Computational Study of the Adaptive Estimation of the Extreme Value Index with Probability Weighted Moments

Frederico Caeiro　and **M. Ivette Gomes**

Abstract　In statistics of extremes, the estimation of the extreme value index (EVI) is an important and central topic of research. We consider the probability weighted moment estimator of the EVI, based on the largest observations. Due to the specificity of the properties of the estimator, a direct estimation of the threshold is not straightforward. In this work, we consider an adaptive choice of the number of order statistics based on the double bootstrap methodology. Computational and empirical properties of the methodology are here provided.

Keywords　Bootstrap · Extreme value index · Heavy tails · Probability weighted moment · Semi-parametric estimation

1　Introduction and Scope of the Article

Let $(X_1, \ldots, X_n)$ denote a random sample of size n from a population with unknown *cumulative distribution function* (CDF) $F(x) = \mathbb{P}(X \leq x)$ and consider the associated sample of ascending *order statistics* (OSs) ($X_{1:n} := \min_{1 \leq i \leq n} X_i \leq \cdots \leq X_{n:n} := \max_{1 \leq i \leq n} X_i$). Further assume that for large values of x, $F(x)$ is a Pareto-type model, i.e., a model with a regular varying right tail with a negative index of regular variation equal to $-1/\xi$ ($\xi > 0$). Consequently,

$$\overline{F}(x) := 1 - F(x) = \mathbb{P}(X > x) = x^{-1/\xi} L(x), \quad \text{as } x \to \infty, \tag{1}$$

F. Caeiro (✉)
NOVA School of Science and Technology (FCT NOVA) and CMA, NOVA University of Lisbon, Campus de Caparica, Lisbon, Portugal
e-mail: fac@fct.unl.pt

M. I. Gomes
Faculty of Science of Lisbon (FCUL/DEIO) and CEAUL, University of Lisbon, Cidade Universitária, Campo Grande, Portugal
e-mail: ivette.gomes@fc.ul.pt

with $L(\cdot)$ a slowly varying function, i.e.

$$\lim_{t \to \infty} \frac{L(tx)}{L(t)} = 1, \quad \forall\, x > 0.$$

Models satisfying the condition (1) are in the domain of attraction for maxima of a non-degenerate distribution. This means that there exist normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{X_{n:n} - b_n}{a_n} \leq x \right) = \lim_{n \to \infty} F^n\left(a_n x + b_n\right) = G(x), \tag{2}$$

with $G(\cdot)$ a non-degenerate CDF. With the appropriate choice of the normalizing constants in (2), and under a general framework, G is the general *extreme value* (EV) distribution,

$$G(x) \equiv \mathrm{EV}_\xi(x) := \begin{cases} \exp\left(-(1 + \xi x)^{-1/\xi}\right), & 1 + \xi x > 0, \;\; \text{if } \xi \neq 0, \\ \exp(-\exp(-x)), & x \in \mathbb{R}, \qquad\quad\;\; \text{if } \xi = 0, \end{cases} \tag{3}$$

given here in the von Mises-Jenkinson form (see [1, 2]). Whenever such a non-degenerate limit exists, we write $F \in \mathcal{D}_\mathcal{M}(EV_\xi)$, and the real parameter ξ is the *extreme value index* (EVI).

As already mentioned, we shall deal with Pareto right-tails, i.e. heavy right-tails or equivalently a model with a positive EVI. Then, the right-tail function is of regular variation with an index of regular variation equal to $-1/\xi$, i.e.

$$F \in \mathcal{D}_\mathcal{M}(EV_\xi)_{\xi > 0} \quad \Longleftrightarrow \quad \overline{F} := 1 - F \in RV_{-1/\xi}, \tag{4}$$

where the notation RV_α stands for the class of *regularly varying* functions at infinity with an *index of regular variation* equal to α, i.e. positive measurable functions g such that $\lim\limits_{t \to \infty} g(tx)/g(t) = x^\alpha$, for all $x > 0$. With the notation

$$U(t) := F^{\leftarrow}(1 - 1/t),\; t \geq 1, \quad F^{\leftarrow}(y) := \inf\left\{ x : F(x) \geq y \right\}, \tag{5}$$

condition (4) is equivalent to $U \in RV_\xi$. Pareto-type models are extremely important in practice due to the frequency and magnitude of extreme values and inference on extreme and large events is usually performed on the basis of the $k + 1$ largest order statistics in the sample, as sketched in Fig. 1.

