under a semi-parametric framework, we consider second-order minimum-variance reduced-bias (MVRB) estimators of a positive extreme value index, the primary parameter in *Statistics of Extremes*, and associated estimation of the *Value at Risk* (VaR) at a level p, the size of the loss occurred with a small probability p. For these MVRB estimators, we propose the use of bootstrap computer-intensive methods for the adaptive choice of *thresholds*. Applications in the fields of insurance and finance, as well as a small-scale simulation study of the bootstrap adaptive estimators' behaviour, will also be provided.

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

Keywords and phrases. *Statistics of Extremes; Value at Risk; semi-parametric estimation; bootstrap methods.*

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

1 Introduction and preliminaries

In this paper, we deal with the semi-parametric estimation of a positive *extreme value index* γ, the primary parameter in *Statistics of Extremes.* We are also interested in the associated estimation of the *Value at Risk* at a level p, denoted VaR_p, the size of the loss occurred with a fixed small probability p, or, in other words, the associated estimation of the (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. Both parameters are of high interest in the most diversified areas, among which we mention insurance, finance and biostatistics.

Let us denote $U(t) := F^{\leftarrow}(1-1/t)$, $t \geq 1$, a (reciprocal) quantile function such that $VaR_p = U(1/p)$. We shall consider parents with

$$U(t) = C\ t^{\gamma}\Big(1 + A(t)/\rho + o(t^{\rho})\Big), \qquad A(t) = \gamma\ \beta\ t^{\rho}, \tag{1.1}$$

as $t \to \infty$, where $\gamma > 0$, $\rho < 0$ and $\beta \neq 0$. For these heavy-tailed parents and given a sample $\underline{\mathbf{X}}_n = (X_1, \cdots, X_n)$, the classical extreme value index estimator is the Hill estimator (Hill, 1975), denoted $H \equiv H(k)$ and given by

$$H(k) := \frac{1}{k}\ \sum_{i=1}^{k}\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}, \tag{1.2}$$

the average of the k log-excesses over a high random threshold $X_{n-k:n}$, an *intermediate* order statistic (o.s.) with rank $n-k$, i.e., an o.s. such that

$$k = k_n \to \infty \quad \text{and} \quad k/n \to 0,\ \text{as } n \to \infty. \tag{1.3}$$

With Q standing for quantile function, the classical Var_p-estimator, based upon the Hill estimator $H(k)$, in (1.2), is given by

$$Q_{p|H}(k) := X_{n-k+1:n}\ c_k^{H(k)}, \quad c_k \equiv c_{k,n,p} := \frac{k}{np}, \tag{1.4}$$

and has been introduced in Weissman (1978).

But the Hill estimator in (1.2), as well as the associated VaR-estimator in (1.4), reveals usually a high asymptotic bias, i.e., $\sqrt{k}\,(H(k) - \gamma)$ is asymptotically normal with variance γ^2

and a non-null mean value, equal to $\lambda/(1-\rho)$, whenever $\sqrt{k}\ A(n/k) \ \rightarrow\ \lambda \neq 0$, finite, with $A(\cdot)$ the function in (1.1). 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. Recently, several authors have been dealing with bias reduction in the field of *extremes* and a simple class of second-order minimum-variance reduced-bias (MVRB) extreme value index estimators is the one in Caeiro *et al.* (2005), used for a semi-parameteric estimation of $\ln VaR_p$ in Gomes and Pestana (2007b). This class, here denoted $\overline{H} \equiv \overline{H}(k)$, depends upon the estimation of the second-order parameters (β, ρ) in (1.1). Its functional form is

$$\overline{H}(k) \equiv \overline{H}_{\hat{\beta},\hat{\rho}}(k) := H(k)\Big(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})\Big), \tag{1.5}$$

where $(\hat{\beta}, \hat{\rho})$ is an adequate consistent estimator of (β, ρ). Algorithms for the estimation of (β, ρ), as well as a heuristic method for the choice of k in the estimation of γ through $\overline{H}(k)$, are provided in Gomes and Pestana (2007a,b), and reformulated in Section 2 of this paper.

We now rephrase Theorem 4.1 in Beirlant *et al.* (2008): for models in (1.1), let $k = k_n$ be an intermediate sequence such that $c_k = k/np \rightarrow \infty$, $\ln c_k/\sqrt{k} \rightarrow 0$ and $\sqrt{k}A(n/k) \rightarrow \lambda$, finite, as $n \rightarrow \infty$. Let $\hat{\gamma}(k)$ be any consistent estimator of the tail index γ, such that

$$\sqrt{k}\big(\hat{\gamma}(k) - \gamma\big) \ \underset{n\rightarrow\infty}{\overset{d}{\longrightarrow}}\ Normal(\lambda\ b_{\hat{\gamma}},\ \sigma^2_{\hat{\gamma}}), \tag{1.6}$$

and let us consider $Q_{p|\hat{\gamma}}(k)$, with $Q_{p|H}(k)$ provided in (1.4). Then,

$$\sqrt{k}\big(Q_{p|\hat{\gamma}}(k)/VaR_p - 1\big)/\ln c_k \ \underset{n\rightarrow\infty}{\overset{d}{\longrightarrow}}\ Normal(\lambda_{_A}\ b_{\hat{\gamma}},\ \sigma^2_{\hat{\gamma}}),$$

even if we work with reduced-bias tail index estimators like the one in (1.5), 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 $n \rightarrow \infty$. In (1.6), we have $b_H = 1/(1-\rho)$ whereas $b_{\overline{H}} = 0$. The asymptotic variance in (1.6) is equal to γ^2 both for the Hill and the corrected-Hill estimators in (1.2) and (1.5), respectively. Consequently, the $\overline{H}$-estimators outperform the H-estimators for all k. A similar remark applies to $Q_{p|\overline{H}}(k)$, comparatively to $Q_{p|H}(k)$, with $Q_{p|H}(k)$ given in (1.4).

In Section 2 of this paper, we discuss the estimation of the second-order parameters β and ρ. In Section 3, we propose the use of bootstrap computer-intensive methods for the adaptive choice of k, not only for the use of $\overline{H}$, as an estimator of the tail index γ, but also the use of $Q_{p|\overline{H}}$ as an estimator of VaR_p. Applications in the fields of insurance and finance will be given in Section 4. A small-scale simulation study of the proposed bootstrap adaptive estimators is provided in Section 5.

2 Estimation of second-order parameters

The reduced-bias tail index estimators require the estimation of the second-order parameters β and ρ in (1.1). Such an estimation will now be briefly discussed.

2.1 Estimation of the shape second-order parameter ρ

We shall consider a class of estimators, parameterized in a tuning real parameter $\tau \in \mathbb{R}$, with the functional expression,

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

dependent on the statistics

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

where

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

Under mild restrictions on k, these statistics converge to $3(1-\rho)/(3-\rho)$, 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 (1.1), for *intermediate* k-values, i.e., k-values such that (1.3) holds, and also such that $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$.

Remark 2.1. *Under adequate general conditions, and for an appropriate tuning parameter* $\tau =: \tau^*$*, 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 patterns of $\hat{\rho}_0(k)$ in Figures 2 and 5).*

Remark 2.2. *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 elect the value of $\tau = \tau^*$ which provides higher stability for large k, by means of any stability criterion, like the one proposed in the Algorithm of Section 2.3.*

We have here decided for the heuristic choice

$$k_1 = \left[n^{1-\epsilon}\right], \quad \epsilon = 0.001, \tag{2.2}$$

where $[x]$ denotes, as usual, the integer part of x. With such a choice of k_1, and provided that $\sqrt{k_1}\ A(n/k_1) \to \infty$, 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 (1.5). Note that with the choice of k_1 in (2.2), we get $\sqrt{k_1}\ A(n/k_1) \to \infty$ if and only if $\rho > 1/2 - 1/(2\epsilon) = -499.5$, an almost irrelevant restriction, from a practical point of view.

2.2 Estimation of the scale second-order parameter β

For the estimation of β we shall consider

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

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

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

with

$$U_i = i\left(\ln \frac{X_{n-i+1:n}}{X_{n-i:n}}\right), \quad 1 \le i \le k,$$

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.* (2008) and Caeiro *et al.* (2009). Consistency is achieved for models in (1.1), k values such that (1.3) holds and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$, 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).

