Semi-Parametric Probability-Weighted Moments Estimation Revisited*

Frederico Caeiro
Universidade Nova de Lisboa, FCT and CMA

M. Ivette Gomes
Universidade de Lisboa, DEIO, CEAUL and FCUL

and

Björn Vandewalle
Universidade Nova de Lisboa, ISEGI, and CEAUL

January 19, 2012

Abstract

In this paper, for heavy-tailed models and through the use of probability weighted moments based on the largest observations, we deal essentially with the semi-parametric estimation of the *Value-at-Risk* at a level p, the size of the loss occurred with a small probability p, as well as the dual problem of estimation of the *probability of exceedance* of a high level x. These estimation procedures depend crucially on the estimation of the *extreme value index*, the primary parameter in *Statistics of Extremes*, also done on the basis of the same weighted moments. Under regular variation conditions on the right-tail of the underlying distribution function F, we prove the consistency and asymptotic normality of the estimators under consideration in this paper, through the usual link of their asymptotic behaviour to the one of the extreme value index estimator they are based on. The performance of these estimators, for finite samples, is illustrated through Monte-Carlo simulations. An adaptive choice of thresholds is put forward. Applications to a real data set in the field of insurance as well as to simulated data are also provided.

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

Keywords and phrases. Heavy tails, value-at-risk or high quantiles, probability of exceedance of a high level, semi-parametric estimation.

*Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2011 and PEst-OE/MAT/UI0297/2011, and PTDC/FEDER.

1 Introduction, preliminaries and scope of the article

Let $X_1, X_2, \ldots, X_n$ be a set of n independent and identically distributed (i.i.d.), or even possibly weakly dependent and stationary, random variables (r.v.'s), from a population with distribution function (d.f.) F. Let us arrange them in ascending order, to get the order statistics (o.s.'s) $X_{1:n} \leq \cdots \leq X_{n:n}$. Suppose that we are interested in the estimation of a *high quantile* of probability $1-p$, or equivalently, in the estimation of the *Value-at-Risk* (VaR) at a level p, the size of the loss occurred with a small probability p, given by

$$\mathrm{VaR}_p \equiv \chi_{1-p} := F^{\leftarrow}(1-p) = \inf\{x : F(x) \geq 1-p\}, \tag{1.1}$$

with the notation $F^{\leftarrow}$ standing thus for the generalized inverse function of F. Moreover, we are also interested in the estimation of the *probability of exceedance* of a high level $x = x_n$,

$$p = p_x := 1 - F(x) =: \overline{F}(x). \tag{1.2}$$

Extreme Value Theory (EVT) provides a great variety of results that enable us to to deal with alternative approaches in the statistical analysis of extreme events. Those approaches are essentially based on the well-established limiting results described in the following.

1.1 Main limiting results in EVT

The main limiting result in EVT can be attributed to Gnedenko (1943), who fully characterized the possible non-degenerate limiting distribution of the linearly normalised maximum, $(X_{n:n} - b_n)/a_n$, $a_n > 0$, $b_n \in \mathbb{R}$. Such a limit is of the type of the general *extreme value distribution* (EVD),

$$EV_\gamma(x) := \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}), & 1+\gamma x > 0 \quad \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)), & x \in \mathbb{R} \quad \text{if} \quad \gamma = 0. \end{cases} \tag{1.3}$$

When such a non-degenerate limit exists, we say that F belongs to the max-domain of attraction of EV_γ and denote this by $F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma)$. The shape parameter γ is related with the heaviness of the right-tail $\overline{F} = 1 - F$ and it is often called the *extreme value index* (EVI).

Another seminal result in the field of EVT is due to Balkema and de Haan (1974) and Pickands (1975). If we properly scale the excesses over a high threshold u, the limit distribution of those scaled excesses is the *Generalized Pareto distribution* (GPD), strongly related with the d.f. $EV_\gamma(x)$, in (1.3), and defined by,

$$GP_\gamma(x) := 1 + \ln EV_\gamma(x) = \begin{cases} 1 - (1+\gamma x)^{-1/\gamma}, & 1+\gamma x > 0, \quad x > 0 \quad \text{if} \quad \gamma \neq 0 \\ 1 - \exp(-x), & x > 0 \quad \text{if} \quad \gamma = 0 \end{cases} \tag{1.4}$$

(see, for instance, Embrechts *et al.*, 1997, Section 3.4, and Reiss and Thomas, 2007, Section 1.4, for more details).

1.2 Most relevant approaches in the field of Statistics of Univariate Extremes

We shall briefly refer the three most important approaches in the area of *Statistics of Univariate Extremes*: the *block maxima* (BM) method, the *peaks-over-threshold* (POT) or even the *peaks-over-random-threshold* (PORT) methods and the *largest observations* (LOB) method. For a more detailed review, with extensive associated references, see Gomes *et al.* (2008) and Beirlant *et al.* (2012).

- The first method, the BM method, is of a parametric nature: we work with a sample of maxima of adequate blocks of observations, and estimate the parameters $(\lambda, \delta, \gamma)$ of the EVD, $EV_\gamma((x-\lambda)/\delta)$, $\lambda \in \mathbb{R}$, $\delta > 0$, $\gamma \in \mathbb{R}$, with $EV_\gamma(x)$ given in (1.3). This method is known to be possibly inefficient, due to the fact that the loss of information in each block can be catastrophic.
- In the second approach, the POT method, inference is performed through the use of the sample of excesses over a high deterministic threshold u. The limiting d.f. of these excesses is, up to a scale factor, the distribution $GP_\gamma(x)$, in (1.4), and the method can be of a parametric or a semi-parametric nature. Note that the high threshold can also be a random value, leading to the PORT methodology, a terminology recently introduced in Araújo Santos *et al.* (2006).
- The third approach, the LOB method, is the one we shall consider in this paper. It uses the *largest k observations* to make inference about the right tail $\overline{F} = 1 - F$, assuming only that F belongs to a wide sub-domain of $\mathcal{D}_{\mathcal{M}}(EV_\gamma)$.

1.3 Estimators under study

Under the largest observations framework, and whenever dealing with heavy-tailed models, the classical semi-parametric EVI and VaR-estimators are the Hill (Hill, 1975) and Weissman-Hill's estimators (Weissman, 1978), with functional expressions

$$\hat{\gamma}^H_{k,n} := \frac{1}{k}\sum_{i=1}^{k}\left(\ln X_{n-i+1:n} - \ln X_{n-k:n}\right) \tag{1.5}$$

and

$$\widehat{Q}^H_{k,n}(p) := X_{n-k:n}\, c_k^{\hat{\gamma}^H_{k,n}}, \quad c_k \equiv c_k(p) := \frac{k}{np}, \quad k = 1, 2, \ldots, n-1, \tag{1.6}$$

respectively, which are pseudo-maximum likelihood estimators, consistent in the whole $\mathcal{D}^+_{\mathcal{M}} := \mathcal{D}_{\mathcal{M}}(EV_\gamma)_{\gamma>0}$, provided that k is intermediate, i.e. if

$$k = k_n \to \infty \quad \text{and} \quad k/n \to 0, \quad \text{as } n \to \infty. \tag{1.7}$$

In a way dual to (1.6), and given a high level $x = x_n$, the probability $p = p_x$ of exceedance of such a level can be estimated by

$$\hat{p}^H_{k,n}(x) := \left(\frac{k}{n}\right)\widetilde{C}_k^{-1/\hat{\gamma}^H_{k,n}}, \quad \widetilde{C}_k \equiv \widetilde{C}_k(x) := \frac{x}{X_{n-k:n}}, \quad k = 1, 2, \ldots, n-1. \tag{1.8}$$

Under further adequate restrictions on k, we can guarantee the asymptotic normality of the estimators $\hat{\gamma}_{k,n}^H$, $\widehat{Q}_{k,n}^H(p)$ and $\hat{p}_{k,n}^H(x)$, in (1.5), (1.6) and (1.8), respectively. But most of the times, these estimators exhibit a large variance for small k, a strong bias for moderate k, sample paths with very short stability regions around the target value and a very peaked mean square error (MSE) structure, as a function of k. This has led researchers to the search of alternative estimators, with a smaller MSE.

Since heavy-tailed models only have mean value if $\gamma < 1$, methods based on sample moments have been rarely considered when we work with such a type of distributions. But in many practical fields like in finance or insurance, for example, we usually have a positive EVI smaller than one, and even smaller than 1/2. In this article, and for the estimation of the above mentioned parameters of extreme events, we now revisit the use of a *probability weighted moments* (PWM) method based on the largest observations, developed in Caeiro and Gomes (2011) for the EVI.

The PWM method is a generalization of the *method of moments*. It also consists in equating sample moments with their corresponding theoretical moments, and then solving those equations in order to obtain estimates of the different parameters under play. The PWM of a r.v. X are defined by

$$M_{p,r,s} := E(X^p(F(X))^r(1 - F(X))^s),$$

where p, r and s are any real numbers (Greenwood *et al.*, 1979). When $r = s = 0$, $M_{p,0,0}$ are the usual noncentral moments of order p. Hosking *et al.* (1985) advise the use of $M_{1,r,s}$, because then the relations between parameters and moments have usually a much simpler form. Also, when r and s are integers, $F^r(1 - F)^s$ can be written as a linear combination of powers of F or $1 - F$. So it is usual to work with the particular case,

$$a_r := M_{1,0,r} = E(X(1 - F(X))^r), \quad r \geq 0,$$

and the associated estimator,

$$\hat{a}_r = \frac{1}{n} \sum_{i=1}^{n-r} \frac{(n-1-r)!(n-i)!}{(n-1)!(n-i-r)!} X_{i:n} = \frac{1}{n} \sum_{i=1}^{n} \frac{(n-i)(n-i-1)\ldots(n-i-r+1)}{(n-1)(n-2)\ldots(n-r)} X_{i:n}. \tag{1.9}$$

For $\gamma < 1$ and for d.f.'s like the EVD, $EV_\gamma((x - \lambda)/\delta)$, with $EV_\gamma(x)$ given in (1.3), the *Pareto* d.f.,

$$P_\gamma(x;\delta) = 1 - (x/\delta)^{-1/\gamma}, \quad x > \delta, \tag{1.10}$$

and the GPD, $GP_\gamma(x/\delta)$, with $GP_\gamma(x)$ defined in (1.4), the PWM have simple expressions, which allow a straightforward estimation of the EVI, γ. For the EVD, see Hosking *et al.* (1985) and the improved versions in Diebolt *et al.* (2007, 2008). As an example, the Pareto PWM (PPWM) and the generalized Pareto PWM (GPPWM) estimators of γ are valid for $\gamma < 1$, and given by

$$\hat{\gamma}^{PPWM} = 1 - \left(\frac{\hat{a}_1}{\hat{a}_0 - \hat{a}_1}\right) \quad \text{and} \quad \hat{\gamma}^{GPPWM} = 1 - \frac{2\hat{a}_1}{\hat{a}_0 - 2\hat{a}_1}, \tag{1.11}$$

respectively, where $\hat{a}_0$ and $\hat{a}_1$ are given in (1.9). The estimator $\hat{\gamma}^{GPPWM}$, in (1.11), was introduced and studied in Hosking and Wallis (1987).

We shall consider in this paper, the PPWM estimators of VaR_p and p_x, the parameters respectively defined in (1.1) and (1.2), associated with the PPWM EVI-estimators studied in Caeiro and Gomes (2011). Those estimators are semi-parametric in nature and, for comparison with the equivalent estimators based on the Hill EVI-estimator, in (1.5), are based on the top $k+1$ largest o.s.'s, $X_{n-k:n} \le X_{n-k+1:n} \le \cdots \le X_{n:n}$. Under such a framework, the estimators $\hat{a}_0$ and $\hat{a}_1$, in (1.9), should be replaced by,

$$\hat{a}_0(k) := \frac{1}{k+1}\sum_{i=1}^{k+1} X_{n-i+1:n} \quad \text{and} \quad \hat{a}_1(k) := \frac{1}{k+1}\sum_{i=1}^{k+1} \frac{i}{k+1} X_{n-i+1:n},$$

respectively. The PPWM EVI, VaR and p-estimators, based on the largest values are

$$\hat{\gamma}_{k,n}^{PPWM} := 1 - \frac{\hat{a}_1(k)}{\hat{a}_0(k)-\hat{a}_1(k)}, \tag{1.12}$$

$$\hat{Q}_{k,n}^{PPWM}(p) := \frac{\hat{a}_0(k)\ \hat{a}_1(k)}{\hat{a}_0(k)-\hat{a}_1(k)}\Big(\frac{k}{np}\Big)^{\hat{\gamma}_{k,n}^{PPWM}} \tag{1.13}$$

and

$$\hat{p}_{k,n}^{PPWM}(x) := \left(\frac{k}{n}\right)\Big(\frac{x(\hat{a}_0(k)-\hat{a}_1(k))}{\hat{a}_0(k)\hat{a}_1(k)}\Big)^{-1/\hat{\gamma}_{k,n}^{PPWM}}, \tag{1.14}$$

respectively, with $k = 1, 2, \ldots, n-1$, and are consistent whenever $\gamma < 1$.

De Haan and Ferreira (2006) considered, also for $\gamma < 1$, the semi-parametric GPPWM EVI-estimator, based on the sample of excesses over the high random level $X_{n-k:n}$, i.e.,

$$\hat{\gamma}_{k,n}^{GPPWM} := 1 - \frac{2\hat{a}_1^\star(k)}{\hat{a}_0^\star(k)-2\hat{a}_1^\star(k)}, \tag{1.15}$$

with $k = 1, 2, \ldots, n-1$, and $\hat{a}_s^\star(k) := \sum_{i=1}^{k} (i/k)^s\,(X_{n-i+1:n} - X_{n-k:n})/k$, $s = 0,\ 1$. For a finite-sample comparison between the PPWM EVI-estimators in (1.12) and the GPPWM EVI-estimators in (1.15), see Caeiro and Gomes (2011).

1.4 Scope of the article

In Section 2, after reviewing a few results already available in the literature, we state a lemma and a theorem related with the asymptotic properties of the PPWM-estimators, defined in (1.13) and (1.14), of the above mentioned parameters of extreme events, the Value-at-Risk at the level p and the probability p_x of exceedance of a high level x, defined in (1.1) and (1.2), respectively. The performance of these estimators, for finite samples, is illustrated, in Section 3, through a Monte-Carlo simulation study. In Section 4, we put forward an adaptive choice of thresholds, again on the basis of bootstrap computer-intensive methods. Applications to a real data set in the field of insurance as well as to a simulated data set are provided in Section 5.