Fig. 1 A Pareto right-tail probability density function
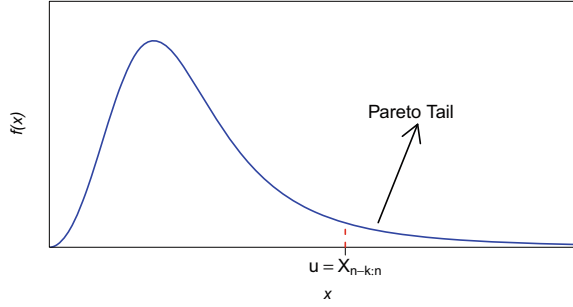
1.1 EVI-Estimators Under Consideration

One of the first classes of semi-parametric estimators of a positive EVI was the class of Hill (H) estimators introduced in [3] and given by

$$\hat{\xi}_{k,n}^{\mathrm{H}} := \frac{1}{k} \sum_{i=1}^{k} \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}, \; k = 1, 2, \ldots, n-1. \tag{6}$$

This estimator can be highly sensitive to the choice of k, especially in the presence of a substantial bias. As an alternative, we shall also consider the *Pareto probability weighted moments* (PPWM) EVI-estimators, introduced in [4]. They are consistent for $0 < \xi < 1$, compare favourably with the Hill estimator, and are given by

$$\hat{\xi}_{k,n}^{\mathrm{PPWM}} := 1 - \frac{\hat{a}_1(k)}{\hat{a}_0(k) - \hat{a}_1(k)}, \tag{7}$$

with

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

For other alternative estimators of the EVI see Refs. [5–7], among others. Consistency of the EVI-estimators in (6) and (7) is achieved if $X_{n-k:n}$ is an *intermediate* OS, i.e. if

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

In order to derive the asymptotic normality of these EVI-estimators, it is often assumed the validity of a second-order condition, like

$$\lim_{t \to \infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \begin{cases} \frac{x^\rho - 1}{\rho}, & \text{if } \rho < 0, \\ \ln x, & \text{if } \rho = 0, \end{cases} \tag{8}$$

Table 1 Asymptotic variance and bias' rulers of H and PPWM EVI–estimators

	H ($\xi > 0$)	PPWM ($0 < \xi < 1/2$)
$\sigma_\bullet^2$	ξ^2	$\xi^2 \ \frac{(1-\xi)(2-\xi)^2}{(1-2\xi)(3-2\xi)}$
$b_\bullet$	$\frac{1}{1-\rho}$	$\frac{(1-\xi)(2-\xi)}{(1-\xi-\rho)(2-\xi-\rho)}$

where $U(\cdot)$ is defined in (5) and $|A| \in RV_\rho, \rho \le 0$. Under such a second-order framework, if $\sqrt{k}A(n/k) \to \lambda_A$, finite, as $n \to \infty$, these EVI-estimators are asymptotically normal. Denoting $\hat{\xi}_{k,n}^\bullet$, any of the estimators above, we have, with $Z_k^\bullet$ an asymptotically standard normal random variable and for adequate $(b_\bullet, \ \sigma_\bullet) \in (\mathbb{R}, \ \mathbb{R}^+)$,

$$\hat{\xi}_{k,n}^\bullet \overset{d}{=} \xi + \frac{\sigma_\bullet Z_k^\bullet}{\sqrt{k}} + b_\bullet \ A(n/k)(1 + o_\mathbb{P}(1)), \quad \text{as} \quad n \to \infty, \tag{9}$$

with $b_\bullet$ the asymptotic bias, and $\sigma_\bullet^2$ the asymptotic standard deviation of the approximation, given in Table 1.

Under the above second-order framework, in (8), but with $\rho < 0$, let us use the parametrization

$$A(t) = \xi\beta t^\rho, \quad \text{with } \beta \ne 0 \text{ and } \rho < 0,$$

where β and ρ are generalized scale and shape second-order parameters, which need to be adequately estimated on the basis of the available sample. Let us denote the optimal level by

$$\tilde{k}_0^\bullet(n) := \arg\min_k \text{MSE}(\hat{\xi}_{k,n}^\bullet),$$

with MSE standing for *mean squared error*. With $\mathbb{E}$ denoting the mean value operator and AMSE standing for *asymptotic* MSE, a possible substitute for $\text{MSE}(\hat{\xi}_{k,n}^\bullet)$ is

$$\text{AMSE}(\hat{\xi}_{k,n}^\bullet) := \mathbb{E}\left(\frac{\sigma_\bullet}{\sqrt{k}}Z_k^\bullet + b_\bullet A(n/k)\right)^2 = \frac{\sigma_\bullet^2}{k} + b_\bullet^2 \xi^2 \beta^2 \left(\frac{n}{k}\right)^{2\rho},$$

cf. Eq. (9). Then, with the notation $k_0^\bullet(n) := \arg\min_k \text{AMSE}(\hat{\xi}_{k,n}^\bullet)$, we get

$$k_0^\bullet(n) = \left(\frac{\sigma_\bullet^2 \ n^{-2\rho}}{(-2\rho) \ b_\bullet^2 \ \xi^2\beta^2}\right)^{1/(1-2\rho)} = \tilde{k}_0^\bullet(n)(1 + o(1)). \tag{10}$$