2.3 An algorithm for the second-order parameter estimation

Based on the above mentioned comments, and on the algorithms proposed before, we now propose the following simplified algorithm:

Algorithm (second-order estimates):

1. Given a sample $(X_1, X_2, \cdots, X_n)$, plot $\hat{\rho}_\tau(k)$ in (2.1), $1 \leq k \leq n-1$, for the tuning parameters $\tau = 0$ and $\tau = 1$.
2. Consider $\{\hat{\rho}_\tau(k)\}_{k \in \mathcal{K}}$, with $\mathcal{K} = ([n^{0.995}],\ [n^{0.999}])$, compute their median, denoted χ_τ, and compute $I_\tau := \sum_{k \in \mathcal{K}} \left(\hat{\rho}_\tau(k) - \chi_\tau\right)^2$, $\tau = 0, 1$. 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 2.3. *If there are negative elements in the sample, the value of n, in the Algorithm, should be replaced by n^+, the number of positive elements in the sample.*

Remark 2.4. *This algorithm leads in almost all situations to the tuning parameter $\tau = 0$ whenever $|\rho| \leq 1$ and $\tau = 1$, otherwise. For details on this and similar algorithms, see* Gomes and Pestana (2007a).

2.4 Asymptotic non-degenerate behaviour of the estimators

The asymptotic normality, as well as the asymptotic behaviour of bias, of the estimators $\hat{\rho}_n^{(\tau)}(k)$ and $\hat{\beta}_{\hat{\rho}}(k)$ in (2.1) and (2.3), respectively, is easier to state if we slightly restrict the class of models in (1.1). In this paper, similarly to what has been done in Gomes *et al.* (2007), we consider a third order framework where we merely make explicit a third order term in (1.1), assuming that

$$U(t) = Ct^\gamma \left(1 + \gamma\beta t^\rho/\rho + \beta' t^{2\rho} + o(t^{2\rho})\right) \tag{2.4}$$

as $t \to \infty$, with $C, \gamma > 0$, $\beta, \beta' \neq 0$, $\rho < 0$. Note that to assume (2.4) is equivalent to say that the more general third order condition

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} - \frac{x^\rho - 1}{\rho}}{B(t)} = \frac{x^{\rho+\rho'} - 1}{\rho + \rho'}, \tag{2.5}$$

holds with $\rho = \rho' < 0$ and that we may choose, in (2.5),

$$A(t) = \alpha\ t^\rho =: \gamma\ \beta\ t^\rho, \qquad B(t) = \beta'\ t^\rho = \frac{\beta' A(t)}{\beta\gamma}, \qquad \beta,\ \beta' \neq 0,$$

with β and β' "scale" second- and third-order parameters, respectively.

Under the third-order framework in (2.4), if (1.3) holds and, with $A(\cdot)$ given in (1.1), $\sqrt{k}A^2(n/k) \to \lambda_{_A}$, finite, as $n \to \infty$, $\sqrt{k}\ A(n/k)\big(\hat\rho_n^{(\tau)}(k) - \rho\big)$ is asymptotically normal with an asymptotic bias proportional to $\lambda_{_A}$. Moreover, if we additionally guarantee that $\hat\rho - \rho = o_p(1/\ln n)$, $\sqrt{k}\ A(n/k)\big(\hat\beta_{\hat\rho}(k) - \beta\big)$ is also asymptotically normal with an asymptotic bias also proportional to $\lambda_{_A}$, as $n \to \infty$ (further details on the subject can be seen in Caeiro *et al.*, 2009).

Remark 2.5. *Despite of the fact that we have not mentioned before the possible dependency on t of the scale second- and third-order parameters β and β' in* (2.4)*, we can indeed have $\beta = \beta(t)$ and $\beta' = \beta'(t)$, with $|\beta(\cdot)|$ and $|\beta'(\cdot)|$ both slowly varying functions, i.e., positive measurable functions, say $g(\cdot)$, such that $\lim_{t\to\infty} g(tx)/g(t) = 1$ for all $x > 0$.*

Remark 2.6. *Several common heavy-tailed models belong to the class in* (2.4)*. Among them we mention:*

1. *the Fréchet model, with d.f. $F(x) = \exp(-x^{-1/\gamma})$, $x \geq 0$, $\gamma > 0$, for which $\rho' = \rho = -1$, $\beta = 0.5$ and $\beta' = 5/6$;*
2. *the Extreme Value* (*EV*) *model, with d.f. $F(x) = \exp\big(-(1+\gamma x)^{-1/\gamma}\big)$, $x > -1/\gamma$, $\gamma > 0$, provided that $\gamma = 1$ or $\gamma \geq 2$. Then $\rho = \rho' = -1$. For $\gamma = 1/2$ we have $\rho = \rho' = -0.5$, and we are also in* (2.4)*. 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 $\gamma > 2$;*
3. *the Generalized Pareto* (*GP*) *model, with d.f. $F(x) = 1 - (1+\gamma x)^{-1/\gamma}$, $x \geq 0$, $\gamma > 0$, for which $\rho' = \rho = -\gamma$ and $\beta = \beta' = 1$;*

4. *the Burr model, with d.f.* $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$*,* $x \geq 0$*,* $\gamma > 0$*,* $\rho' = \rho < 0$ *and, as for the GP model,* $\beta = \beta' = 1$*;*
5. *the Student's* t_ν*-model with* ν *degrees of freedom, with a probability density function*

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

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

Remark 2.7. *The EV model in* Remark 2.6*, item 2., does not belong to the class in* (2.4) *if* $\gamma \in \left(0, 2\right) \setminus \left\{\frac{1}{2}, 1\right\}$*, but it belongs to the class in* (2.5)*. If* $0 < \gamma < 1$*, we get* $\rho = -\gamma$*,* $\beta = 1$*,* $\rho' = \gamma - 1$ *and* $\beta' = -\gamma/2$*. If* $1 < \gamma < 2$*, we get* $\rho = -1$*,* $\beta = 1/2$*,* $\rho' = 1 - \gamma$ *and* $\beta' = -1$*.*

3 The bootstrap methodology and adaptive reduced-bias tail index estimation

Let $\{E_i\}$ denote a sequence of independent, identically distributed standard exponential random variables, and define

$$Z_k := \frac{1}{k}\sum_{i=1}^{k} E_i \qquad \text{and} \qquad \overline{Z}_k := \sqrt{k}\,\left(Z_k - 1\right). \tag{3.1}$$

3.1 Asymptotic behaviour of the MVRB estimator

We now refer the following result, a particular case of Theorem 2.1 in Caeiro *et al.* (2009): for models in (2.4), with $A(\cdot)$ and $\overline{Z}_k$ given in (1.1) and (3.1), respectively, and under mild restrictions on k, we can guarantee that

$$\overline{H}_{\hat\beta,\hat\rho}(k) \;\stackrel{d}{=}\; \gamma + \frac{\gamma\,\overline{Z}_k}{\sqrt{k}} + \Big(b_{\overline{H}}\;A^2(n/k) + O_p\Big(\frac{A(n/k)}{\sqrt{k}}\Big)\Big)(1 + o_p(1)), \tag{3.2}$$

where, with $\xi := \beta'/\beta$,

$$b_{\overline{H}} = \frac{1}{\gamma}\Big(\frac{\xi}{1-2\rho} - \frac{1}{(1-\rho)^2}\Big). \tag{3.3}$$

Consequently, even if $\sqrt{k}\ A(n/k) \to \infty$, with $\sqrt{k}\ A^2(n/k) \to \lambda_{_A}$, finite, the type of levels k where the mean squared error of $\overline{H}(k)$ is minimum,

$$\sqrt{k}\ \big(\overline{H}_{\hat\beta,\hat\rho}(k) - \gamma\big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \text{Normal}\big(\lambda_{_A}\ b_{_{\overline{H}}},\ \sigma^2_{_{\overline{H}}} = \gamma^2\big).$$

If $\sqrt{k}\ A^2(n/k) \to \infty$, $\overline{H}_{\hat\beta,\hat\rho}(k) - \gamma$ is $O_p(A^2(n/k))$.

Regarding the VaR_p-estimation through $Q_{p|\overline{H}}$, with $Q_{p|H}$ and c_k given in (1.4), we have

$$\frac{\sqrt{k}}{\ln c_k}\ \Big(\frac{Q_{p|\overline{H}}(k)}{VaR_p} - 1\Big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \text{Normal}\big(\lambda_{_A}\ b_{_{\overline{H}}},\ \sigma^2_{_{\overline{H}}} = \gamma^2\big). \tag{3.4}$$

3.2 The bootstrap methodology in the estimation of the optimal sample fraction

The bootstrap methodology can enable us to estimate the optimal sample fraction (OSF) $k_0^{\overline{H}}(n)/n := \arg\min_k MSE(\overline{H}(k))/n$, in a way similar to the one used for the classical tail index estimators, now through the use of the auxiliary statistic,

$$T_n^{\overline{H}}(k) := \overline{H}([k/2]) - \overline{H}(k), \quad k = 2, \cdots, n-1, \tag{3.5}$$

which converges in probability to zero, for intermediate k, i.e. whenever (1.3) holds.