2 Asymptotic Behaviour of the Estimators

2.1 Most common first and second-order frameworks for heavy tails

For heavy-tailed models, i.e., models with a positive EVI, we assume that F has a Pareto-type right-tail, i.e., with the notation $a(t) \sim b(t)$ if and only if $a(t)/b(t) \to 1$, as $t \to \infty$,

$$\overline{F}(x) = 1 - F(x) \ \sim \ (x/C)^{-1/\gamma} \quad \Longleftrightarrow \quad U(t) \sim C\ t^\gamma, \quad C > 0,\ \gamma > 0, \tag{2.1}$$

where C is a scale parameter and $U(t) := F^{\leftarrow}(1 - 1/t)$, $t > 1$. Note that models with the Pareto-type right-tail, in (2.1), have a regularly varying (RV) right-tail with a negative index of regular variation equal to $-1/\gamma$ (with the notation $\overline{F} \in RV_{-1/\gamma}$), and belong to the max-domain of attraction $\mathcal{D}^+_{\mathcal{M}}$. Indeed, more specifically and for all $x > 0$, we have,

$$F \in \mathcal{D}^+_{\mathcal{M}} \quad \Longleftrightarrow \quad \lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma} \ \big(\overline{F} \in RV_{-1/\gamma}\big) \quad \Longleftrightarrow \quad \lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\gamma \ \big(U \in RV_\gamma\big), \tag{2.2}$$

i.e. $\overline{F}(x) = x^{-1/\gamma} L_{\overline{F}}(x)$ and $U(t) = x^\gamma L_U(t)$, where $L_{\overline{F}}$ and L_U are both slowly varying functions (i.e. they both belong to RV_0), not necessarily converging to constants, as happens if condition (2.1) holds.

To guarantee the consistency of many semi-parametric estimators, we usually need to assume that k is intermediate, i.e., that k is a sequence of integers in $[1, n[$, such that (1.7) holds. To obtain information on the non-degenerate distributional behaviour of semi-parametric estimators of parameters of extreme events, we often assume a second-order condition, like

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \quad \Longleftrightarrow \quad \lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma\ \frac{x^\rho - 1}{\rho}, \tag{2.3}$$

valid for all $x > 0$, where $\rho \le 0$ is a second-order parameter controlling the speed of convergence of $U(tx)/U(t)$ to x^γ, as $t \to \infty$. We then say that U is of second-order regular variation, with parameters γ and ρ, and denote such a fact by $U \in 2RV^\rho_\gamma$. If the limits in (2.3) exist, they are necessarily of the above mentioned types and $|A| \in RV_\rho$ (Geluk and de Haan, 1987; see also de Haan and Ferreira, 2006). Moreover, we have

$$U \in 2RV^\rho_\gamma \quad \Longleftrightarrow \quad \overline{F} \in 2RV^{\tilde\rho}_{-1/\gamma}, \quad \text{with} \quad \tilde\rho = \rho/\gamma$$

and a rate function $\widetilde{A}$ related with A, in (2.3), through the relation $\widetilde{A}(t) = A(1/\overline{F}(t))/\gamma^2$. The validity of condition (2.3), with $\rho < 0$, is equivalent to condition (2.1).

2.2 Auxiliary results on intermediate order statistics

In the sequel, let us denote $(Y_1, \ldots, Y_n)$ a random sample of size n from a strict Pareto model, with d.f. $P_\gamma(x; 1)$, $P_\gamma(x; \delta)$ given in (1.10). Let $(Y_{1:n}, \ldots, Y_{n:n})$ denote the sample of associated ascending o.s.'s. We first state without proof the following results on the asymptotic behaviour of intermediate o.s.'s (see de Haan and Ferreira, 2006, for details).

Lemma 2.1. *If* (1.7) *holds, then for the intermediate Pareto o.s.* $Y_{n-k:n}$, $k\, Y_{n-k:n}/n \overset{p}{\to} 1$ *and consequently,* $Y_{n-k:n} \overset{p}{\to} \infty$, *as* $n \to \infty$. *Moreover,*

$$B_{k,n} := \sqrt{k}\big(k\; Y_{n-k:n}/n - 1\big) \overset{d}{\longrightarrow} \mathcal{B}, \quad \text{a standard normal r.v.,} \tag{2.4}$$

and $Cov(B_{r,n}, B_{s,n}) = \sqrt{r\,s}\;(1 - s/n)/(s-1),\; r < s$.

Lemma 2.2. *Under the first-order framework in* (2.2), *let us assume that* k *is intermediate, i.e.,* (1.7) *holds. Then, for the intermediate o.s.* $X_{n-k:n}$ $\big(\overset{d}{=} U(Y_{n-k:n})\big)$, $X_{n-k:n}/U(n/k) \overset{p}{\to} 1$, *and consequently,* $X_{n-k:n} \overset{p}{\to} \infty$, *as* $n \to \infty$. *If we further assume the validity of the second-order framework in* (2.3), *then, as* $n \to \infty$, *the asymptotic distributional representation*

$$X_{n-k:n}/U(n/k) \overset{d}{=} 1 + \gamma\; B_{k,n}/\sqrt{k} + o_p(A(n/k))$$

holds, where $B_{k,n}$ *is the asymptotically standard normal sequence of r.v.'s in* (2.4).

2.3 Asymptotic behaviour of the EVI-estimators under play

Regarding the Hill estimator, $\hat{\gamma}^{H}_{k,n}$, in (1.5):

Proposition 2.1 (de Haan and Peng, 1998, Theorem 1)**.** *Under the second-order framework in* (2.3), *and for intermediate* k, *i.e. if* (1.7) *holds, the asymptotic distributional representation*

$$\hat{\gamma}^{H}_{k,n} \overset{d}{=} \gamma + \frac{\sigma_{_H} Z^{H}_{k}}{\sqrt{k}} + b_{_H} A(n/k)(1 + o_p(1)), \qquad \sigma_{_H} = \gamma, \quad b_{_H} = \frac{1}{1-\rho},$$

holds, with Z^{H}_{k} *asymptotically standard normal. Consequently, if we choose* k *such that* $\sqrt{k}\, A(n/k) \to \lambda_{_A}$, *finite and not necessarily null,*

$$\sqrt{k}(\hat{\gamma}^{H}_{k,n} - \gamma) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N\left(\frac{\lambda_{_A}}{1-\rho}, \gamma^2\right).$$

We next refer the following results related with the asymptotic behaviour of the PPWM EVI-estimator, $\hat{\gamma}^{PPWM}_{k,n}$, in (1.12):

Proposition 2.2 (Caeiro and Gomes, 2011, Proposition 2.4)**.** *Under the conditions of* Proposition 2.1, *and with* Z^{PPWM}_{k} *asymptotically standard normal, the asymptotic distributional representation*

$$\hat{\gamma}^{PPWM}_{k,n} \overset{d}{=} \gamma + \frac{\sigma_{_{PPWM}} Z^{PPWM}_{k}}{\sqrt{k}} + b_{_{PPWM}} A(n/k)(1 + o_p(1))$$

holds for any $\gamma < 1/2$, *as* $n \to \infty$, *where*

$$\sigma^2_{_{PPWM}} := \frac{\gamma^2(1-\gamma)(2-\gamma)^2}{(1-2\gamma)(3-2\gamma)} \quad \text{and} \quad b_{_{PPWM}} := \frac{(1-\gamma)(2-\gamma)}{(1-\gamma-\rho)(2-\gamma-\rho)}. \tag{2.5}$$

Consequently, if we choose k *such that* $\sqrt{k}\, A(n/k) \to \lambda_{_A}$, *finite and not necessarily null,*

$$\sqrt{k}\Big(\hat{\gamma}^{PPWM}_{k,n} - \gamma\Big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N\left(\lambda_{_A}\; b_{_{PPWM}}, \sigma^2_{_{PPWM}}\right).$$

Remark 2.1. *It is obvious that* $\sigma^2_{_H} := \gamma^2 < \sigma^2_{_{PPWM}}$, *for every* $0 < \gamma < 1/2$. *On the other hand,* $b_{_{PPWM}} < b_{_H} = 1/(1-\rho)$, *unless* $\rho = 0$. *An asymptotic comparison of H and PPWM EVI-estimators at optimal levels can be seen in* Caeiro and Gomes (2011)*.*

2.4 Main asymptotic results

If we combine the results of Lemma 2.1 with the fact that $X_{n-k:n} \stackrel{d}{=} U(Y_{n-k:n})$ and the results in Drees (1998), we first state, without the need of a proof, the following lemma:

Lemma 2.3. *For intermediate k, i.e., whenever* (1.7) *holds, and under the second-order framework in* (2.3)*, but with* $\rho < 0$*,* (2.1) *holds and, if* $p = p_n$ *is a sequence of probabilities such that*

$$c_k \equiv c_k(p) = k/(np) \to \xi \in (0, \infty], \quad \textit{as} \quad n \to \infty, \tag{2.6}$$

then, with

$$C_k \equiv C_k(p) := \left(pY_{n-k:n}\right)^{-1} = \frac{nc_k}{kY_{n-k:n}},$$

and $VaR_p = U(1/p)$*,*

$$\frac{\frac{\mathrm{VaR}_p}{X_{n-k:n}}/C_k^\gamma - 1}{A(n/k)} \xrightarrow{p} \frac{\xi^\rho - 1}{\rho} =: h_\rho(\xi), \tag{2.7}$$

as $n \to \infty$*. By continuity arguments,* $h_\rho(\xi) = -1/\rho$*, if* $\xi = \infty$*. If we furthermore assume that* $x = x_n$ *is a sequence of high levels such that*

$$\widetilde{C}_k \equiv \widetilde{C}_k(x) = x/X_{n-k:n} \xrightarrow{p} \tilde{\xi} \in (0, \infty], \quad \textit{as} \quad n \to \infty, \tag{2.8}$$

and the rate function $\widetilde{A}$ *satisfies* $\gamma^2 \widetilde{A}(t) = A(1/\overline{F}(t))$*, with* A *given in* (2.3)*, then, as* $n \to \infty$*,*

$$\frac{\frac{\overline{F}(x)}{1/Y_{n-k:n}}/\widetilde{C}_k^{-1/\gamma} - 1}{\widetilde{A}(X_{n-k:n})} \xrightarrow{p} \frac{\tilde{\xi}^{\tilde{\rho}} - 1}{\tilde{\rho}} = h_{\tilde{\rho}}(\tilde{\xi})$$

with $\tilde{\rho} = \rho/\gamma$ *and* $h_\rho(\xi)$ *defined in* (2.7)*.*

Corollary 2.1. *Under the conditions of* Lemma 2.3*, if* $p = p_n$ *is a sequence of probabilities and* $x = x_n$ *a sequence of high thresholds such that* (2.6) *and* (2.8) *hold, then, as* $n \to \infty$*,*

$$\frac{\widetilde{C}_k(\mathrm{VaR}_p)}{c_k^\gamma(p)} = 1 + o_p(1) \qquad \textit{as well as} \qquad \frac{\widetilde{C}_k(x)}{c_k^\gamma(\overline{F}(x))} = 1 + o_p(1).$$

Just as happens with the semi-parametric EVI-estimators $\hat{\gamma}_{k,n}^H$ and $\hat{\gamma}_{k,n}^{PPWM}$, in (1.5) and (1.12), respectively, let generally $\hat{\gamma}_{k,n}^\bullet$ be any semi-parametric estimator of γ, such that, with $Z_k^\bullet$ asymptotically standard normal, $\sigma_\bullet > 0$ and $b_\bullet \in \mathbb{R}$,

$$\hat{\gamma}_{k,n}^\bullet \stackrel{d}{=} \gamma + \frac{\sigma_\bullet Z_k^\bullet}{\sqrt{k}} + b_\bullet \, A(n/k)(1 + o_p(1)), \quad \text{as} \quad n \to \infty. \tag{2.9}$$

Consequently, if $\sqrt{k}A(n/k) \to \lambda_{_A}$, finite

$$\sqrt{k}\left(\hat{\gamma}_{k,n}^\bullet - \gamma\right) \underset{n\to\infty}{\xrightarrow{d}} N\left(\lambda_{_A}\, b_\bullet, \sigma_\bullet^2\right) =: \mathcal{G}_\bullet. \tag{2.10}$$

More generally than Theorem 4.1 in Beirlant *et al.* (2008), and Proposition 5 in Caeiro and Gomes (2009), we now state the following theorem, a generalization to the case $\xi < \infty$ of Theorem 4.3.8 and 4.4.7 in de Haan and Ferreira (2006).

Theorem 2.1. *Under the conditions of* Lemma 2.3, *let us further assume that* $p = p_n$ *is a sequence of probabilities such that, as* $n \to \infty$,

(A1) $c_k \equiv c_k(p) = k/(np) \to \xi,\ \xi \in (0,\infty], \qquad \ln c_k = o(\sqrt{k}),$

(A2) $r_k A(n/k) \to \eta_{_A} \in \mathbb{R},$

with

$$r_k = \begin{cases} \sqrt{k} & \text{if } \ \xi \in (0,\infty) \\ \sqrt{k}/\ln c_k & \text{if } \ \xi = \infty. \end{cases}$$

Then, with $\mathrm{VaR}_p = F^{\leftarrow}(1-p) = U(1/p)$, $\widehat{Q}^{H}_{k,n}(p)$, $\widehat{Q}^{PPWM}_{k,n}(p)$, *generally denoted* $\widehat{Q}^{\bullet}_{k,n}(p)$, $\mathcal{B}$ *and* $\mathcal{G}_\bullet$ *given in* (1.1), (1.6), (1.13), (2.4) *and* (2.10), *respectively,* $\rho < 0$ *and* $h_\rho(\xi)$ *given in* (2.7),

$$r_k\big(\ln \widehat{Q}^{\bullet}_{k,n}(p) - \ln \mathrm{VaR}_p\big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{N}_p = \begin{cases} (\ln\xi)\mathcal{G}_\bullet + \gamma\mathcal{B} - \eta_{_A} h_\rho(\xi) & \text{if } \ \xi \in (0,\infty) \\ \mathcal{G}_\bullet + \eta_{_A}/\rho & \text{if } \ \xi = \infty. \end{cases}$$

Indeed, we have the validity of the asymptotic distributional representation

$$\ln \hat{Q}^{\bullet}_{k,n}(p) - \ln \mathrm{VaR}_p \overset{d}{=} \ln c_k\big(\hat{\gamma}^{\bullet}_{k,n} - \gamma\big) + \gamma\mathcal{B} - A(n/k)h_\rho(\xi) + o_p(A(n/k)).$$

On the other hand, let $x = x_n$ *be a sequence of levels such that, as* $n \to \infty$,

(B1) $\widetilde{C}_k \equiv \widetilde{C}_k(x) := x/X_{n-k:n} \overset{p}{\to} \tilde{\xi} \in (0,\infty], \qquad \ln \widetilde{C}_k = o_p(\sqrt{k}),$

(B2) $\widetilde{R}_k\ A(n/k) \overset{p}{\to} \tilde{\eta}_{_A} \in \mathbb{R},$

with

$$\widetilde{R}_k = \begin{cases} \sqrt{k} & \text{if } \ \tilde{\xi} \in (0,\infty) \\ \sqrt{k}/\ln \widetilde{C}_k & \text{if } \ \tilde{\xi} = \infty. \end{cases}$$

Then, similarly as before, with p_x, $\hat{p}^{H}_{k,n}(x)$ *and* $\hat{p}^{PPWM}_{k,n}(x)$, *generally denoted* $\hat{p}^{\bullet}_{k,n}(x)$, *defined in* (1.2), (1.8) *and* (1.14), *respectively,*

$$\widetilde{R}_k\big(\ln \hat{p}^{\bullet}_{k,n}(x) - \ln p_x\big) \underset{n\to\infty}{\overset{d}{\to}} \mathcal{N}_x = \begin{cases} (\ln\tilde{\xi})\mathcal{G}_\bullet/\gamma^2 + \mathcal{B} - \tilde{\eta}_{_A} h_{\tilde{\rho}}(\tilde{\xi})/\gamma^2 & \text{if } \ \tilde{\xi} \in (0,\infty) \\ (\mathcal{G}_\bullet + \tilde{\eta}_{_A}/\tilde{\rho})/\gamma^2 & \text{if } \ \tilde{\xi} = \infty. \end{cases}$$

If $\tilde{\xi} = \infty$ *and* $\sqrt{k}A(n/k) \to \lambda_{_A}$, *finite,*

$$\frac{\gamma^2\sqrt{k}}{\ln \widetilde{C}_k}\left(\frac{\hat{p}^{\bullet}_{k,n}(x)}{p_x} - 1\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{G}_\bullet. \tag{2.11}$$

Indeed, as $n \to \infty$, we have the validity of the asymptotic distributional representation,

$$\hat{p}^{\bullet}_{k,n}(x) \stackrel{d}{=} p\left\{1 + \frac{\ln \widetilde{C}_k}{\gamma^2}\,(\hat{\gamma}^{\bullet}_{k,n} - \gamma)(1 + o_p(1)) + \frac{\mathcal{B}}{\sqrt{k}} + \frac{\widetilde{A}(n/k)(1 + o_p(1))}{\tilde{\rho}}\right\},$$

with $\widetilde{A}(t) = A(1/\overline{F}(t))/\gamma^2$.