For the Hill estimator in (6), and as can be seen in Table 1, we have $(b_H, \sigma_H) = (1/(1 - \rho), \xi)$. Consequently, with $(\hat{\beta}, \hat{\rho})$ a consistent estimator of (β, ρ) and $[x]$ denoting the integer part of x, we have an asymptotic justification for the estimator

$$\hat{k}_0^{\mathrm{H}} := \left[\left(\frac{(1 - \hat{\rho})^2 n^{-2\hat{\rho}}}{(-2\hat{\rho}\hat{\beta}^2)} \right)^{1/(1-2\hat{\rho})} \right] + 1.$$

The same does not happen with the PPWM EVI-estimators, due to the fact that σ_{PPWM}, b_{PPWM} and consequently k_0^{PPWM} depend on the value of ξ (see Table 1, again). It is thus sensible to use the bootstrap methodology for the adaptive choice of the threshold associated to the PPWM EVI-estimation.

1.2 Scope of the Article

The main goal is the adaptive estimation of the EVI. For that purpose, the choice of the threshold is crucial and we study computationally a recent bootstrap algorithm. After a review, in Sect. 2, of the role of the bootstrap methodology in the estimation of optimal sample fractions, we provide an algorithm for the adaptive estimation through the Hill and the PPWM EVI-estimators. In Sect. 3 we provide results from a Monte Carlo simulation study. In Sect. 4, as an illustration, we apply such methodology to a data set in the field of insurance. Section 5 concludes the paper.

2 Adaptive EVI-Estimation and the Bootstrap Methodology

Similarly to what has been done in Gomes and Oliveira [8], for the H estimator, and in Gomes et al. [9], for adaptive reduced-bias estimation, we can use the algorithm in Caeiro et al. [10] (see also [4]), considering the auxiliary statistic,

$$T_{k,n}^{\bullet} := \hat{\xi}_{[k/2],n}^{\bullet} - \hat{\xi}_{k,n}^{\bullet}, \quad k = 2, \ldots, n - 1, \tag{11}$$

which converges to the known value zero, and double-bootstrap it adequately, in order to estimate $k_0^{\bullet}(n)$, through a bootstrap estimate $\hat{k}_0^{\bullet,*}$. Indeed, again under the second-order framework, in (8), we get, for the auxiliary statistic $T_{k,n}^{\bullet}$, in (11), the asymptotic distributional representation,

$$T_{k,n}^{\bullet} \overset{d}{=} \frac{\sigma_{\bullet} Q_k^{\bullet}}{\sqrt{k}} + b_{\bullet}(2^{\rho} - 1) \, A(n/k) + o_{\mathbb{P}}(A(n/k)),$$

with $Q_k^{\bullet}$ asymptotically standard normal, and $(b_{\bullet}, \sigma_{\bullet})$ given in Table 1. The AMSE of $T_{k,n}^{\bullet}$ is thus minimal at a level $k_{0|T}^{\bullet}(n)$ such that $\sqrt{k} \, A(n/k) \to \lambda_A' \neq 0$, i.e. a level of the type of the one in (10), with $b_{\bullet}$ replaced by $b_{\bullet}(2^{\rho} - 1)$, and we consequently have

$$k_0^{\bullet}(n) = k_{0|T}^{\bullet}(n) \, (1 - 2^{\rho})^{\frac{2}{1-2\rho}} \, (1 + o(1)).$$

2.1 The Bootstrap Methodology in Action

Given the sample $\underline{X}_n = (X_1, \ldots, X_n)$ from an unknown model F, consider for any $n_1 = O(n^{1-\epsilon})$, with $0 < \epsilon < 1$, the bootstrap sample $\underline{X}_{n_1}^* = (X_1^*, \ldots, X_{n_1}^*)$, from the model $F_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i \leq x\}}$, the empirical CDF associated with the original sample $\underline{X}_n$. We choose the resample size n_1 to be less than the original sample size to avoid underestimation of the bias (see Hall [11]). Next, associate to that bootstrap sample the corresponding bootstrap auxiliary statistic, denoted $T_{k_1,n_1}^{\bullet,*}$, $1 < k_1 < n_1$. Then, with the notation

$$k_{0|T}^{\bullet,*}(n_1) = \arg\min_{k_1} \mathrm{AMSE}\big(T_{k_1,n_1}^{\bullet,*}\big),$$

we have that

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

Consequently, for another sample size $n_2 = n_1^2/n$,

$$\frac{\big(k_{0|T}^{\bullet,*}(n_1)\big)^2}{k_{0|T}^{\bullet,*}(n_2)} = k_{0|T}^{\bullet}(n)(1 + o(1)), \quad \text{as } n \to \infty.$$