Under a semi-parametric framework, with $\mathbb{E}$ denoting the mean value operator, a possible substitute for the MSE of $\overline{H}(k)$ is, cf. equation (3.2),

$$AMSE(\overline{H}(k)) := \mathbb{E}\left(\frac{\gamma\ \overline{Z}_k}{\sqrt{k}} + b_{_{\overline{H}}}\ A^2(n/k)\right)^2 = \frac{\gamma^2}{k} + b^2_{_{\overline{H}}}\ A^4(n/k), \tag{3.6}$$

depending on n and k, and with $AMSE$ standing for *asymptotic mean squared error.* We get

$$\begin{aligned} k_{0|\overline{H}}(n) &:= \arg\min_k AMSE\big(\overline{H}(k)\big) \\ &= \big((-4\rho)\ b^2_{_{\overline{H}}}\ \gamma^2\beta^4\ n^{4\rho}\big)^{-1/(1-4\rho)} = k_0^{\overline{H}}(n)(1+o(1)). \end{aligned} \tag{3.7}$$

The results in Gomes *et al.* (2000) and Gomes and Oliveira (2001) enable us to get, now under the third order framework in (2.4), the following asymptotic distributional representation for $T_n^{\overline{H}}(k)$ in (3.5):

$$T_n^{\overline{H}}(k) \overset{d}{=} \frac{\gamma\ P_k}{\sqrt{k}} + b_{_{\overline{H}}}\ (2^{2\rho} - 1)\ A^2(n/k) + O_p(A(n/k)/\sqrt{k}), \tag{3.8}$$

with P_k asymptotically standard normal, and $b_{\overline{H}}$ given in (3.3). The *AMSE* of $T_n^{\overline{H}}(k)$ is thus minimal at a level $k_{0|T}(n)$ such that $\sqrt{k}\ A^2(n/k) \to \lambda'_A \neq 0$, i.e. a level of the type of the one in (3.7), with $b_{\overline{H}}$ replaced by $b_{\overline{H}}(2^{2\rho}-1)$, and we consequently have

$$k_{0|\overline{H}}(n) = k_{0|T}(n) \left(1 - 2^{2\rho}\right)^{\frac{2}{1-4\rho}} (1 + o(1)). \tag{3.9}$$

Remark 3.1. *Note that a similar procedure works for the MVRB estimators in* Gomes *et al.* (2007, 2008).

How does the bootstrap methodology then work?

Given the sample $\underline{X}_n = (X_1, \cdots, X_n)$ from an unknown model F, and the functional $T_n^{\overline{H}}(k) =: \phi_k(\underline{X}_n)$, $1 \le k < n$, define I_A as the indicator function of the event A, and consider for any $n_1 = O(n^{1-\epsilon})$, with $0 < \epsilon < 1$, the bootstrap sample $\underline{X}^*_{n_1} = (X^*_1, \cdots, X^*_{n_1})$, from the model $F^*_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}}$, the empirical d.f. associated with the original sample $\underline{X}_n$. Next, associate to that bootstrap sample the corresponding bootstrap auxiliary statistic, related with $T_n^{\overline{H}}$, in (3.5), denoted $T^*_{n_1}(k_1) := \phi_{k_1}(\underline{X}^*_{n_1})$, $1 \le k_1 < n_1$. Then, with the obvious notation $k^*_{0|T}(n_1) \equiv k^*_{0|T_n^{\overline{H}}}(n_1) = \arg\min_{k_1} AMSE\big(T^*_{n_1}(k_1)\big)$, we get

$$\frac{k^*_{0|T}(n_1)}{k_{0|T}(n)} = \left(\frac{n_1}{n}\right)^{-\frac{4\rho}{1-4\rho}} (1 + o_p(1)), \text{ as } n \to \infty.$$

Consequently, for another sample size n_2, and for every $\alpha > 1$, we have

$$\frac{\left[k^*_{0|T}(n_1)\right]^{\alpha}}{k^*_{0|T}(n_2)} \left(\frac{n_1^{\alpha}}{n^{\alpha}}\frac{n}{n_2}\right)^{-\frac{4\rho}{1-4\rho}} = \left\{k_{0|T}(n)\right\}^{\alpha-1} (1 + o_p(1)).$$

It is then enough to choose $n_2 = n\left(\frac{n_1}{n}\right)^{\alpha}$ in the expression above in order to have independence of ρ. The simplest choice is $n_2 = n_1^2/n$, i.e. $\alpha = 2$. We then have

$$\frac{\left[k^*_{0|T}(n_1)\right]^2}{k^*_{0|T}(n_2)} = k_{0|T}(n)(1 + o_p(1)), \text{ as } n \to \infty. \tag{3.10}$$

On the basis of (3.10) we are thus able to estimate $k_{0|T}$, and next, equation (3.9) enables us to estimate $k_{0|\overline{H}}(n)$ provided that we know how to estimate ρ. We shall use here the estimator $\hat{\rho}^*$ proposed in Step 3. of the Algorithm in Section 2.3. Then, with $\hat{k}^*_{0|T}$ denoting the sample counterpart of $k^*_{0|T}$, the usual notation $[x]$ for the integer part of x, and taking into account (3.7), we have the k_0-estimate,

$$\hat{k}_0^{\overline{H}} \equiv \hat{k}_0^{\overline{H}}(n; n_1) := \min\left(n-1,\ \left[\frac{\left(1-2^{2\hat{\rho}^*}\right)^{\frac{2}{1-4\hat{\rho}^*}}\left(\hat{k}^*_{0|T}(n_1)\right)^2}{\hat{k}^*_{0|T}([n_1^2/n]+1)}\right]+1\right), \tag{3.11}$$

and the γ-estimate

$$\overline{H}^* \equiv \overline{H}^*(n, n_1) := \overline{H}_{\hat{\beta}^*,\ \hat{\rho}^*}(\hat{k}_0^{\overline{H}}(n; n_1)). \tag{3.12}$$

Note also that, on the basis of (3.4), and with c_k given in (1.4), we get

$$\begin{aligned} k_{0|Q_p} &\equiv k_{0|Q_{p|\overline{H}}} := \arg\min_k AMSE\big(Q_{p|\overline{H}}(k)\big) \\ &= (\ln c_k)^2 AMSE\big(\overline{H}(k)\big). \end{aligned} \tag{3.13}$$

A few practical questions may be raised under the set-up developed: How does the asymptotic method work for moderate sample sizes? What is the type of the sample path of the new estimator for different values of n_1? Is the method strongly dependent on the choice of n_1? What is the sensitivity of the method with respect to the choice of the ρ-estimate? Although aware of the theoretical need to have $n_1 = o(n)$, what happens if we choose $n_1 = n$? Answers to these questions are similar to the ones given for classical estimation (see Hall, 1990; Draisma *et al.*, 1999; Danielsson *et al.* 2001; Gomes and Oliveira, 2001), and some of them will be given in Sections 4 and 5 of this article.

3.3 An algorithm for the adaptive MVRB estimation

The algorithm goes on:

4. Compute $\overline{H}_{\hat{\beta}^*,\ \hat{\rho}^*}(k)$, $k = 1, 2, \cdots, n-1$, with $\overline{H}_{\hat{\beta},\ \hat{\rho}}(k)$ given in (1.5).
5. Next, consider sub-sample sizes $n_1 = o(n)$ and $n_2 = [n_1^2/n] + 1$.
6. For l from 1 until $B = 250$, generate independently B bootstrap samples $(x_1^*, \cdots, x_{n_2}^*)$ and $(x_1^*, \cdots, x_{n_2}^*, x_{n_2+1}^*, \cdots, x_{n_1}^*)$, of sizes n_2 and n_1, respectively, from the empirical d.f. $F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}}$ associated with the observed sample $(x_1, \cdots, x_n)$.
7. Denoting $T_n^*(k)$ the bootstrap counterpart of $T_n^{\overline{H}}(k)$, in (3.5), obtain $(t^*_{n_1,l}(k),\ t^*_{n_2,l}(k))$, $1 \le l \le B$, the observed values of the statistic $T^*_{n_i}(k),\ i = 1, 2$. For $k = 1, 2, \cdots, n_i - 1$,

compute

$$MSE_1^*(n_i,k) = \frac{1}{B}\sum_{l=1}^{B}\left(t^*_{n_i,l}(k)\right)^2,\ i=1,2, \tag{3.14}$$

as well as

$$MSE_2^*(n_i,k) = \ln^2 c_k \times MSE_1^*(n_i,k),\ i=1,2.$$

8. Obtain $\hat{k}^*_{0|T}(n_i) := \arg\min_{1\le k\le n_i-1} MSE_1^*(n_i,k)$ and $\hat{k}^*_{0|Q_p}(n_i) := \arg\min_{1\le k\le n_i-1} MSE_2^*(n_i,k)$, $i=1,2$.
9. Compute the threshold estimate $\hat{k}_0^{\overline{H}}$, in (3.11).
10. Obtain $\overline{H}^* \equiv \overline{H}^*(n,n_1) := \overline{H}_{\hat{\beta}^*,\hat{\rho}^*}\big(\hat{k}_0^{\overline{H}}\big)$, already provided in (3.12).
11. Compute also

$$\hat{k}_0^{Q_{p|\overline{H}}} \equiv \hat{k}_0^{Q_{p|\overline{H}}}(n;n_1) := \min\left(n-1,\ \left[\frac{\left(1-2^{2\hat{\rho}^*}\right)^{\frac{2}{1-4\hat{\rho}^*}}\left(\hat{k}^*_{0|Q_p}(n_1)\right)^2}{\hat{k}^*_{0|Q_p}([n_1^2/n]+1)}\right]+1\right). \tag{3.15}$$

12. Finally, compute $\overline{Q}_p^* \equiv \overline{Q}^*_{p|\overline{H}} := Q_{p|\overline{H}}\left(\hat{k}_0^{Q_{p|\overline{H}}}\right)$.