Proof. The proof follows straightforwardly from the results in Lemma 2.3, and is left to the reader. □

Remark 2.2. *With $x = U(1/p)$, and since $\gamma > 0$, we can replace* (B1) *by* (A1)*. And we have $\tilde{\xi} = \xi^{\gamma}$. Also, if $\tilde{\xi} = \infty$, we can choose $\widetilde{R}_k = \sqrt{k}/(\gamma \ln c_k)$ (or equivalently, $\widetilde{C}_k = c_k^{\gamma}$).*

Corollary 2.2 (de Haan and Ferreira, 2006, Theorem 4.3.8)**.** *For intermediate k, i.e., whenever* (1.7) *holds, and under the second-order framework in* (2.3)*, but with $\rho < 0$, if $p = p_n$ is a sequence of probabilities such that* (A1) *and* (A2) *hold with $\xi = \infty$, and if $\sqrt{k}A(n/k) \to \lambda_{_A}$, finite, then,*

$$\frac{\sqrt{k}}{\ln c_k}\left(\frac{\widehat{Q}^{\bullet}_{k,n}(p)}{\mathrm{VaR}_p} - 1\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{G}_{\bullet},$$

given in (2.10)*. We can indeed write the asymptotic distributional representation,*

$$\widehat{Q}^{\bullet}_{k,n}(p) \stackrel{d}{=} \mathrm{VaR}_p\left\{1 + \ln c_k(\hat{\gamma}^{\bullet}_{k,n} - \gamma)(1 + o_p(1)) + \frac{\gamma\mathcal{B}}{\sqrt{k}} + \frac{A(n/k)(1 + o_p(1))}{\rho}\right\}.$$

Corollary 2.3 (de Haan and Ferreira, 2006, Theorem 4.4.7)**.** *Under the conditions of* Corollary 2.2*, if we consider that $x = U(1/p)$, then*

$$\frac{\gamma\sqrt{k}}{\ln c_k}\left(\frac{\hat{p}^{\bullet}_{k,n}(x)}{p} - 1\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{G}_{\bullet}.$$

Remark 2.3. *For $\gamma > 0$, the asymptotic dominant behaviour of $\widehat{Q}^{\bullet}_{k,n}(p)$ and $\hat{p}^{\bullet}_{k,n}(x)$ is thus fully determined by the asymptotic behaviour of $\hat{\gamma}^{\bullet}_{k,n}$.*

Remark 2.4. *In a certain dualistic way, the above conditions on the sequence of probabilities $p = p_x$ and the Value-at-Risk,* VaR_p*, regarded as a level $\mathrm{VaR}_p = x = x_n$ can be interchanged. From the second half of* Corollary 2.1*, for instance, when $x = x_n$ is such that $\widetilde{C}_k = \widetilde{C}_k(x)$ satisfies the conditions of* Theorem 2.1*, it is readily understood that the sequence $c_k \equiv c_k(\overline{F}(x)) = k/(n\ \overline{F}(x))$ satisfies the condition $\widetilde{C}_k^{1/\gamma}(x)/c_k(\overline{F}(x)) \overset{p}{\to} 1$, as $n \to \infty$, and consequently, also the conditions on c_k assumed in* Theorem 2.1*. Alternatively, when $p = p_n$ is such that $c_k = c_k(p)$ satisfies the conditions in* Theorem 2.1*, from the first half of* Corollary 2.1 *we see that $\widetilde{C}_k(U(1/p))/c_k^{\gamma}(p) \overset{p}{\to} 1$. Consequently, the conditions on $\widetilde{C}_k(U(1/p))$ assumed in* Theorem 2.1 *also hold.*

Remark 2.5. *Because of this dualistic interpretation, i.e. more specifically from the fact that $\widetilde{C}_k(x)/c_k^{\gamma}(\overline{F}(x)) \overset{p}{\to} 1$, as $n \to \infty$, it is also easily understood how, for the exceedance probability estimator, $\hat{p}^{\bullet}_{k,n}(x)$, and with*

$$r_k(x) = \begin{cases} \sqrt{k} & \text{if } \tilde{\xi} \in (0, \infty) \\ \sqrt{k}/\ln c_k(\overline{F}(x)) & \text{if } \tilde{\xi} = \infty, \end{cases}$$

we can state in an equivalent way that, as $n \to \infty$, $r_k(x)\big(\ln \hat p^{\bullet}_{k,n} - \ln p\big) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{N}_x/\gamma$.

3 Simulation results

In this section, we have implemented a multi-sample Monte Carlo simulation experiment of size 5000×10, to obtain the distributional behaviour of the new semi-parametric VaR-estimators $\widehat{Q}^{PPWM}_{k,n}(p)$, in (1.13), comparatively with the behaviour of the classical estimators $\hat{Q}^{H}_{k,n}(p)$, in (1.6), for $p = 1/n$ ($\xi = \infty$), $p = n^{-1/(1-2\rho)}/10$ (ξ finite and non-null at optimal levels) and $p = 0.01$ ($\xi = 0$, a value out of the scope of Theorem 2.1), for n = 50, 100, 200, 500, 1000, 2000, 5000, 10000 and 20000, and for the following underlying parents: the Fréchet(γ) parent, with d.f. $F(x) = \exp(-x^{-1/\gamma})$, $x > 0$, $\gamma > 0$ ($\rho = -1$) and the Burr(γ, ρ) parent, with d.f. $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x > 0$. For the same models, we have also dealt with the dual problem of estimation of p (again known to be equal to the same values as before), the probability of exceedance of a high level $x = \mathrm{VaR}_p$, through the new estimator $\hat{p}^{PPWM}_{k,n}(x)$, in (1.14), comparatively with the behaviour of the classical estimators $\hat{p}^{H}_{k,n}(x)$, in (1.8). We have further considered the non-parametric (NP) Var_p-estimator, $X_{[n(1-p)]+1:n}$, and p_x-estimator, $\hat{p}_x = \{\#X_i > x, 1 \le i \le n\}/(n + 1)$.

3.1 Simulated mean values and root mean squared errors of the VaR-estimators

To ilustraste the finite sample behaviour of the VaR-estimators, we present, in Figures 1 and 2, the simulated mean value (E) and root mean square error (RMSE) patterns of the normalised values of $\widehat{Q}^{H}_{k,n}(p)$ and $\hat{Q}^{PPWM}_{k,n}(p)$, denoted $\overline{Q}^{H} := \widehat{Q}^{H}_{k,n}(p)/\mathrm{VaR}_p$ and $\overline{Q}^{PPWM} := \hat{Q}^{PPWM}_{k,n}(p)/\mathrm{VaR}_p$, respectively, as functions of k, the number of top o.s. used, for Fréchet(0.25) and Burr($0.25, -0.75$) parents, respectively, and sample size $n = 500$. Similar patterns were obtained for the other simulated models and other values of n.

3.1.1 Finite sample behaviour of the VaR-estimators at simulated optimal levels

In Table 1, we present the simulated mean values of the above mentioned normalised VaR_p-estimators, denoted H and PPWM, for the sake of simplicity, at their simulated optimal levels k^{H}_0 and k^{PPWM}_0, respectively, for $p = 1/n$. We also present the simulated mean values of the normalised NP estimator of VaR_p. In Table 2, we present, in the first row, the simulated relative efficiencies ($REFF$) of $\widehat{Q}^{PPWM}_{k,n}(p)$, $p = 1/n$, comparatively with the Weissman-Hill estimator, whenever computed at their simulated optimal levels, i.e., the simulated values of

$$REFF_{PPWM|H} := RMSE\big(\widehat{Q}^{H}_{n,k^H_0}(p)\big)/RMSE\big(\widehat{Q}^{PPWM}_{n,k^{PPWM}_0}(p)\big).$$

We also present, in the second row of each entry, the equivalent REFF indicator of the NP estimator of VaR_p. In the third row, we present the simulated value of $RMSE\big(\widehat{Q}^{H}_{n,k^H_0}(p)\big)$, also for $p = 1/n$, so that we can easily recover the RMSE of the other estimators. In all tables the “best” values are written in **bold**.

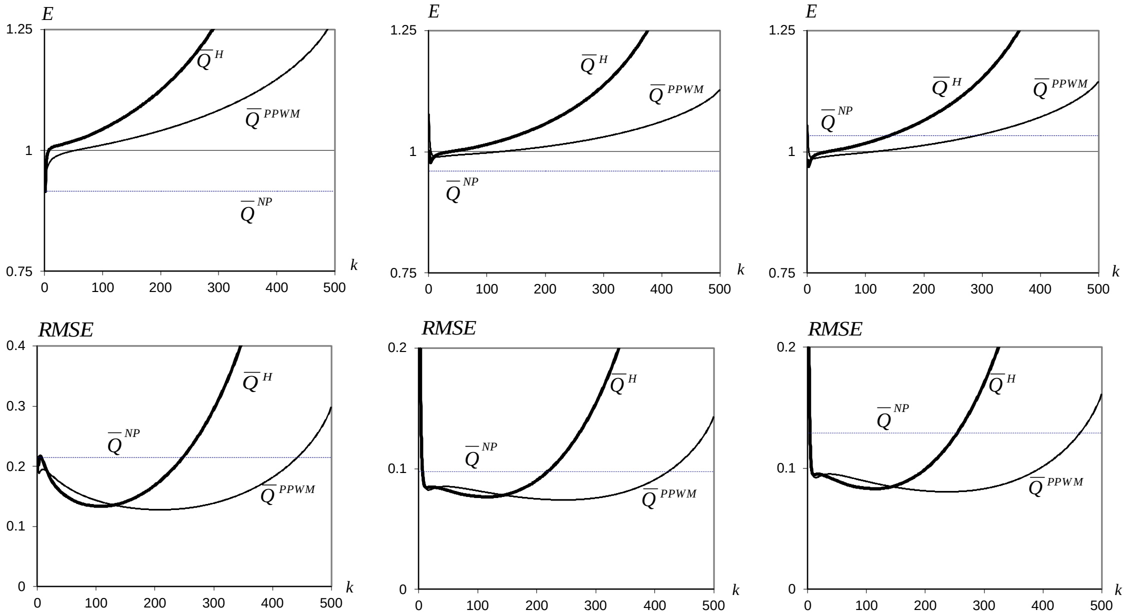

Figure 1: Simulated mean values (*above*) and root mean squared errors (*below*) for the Fréchet(γ) model, $\gamma = 0.25$, $n = 500$ and $p = 1/n$ (*left*), $p = n^{-1/(1-2\rho)}/10$ (*center*) and $p = 0.01$ (*right*).

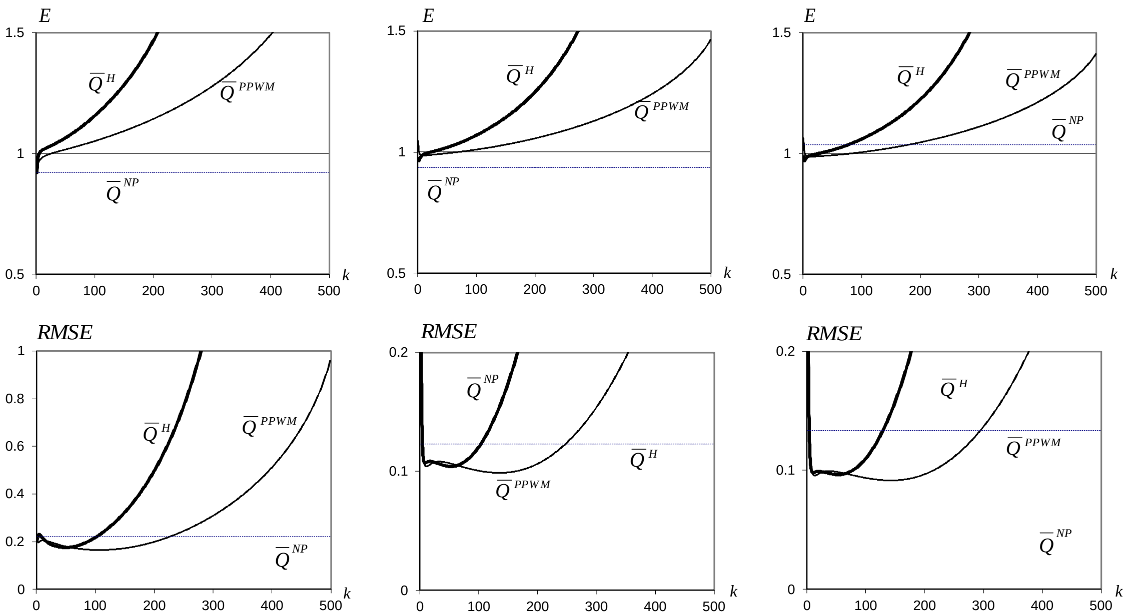

Figure 2: Simulated mean values (*above*) and root mean squared errors (*below*) for the Burr(γ, ρ) model, with $(\gamma, \rho) = (0.25, -0.75)$, $n = 500$ and $p = 1/n$ (*left*), $p = n^{-1/(1-2\rho)}/10$ (*center*) and $p = 0.01$ (*right*).

Table 1: Simulated mean values of the normalised NP and semi-parametric VaR_p-estimators under consideration, at their simulated optimal levels, for $p = 1/n$.

	n	50	100	200	500	1000	2000	5000	10000	20000
	H	**0.921**	0.931	0.938	0.945	0.948	0.950	**0.980**	1.083	1.087
Burr(0.25, −0.2)	PPWM	0.824	**0.985**	**1.036**	**1.032**	**1.040**	**1.050**	1.062	**1.068**	**1.072**
	NP	0.864	0.874	0.882	0.892	0.896	0.899	0.901	0.905	0.907
	H	**0.932**	**0.928**	0.932	**1.017**	1.175	1.164	1.148	1.137	1.125
Burr(0.5, −0.5)	PPWM	0.770	0.857	**1.034**	1.079	**1.049**	**1.005**	**1.031**	**1.019**	**1.104**
	NP	0.873	0.878	0.880	0.886	0.885	0.886	0.883	0.884	0.885
	H	0.959	**0.962**	1.058	1.066	1.062	1.057	1.050	1.046	1.042
Burr(0.25, −0.75)	PPWM	**1.013**	1.041	**1.053**	**1.057**	**1.057**	**1.055**	**1.049**	**1.045**	**1.040**
	NP	0.914	0.917	0.918	0.920	0.920	0.919	0.918	0.919	0.919
	H	**1.001**	1.044	1.050	1.050	1.045	1.042	1.036	**1.030**	**1.027**
Fréchet(0.25)	PPWM	1.031	**1.038**	**1.042**	**1.044**	**1.043**	**1.041**	**1.035**	1.031	1.027
	NP	0.916	0.913	0.912	0.915	0.917	0.919	0.920	0.919	0.919
	H	1.035	1.051	1.052	1.045	1.040	**1.034**	**1.028**	**1.023**	**1.020**
Burr(0.25, −1.5)	PPWM	**1.033**	**1.039**	**1.044**	**1.044**	**1.040**	1.036	1.029	1.024	1.021
	NP	0.919	0.918	0.919	0.921	0.920	0.920	0.918	0.919	0.919
	H	**0.990**	1.247	1.220	1.168	1.144	1.117	**1.090**	1.075	**1.063**
Burr(0.75, −1.5)	PPWM	0.724	0.769	0.823	0.900	**0.962**	**1.024**	1.109	**1.069**	0.896
	NP	0.900	**0.897**	**0.898**	**0.913**	0.911	0.912	0.903	0.904	0.904

Table 2: Simulated values of the indicators $REFF_{PPWM|H}$ (*first row*), $REFF_{NP|H}$ (*second row*) and $RMSE_H := RMSE\big(\widehat{Q}^H_{n,k_0^H}(p)\big)$ (*third row*) for the VaR_p estimation, $p = 1/n$.

	n	50	100	200	500	1000	2000	5000	10000	20000
	PPWM|H	**1.381**	**1.999**	**1.299**	**1.207**	**1.181**	**1.181**	**1.219**	**1.180**	**1.158**
Burr(0.25, −0.2)	NP|H	0.931	0.933	0.947	0.947	0.951	0.949	0.960	0.905	0.859
	RMSE_H	0.372	0.340	0.316	0.292	0.279	0.268	0.254	0.235	0.217
	PPWM|H	**1.659**	**2.181**	**2.535**	**1.516**	**1.310**	**1.144**	0.977	0.873	0.800
Burr(0.5, −0.5)	NP|H	0.965	0.963	0.992	0.972	0.903	0.790	0.688	0.610	0.544
	RMSE_H	0.537	0.516	0.505	0.485	0.444	0.391	**0.329**	**0.290**	**0.255**
	PPWM|H	**1.107**	**1.121**	**1.102**	**1.067**	**1.046**	**1.032**	**1.027**	**1.024**	**1.021**
Burr(0.25, −0.75)	NP|H	0.952	0.953	0.914	0.790	0.701	0.614	0.514	0.447	0.388
	RMSE_H	0.220	0.217	0.204	0.175	0.155	0.136	0.113	0.098	0.085
	PPWM|H	**1.118**	**1.091**	**1.064**	**1.041**	**1.019**	**1.005**	0.998	0.996	0.995
Fréchet(0.25)	NP|H	0.944	0.857	0.762	0.623	0.526	0.448	0.363	0.306	0.257
	RMSE_H	0.204	0.182	0.160	0.133	0.115	0.098	**0.079**	**0.067**	**0.057**
	PPWM|H	**1.103**	**1.071**	**1.029**	**1.001**	0.992	0.980	0.976	0.974	0.971
Burr(0.25, −1.5)	NP|H	0.928	0.828	0.726	0.585	0.496	0.412	0.324	0.268	0.221
	RMSE_H	0.200	0.178	0.156	0.128	**0.109**	**0.091**	**0.071**	**0.058**	**0.048**
	PPWM|H	**2.209**	**2.176**	**1.945**	**1.586**	**1.289**	0.995	0.679	0.491	0.393
Burr(0.75, −1.5)	NP|H	0.962	0.861	0.740	0.522	0.421	0.320	0.261	0.211	0.178
	RMSE_H	0.810	0.729	0.603	0.466	0.383	**0.309**	**0.234**	**0.189**	**0.154**

Tables 3 and 4 are similar to Tables 1 and 2, respectively, but for $p = n^{-1/(1-2\rho)}/10$. Finally, Tables 5 and Table 6 are similar to Tables 1 and 2, respectively, but for a fixed value $p = 0.01$.

Table 3: Simulated mean values of the normalised NP and semi-parametric VaR_p-estimators under consideration, at their simulated optimal levels, for $p = n^{-1/(1-2\rho)}/10$.

	n	50	100	200	500	1000	2000	5000	10000	20000
	H	**0.995**	0.944	0.919	0.902	0.899	0.899	0.948	1.064	1.061
Burr(0.25,-0.2)	PPWM	1.176	**1.039**	**1.031**	**1.041**	**1.051**	**1.056**	**1.057**	**1.055**	**1.053**
	NP	0.523	0.594	0.660	0.744	0.804	0.862	0.793	0.848	0.903
	H	**0.887**	**0.928**	0.941	**0.975**	**0.982**	1.051	1.043	1.038	1.030
Burr(0.5,-0.5)	PPWM	0.790	0.857	**0.946**	1.075	1.059	**1.022**	**0.965**	0.974	**0.981**
	NP	0.715	0.878	0.773	0.818	0.856	0.920	0.924	**0.987**	0.962
	H	0.958	0.971	**0.993**	**1.021**	**1.018**	**1.015**	**1.012**	**1.010**	**1.009**
Burr(0.25,-0.75)	PPWM	**1.020**	**1.028**	1.027	1.023	1.020	1.017	1.013	1.011	1.009
	NP	0.799	0.897	0.915	0.936	0.961	0.980	0.988	0.988	0.992
	H	0.955	**1.022**	1.024	1.020	1.016	**1.013**	**1.010**	**1.008**	**1.007**
Fréchet(0.25)	PPWM	**1.024**	1.023	**1.021**	**1.018**	**1.016**	1.014	1.011	1.009	1.007
	NP	0.866	0.891	0.938	0.960	0.990	0.992	0.991	0.995	0.998
	H	0.977	1.020	1.017	1.014	1.010	1.008	1.006	1.005	1.004
Burr(0.25, −1.5)	PPWM	**1.012**	**1.018**	**1.014**	**1.013**	1.011	1.009	1.006	1.005	1.004
	NP	0.942	0.922	0.956	0.982	**0.992**	**0.996**	**0.996**	**0.999**	**0.998**
	H	**0.984**	**0.995**	**1.039**	1.037	1.030	1.023	1.017	1.013	1.010
Burr(0.75, −1.5)	PPWM	0.777	0.844	0.891	0.936	0.957	0.972	0.984	0.989	0.993
	NP	0.909	0.818	0.899	**0.964**	**0.987**	**0.995**	**0.992**	**0.999**	**0.996**

Table 4: Simulated values of the indicators $REFF_{PPWM|H}$ (*first row*), $REFF_{NP|H}$ (*second row*) and $RMSE_0 := RMSE\big(\widehat{Q}^H_{n,k_0^H}(p)\big)$ (*third row*) for the VaR_p estimation, $p = n^{-1/(1-2\rho)}/10$.

	n	50	100	200	500	1000	2000	5000	10000	20000
	PPWM|H	**2.547**	**1.627**	**1.354**	**1.171**	**1.138**	**1.186**	**1.214**	**1.173**	**1.150**
Burr(0.25 − 0.2)	NP|H	1.657	1.297	1.135	1.002	0.976	0.984	0.913	0.894	0.873
	RMSE_H	0.875	0.606	0.468	0.352	0.307	0.283	0.237	0.206	0.179
	PPWM|H	**1.963**	**2.181**	**2.283**	**1.922**	**1.276**	**1.125**	**1.023**	0.990	0.961
Burr(0.5, −0.5)	NP|H	1.149	0.963	1.051	0.989	0.956	0.911	0.858	0.814	0.815
	RMSE_H	0.608	0.516	0.405	0.314	0.261	0.213	0.161	**0.132**	**0.108**
	PPWM|H	**1.099**	**1.089**	**1.079**	**1.052**	**1.033**	**1.023**	**1.017**	**1.011**	**1.006**
Burr(0.25, −0.75)	NP|H	0.868	0.925	0.897	0.846	0.834	0.818	0.800	0.780	0.776
	RMSE_H	0.214	0.174	0.141	0.104	0.083	0.066	0.049	0.039	0.032
	PPWM|H	**1.093**	**1.080**	**1.053**	**1.028**	**1.010**	0.997	0.989	0.987	0.985
Fréchet(0.25)	NP|H	0.901	0.839	0.827	0.784	0.771	0.760	0.748	0.740	0.736
	RMSE_H	0.177	0.137	0.106	0.076	0.060	**0.047**	**0.034**	**0.027**	**0.021**
	PPWM|H	**1.071**	**1.054**	**1.019**	0.994	0.987	0.978	0.973	0.970	0.966
Burr(0.25, −1.5)	NP|H	0.910	0.848	0.832	0.821	0.810	0.802	0.795	0.785	0.781
	RMSE_H	0.149	0.114	0.087	**0.062**	**0.047**	**0.036**	**0.025**	**0.020**	**0.015**
	PPWM|H	**1.522**	**1.323**	**1.207**	**1.083**	**1.028**	0.988	0.957	0.940	0.928
Burr(0.75, −1.5)	NP|H	0.939	1.021	0.944	0.870	0.839	0.818	0.813	0.794	0.791
	RMSE_H	0.512	0.379	0.288	0.196	0.148	**0.112**	**0.078**	**0.059**	**0.046**

Table 5: Simulated mean values of the normalised semi-parametric and NP VaR_p-estimators under consideration, at their simulated optimal levels, for $p = 0.01$.

		n								
	n	50	100	200	500	1000	2000	5000	10000	20000
	H	0.893	0.931	0.955	0.972	0.976	0.989	0.996	0.998	0.999
Burr(0.25, −0.2)	PPWM	**0.975**	**0.985**	**0.990**	**0.993**	**0.994**	**0.995**	**0.998**	**0.999**	**0.999**
	NP	0.638	1.400	1.162	1.059	1.029	1.015	1.006	1.003	1.001
	H	**0.955**	**0.928**	**0.964**	**0.986**	**0.988**	**0.991**	**0.996**	**0.999**	**1.000**
Burr(0.5, −0.5)	PPWM	0.841	0.857	0.879	0.902	0.914	0.973	0.984	0.989	0.992
	NP	0.589	1.870	1.285	1.095	1.046	1.023	1.009	1.004	1.002
	H	0.932	**0.962**	**1.004**	**1.017**	**1.010**	**1.006**	**1.003**	1.002	**1.001**
Burr(0.25, −0.75)	PPWM	**0.944**	1.041	1.032	1.020	1.013	1.009	1.004	1.002	1.001
	NP	0.763	1.234	1.096	1.035	1.017	1.009	1.003	**1.002**	1.001
	H	1.071	1.044	1.035	1.023	1.016	1.011	1.007	1.004	1.003
Fréchet(0.25)	PPWM	**1.050**	**1.038**	**1.030**	**1.021**	1.016	1.012	1.007	1.005	1.003
	NP	0.769	1.193	1.086	1.033	**1.016**	**1.008**	**1.003**	**1.002**	**1.001**
	H	1.070	1.051	1.036	1.023	**1.016**	1.011	1.007	1.005	1.003
Burr(0.25, −1.5)	PPWM	**1.052**	**1.039**	**1.032**	**1.022**	1.016	1.012	1.007	1.005	1.003
	NP	0.773	1.212	1.093	1.034	1.017	**1.009**	**1.003**	**1.002**	**1.001**
	H	1.459	**1.247**	**1.142**	**1.072**	**1.050**	**1.032**	1.020	1.013	1.009
Burr(0.75, −1.5)	PPWM	**0.736**	0.769	0.799	0.842	0.934	0.962	0.982	0.989	0.994
	NP	0.535	2.938	1.510	1.157	1.074	1.036	**1.014**	**1.007**	**1.003**

Table 6: Simulated values of the indicators $REFF_{PPWM|H}$ (*first row*), $REFF_{NP|H}$ (*second row*) and $RMSE_0 := RMSE\big(\widehat{Q}^H_{n,k_0^H}(p)\big)$ (*third row*) for the VaR_p estimation, $p = 0.01$.

		n								
		50	100	200	500	1000	2000	5000	10000	20000
	PPWM|H	**1.144**	**1.091**	**1.059**	**1.030**	**1.010**	0.996	0.986	0.984	0.980
Burr(0.25, −0.2)	NP|H	0.939	0.360	0.529	0.644	0.707	0.749	0.785	0.803	0.818
	RMSE_H	0.487	0.340	0.243	0.154	0.110	**0.078**	**0.049**	**0.035**	**0.025**
	PPWM|H	**3.129**	**2.181**	**1.768**	**1.396**	**1.167**	**1.013**	0.987	0.969	0.954
Burr(0.5, −0.5)	NP|H	1.666	0.168	0.445	0.669	0.751	0.797	0.825	0.831	0.836
	RMSE_H	0.917	0.516	0.345	0.213	0.150	0.105	**0.066**	**0.047**	**0.033**
	PPWM|H	**1.173**	**1.121**	**1.087**	**1.050**	**1.029**	**1.018**	**1.011**	**1.006**	**1.002**
Burr(0.25, −0.75)	NP|H	1.038	0.378	0.579	0.722	0.774	0.807	0.829	0.832	0.837
	RMSE_H	0.309	0.217	0.154	0.096	0.068	0.048	0.030	0.022	0.015
	PPWM|H	**1.144**	**1.091**	**1.059**	**1.030**	**1.010**	0.996	0.986	0.984	0.980
Fréchet(0.25)	NP|H	0.939	0.360	0.529	0.644	0.707	0.749	0.785	0.803	0.818
	RMSE_H	0.268	0.182	0.129	0.083	0.060	**0.043**	**0.028**	**0.020**	**0.015**
	PPWM|H	**1.117**	**1.071**	**1.026**	0.998	0.989	0.979	0.973	0.970	0.966
Burr(0.25, −1.5)	NP|H	0.918	0.351	0.496	0.628	0.687	0.727	0.762	0.778	0.794
	RMSE_H	0.259	0.178	0.126	**0.081**	**0.059**	**0.042**	**0.027**	**0.020**	**0.014**
	PPWM|H	**3.384**	**2.176**	**1.566**	**1.104**	**1.007**	0.961	0.943	0.940	0.944
Burr(0.75, −1.5)	NP|H	1.948	0.065	0.313	0.553	0.655	0.715	0.764	0.781	0.798
	RMSE_H	1.326	0.729	0.455	0.269	0.188	**0.132**	**0.084**	**0.059**	**0.042**