Remarks:

(i) If there are negative elements in the sample, and apart from the computation of $c_k = k/(np)$, used in the quantile estimator $Q_{p|\overline{H}}$, with $Q_{p|H}$ given in (1.4), the value of n should be replaced by $n^+ = \sum_{i=1}^{n} I_{\{X_i>0\}}$ (the number of positive values in the sample).

(ii) The use of the sample of size n_2, $(x_1^*,\cdots,x_{n_2}^*)$, and of the extended sample of size n_1, $(x_1^*,\cdots,x_{n_2}^*,x_{n_2+1}^*,\cdots,x_{n_1}^*)$, lead us to increase the precision of the result with a smaller B, the number of bootstrap samples generated in Step 6. of the Algorithm. This is quite similar to the use of the simulation technique of *"Common Random Numbers"* in comparison problems, when we want to decrease the variance of a final answer to $z=y_1-y_2$, inducing a positive dependence between y_1 and y_2.

(iii) The Monte Carlo procedure in Steps 6., 7. and 8. of the Algorithm in this Section can be replicated if we want to associate easily a standard error to the estimated parameters, i.e., we can repeat steps from 6. until 12. r times, obtain $\big(\hat{k}_{0,j}^{\overline{H}}, \overline{H}_j^*, \hat{k}_{0,j}^{Q_{p|\overline{H}}}, \overline{Q}_{p,j}^*\big)$, $1 \le j \le r$, and take as overall estimates the averages

$$\Big(\overline{\hat{k}_{0,j}^{\overline{H}}} := \tfrac{1}{r}\sum_{j=1}^{r} \hat{k}_{0,j}^{\overline{H}}, \quad \overline{\overline{H}_j^*} := \tfrac{1}{r}\sum_{j=1}^{r} \overline{H}_j^*, \overline{\hat{k}_{0,j}^{Q_{p|\overline{H}}}} := \tfrac{1}{r}\sum_{j=1}^{r} \hat{k}_{0,j}^{Q_{p|\overline{H}}}, \quad \overline{\overline{Q}_{p,j}^*} := \tfrac{1}{r}\sum_{j=1}^{r} \overline{Q}_{p,j}^*\Big).$$

(iv) Although aware of the fact that k must be intermediate, we have used the entire region of search of $\arg\min_k MSE(k)$, the region $1 \le k \le n-1$, with n possibly denoting n_1 or n_2 and $MSE(k)$ standing for any mean squared error structure dependent on k. Indeed, whenever working with small values of n, $k = 1$ or $k = n-1$ are also possible candidates to the value of an intermediate sequence for such n. Here we have again the old controversy between theoreticians and practioners — $k_n = [c \ln n]$ is intermediate for every constant c, and if we take for instance $c = 1/5$, $k_n = 1$ for every $n \le 22026$. Also, Hall's formula of the asymptotically optimal level, given by

$$k_0^H(n) = \big((1-\rho)^2 n^{-2\rho}/\big(-2\ \rho\ \beta^2\big)\big)^{1/(1-2\rho)},$$

valid for Hill's estimator and for models in (1.1), may lead, for a fixed n, and for several choices of the parameters β and ρ, to values of $k_0^H(n)$ either equal to 1 or to $n-1$ according as ρ is close to 0 or quite small, respectively.

(v) Although aware of the need to have $n_1 = o(n)$, it seems that, once again, like in the classical semi-parametric estimation, we get good results up till n. A possible choice seems then to be for instance $n_1 = n^{0.95}$ or $n_1 = n^{0.955}$, the one used later on in Sections 4.1 and 4.2.

4 Applications in the fields of finance and insurance

4.1 An application to financial data

We shall now consider an illustration of the performance of the adaptive MVRB estimates under study, through the analysis of one of the sets of financial data considered in Gomes and Pestana (2007b), the log-returns of the daily closing values of the Microsoft Corp. (MSFT) stock, collected from January 4, 1999 through November 17, 2005. This corresponds to a sample of size $n = 1762$, with $n^+ = 882$ positive values. Figure 1 shows a box-plot (left) and a histogram (right) of the available data. Particularly the box-plot clearly provides evidence on

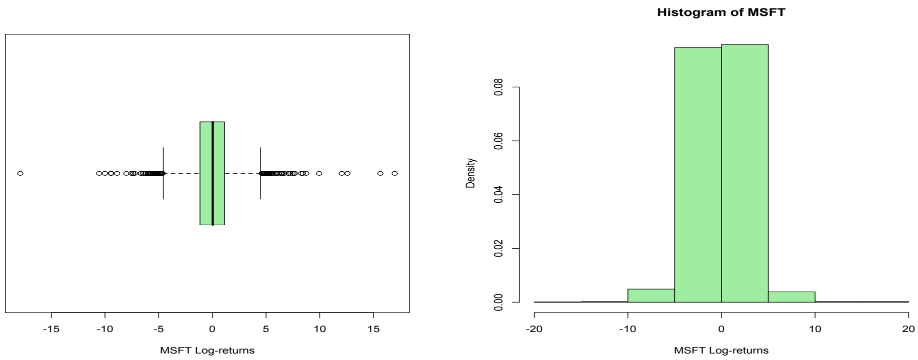

Figure 1: Box-plot (*left*) and Histogram (*right*) of the MSFT log-returns.

the heaviness of both left and right tails.

In Figure 2, we present the sample path of $\hat{\rho}_\tau(k)$ in (2.1), as a 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 $\tau = 1$.

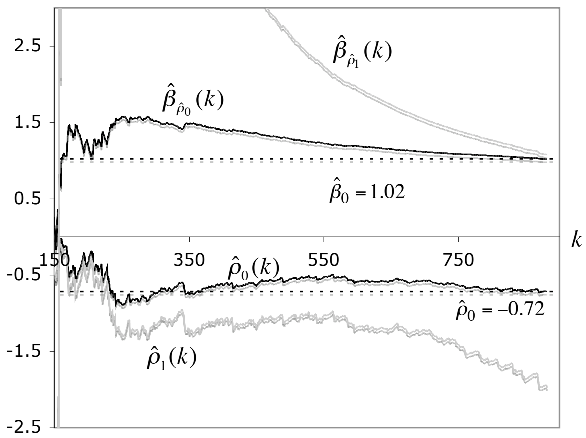

Figure 2: Estimates of the shape second-order parameter ρ and of the scale second-order parameter β for the MSFT log-returns.

Note that the ρ-estimates associated with $\tau = 0$ and $\tau = 1$ lead us to choose, on the basis of any stability criterion for large k, the estimate associated with $\tau = 0$. The algorithm here presented led us to the ρ-estimate $\hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [(n^+)^{0.999}] = 876$. The associated β-estimate is $\hat{\beta}_0 = 1.02$. The methodology is quite resistant to different choices of k_1, leading practically to no changes in the sample paths of the MVRB estimators.

In Figure 3, we present, at the left, the estimates of the tail index γ, provided by the H and the $\overline{H}$ estimators, in (1.2) and (1.5), respectively. At the right we present the corresponding quantile estimates associated to $p = 1/(2n)$ in the quantile estimators $Q_H \equiv Q_{p|H}$ and $Q_{\overline{H}} \equiv Q_{p|\overline{H}}$, with $Q_{p|H}(k)$ provided in (1.4).