Some remarks:

1. Regarding mean values at optimal levels, and whenever $\xi = \infty$, the PPWM VaR_p estimator performs better than the classical Weissman-Hill estimator, in a large variety of situations, beating always the NP-estimator, unless we work with a Burr$(0.75, -1.5)$ model, for which the PPWM estimators are not valid. If ξ is finite the outperformance of the PPWM VaR_p-estimator becomes clear when $\gamma + \rho > 0$, and for small values of n whenever $\gamma + \rho \leq 0$. For large n and $\gamma + \rho < 0$, the NP-estimator appears then as an interesting alternative.
2. Regarding RMSE, or equivalently REFF-indicators, the PPWM VaR_p estimators performs better than the Weissman-Hill quantile estimator at optimal levels, for all n, provided that $\gamma + \rho > 0$ and $\xi = \infty$ or ξ finite and non-null. But in this region of the (γ, ρ)-plane, even for finite ξ (possibly null), the H (the best for large n) and the PPWM (the best for small n) estimators are never beaten by the NP estimator.
3. The new PPWM quantile estimator can be used as an alternative to the Weissman-Hill estimator, specially for small to moderate sample sizes, but also for large samples, in most situations. All the simulated results suggest that its RMSE is never much bigger than the one of Weissman-Hill's estimator RMSE and it has often a smaller bias than the Weissman-Hill estimator. At their optimal levels and for small sample sizes, the PPWM VaR_p estimator is usually much more efficient than Weissman-Hill's estimator.
4. The asymptotic properties of the new PPWM quantile estimators do not hold for the Burr model with $(\gamma, \rho) = (0.75, -1.5)$. But even in this example, the PPWM estimators are more efficient than the Weissman-Hill estimators at their simulated optimal levels, for small up to moderate sample sizes.

3.2 Simulated mean values and root mean squared errors of the p_x-estimators

Figures 3 and 4 are equivalent to Figures 1 and 2, respectively, but for the estimation of $p = p_x = 1 - F(x) = 1/n$, i.e. $x = \big(-\ln(1 - 1/n)\big)^{-\gamma}$ for the Fréchet(γ) parent, and $x = \big(n^{-\rho} - 1\big)^{-\gamma/\rho}$ for the Burr(γ, ρ) parent. In Figure 3 and Figure 4, we have considered a sample size $n = 500$, again from Fréchet(0.25) and Burr$(0.5, -0.5)$ parents, respectively.

3.2.1 Finite sample behaviour of the p_x-estimators at simulated optimal levels

Tables 7 and 8 are similar to Tables 1 and 2, respectively, for the estimation of p_x, with $x = F^{\leftarrow}(1-1/n)$. Tables 9 and 10 are equivalent to Tables 7 and 8, respectively, now for the estimation of p_x, with $x = F^{\leftarrow}(1 - n^{-1/(1-2\rho)}/10)$. Also, Tables 11 and 12 are equivalent to Tables 7 and 8, respectively, for the estimation of p_x, with $x = F^{\leftarrow}(1 - 0.01)$. We have decided to work with the normalised semi-parametric estimators $\bar{p}^H = \hat{p}^H_{n,k_0^H}(x)/p_x$ and $\bar{p}^{PPWM} = \hat{p}^{PPWM}_{n,k_0^{PPWM}}(x)/p_x$, as well as with the normalised NP estimator of p_x. For the sake of simplicity, these estimators are again denoted respectively H, PPWM and NP.

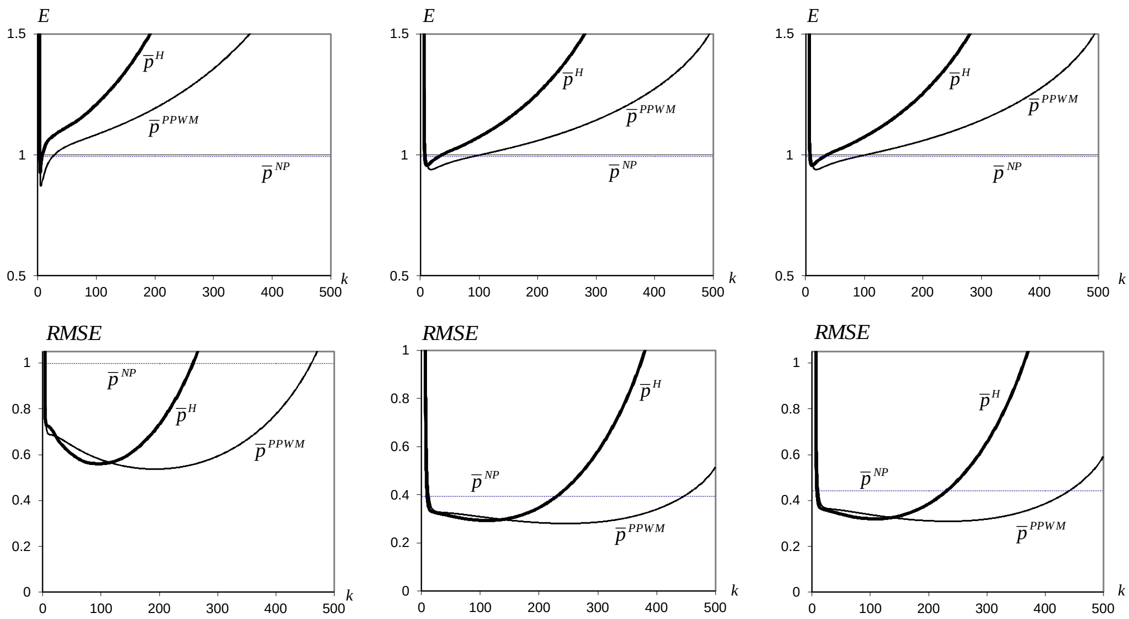

Figure 3: Simulated mean values (above) and root mean squared errors (below) for the Fréchet(γ) model, with $\gamma = 0.25$, $n = 500$ and $p = 1/n$ (left), $p = n^{-1/(1-2\rho)}/10$ (center) and $p = 0.01$ (right).

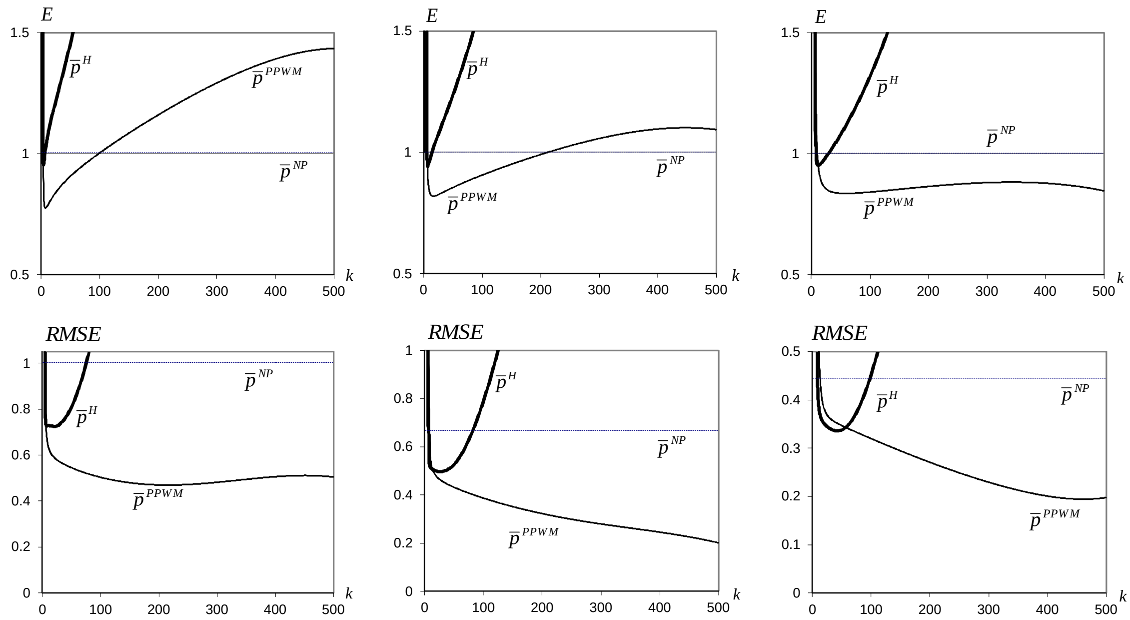

Figure 4: Simulated mean values (above) and root mean squared errors (below) for the Burr(γ, ρ) model, with $(\gamma, \rho) = (0.5, -0.5)$, $n = 500$ and $p = 1/n$ (left), $p = n^{-1/(1-2\rho)}/10$ (center) and $p = 0.01$ (right).

Table 7: Simulated mean values of the normalised NP and semi-parametric p_x-estimators under consideration, at their simulated optimal levels, for $x = F^{\leftarrow}(1 - 1/n)$.

		n								
	n	50	100	200	500	1000	2000	5000	10000	20000
Burr(0.25, −0.2)	H	**1.046**	1.044	**1.032**	**1.027**	**1.025**	**1.023**	**1.031**	1.189	1.261
	PPWM	0.737	**0.984**	1.156	1.122	1.123	1.134	1.159	**1.183**	**1.212**
	NP	0.981	0.996	1.001	1.005	1.004	0.997	0.994	0.996	1.003
Burr(0.5, −0.5)	H	**1.039**	**1.040**	**1.027**	**1.154**	1.230	1.256	1.252	1.252	1.237
	PPWM	0.671	0.823	1.039	1.182	**1.139**	**1.127**	**1.138**	**1.152**	**1.176**
	NP	0.981	0.996	1.001	1.005	1.004	0.997	0.994	0.996	1.003
Burr(0.25, −0.75)	H	**1.037**	**1.053**	1.176	1.230	1.237	1.236	1.211	1.197	1.180
	PPWM	1.154	1.146	**1.170**	**1.209**	**1.229**	**1.228**	**1.211**	**1.194**	**1.175**
	NP	0.981	0.996	1.001	1.005	1.004	0.997	0.994	0.996	1.003
Fréchet(0.25)	H	**1.069**	1.171	1.192	1.196	1.185	1.177	**1.150**	**1.130**	1.117
	PPWM	1.125	**1.136**	**1.160**	**1.184**	**1.185**	**1.176**	1.154	1.132	**1.114**
	NP	0.968	0.969	0.970	0.992	0.990	1.002	0.999	0.998	0.999
Burr(0.25, −1.5)	H	**1.132**	1.176	1.197	**1.185**	**1.169**	**1.145**	**1.116**	**1.098**	**1.083**
	PPWM	1.145	**1.153**	**1.176**	1.187	1.176	1.156	1.126	1.104	1.087
	NP	0.977	0.989	0.997	1.004	1.004	0.997	0.994	0.996	1.003
Burr(0.75, −1.5)	H	**1.132**	**1.176**	**1.197**	1.185	1.169	1.145	**1.116**	**1.098**	**1.083**
	PPWM	0.648	0.704	0.772	**0.867**	**0.944**	**1.022**	1.129	1.187	1.091
	NP	0.977	0.989	0.997	1.004	1.004	0.997	0.994	0.996	1.003

Table 8: Simulated values of the indicators $REFF_{PPWM|H}$ (*first row*), $REFF_{NP|H}$ (*second row*) and $RMSE_H := RMSE\big(\hat{p}^H_{n,k_0^H}(x)\big)$ (*third row*) for the p_x-estimation, with $x = F^{\leftarrow}(1 - 1/n)$.

		n								
	n	50	100	200	500	1000	2000	5000	10000	20000
Burr(0.25, −0.2)	PPWM|H	**1.809**	**3.054**	**1.584**	**1.325**	**1.252**	**1.206**	**1.181**	**1.162**	**1.148**
	NP|H	0.659	0.675	0.690	0.696	0.705	0.711	0.716	0.709	0.698
	$RMSE_H$	0.635	0.665	0.685	0.697	0.710	0.711	0.710	0.703	0.696
Burr(0.5, −0.5)	PPWM|H	**1.638**	**2.391**	**2.935**	**1.538**	**1.323**	**1.183**	**1.038**	0.938	0.861
	NP|H	0.698	0.710	0.723	0.721	0.702	0.668	0.605	0.553	0.498
	$RMSE_H$	0.673	0.700	0.718	0.722	0.707	0.668	0.603	**0.549**	**0.497**
Burr(0.25, −0.75)	PPWM|H	**1.187**	**1.128**	**1.097**	**1.065**	**1.042**	**1.024**	**1.017**	**1.015**	**1.016**
	NP|H	0.719	0.726	0.727	0.683	0.630	0.570	0.485	0.423	0.367
	$RMSE_H$	0.693	0.715	0.722	0.683	0.634	0.570	0.482	0.420	0.366
Fréchet(0.25)	PPWM|H	**1.122**	**1.098**	**1.070**	**1.038**	**1.011**	0.998	0.990	0.989	0.991
	NP|H	0.730	0.700	0.655	0.560	0.494	0.423	0.343	0.288	0.241
	$RMSE_H$	0.710	0.688	0.641	0.558	0.488	**0.423**	**0.341**	**0.287**	**0.241**
Burr(0.25, −1.5)	PPWM|H	**1.105**	**1.076**	**1.034**	0.994	0.983	0.968	0.966	0.966	0.966
	NP|H	0.729	0.696	0.641	0.541	0.464	0.390	0.305	0.250	0.203
	$RMSE_H$	0.703	0.684	0.635	0.542	**0.467**	**0.390**	**0.303**	**0.248**	**0.203**
Burr(0.75, −1.5)	PPWM|H	**1.505**	**1.624**	**1.657**	**1.539**	**1.341**	**1.085**	0.760	0.550	0.420
	NP|H	0.729	0.696	0.641	0.541	0.464	0.390	0.305	0.250	0.203
	$RMSE_H$	0.703	0.684	0.635	0.542	0.467	0.390	**0.303**	**0.248**	**0.203**

Table 9: Simulated mean values of the normalised NP and semi-parametric p_x-estimators under consideration, at their simulated optimal levels, for $x = F^{\leftarrow}(1 - n^{-1/(1-2\rho)}/10)$.