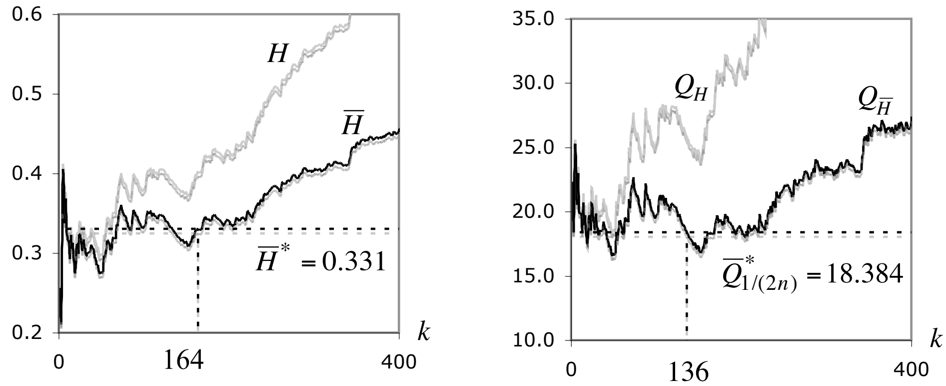

Figure 3: Estimates of the tail index γ (*left*) and of the quantile $\chi_{1-p} \equiv VaR_p$, associated to $p = 1/(2n) \approx 0.00028$ (*right*) for the MSFT log-returns.

Regarding the tail index estimation, note that whereas the Hill estimator is unbiased for the estimation of the tail index γ when the underlying model is a strict Pareto model, it exhibits a relevant bias when we have only Pareto-like tails, as happens here, and may be seen from Figure 3 (*left*). A similar comment applies to the VaR-estimation, as can be seen from Figure 3 (*right*). The MVRB estimators, which are "asymptotically unbiased", have a smaller bias, exhibit more stable sample paths as functions of k, and enable us to take a decision upon the estimate of γ and VaR, even with the help of any heuristic stability criterion, like the "largest run" method suggested in Gomes and Figueiredo (2006). For the Hill estimator, as we know how to estimate the second-order parameters β and ρ, we have simple techniques to estimate the OSF. Indeed, we get $\hat{k}_0^H = \left[\left((1-\hat{\rho}_0)^2 (n^+)^{-2\hat{\rho}_0} / \left(-2\,\hat{\rho}_0\,\hat{\beta}_0^2\right)\right)^{1/(1-2\hat{\rho}_0)}\right] + 1 = 73$, and an associated γ-estimate equal to 0.401. Up to now, we did not have the possibility of adaptively estimate the OSF associated to the MVRB estimates. The algorithm presented in this paper helps us to provide such an adaptive choice. For a sub-sample size $n_1 = [(n^+)^{0.955}] = 650$, and B=250 bootstrap generations, we have got $\hat{k}_0^{\overline{H}} = 164$ and the MVRB tail index estimate $\overline{H}^* = 0.331$, the value pictured in Figure 3 (*left*). For the estimation of $VaR_{1/(2n)}$, the estimated threshold was 136, and $\overline{Q}^*_{1/(2n)} = 18.384$, the value pictured in Figure 3 (*right*).

4.2 An application to insurance data

We shall next consider an illustration of the performance of the adaptive MVRB estimates under study, through the analysis of automobile claim amounts exceeding 1,200,000 Euro over the period 1988-2001, gathered from several European insurance companies co-operating with the same re-insurer (Secura Belgian Re). This data set was already studied in Beirlant *et al.* (2004), Vandewalle and Beirlant (2006) and Beirlant *et al.* (2008) as an example to excess-of-loss reinsurance rating and heavy-tailed distributions in car insurance. First, we have performed a preliminary graphical analysis of the data, x_i, $1 \le i \le n$, $n = 371$. Figure 4 is equivalent to Figure 1, but for the SECURA Belgian Re data.

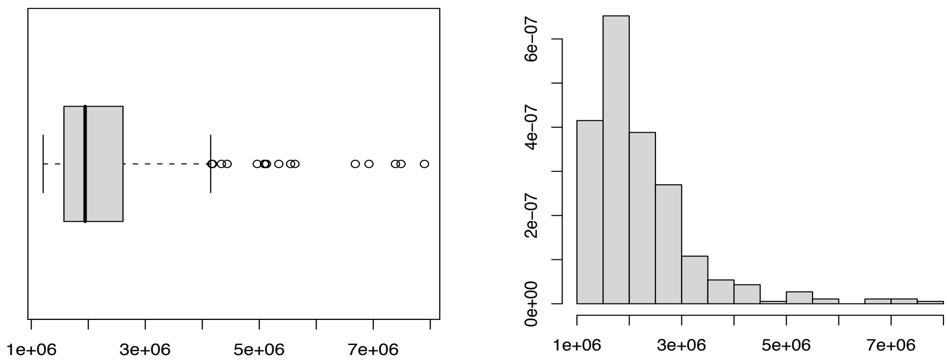

Figure 4: Box-plot (*left*) and Histogram (*right*) of the Secura data.

It is clear from the graph in Figure 4, that data have been censored to the left and that the right tail of the underlying model is quite heavy. Figures 5 and 6 are equivalent to Figures 2 and 3, respectively, but for the Secura Belgian Re data.

Once again, the sample paths of the ρ-estimates associated with $\tau = 0$ and $\tau = 1$ lead

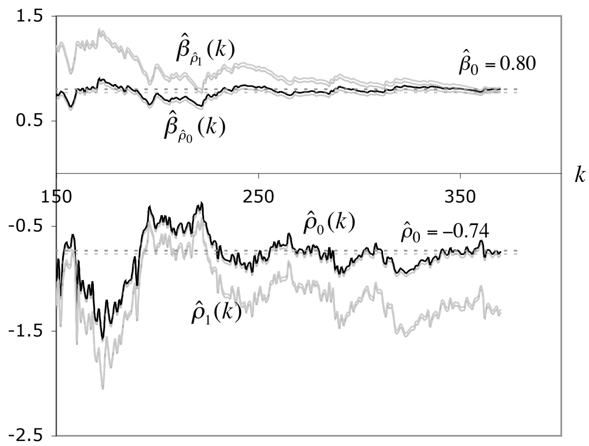

Figure 5: Estimates of the shape second-order parameter ρ and of the scale second-order parameter β for the Secura Belgian Re data.

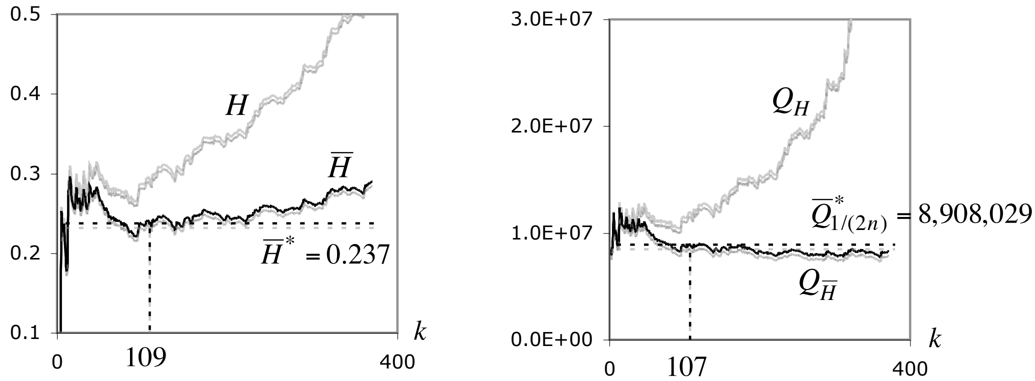

Figure 6: Estimates of the tail index γ (*left*) and of the quantile $\chi_{1-p} \equiv VaR_p$, associated to $p = 1/(2n) \approx 0.00135$ (*right*) for the Secura Belgian Re data.

us to choose, on the basis of any stability criterion for large k, the estimate associated with $\tau = 0$. The algorithm here presented led us to the ρ-estimate $\hat{\rho}_0 = -0.74$, obtained at the level $k_1 = [n^{0.999}] = 368$. The associated β-estimate is $\hat{\beta}_0 = 0.80$. The comments made in Section 4.1 on the classical versus reduced-bias tail index and VaR estimation apply here as well. For the Hill estimator, we get the estimate $\hat{k}_0^H = 55$, and an associated γ-estimate equal to 0.291. The application of the algorithm presented in this paper, with a sub-sample size $n_1 = [n^{0.955}] = 284$, and $B = 250$ bootstrap generations, led us to $\hat{k}_0^{\overline{H}} = 109$ and the to the adaptive MVRB tail index estimate $\overline{H}^* = 0.237$, the value pictured in Figure 6 (*left*). For the estimation of $VaR_{1/(2n)}$, the estimated threshold was 107, and $\overline{Q}^*_{1/(2n)} = 8,908,029$, the value pictured in Figure 6 (*right*).

Remark 4.1. *Note that the estimates obtained are not a long way from the ones obtained before in Beirlant et al.* (2008)*. Note also that bootstrap confidence intervals are easily associated with the estimates presented.*

4.3 Resistance of the methodology to changes in the sub-sample size n_1

Figures 7 and 8 are associated with the data in Sections 4.1 and 4.2, respectively. In these figures, we picture at the left, and as a function of the sub-sample size n_1, ranging, in Figure 7, from $n_1 = (n^+)^{0.95} = 628$ until $n_1 = (n^+)^{0.9999} = 881$, the estimates of the OSF for the adaptive bootstrap estimation of γ through the MVRB estimator $\overline{H}$ in (1.5). The associated γ-estimates are pictured at the right.

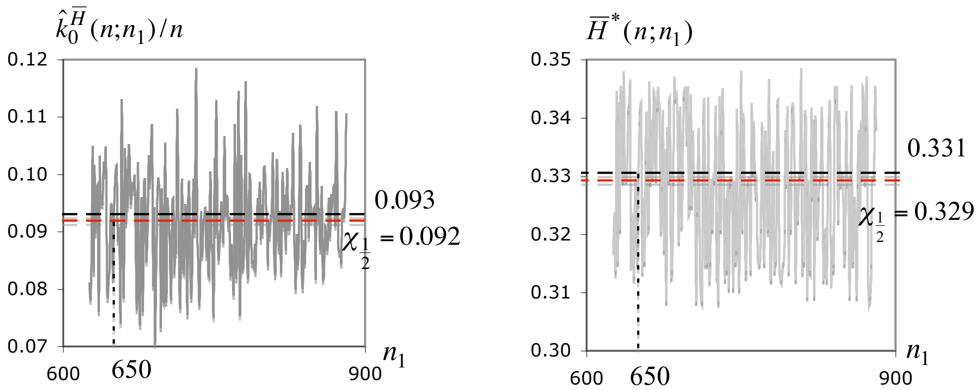

Figure 7: Estimates of the OSF's $\hat{k}_0^{\overline{H}}/n$ (*left*) and the bootstrap adaptive extreme value index estimates $\overline{H}^*$ (*right*), as functions of the sub-sample size n_1, for the MSFT log-returns.

For the MSFT log-returns, the bootstrap tail index estimates are quite stable as a function of the sub-sample size n_1 (see Figure 7, at the right), varying from a minimum value equal to 0.309 until 0.348, with a median equal to 0.329, not a long way from the γ-estimate obtained in Section 4.1, equal to 0.331 and associated with the arbitrarily chosen sub sample size $n_1 = (n^+)^{0.955} = 650$. The bootstrap estimates of the OSF for the estimation of γ through $\overline{H}$ in (1.5) vary from 7% until 12%.

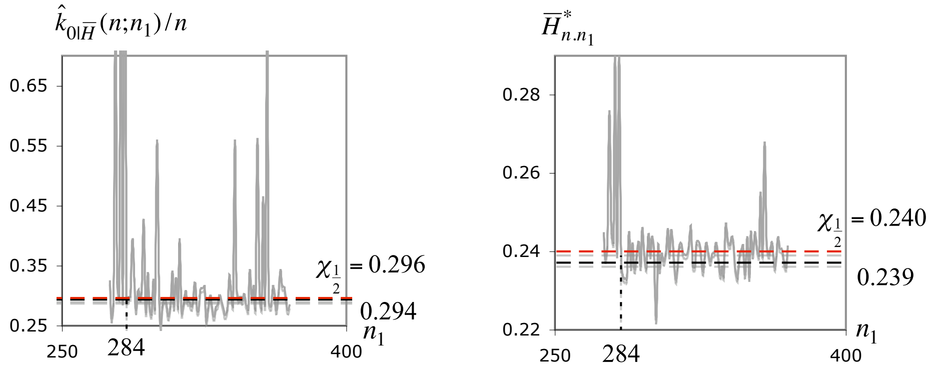

Figure 8: Estimates of the OSF's $\hat{k}_0^{\overline{H}}/n$ (*left*) and the bootstrap adaptive extreme value index estimates $\overline{H}^*$ (*right*), as functions of the sub-sample size n_1, for the Secura Belgian Re data.

For the Secura Belgian Re data, and for sub-sample sizes n_1, ranging from $n_1 = n^{0.95} = 275$ until $n_1 = n^{0.9999} = 370$, the estimates of the OSF are extremely volatile, ranging from 24.8% until 99.7%. This obviously gives rise to estimates of γ ranging from 0.223 until 0.291, with a median equal to 0.240, close by chance to the γ-estimate 0.239 obtained in Section 4.2, at the arbitrarily chosen sub-sample size $n_1 = n^{0.955} = 284$. The above mentioned volatility of the $k_0^{\overline{H}}$-estimates is due to the behaviour of the estimates $MSE_1^*(n_i, k),\ k = 1, 2, \cdots, n_i - 1,\ i = 1, 2,$ in (3.14), with two minima, for a few values of either n_1 or n_2 (11 pairs (n_1, n_2) out of 96), being the global minima achieved at the largest k-value. In Figure 9, left, we can see that only 11 pairs $\big(\hat{k}^*_{0|T}(n_1),\ \hat{k}^*_{0|T}(n_2)\big)$, the ones associated with peaks, are neatly wrong estimates of $\big(k^*_{0|T}(n_1),\ k^*_{0|T}(n_2)\big)$.

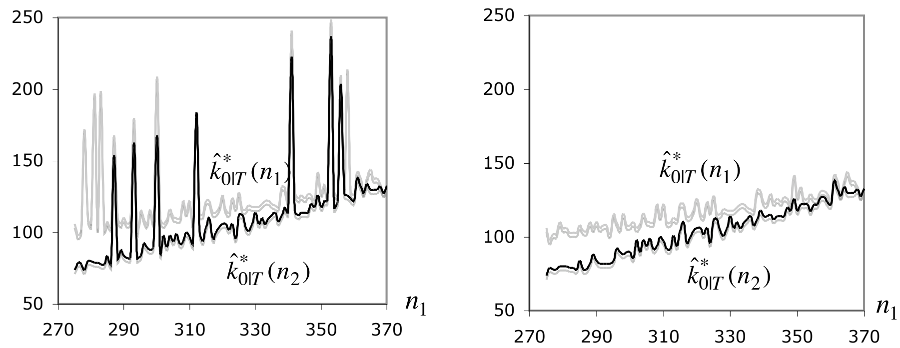

Figure 9: Estimates $\big(\hat{k}^*_{0|T}(n_1),\ \hat{k}^*_{0|T}(n_2)\big)$ for the Secura Belgian Re data, associated to the global minima of $MSE_1^*(n_i),\ i = 1, 2,$ in (3.14) (*left*) and to the first minimum of both MSE's (*right*).

If we restrict the search to the first minimum (something easy to perform for a single data set, but not quite feasible in a simulation), we get the graph in Figure 9, right. The associated estimates of $k_0^{\overline{H}}/n$ vary then from 24.8% until 35%, with a median equal to 0.29 and the adaptive bootstrap estimates $\overline{H}^*(n, n_1)$ range from 0.223 until 0.243, with a median equal to 0.24, as can be seen in Figure 10, where we have chosen the same scales as in Figure 8, for easier comparison.

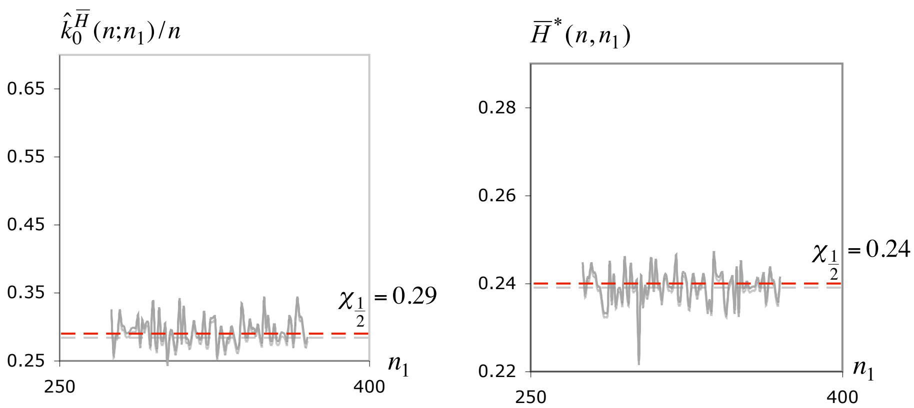

Figure 10: Estimates of the OSF's $\hat{k}_0^{\overline{H}}/n$ (*left*) and the bootstrap adaptive extreme value index estimates $\overline{H}^*$ (*right*), as functions of the sub-sample size n_1, for the Secura Belgian Re data, and a search of the first minima in the bootstrap estimates $MSE_1^*(n_i)$, $i = 1, 2$, in (3.14).

The procedure needs thus to be used with care and we think it is not sensible to choose an "ad hoc" value of n_1, as we did before in Sections 4.1 and 4.2. We suggest here the replacement of Step 5. in the Algorithm by

5'. Consider values of n_1 ranging, for instance, in the above mentioned region, $\mathcal{S} = [(n^+)^{0.95}, (n^+)^{0.9999}]$, and put $n_2 = [n_1^2/n] + 1$.