		n								
		50	100	200	500	1000	2000	5000	10000	20000
Burr(0.25, −0.2)	H	1.197	1.138	1.078	1.037	1.028	1.024	1.053	1.137	1.168
	PPWM	1.236	1.237	1.168	1.134	1.133	1.140	1.148	1.150	1.154
	NP	**0.990**	**0.995**	**1.004**	**1.002**	**1.002**	**0.996**	**0.995**	**0.996**	**0.998**
Burr(0.5, −0.5)	H	1.044	1.040	1.058	1.084	1.100	1.086	1.078	1.070	1.059
	PPWM	0.727	0.823	0.930	1.093	1.234	1.089	1.065	1.056	1.050
	NP	**0.981**	**0.996**	**1.003**	**1.001**	**1.003**	**1.002**	**1.002**	**1.001**	**1.001**
Burr(0.25, −0.75)	H	1.050	1.077	1.079	1.078	1.073	1.058	1.046	1.039	1.036
	PPWM	1.156	1.123	1.108	1.091	1.078	1.066	1.053	1.044	1.037
	NP	**0.980**	**0.990**	**0.999**	**1.003**	**1.002**	**1.001**	**1.001**	**1.000**	**1.001**
Fréchet(0.25)	H	1.088	1.083	1.085	1.077	1.062	1.052	1.041	1.032	1.026
	PPWM	1.123	1.098	1.085	1.073	1.063	1.054	1.042	1.034	1.028
	NP	**0.973**	**0.983**	**0.988**	**0.994**	**0.995**	**0.999**	**1.001**	**1.001**	**1.000**
Burr(0.25, −1.5)	H	1.083	1.079	1.069	1.060	1.047	1.032	1.024	1.019	1.014
	PPWM	1.136	1.097	1.077	1.058	1.047	1.037	1.026	1.020	1.016
	NP	**0.977**	**0.989**	**0.994**	**1.000**	**1.000**	**1.001**	**1.001**	**1.000**	**1.000**
Burr(0.75, −1.5)	H	1.083	1.079	1.069	1.060	1.047	1.032	1.024	1.019	1.014
	PPWM	0.641	0.691	0.741	0.874	0.919	0.949	0.972	0.982	0.988
	NP	**0.977**	**0.989**	**0.994**	**1.000**	**1.000**	**1.001**	**1.001**	**1.000**	**1.000**

Table 10: Simulated values of the indicators $REFF_{PPWM|H}$ (*first row*), $REFF_{NP|H}$ (*second row*) and $RMSE_H := RMSE\big(\hat{p}^H_{n,k_0^H}(x)\big)$ (*third row*) for the p_x-estimation, with $x = F^{\leftarrow}(1 - n^{-1/(1-2\rho)}/10)$.

		n								
	n	50	100	200	500	1000	2000	5000	10000	20000
Burr(0.25, −0.2)	PPWM|H	**2.541**	**1.585**	**1.387**	**1.278**	**1.233**	**1.202**	**1.181**	**1.158**	**1.140**
	NP|H	0.656	0.665	0.667	0.676	0.695	0.707	0.720	0.719	0.719
	RMSE$_H$	1.164	1.078	0.983	0.883	0.821	0.754	0.670	0.605	0.549
Burr(0.5, −0.5)	PPWM|H	**2.085**	**2.391**	**2.795**	**2.471**	**1.491**	**1.223**	**1.057**	0.966	0.904
	NP|H	0.689	0.710	0.728	0.746	0.750	0.749	0.744	0.748	0.748
	RMSE$_H$	0.797	0.700	0.612	0.497	0.423	0.356	0.281	**0.236**	**0.199**
Burr(0.25, −0.75)	PPWM|H	**1.193**	**1.146**	**1.106**	**1.069**	**1.047**	**1.035**	**1.024**	**1.017**	**1.011**
	NP|H	0.721	0.733	0.748	0.757	0.756	0.765	0.763	0.759	0.759
	RMSE$_H$	0.678	0.573	0.479	0.370	0.303	0.247	0.187	0.151	0.123
Fréchet(0.25)	PPWM|H	**1.129**	**1.097**	**1.069**	**1.037**	**1.016**	**1.002**	0.992	0.990	0.988
	NP|H	0.738	0.741	0.747	0.745	0.741	0.741	0.735	0.728	0.726
	RMSE$_H$	0.617	0.495	0.395	0.292	0.232	0.184	**0.136**	**0.107**	**0.085**
Burr(0.25, −1.5)	PPWM|H	**1.136**	**1.089**	**1.047**	**1.011**	0.999	0.989	0.979	0.975	0.970
	NP|H	0.742	0.753	0.762	0.775	0.777	0.775	0.779	0.775	0.775
	RMSE$_H$	0.521	0.412	0.324	0.234	**0.183**	**0.141**	**0.100**	**0.077**	**0.060**
Burr(0.75, −1.5)	PPWM|H	**1.170**	**1.054**	0.944	0.853	0.838	0.834	0.838	0.843	0.850
	NP|H	0.742	0.753	0.762	0.775	0.777	0.775	0.779	0.775	0.775
	RMSE$_H$	0.521	0.412	**0.324**	**0.234**	**0.183**	**0.141**	**0.100**	**0.077**	**0.060**

Table 11: Simulated mean values of the normalised NP and semi-parametric p_x-estimators under consideration, at their simulated optimal levels, for $x = F^{\leftarrow}(1 - 0.01)$.

		n								
	n	50	100	200	500	1000	2000	5000	10000	20000
	H	1.093	1.044	1.037	1.031	1.019	1.010	1.005	1.003	1.001
Burr(0.25, −0.2)	PPWM	0.975	0.984	0.989	0.992	0.992	0.993	0.996	0.998	0.999
	NP	**0.986**	**0.996**	**1.001**	**1.002**	**1.002**	**1.001**	**1.001**	**1.000**	**1.000**
	H	1.067	1.040	1.065	1.051	1.038	1.024	1.011	1.007	1.003
Burr(0.5, −0.5)	PPWM	0.803	0.823	0.836	0.863	0.878	0.889	0.896	0.976	0.984
	NP	**0.986**	**0.996**	**1.001**	**1.002**	**1.002**	**1.001**	**1.001**	**1.000**	**1.000**
	H	1.052	1.053	1.092	1.062	1.048	1.035	1.017	1.010	1.006
Burr(0.25, −0.75)	PPWM	1.132	1.146	1.120	1.081	1.057	1.038	1.020	1.012	1.007
	NP	**0.986**	**0.996**	**1.001**	**1.002**	**1.002**	**1.001**	**1.001**	**1.000**	**1.000**
	H	**0.990**	1.171	1.127	1.089	1.062	1.044	1.029	1.018	1.013
Fréchet(0.25)	PPWM	1.152	1.136	1.112	1.084	1.063	1.047	1.030	1.020	1.013
	NP	0.948	**0.969**	**0.982**	**0.994**	**0.995**	**0.999**	**1.000**	**1.000**	**1.000**
	H	1.108	1.176	1.132	1.089	1.063	1.042	1.028	1.019	1.013
Burr(0.25, −1.5)	PPWM	1.131	1.153	1.125	1.090	1.066	1.047	1.030	1.020	1.014
	NP	**0.974**	**0.989**	**0.999**	**1.002**	**1.002**	**1.001**	**1.001**	**1.000**	**1.000**
	H	1.108	1.176	1.132	1.089	1.063	1.042	1.028	1.019	1.013
Burr(0.75, −1.5)	PPWM	0.660	0.704	0.744	0.784	0.806	0.927	0.968	0.982	0.990
	NP	**0.974**	**0.989**	**0.999**	**1.002**	**1.002**	**1.001**	**1.001**	**1.000**	**1.000**

Table 12: Simulated values of the indicators $REFF_{PPWM|H}$ (*first row*), $REFF_{NP|H}$ (*second row*) and $RMSE_H := RMSE\big(\hat{p}^H_{n,k_0^H}(x)\big)$ (*third row*) for the p_x-estimation, with $x = F^{\leftarrow}(1 - 0.01)$.

		n								
	n	50	100	200	500	1000	2000	5000	10000	20000
	PPWM|H	**2.976**	**3.054**	**3.122**	**3.155**	**3.163**	**3.176**	**3.163**	**3.144**	**3.155**
Burr(0.25 − 0.2)	NP|H	0.654	0.675	0.695	0.714	0.721	0.725	0.730	0.726	0.727
	RMSE$_H$	0.904	0.665	0.489	0.317	0.227	0.161	0.102	0.072	0.051
	PPWM|H	**2.578**	**2.391**	**2.125**	**1.728**	**1.447**	**1.191**	0.881	0.835	0.832
Burr(0.5, −0.5)	NP|H	0.680	0.710	0.733	0.756	0.766	0.774	0.783	0.780	0.783
	RMSE$_H$	0.940	0.700	0.516	0.336	0.241	0.172	**0.110**	**0.078**	**0.055**
	PPWM|H	**1.137**	**1.128**	**1.101**	**1.073**	**1.056**	**1.047**	**1.040**	**1.035**	**1.032**
Burr(0.25, −0.75)	NP|H	0.692	0.726	0.743	0.764	0.778	0.789	0.802	0.802	0.808
	RMSE$_H$	0.957	0.715	0.523	0.339	0.245	0.176	0.112	0.080	0.057
	PPWM|H	**1.108**	**1.098**	**1.068**	**1.037**	**1.016**	**1.002**	0.993	0.991	0.987
Fréchet(0.25)	NP|H	0.693	0.700	0.706	0.725	0.741	0.757	0.777	0.794	0.805
	RMSE$_H$	0.950	0.688	0.492	0.320	0.232	0.169	**0.110**	**0.079**	**0.057**
	PPWM|H	**1.108**	**1.076**	**1.036**	**1.004**	0.995	0.985	0.979	0.975	970
Burr(0.25, −1.5)	NP|H	0.695	0.696	0.698	0.717	0.732	0.745	0.765	0.775	789
	RMSE$_H$	0.957	0.684	0.490	0.318	**0.231**	**0.166**	**0.107**	**0.077**	**0.055**
	PPWM|H	**1.934**	**1.624**	**1.353**	**1.055**	0.866	0.802	0.821	0.843	868
Burr(0.75, −1.5)	NP|H	0.695	0.696	0.698	0.717	0.732	0.745	0.765	0.775	789
	RMSE$_H$	0.957	0.684	0.490	0.318	**0.231**	**0.166**	**0.107**	**0.077**	**0.055**

Despite of the fact that there is not a clear agreement between the behaviour of the PPWM p_x and VaR_p-estimators, the final conclusions related with RMSEs do not differ too much from the ones we have previously drawn for the PPWM VaR_p-estimator. In what concerns the pattern of mean values, and for ξ finite, the NP-estimator outperforms the semi-parametric ones for all simulated models and for almost all values of n. Just as before, the new quantile estimator $\hat{p}_x^{PPWM}$ can be used as an alternative to $\hat{p}_x^{H}$, specially for small to moderate sample sizes, but also for large samples, in most situations. All the simulated results suggest that the RMSE of $\hat{p}_x^{PPWM}$ is never much bigger than the one of $\hat{p}_x^{H}$. At their optimal levels and for small sample sizes, $\hat{p}_x^{PPWM}$ is usually much more efficient than $\hat{p}_x^{H}$. The variance of the NP-estimator is terribly high. Regarding RMSE, the NP-estimator can never beat the semi-parametric ones, at optimal levels, even when ξ is finite.

4 The bootstrap methodology in the estimation of optimal thresholds

We now put forward an adaptive choice of thresholds, again on the basis of bootstrap computer-intensive methods. Indeed, due to the specificity of these PPWM estimators, and contrarily to what happens with the Hill estimators, in (1.5), the most common estimators of a positive EVI, a direct estimation of the optimal sample fraction (OSF), done on the basis of estimates of scale and shape second-order parameters, is problematic. The use of bootstrap computer intensive methods helps us to provide an adaptive choice of the optimal number of o.s.'s to be used in the estimation, and will be the topic of discussion in this section.

Again with $\hat{\gamma}^\bullet_{k,n}$ denoting either the Hill or the PPWM EVI-estimators, in (1.5) and (1.12), respectively, let us use the notation

$$k_0 \equiv k_0^\bullet(n) := \arg\min_k MSE(\hat{\gamma}^\bullet_{k,n}).$$

Under a semi-parametric framework, and under the validity of the second-order condition, in (2.3), with $\rho < 0$, let us parameterize the function $A(\cdot)$ as $A(t) = \gamma\beta t^\rho$, where $\beta \neq 0$ and $\rho < 0$ are generalized scale and shape second-order parameters. Then, with $\mathbb{E}$ denoting the mean value operator, a possible substitute for the MSE of any classical EVI-estimator $\hat{\gamma}^\bullet_{k,n}$ is, cf. equation (2.9),

$$AMSE(\hat{\gamma}^\bullet_{k,n}) := \mathbb{E}\big(\sigma_\bullet Z_k^\bullet/\sqrt{k} + b_\bullet A(n/k)\big)^2 = \sigma_\bullet^2/k + b_\bullet^2\ \gamma^2\ \beta^2\ (n/k)^{2\rho},$$

depending on n and k, and with $AMSE$ standing for *asymptotic mean squared error*. We get (Dekkers and de Haan, 1993),

$$k_{0|\bullet}(n) := \arg\min_k AMSE\big(\hat{\gamma}^\bullet_{k,n}\big) = \big((-2\rho)\ b_\bullet^2\ \gamma^2\beta^2\ n^{2\rho}/\sigma_\bullet^2\big)^{-1/(1-4\rho)} = k_0^\bullet(n)(1+o(1)).$$

For the Hill estimator, we have, in (2.9), $\sigma_{_H} = \gamma$ and $b_{_H} = 1/(1-\rho)$. Consequently, with $(\hat{\beta}, \hat{\rho})$ any consistent estimator of the vector (β, ρ) of second-order parameters, we have an asymptotic justification for the estimator

$$\hat{k}_0^H := \big[\big((1-\hat{\rho})^2 n^{-2\hat{\rho}}/(-2\hat{\rho}\hat{\beta}^2)\big)^{1/(1-2\hat{\rho})}\big], \tag{4.1}$$

where, as usual, $[x]$ denotes the integer part of x. Moreover, provided that $\sqrt{k}\ (n/k)^\rho \to \lambda$, finite, and with $b_{k,n,\rho} = 1 + \beta(n/k)^\rho/(1-\rho)$, $\sqrt{k}\{\hat\gamma^H_{k,n}/\gamma - b_{k,n,\rho}\}$ is approximately $\mathcal{N}(0,1)$. We may then get approximate $100(1-\alpha)\%$ confidence intervals (CI's) for γ,

$$\left(\hat\gamma^H_{k,n}/\big(b_{k,n,\rho} + \xi_{1-\alpha/2}/\sqrt{k}\big),\ \hat\gamma^H_{k,n}/\big(b_{k,n,\rho} - \xi_{1-\alpha/2}/\sqrt{k}\big)\right), \tag{4.2}$$

where ξ_p is the p-quantile of a $\mathcal{N}(0,1)$ d.f. If $\lambda = 0$, we may replace in (4.2) the bias summand $\beta(n/k)^\rho/(1-\rho)$ by 0, i.e., $b_{k,n,\rho} = 1$.