After running Steps 6., 7. and 8. for the different pairs (n_1, n_2) in Step 5', replace Steps 9., 10., 11. and 12. by

9'. Compute the median, denoted $\hat{k}_{0|M}^{\overline{H}}$, of the estimates $\{\hat{k}_0^{\overline{H}}(n; n_1),\ n_1 \in \mathcal{S}\}$, with $\hat{k}_0^{\overline{H}}$ given in (3.11).

10'. Consider as overall γ-estimate, $\overline{H}^{**} = \overline{H}\big(\hat{k}_{0|M}^{\overline{H}}\big)$, with $\overline{H}(k)$ provided in (1.5).

11'. Compute the median, denoted $\hat{k}_{0|M}^{Q_{p|\overline{H}}}$, of $\{\hat{k}_0^{Q_{p|\overline{H}}}(n; n_1),\ n_1 \in \mathcal{S}\}$, with $\hat{k}_0^{Q_{p|\overline{H}}}$ given in (3.15).

12'. Finally, compute $\overline{Q}_p^{**} := Q_{p|\overline{H}}\left(\hat{k}_{0|M}^{Q_{p|\overline{H}}}\right)$.

In Table 1, we present, jointly with the set $\mathcal{S}$ in Step 5'., the above mentioned estimates, in Steps 9'., 10'., 11'. and 12'., for the two sets of data under analysis in Sections 4.1 and 4.2.

Data	$\mathcal{S}$	$\hat{k}_{0\|M}^{\overline{H}}$	$\overline{H}^{**}$	$\hat{k}_{0\|M}^{Q_{p\|\overline{H}}}$	$\overline{Q}_{(1/2n)}^{**}$
MSFT	[628, 881]	162	0.329	157	18.18
SECURA	[275, 370]	110	0.240	106	8,869,911

Table 1: Adaptive bootstrap estimates of $k_0^{\overline{H}}$, γ, $k_0^{Q_{(1/(2n))|\overline{H}}}$ and $VaR_{1/(2n)}$.

The running of the above mentioned algorithm 100 times provided the average estimates and the 95% bootstrap confidence intervals for γ and $Var_{1/(2n)}$ presented in Table 2. The 95% confidence intervals are based on the quantiles with probability 0.025 and 0.975 of the 100 replicates.

Data	$\hat{\gamma}$	95% CI's for γ	$\widehat{VaR_{1/(2n)}}$	95% CI's for $VaR_{1/(2n)}$
MSFT	0.326	(0.308, 0.346)	18.318	(16.75, 20.79)
SECURA	0.245	(0.225, 0.291)	9,158,849	(8,381,519, 11,696,720)

Table 2: Adaptive bootstrap estimates and 95% confidence intervals for γ and $VaR_{1/(2n)}$.

5 A small-scale simulation study

We have also run the above mentioned algorithm $r = 100$ times for Burr models,

$$F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}, \ x \geq 0, \ \gamma > 0, \ \rho < 0,$$

with $\gamma = 0.5$ and $\rho = -0.5, -0.75$ and -1, for sample sizes $n = 200$, 500 and 1000. An equivalent algorithm, with the adequate modifications, was run for the classical estimator H, in (1.2), and associated $VaR_{1/(2n)}$ estimator, $Q_{1/(2n)|H}$, with $Q_{p|H}$ given in (1.4). The overall estimates of the OSF's $\left(k_0^H/n, k_0^{\overline{H}}/n, k_0^{Q_{p|H}}/n, k_0^{Q_{p|\overline{H}}}/n\right)$ and of $\left(\gamma, VaR_{1/(2n)}\right)$ are the averages of the corresponding r partial estimates and are provided in Tables 3 and 4, respectively. Associated standard errors are provided between parenthesis close to those estimates, both presented at the first row of each entry. In the second row of each entry, we present the 95% confidence intervals based on the quantiles with probability 0.025 and 0.0975 of the 100 partial estimates.

ρ	H	$\overline{H}$	$Q_{1/(2n)\|H}$	$Q_{1/(2n)\|\overline{H}}$
		$n = 200$		
-0.50	0.111 (0.052)	0.361 (0.125)	0.078 (0.048)	0.278 (0.144)
	(0.030, 0.215)	(0.100, 0.600)	(0.015, 0.185)	(0.025, 0.540)
-0.75	0.140 (0.066)	0.431 (0.112)	0.109 (0.063)	0.362 (0.149)
	(0.030, 0.255)	(0.220, 0.630)	(0.015, 0.240)	(0.045, 0.620)
-1.00	0.176 (0.073)	0.490 (0.118)	0.144 (0.079)	0.439 (0.137)
	(0.050, 0.335)	(0.270, 0.735)	(0.020, 0.300)	(0.185, 0.680)
		$n = 500$		
-0.50	0.086 (0.038)	0.268 (0.109)	0.060 (0.037)	0.222 (0.119)
	(0.028, 0.172)	(0.086, 0.486)	(0.006, 0.126)	(0.022, 0.436)
-0.75	0.106 (0.043)	0.369 (0.112)	0.090 (0.050)	0.328 (0.127)
	(0.030, 0.194)	(0.160, 0.548)	(0.006, 0.194)	(0.086, 0.542)
-1.00	0.129 (0.048)	0.453 (0.108)	0.114 (0.051)	0.421 (0.124)
	(0.038, 0.222)	(0.216, 0.612)	(0.012, 0.220)	(0.176, 0.612)
		$n = 1000$		
-0.50	0.068 (0.032)	0.240 (0.096)	0.054 (0.034)	0.217 (0.101)
	(0.022, 0.139)	(0.078, 0.415)	(0.003, 0.138)	(0.036, 0.415)
-0.75	0.092 (0.037)	0.334 (0.100)	0.075 (0.039)	0.298 (0.109)
	(0.029, 0.167)	(0.146, 0.514)	(0.015, 0.149)	(0.099, 0.513)
-1.00	0.119 (0.044)	0.423 (0.097)	0.106 (0.048)	0.403 (0.109)
	(0.045, 0.204)	(0.180, 0.583)	(0.024, 0.200)	(0.170, 0.578)

Table 3: OSF's for the proposed bootstrap adaptive estimation through the MVRB estimator $\overline{H}_\tau$ in (1.5), with $\tau = 0$, and the MVRB estimator $Q_{p=1/(2n)|\overline{H}}$, with $Q_{p|H}$ given in in (1.4), for underlying Burr(γ, ρ) parents, with $\gamma = 0.5$.

Tables 5 and 6 are equivalent to Tables 3 and 4, respectively, but for the *Extreme Value* model,

$$EV_\gamma(x) = \exp(-(1+\gamma x)^{-1/\gamma}),\ 1+\gamma x > 0,$$

with $\gamma = 0.50,\ 0.75$ and 1.00. Note that the d.f. $EV_{0.75}$ does not belong to the class of models

ρ	H	$\overline{H}$	$Q_{1/(2n)\|H}$	$Q_{1/(2n)\|\overline{H}}$	$VaR_{1/(2n)}$
$n = 200$					
-0.50	0.671 (0.162) (0.367, 0.953)	0.619 (0.101) (0.334, 0.776)	26.04 (16.50) (7.28, 65.44)	15.59 (6.38) (6.76, 32.26)	19.00
-0.75	0.585 (0.131) (0.308, 0.846)	0.518 (0.062) (0.411, 0.645)	24.19 (12.73) (8.06, 52.44)	14.36 (5.04) (7.81, 31.43)	19.85
-1.00	0.547 (0.113) (0.297, 0.765)	0.468 (0.055) (0.359, 0.599)	23.09 (11.24) (9.37, 54.60)	13.16 (3.95) (7.79, 25.02)	19.98
$n = 500$					
-0.50	0.634 (0.124) (0.424, 0.829)	0.592 (0.077) (0.418, 0.718)	38.88 (19.99) (9.66, 76.22)	26.70 (8.72) (9.93, 43.46)	30.62
-0.75	0.556 (0.097) (0.365, 0.710)	0.511 (0.050) (0.386, 0.587)	37.14 (16.05) (10.45, 74.83)	22.97 (5.50) (12.40, 36.10)	31.50
-1.00	0.520 (0.081) (0.350, 0.638)	0.467 (0.038) (0.398, 0.551)	34.08 (12.49) (12.37, 61.44)	19.94 (4.04) (14.11, 30.55)	31.61
$n = 1000$					
-0.50	0.615 (0.091) (0.441, 0.762)	0.592 (0.066) (0.424, 0.706)	57.39 (26.46) (17.69, 115.5)	42.38 (12.10) (19.61, 70.55)	43.72
-0.75	0.554 (0.071) (0.398, 0.664)	0.513 (0.039) (0.419, 0.585)	51.37 (17.08) (23.31, 87.94)	33.93 (6.41) (23.45, 48.841)	44.62
-1.00	0.529 (0.058) (0.388, 0.622)	0.469 (0.030) (0.412, 0.519)	49.57 (15.65) (22.04, 80.28)	28.37 (4.25) (21.59, 36.98)	44.71

Table 4: Bootstrap adaptive estimates of γ through the MVRB estimator $\overline{H}_\tau$ in (1.5), with $\tau = 0$, and of $VaR_{1/(2n)}$, through the MVRB estimator $Q_{p=1/(2n)|\overline{H}}$, with $Q_{p|H}$ given in in (1.4), for underlying Burr(γ, ρ) parents, with $\gamma = 0.5$.

in (2.4). This was done in order to understand the robustness of the method to models with $\rho \neq \rho'$.