The same does not happen with the PPWM EVI-estimators, with an asymptotic variance (σ^2_{PPWM}) and a dominant component of bias (b_{PPWM}) dependent on γ (see equation (2.5)). In this situation, it is sensible to use the bootstrap methodology for the adaptive PPWM EVI-estimation. Just as in Gomes and Oliveira (2001), for the estimation of γ through the Hill estimator, and in Gomes *et al.* (2011, 2012), for adaptive reduced-bias estimation, let us more generally consider the auxiliary statistics,

$$T^\bullet_{k,n} := \hat\gamma^\bullet_{[k/2],n} - \hat\gamma^\bullet_{k,n}, \quad k = 2, \ldots, n-1, \tag{4.3}$$

which converge in probability to the known value zero, for any intermediate k, enabling thus easily the simulation of their MSE through the non-central moment of order two. On the basis of results similar to the ones in Gomes *et al.* (2000) and Gomes and Oliveira (2001), we can get, for the auxiliary statistic $T_{k,n}$, in (4.3), the asymptotic distributional representation,

$$T_{k,n} \equiv T^\bullet_{k,n} \overset{d}{=} \sigma_\bullet\ Q^\bullet_k/\sqrt{k} + b_\bullet(2^\rho - 1)\ A(n/k) + o_p(A(n/k)),$$

with Q_k asymptotically standard normal, and $(b_\bullet,\ \sigma_\bullet)$ given in (2.9). The $AMSE$ of $T_{k,n}$ is thus minimal at a level $k_{0|T}(n)$ such that $\sqrt{k}\ A(n/k) \to \lambda'_A \neq 0$, i.e. a level such that

$$k_{0|\bullet}(n) = k_{0|T}(n)\,(1-2^\rho)^{\frac{1}{1-2\rho}}\,(1+o(1)).$$

Then (see the above mentioned papers for further details), given the sample $\underline{X}_n = (X_1, \ldots, X_n)$ from an unknown model F, for any $n_1 = o(n)$, and a bootstrap sample $\underline{X}^*_{n_1} = (X^*_1, \ldots, X^*_{n_1})$, denote $T^*_{k_1,n_1}$, $1 < k_1 < n_1$, the corresponding bootstrap auxiliary statistic. With the obvious notation $k^*_{0|T}(n_1) = \arg\min_{k_1} AMSE\big(T^*_{k_1,n_1}\big)$,

$$k^*_{0|T}(n_1)/k_{0|T}(n) = (n_1/n)^{-\frac{2\rho}{1-2\rho}}\,(1+o(1)).$$

Consequently, for another sample size $n_2 = n_1^2/n$, we have $\big(k^*_{0|T}(n_1)\big)^2/k^*_{0|T}(n_2) = k_{0|T}(n)(1+o(1))$, as $n \to \infty$. On the basis of the estimation of $k^*_{0|T}$, we are now able to estimate $k_{0|T}$, and next $k^\bullet_0(n)$. With $\hat k^*_{0T}$ denoting the sample counterpart of k^*_{0T}, we can build the k_0-estimate,

$$\hat k^{\bullet *}_0 \equiv \hat k^{\bullet *}_0(n; n_1) := \min\left(n-1,\ \left[\frac{(1-2^{\hat\rho})^{\frac{1}{1-2\hat\rho}}\big(\hat k^*_{0|T}(n_1)\big)^2}{\hat k^*_{0|T}([n_1^2/n]+1)}\right]+1\right), \tag{4.4}$$

and the γ-estimate

$$\hat{\gamma}^{\bullet *} \equiv \hat{\gamma}^{\bullet *}(n; n_1) := \hat{\gamma}^{\bullet}_{\hat{k}_0^{\bullet *}(n;n_1),n}. \tag{4.5}$$

Note also that, with $c_k(p)$ defined in (1.6), on the basis of (2.2), for the estimation of VaR_p, and on the basis of (2.11) with $\tilde{c}_k$ the observed value of $\widetilde{C}_k$, in (1.8), for the estimation of $p_x = \overline{F}(x)$, we get

$$k^{\bullet}_{0|p} := \arg\min_k \; AMSE\big(\hat{Q}^{\bullet}_{k,n}(p)\big) = \arg\min_k \; (\ln c_k)^2 AMSE\big(\hat{\gamma}^{\bullet}_{k,n}\big),$$

and

$$k^{\bullet}_{0|x} := \arg\min_k \; AMSE\big(\hat{p}^{\bullet}_{k,n}(x)\big) = \arg\min_k \; (\ln \tilde{c}_k)^2 AMSE\big(\hat{\gamma}^{\bullet}_{k,n}\big).$$

A few practical questions, some of them with answers out of the scope of this paper, 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? 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 expected not to be a long way from the ones given in previous papers (see Hall, 1990; Draisma *et al.*, 1999; Danielsson *et al.* 2001; Gomes and Oliveira, 2001; Gomes *et al.*, 2009, 2011, 2012), and some of them will be given in Section 5 of this article, on the basis of the analysis of real and simulated data.

4.1 An algorithm for adaptive estimation of parameters

The estimates $\hat{\rho}$, of the second-order parameter ρ, are the ones already used in previous papers. See for instance the algorithm provided in Gomes and Pestana (2007). Now, with $\bullet$ denoting either H or $PPWM$, the algorithm is the following:

1. Given an observed sample $(x_1, x_2, \ldots, x_n)$, compute the estimates $\hat{\gamma}^{\bullet}_{k,n}$, $\widehat{Q}^{\bullet}_{k,n}(p)$ and $\hat{p}^{\bullet}_{k,n}(x)$, $k = 1, 2, \ldots, n-1$.

2. Compute, for the tuning parameters $\tau = 0$ and $\tau = 1$, the observed values of the most simple ρ-estimators introduced and studied in Fraga Alves *et al.* (2003), and also used in the algorithm of Gomes and Pestana (2007). Such estimators depend on the statistics

$$V^{(\tau)}_{k,n} := \begin{cases} \dfrac{\left(M^{(1)}_{k,n}\right) - \left(M^{(2)}_{k,n}/2\right)^{1/2}}{\left(M^{(2)}_{k,n}/2\right)^{1/2} - \left(M^{(3)}_{k,n}/6\right)^{1/3}} & \text{if} \quad \tau = 1 \\[2ex] \dfrac{\ln\left(M^{(1)}_{k,n}\right) - \frac{1}{2}\ln\left(M^{(2)}_{k,n}/2\right)}{\frac{1}{2}\ln\left(M^{(2)}_{k,n}/2\right) - \frac{1}{3}\ln\left(M^{(3)}_{k,n}/6\right)} & \text{if} \quad \tau = 0, \end{cases}$$

where

$$M^{(j)}_{k,n} := \frac{1}{k}\sum_{i=1}^{k} \big(\ln X_{n-i+1:n} - \ln X_{n-k:n}\big)^j, \; j \geq 1,$$

$M^{(1)}_{k,n} \equiv H(k)$, in (1.5). They have the functional form

$$\hat\rho_\tau(k) := \min\left(0,\ \frac{3(V^{(\tau)}_{k,n}-1)}{V^{(\tau)}_{k,n}-3}\right).$$

3. Consider $\{\hat\rho_\tau(k)\}_{k\in\mathcal{K}}$, with $\mathcal{K} = ([n^{0.995}],\ [n^{0.999}])$, compute their median, denoted η_τ, and compute $I_\tau := \sum_{k\in\mathcal{K}} \left(\hat\rho_\tau(k) - \eta_\tau\right)^2$, $\tau = 0, 1$. Next choose the *tuning parameter* $\tau^* = 0$ if $I_0 \le I_1$; otherwise, choose $\tau^* = 1$.

4. Work with $\hat\rho \equiv \hat\rho_{\tau^*} = \hat\rho_{\tau^*}(k_1)$, $k_1 = [n^{0.999}]$.

5. Next, consider a sub-sample size $n_1 = o(n)$, and $n_2 = [n_1^2/n] + 1$.

6. For l from 1 until $B = 250$, generate independently B bootstrap samples $(x_1^*, \ldots, x_{n_2}^*)$ and $(x_1^*, \ldots, x_{n_2}^*, x_{n_2+1}^*, \ldots, x_{n_1}^*)$, 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 $(x_1, \ldots, x_n)$.

7. Denoting $T^*_{k,n} \equiv T^{\bullet *}_{k,n}$ the bootstrap counterpart of $T^\bullet_{k,n}$, defined in (4.3), obtain $t^*_{k,n_1,l}$, $1 < k < n_1$, $t^*_{k,n_2,l}$, $1 < k < n_2$, $1 \le l \le B$, the observed values of the statistic T^*_{k,n_i}, $i = 1, 2$.
For $k = 2, \ldots, n_i - 1$, compute

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

as well as

$$MSE_p^*(n_i,k) = (\ln c_k)^2 MSE_1^*(n_i,k), \quad MSE_x^*(n_i,k) = (\ln \tilde c_k^2) MSE_1^*(n_i,k), \quad i = 1, 2.$$

8. Obtain, for $i = 1, 2$,

$$\hat k^*_{0|T}(n_i) := \arg\min_{1<k<n_i} MSE_1^*(n_i,k), \tag{4.6}$$

$$\hat k^*_{0|p}(n_i) := \arg\min_{1<k<n_i} MSE_p^*(n_i,k), \tag{4.7}$$

$$\hat k^*_{0|x}(n_i) := \arg\min_{1<k<n_i} MSE_x^*(n_i,k), \tag{4.8}$$

and return to Step **6.** whenever $\hat k^*_{0|\bullet}(n_2) \ge \hat k^*_{0|\bullet}(n_1)$.

9. On the basis of (4.6), compute the threshold estimate $\hat k_0^{\bullet *} \equiv \hat k_0^{\bullet *}(n; n_1)$, in (4.4), for the adaptive EVI-estimation. Compute also the threshold estimates, $\hat k^{\bullet *}_{0|p} \equiv \hat k^{\bullet *}_{0|p}(n; n_1)$ and $\hat k^{\bullet *}_{0|x} \equiv \hat k^{\bullet *}_{0|x}(n; n_1)$, on the basis of an equation similar to (4.4), but with $\hat k^*_{0|T}$ replaced by $\hat k^*_{0|p}$ and $\hat k^*_{0|x}$, in (4.7) and (4.8), according as we are interested in the adaptive VaR_p-estimation or the p_x-estimation, respectively. Return also to Step **6.** whenever these k-estimates are equal to $n - 1$.

10. Obtain the adaptive EVI-estimate, $\hat\gamma^\bullet_* \equiv \hat\gamma^\bullet_*(n, n_1) := \hat\gamma^\bullet_{\hat k_0^{\bullet *}, n}$, already provided in (4.5).

11. Finally, compute the adaptive VaR_p- and p_x-estimates,

$$\widehat{Q}_p^{\bullet *} \equiv \widehat{Q}_{p|\hat{\gamma}}^{\bullet *} := \widehat{Q}^{\bullet}_{\hat{k}^{\bullet *}_{0|p},n}, \qquad \hat{p}_x^{\bullet *} \equiv \hat{p}_{x|\hat{\gamma}}^{\bullet *} := \hat{p}^{\bullet}_{\hat{k}^{\bullet *}_{0|x},n},$$

with $\hat{k}^{\bullet *}_{0|p}$ and $\hat{k}^{\bullet *}_{0|x}$ obtained in Step 9.

Remarks:

(i) If there are negative elements in the sample, n should be replaced by $n^+ = \sum_{i=1}^n I_{[X_i>0]}$ (the number of positive values in the sample). Analogously for n_1 and for n_2.

(ii) The Monte-Carlo procedure in Steps **6.** to **11.** of the Algorithm can be replicated r times if we want to associate standard errors to the OSF's estimates and to the γ, VaR_p and p_x bootstrap estimates. The value of B can also be adequately chosen.

5 Applications to simulated and real data

5.1 A simulated data set

In order to have an indication on the way the algorithm in Section 4.1 performs, and to further motivate its use, prior to its validation through a surely highly time-consuming simulation, a topic out of the scope of this paper, we first apply it to an arbitrarily simulated sample, with size $n = 500$, from a Fréchet model, with $\gamma = 0.25$. The algorithm here presented led to the ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -1.173$, obtained at the level $k_1 = [n_0^{0.999}] = 496$. The associated β-estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 0.901$. Then, we got $\hat{k}_0^H = 102$, with $\hat{k}_0^H$ provided in (4.1), and an associated EVI estimate equal to 0.285, clearly over-estimatimating the true value $\gamma = 0.25$. The associated approximate 95% confidence interval, in (4.2), is $(0.2266, 0.3277)$, with a size 0.1011.

The application of the algorithm presented in Section 4.1 of this paper, with a sub-sample size $n_1 = [n^{0.955}] = 378$, and $B = 250$ bootstrap generations, led us to $\hat{k}_0^{PPWM*} = 52$ and to the adaptive PPWM EVI-estimate $PPWM^* \equiv \hat{\gamma}_*^{PPWM} = 0.254$. This same algorithm applied to the Hill estimators leads us to $\hat{k}_0^{H*} = 167$ and to the adaptive Hill EVI-estimate $H^* \equiv \hat{\gamma}_*^H = 0.284$. These values are pictured in Figure 5, where we also present the estimates under study as a function of k. The most adequate estimate seems neatly to be the one associated with the PPWM methodology, particularly due to the smoothness of the stability region of the estimates as a function of k. Similar conclusions can be drawn for the VaR_p and p_x-estimates.

For the estimation of $\mathrm{VaR}_{1/(2n)} = \mathrm{VaR}_{0.001}$, equal to 5.623 for this Fréchet sample, and through the PPWM estimators, we were led to a choice of k equal to 52 and to the $\mathrm{VaR}_{0.001}$ estimate, $\hat{Q}_{0.001} = 5.987$, slightly overestimating the true value of $\mathrm{VaR}_{0.001} = 5.623$. For the estimation of $p = 1/(2n) = 0.001$, the PPWM methodology led us to a k-value equal to 51 and $\hat{p} = 0.0013$, quite close to p. Regarding the classical estimators of VaR_p and p, we were led to a k-value equal to 167, in both cases, and to the estimates 5.638 and 0.002, respectively.

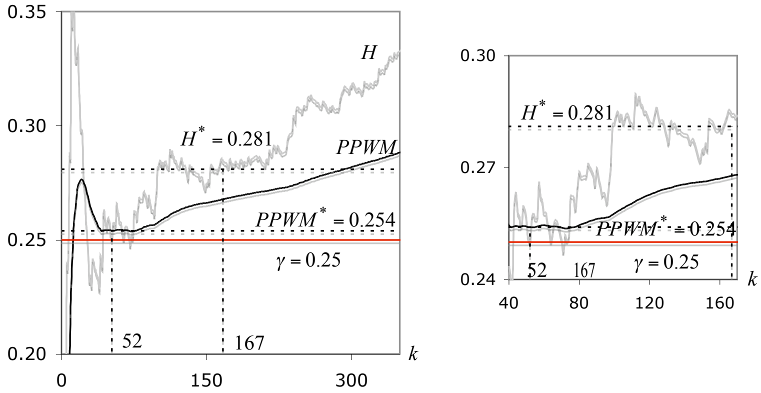

Figure 5: Non-adaptive (as a function of k) and bootstrap adaptive Hill (H) and PPWM EVI-estimates, for a Fréchet sample of size $n = 500$ and $\gamma = 0.25$ (*left*), and a zoomed figure for $40 \leq k \leq 170$ (*right*).

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

In Figure 6, we picture at the left, as a function of the sub-sample size n_1, ranging from $n_1 = [n^{0.9}] = 268$ until $n_1 = [n^{0.9999}] = 499$, the bootstrap EVI-estimates associated with the Hill and the PPWM estimators, in (1.5) and (1.12), respectively. The pictures in the center and the right are similar to the picture at the left, but related with the corresponding adaptive bootstrap estimation of $\text{VaR}_{0.001} = 5.623$ and $p = 0.001$, respectively.

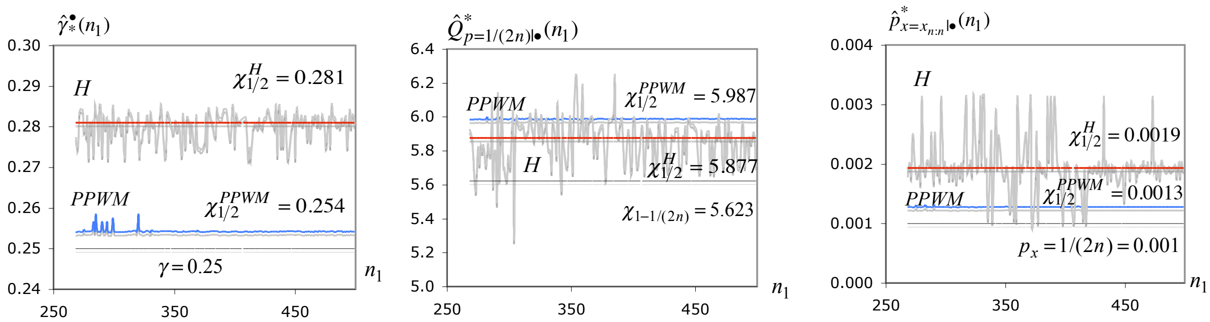

Figure 6: Bootstrap Hill (H) and PPWM EVI-estimates of γ (*left*), $Var_{0.001} = 5.623$ (*center*) and $p = 1/(2n) = 0.001$ (*right*), as functions of n_1.

The low sensitivity of the PPWM estimates to changes of the subsample size n_1 seems also to be a point in favour of these new estimators. Whereas the bootstrap PPWM EVI, VaR_p and p-estimates lie in the intervals (0.254, 0.259), (5.977, 6.002) and (0.00127, 0.00131), respectively, the bootstrap Hill-type EVI, VaR_p and p-estimates lie in the intervals (0.272, 0.285), (5.273, 6.254) and (0.0010, 0.0029), respectively. Note however that the reasonably high volatility of the bootstrap Hill VaR_p and p_x-estimates as a function of n_1 can lead us to estimates closer to the target for some of the values of the sub-sample size n_1.

5.2 A case study in the field of insurance

We shall next consider an illustration of the performance of the adaptive PPWM EVI-estimates under study, comparatively with the same methodology applied to the Hill EVI-estimates, again 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. See also Gomes *et al.* (2009). A preliminary graphical analysis of the data, x_i, $1 \le i \le n$, $n = 371$, leads us to an immediate conclusion that data have been censored to the left and that the right-tail of the underlying model is quite heavy. The sample paths of 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 $\hat\rho_0 = -0.74$, obtained at the level $k_1 = [n^{0.999}] = 368$. The associated β-estimate, based on the β-estimator in Gomes and Martins (2002), was $\hat\beta_0 = 0.80$. Then, we got the estimate $\hat k_0^H = 55$, with $\hat k_0^H$ provided in (4.1), and an associated adaptive EVI estimate equal to 0.291. The associated approximate 95% confidence interval, in (4.2), is $(0.2115, 0.3432)$, with a size 0.1317.

The application of the algorithm presented in Section 4.1 of this paper, with a sub-sample size $n_1 = [n^{0.955}] = 284$, and $B = 250$ bootstrap generations, led us to $\hat k_0^{PPWM*} = 58$ and to the adaptive PPWM EVI-estimate $PPWM^* \equiv \hat\gamma_*^{PPWM} = 0.272$. This same algorithm applied to the Hill estimates leads us to $\hat k_0^{H*} = 52$ and to the adaptive Hill EVI-estimate $H^* \equiv \hat\gamma_*^{H} = 0.299$. These values are pictured in Figure 7, where we also present the EVI-estimates under study as a function of k. Again, similar results have been obtained for the Var_p and p_x-estimates.

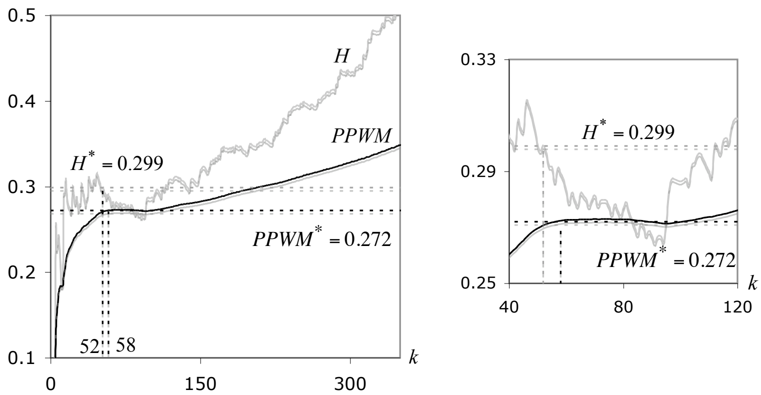

Figure 7: Hill (H) and PPWM EVI-estimates for the SECURA data, as a function of k (*left*), and a zoomed figure for $40 \le k \le 120$ (*right*).

For the estimation of $\mathrm{VaR}_{1/(2n)} = \mathrm{VaR}_{0.0013}$, and through the PPWM estimators, we were led to a choice of k equal to 59 and to the VaR estimate, $\widehat{Q}_{1/(2n)} = 10,658,360$. For the estimation of $p = \mathbb{P}(X > X_{n:n}) = 1/(n+1) = 0.0027$, the PPWM methodology led us to a k-value also equal to 59 and $\hat{p} = 0.0040$, reasonably above the true value of p. Regarding the classical estimators of VaR_p and p, we were led to k-values respectively equal to 52 and 53, and to the estimates $11,356,460$ and 0.0053, respectively. The overestimation of the p-estimate is now more flagrant.

Figure 8 is similar to Figure 6, but for the insurance data.

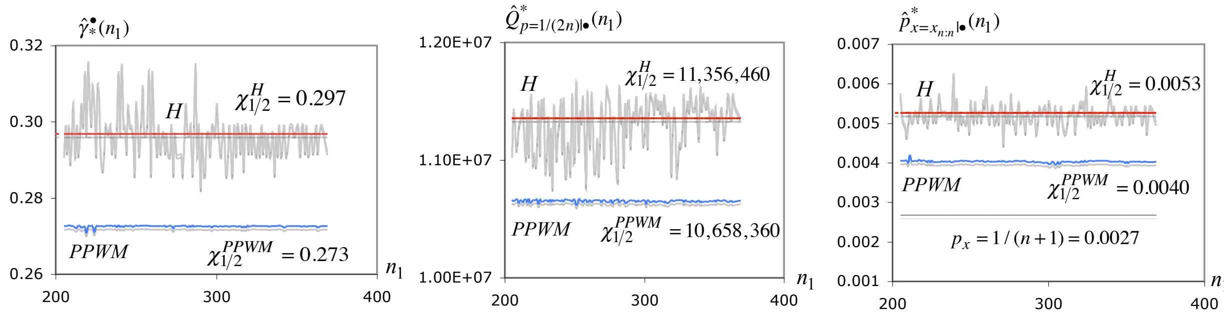

Figure 8: Bootstrap Hill (H) and PPWM EVI-estimates of γ (*left*), $Var_{1/(2n)}$ (*center*) and $p = 1/(n+1) = 0.0027$ (*right*), as functions of n_1.

The bootstrap PPWM EVI, VaR_p and p-estimates are indeed quite stable as a function of the sub-sample size n_1. The bootstrap PPWM EVI-estimates vary from a minimum value equal to 0.271 until 0.273, with a median equal to 0.273, equal to the value obtained for the bootstrap EVI estimate associated to the arbitrarily chosen sub-sample size $n_1 = [n^{0.955}] = 284$. We can indeed guarantee the two decimal figures, i.e. the EVI-estimate 0.27. Up to three decimal figures, all bootstrap p_x-estimates were equal to 0.004. Regarding the bootstrap VaR_p-estimates, we got values between a minimum equal to 10618780 and a maximum equal to 10673260, with a median 10658360. Again, the low sensitivity of the PPWM estimates to changes of the subsample size n_1 seems again to be a point in favour of these new estimators. Whereas the bootstrap PPWM EVI, VaR_p and p-estimates lie in the intervals (0.271, 0.273), (10618780, 10673260) and (0.0040, 0.0042), respectively, the bootstrap Hill-type EVI, VaR_p and p-estimates lie in the intervals (0.283, 0.315), (10764560, 11660710) and (0.0045, 0.0063), respectively. As already detected in previous papers, and in the most diversified comparisons, the Hill estimates are clearly over-estimating the true value of the EVI. As mentioned before, the most adequate estimate seems neatly to be the one associated with the PPWM methodology, particularly due to the smoothness of the stability region of the estimates as a function of k.

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] Balkema, A.A. and de Haan, L. (1974). Residual Life Time at Great Age. *Annals of Probability*, *2*, 792–804.

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

[4] 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.

[5] Beirlant, J., Caeiro, F. and Gomes, M.I. (2012). *An Overview and Open Research Topics in the Field of Statistics of Univariate Extremes*. Notas e Comunicações CEAUL 01/2012.

[6] Caeiro, F. and Gomes, M.I. (2009). Semi-Parametric Second-Order Reduced-Bias High Quantile Estimation. *Test*, *18*(2), 392–413.

[7] Caeiro, F. and Gomes, M.I. (2011) . Semi-Parametric Tail Inference through Probability-Weighted Moments. *J. Statist. Plann. Inference*, *141*(2), 937–950.

[8] Danielsson, J., Haan, L. de, 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.

[9] Dekkers, A., and de Haan, L. (1993). Optimal Sample Fraction in Extreme Value Estimation. *J. Multivariate Analysis* *47*(2), 173-195.

[10] Diebolt, J., Guillou, A. and Rached, I. (2007). Approximation of the Distribution of Excesses through a Generalized Probability-Weighted Moments Method. *J. Statist Plann. Inference*, *137*, 841–857.

[11] Diebolt, J., Guillou, A., Naveau, P. and Ribereau, P. (2008). Improving Probability-Weighted Moment Methods for the Generalized Extreme Value Distribution. *Revstat*, *6*(1), 33–50.

[12] Draisma, G., de Haan, L., Peng, L. and Pereira, M.T. (1999). A Bootstrap-Based Method to Achieve Optimality in Estimating the Extreme Value Index. *Extremes*, *2* (4), 367-404.

[13] Drees, H. (1998). On Smooth Statistical Tail Functions. *Scand. J. Statist.*, *25*, 187–210.

[14] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, Heidelberg.

[15] 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.

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

[17] Gnedenko, B.V. (1943). Sur la Distribution Limite du Terme Maximum d'une Série Aléatoire. *Ann. Math.*, *44*, 423–453.

[18] 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.

[19] 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.

[20] Gomes, M.I. and Pestana, D. (2007). A Sturdy Reduced-Bias Extreme Quantile (VaR) Estimator. *J. American Statistical Association*, *102*(477), 280–292.

[21] 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.

[22] Gomes, M.I., Canto e Castro, L., Fraga Alves, M.I. and Pestana, D.D. (2008). Statistics of Extremes for IID Data and Breakthroughs in the Estimation of the Extreme Value Index: Laurens de Haan Leading Contributions. *Extremes*, *11*(1), 3–34.

[23] Gomes, M.I., Mendonça, S. and Pestana D. (2009). Adaptive Reduced-Bias Tail Index and Value-at-Risk Estimation. In Sakalauskas L., Skiadas C. and Zavadskas E.K. (eds.). *Applied Stochastic Models and Data Analysis ASMDA* 2009, IMI & VGTU Editions, 41–44.

[24] Gomes, M.I., Mendonça, S. and Pestana D. (2011). Adaptive Reduced-Bias Tail Index and VaR Estimation via the Bootstrap Methodology. *Comm. in Statistics – Theory and Methods*, *40*(16), 2946-2968.

[25] Gomes, M.I., Figueiredo, F. and Neves, M.M. (2012). Adaptive Estimation of Heavy Right Tails: the Bootstrap Methodology in Action. *Extremes*, DOI: 10.1007/s10687-011-0146-6

[26] Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R. (1979). Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. *Water Resources Research*, *15*, 1049–1054.

[27] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: an Introduction.* Springer Science+Business Media, LLC, New York.

[28] de Haan, L. and Peng, L. (1998). Comparison of Tail Index Estimators. *Statistica Neerlandica*, *52*, 60–70.

[29] Hall, P. (1990). Using the Bootstrap to Estimate Mean-Square Error and Selecting Smoothing Parameter in Non-Parametric Problems. *J. Multivariate Anal.*, *32*, 177-203.

[30] Hill, B.M. (1975). A Simple General Approach to Inference about the Tail of a Distribution. *Ann. Statist.*, *3*(5), 1163–1174.

[31] Hosking, J., Wallis, J. and Wood E. (1985). Estimation of the Generalized Extreme Value Distribution by the Method of Probability-Weighted Moments. *Technometrics*, *27*(3), 251–261.

[32] Hosking, J. and Wallis, J. (1987). Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics*, *29*(3), 339–349.

[33] Pickands III, J. (1975). Statistical Inference Using Extreme Order Statistics. *Ann. Statist.*, *3*, 119–131.

[34] Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values, with Application to Insurance, Finance, Hydrology and Other Fields*, 3rd edition. Birkhäuser Verlag.

[35] Vandewalle B., Beirlant J. (2006). On Univariate Extreme Value Statistics and the Estimation of Reinsurance Premiums. *Insurance Mathematics and Economics*, *38*:3, 444–459.

[36] Weissman, I. (1978). Estimation of Parameters of Large Quantiles Based on the k Largest Observations. *J. Amer. Statist. Assoc.*, *73*, 812–815.