Some overall remarks:

- Note first that the results for the Burr model hold true for any value γ, provided that we scale the estimates adequately.
- Regarding the OSF's, and as expected, the optimal number of o.s.'s involved in the estimation of any of the parameters when we use the MVRB estimator $\overline{H}$ is always bigger than the corresponding OSF associated with the classical H-estimator. Also, the OSF associated with VaR-estimation is always smaler than the corresponding one for the

γ	H	$\overline{H}$	$Q_{1/(2n)\vert H}$	$Q_{1/(2n)\vert \overline{H}}$
		$n = 200$		
0.50	0.068 (0.035) (0.015, 0.140)	0.228 (0.092) (0.065, 0.395)	0.047 (0.030) (0.015, 0.120)	0.159 (0.101) (0.020, 0.385)
0.75	0.092 (0.044) (0.020, 0.205)	0.268 (0.098) (0.100, 0.440)	0.062 (0.038) (0.015, 0.145)	0.206 (0.118) (0.025, 0.425)
1.00	0.113 (0.054) (0.030, 0.255)	0.306 (0.100) (0.120, 0.495)	0.077 (0.051) (0.015, 0.220)	0.258 (0.122) (0.045, 0.475)
		$n = 500$		
0.50	0.064 (0.031) (0.010, 0.130)	0.205 (0.071) (0.066, 0.326)	0.048 (0.032) (0.006, 0.112)	0.165 (0.084) (0.010, 0.308)
0.75	0.081 (0.034) (0.016, 0.160)	0.254 (0.070) (0.112, 0.366)	0.066 (0.040) (0.006, 0.138)	0.229 (0.084) (0.010, 0.364)
1.00	0.099 (0.039) (0.024, 0.176)	0.297 (0.068) (0.142, 0.398)	0.076 (0.046) (0.006, 0.170)	0.271 (0.079) (0.114, 0.396)
		$n = 1000$		
0.50	0.055 (0.026) (0.015, 0.123)	0.175 (0.065) (0.056, 0.299)	0.040 (0.027) (0.003, 0.096)	0.155 (0.074) (0.026, 0.299)
0.75	0.072 (0.031) (0.019, 0.152)	0.237 (0.067) (0.112, 0.352)	0.058 (0.034) (0.005, 0.134)	0.218 (0.077) (0.054, 0.352)
1.00	0.086 (0.034) (0.031, 0.164)	0.287 (0.062) (0.168, 0.390)	0.077 (0.038) (0.014, 0.164)	0.269 (0.071) (0.118, 0.383)

Table 5: OSF's for the proposed bootstrap adaptive estimation through the MVRB estimator $\overline{H}_\tau$ in (1.5), with $\tau = 0$, and the MVRB estimator $Q_{p=1/(2n)|\overline{H}}$, with $Q_{p|H}$ given in in (1.4), for underlying EV_γ parents.

extreme value index estimation.

- Also, as expected, the confidence intervals based on $\overline{H}$ are always of a smaller size than the ones based on H, both for the extreme value index and for the high quantile. However, the estimates of the quantiles associated with the MVRB estimator are dangerously under estimating the true value of the quantile when ρ becomes small, say close to -1 (see the entries $\rho = -1$ ($n = 500$ and $n = 1000$) for the Burr model and the entries $\gamma = 1$ ($n = 500$ and $n = 1000$) for the EV model), providing bootstrap confidence intervals with an upper limit below the true value of the parameter.
- For small values of $|\rho|$ the estimates of γ based on the MVRB estimator overestimate the

γ	H	$\overline{H}$	$Q_{1/(2n)\|H}$	$Q_{1/(2n)\|\overline{H}}$	$VaR_{1/(2n)}$
			$n = 200$		
0.50	0.596 (0.179) (0.200, 0.903)	0.581 (0.120) (0.316, 0.801)	44.46 (31.37) (13.25, 138.00)	29.52 (13.66) (11.61, 64.23)	37.98
0.75	0.833 (0.217) (0.427, 1.189)	0.745 (0.144) (0.435, 0.991)	147.4 (133.5) (29.48, 670.2)	75.64 (54.62) (25.67, 218.7)	117.81
1.00	1.069 (0.265) (0.522, 1.533)	0.924 (0.160) (0.571, 1.230)	546.2 (613.7) (64.80, 2573.5)	202.8 (156.4) (46.65, 663.2)	398.5
			$n = 500$		
0.50	0.622 (0.155) (0.175, 0.922)	0.583 (0.092) (0.368, 0.710)	81.042 (48.00) (19.60, 169.7)	50.34 (17.49) (20.75, 91.01)	61.23
0.75	0.843 (0.170) (0.427, 1.153)	0.759 (0.084) (0.565, 0.897)	333.0 (256.5) (45.96, 1236.5)	147.1 (57.60) (47.32, 262.2)	235.68
1.00	1.074 (0.200) (0.565, 1.410)	0.942 (0.093) (0.774, 1.116)	1418.0 (1415.8) (115.5, 5482.3)	461.0 (227.8) (150.4, 992.2)	998.5
			$n = 1000$		
0.50	0.618 (0.110) (0.392, 0.853)	0.569 (0.068) (0.385, 0.679)	114.88 (67.33) (33.98, 339.9)	76.02 (20.50) (37.09, 114.0)	87.43
0.75	0.840 (0.127) (0.578, 1.071)	0.749 (0.068) (0.603, 0.861)	573.1 (417.2) (98.11, 1667.0)	244.4 (80.30) (127.8, 460.3)	397.4
1.00	1.065 (0.141) (0.743, 1.276)	0.937 (0.066) (0.809, 1.071)	2987.1 (2370.6) (312.7, 7900.0)	845.6 (371.2) (377.5, 1828.2)	1998.5

Table 6: Bootstrap adaptive estimates of γ through the MVRB estimator $\overline{H}_\tau$ in (1.5), with $\tau = 0$, and of $VaR_{1/(2n)}$, through the MVRB estimator $Q_{p=1/(2n)|\overline{H}}$, with $Q_{p|H}$ given in in (1.4), for underlying EV_γ parents.

true γ (as happens also, and usually even more drastically, with the estimates associated with H). Possibly, the choice $\tau = 1$ in (1.5) would be more sensible in these cases. The associated bootstrap confidence intervals would then cover the true value of $VaR_{1/(2n)}$.

- Despite of the fact that $\rho \neq \rho'$ for the EV model, with $\gamma = 0.75$ (we have then $\rho = -0.75$ and $\rho' = -0.25$), we could not find any difference in the pattern of the adaptive estimates. This shows some resistance of the methodology to changes in the model in (2.4).

References

[1] 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.

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

[3] Caeiro, F. and Gomes, M.I. (2006). A new class of estimators of a "scale" second order parameter *Extremes* **9**, 193-211, 2007.

[4] Caeiro, F. and Gomes, M.I. (2008). Minumum-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] Danielsson, J., de Haan, L., Peng, L. and de Vries, C.G. (2001). Using a bootstrap method to choose the sample fraction in the tail index estimation. *J. Multivariate Analysis* **76**, 226-248.

[8] Draisma, G., de Haan, L., Peng, L. and Pereira, T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme value index. *Extremes* **2**:4, 367-404.

[9] 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.

[10] Gomes, M.I. and Figueiredo, F. (2006). Bias reduction in risk modelling: semi-parametric quantile estimation. *Test* **15**:2, 375-396.

[11] Gomes, M. I., de Haan, L. and Henriques-Rodrigues, L. (2008). 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. 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.

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

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] Hall, P. (1990). Using bootstrap to estimate mean squared error and selecting parameter in non-parametric problems. *J. Multivariate Analysis* **32**, 177-203.

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

[20] Vandewalle, B. and Beirlant, J. (2006). On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance Mathematics and Economics* **38**:3, 444-459.

[21] Weissman, I. (1978). Estimation of parameters and large quantilesbased on the k largest observations. *J. American Statistical Association*, Vol. **73**, 812-815.