We are now able to estimate $k_0^{\bullet}(n)$, on the basis of any estimate $\hat{\rho}$ of ρ. With $\hat{k}_{0|T}^{\bullet,*}$ denoting the sample counterpart of $k_{0|T}^{\bullet,*}$, $\hat{\rho}$ the ρ-estimate and taking into account (10), we can build the k_0-estimate,

$$\hat{k}_0^{\bullet,*} \equiv \hat{k}_0^{\bullet,*}(n; n_1) := \min\left(n-1, \left[\frac{\left(1 - 2^{\hat{\rho}}\right)^{\frac{2}{1-2\hat{\rho}}} \left(\hat{k}_{0|T}^{\bullet,*}(n_1)\right)^2}{\hat{k}_{0|T}^{\bullet,*}([n_1^2/n] + 1)}\right] + 1\right), \qquad (12)$$

and the ξ-estimate

$$\hat{\xi}^{\bullet,*} \equiv \hat{\xi}^{\bullet,*}(n; n_1) := \hat{\xi}_{\hat{k}_{0|T}^{\bullet,*}(n;n_1),n}. \qquad (13)$$

A few questions, some of them with answers outside the scope of this paper, may be raised: How does the bootstrap method work for small or moderate sample sizes? Is the method strongly dependent on the choice of n_1? What is the type of the sample path of the EVI-estimator, as a function of n_1? What is the sensitivity of the bootstrap 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 $n_1 = n$?

2.2 An Algorithm for the Adaptive EVI-Estimation

The estimates $(\hat{\beta}, \hat{\rho})$, of the vector (β, ρ) of second-order parameters, are the ones already used in previous papers:

1. Given a sample $(x_1, \ldots, x_n)$, consider the observed values of the ρ-estimators $\hat{\rho}_\tau(k)$, introduced and studied in Fraga Alves et al.[12], for tuning parameters $\tau = 0$ and $\tau = 1$.
2. Select $\left\{\hat{\rho}_\tau(k)\right\}_{k \in \mathcal{K}}$, with $\mathcal{K} = ([n^{0.995}], [n^{0.999}])$, and compute their median, denoted by η_τ, $\tau = 0, 1$.
3. Next compute $I_\tau := \sum_{k \in \mathcal{K}} \left(\hat{\rho}_\tau(k) - \eta_\tau\right)^2$, $\tau = 0, 1$, and choose the *tuning parameter* $\tau^* = 0$ if $I_0 \leq I_1$; otherwise, choose $\tau^* = 1$.
4. Work with $\hat{\rho} \equiv \hat{\rho}_{\tau^*} = \hat{\rho}_{\tau^*}(k_1)$ and $\hat{\beta} \equiv \hat{\beta}_{\tau^*} := \hat{\beta}_{\hat{\rho}_{\tau^*}}(k_1)$, $k_1 = [n^{0.999}]$ and $\hat{\beta}_{\hat{\rho}}(k)$ given in Gomes and Martins [13].

Now, and with $\hat{\xi}_{k,n}^{\mathrm{H}}$ and $\hat{\xi}_{k,n}^{\mathrm{PPWM}}$ respectively defined in (6) and (7), the algorithm goes on with the following steps:

5. Compute $\hat{\xi}_{k,n}^{\bullet}$, $k = 1, \ldots, n-1$, $\bullet = $ H and/or PPWM.
6. Next, consider a sub-sample size $n_1 = o(n)$, and $n_2 = [n_1^2/n] + 1$.
7. For l from 1 until B, 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 CDF, $F_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i \leq x\}}$, associated with the observed sample $(x_1, \ldots, x_n)$.
8. Denoting by $T_{k,n}^{\bullet,*}$ the bootstrap counterpart of $T_{k,n}^{\bullet}$, defined in (11), obtain $(t_{k,n_1,l}^{\bullet,*}, \ t_{k,n_2,l}^{\bullet,*})$, $1 \leq l \leq B$, the observed values of the statistic $T_{k,n_i}^{\bullet,*}$, $i = 1, 2$. For $k = 2, \ldots, n_i - 1$, compute

$$\mathrm{MSE}^{\bullet,*}(n_i, k) = \frac{1}{B} \sum_{l=1}^{B} \left(t_{k,n_i,l}^{\bullet,*}\right)^2,$$

and obtain

$$\hat{k}_{0|T}^{\bullet,*}(n_i) := \arg \min_{1 < k < n_i} \mathrm{MSE}^{\bullet,*}(n_i, k), \quad i = 1, 2.$$

9. Compute the threshold estimate $\hat{k}_0^{\bullet,*}$, in (12).
10. Finally obtain

$$\hat{\xi}^{\bullet,*} \equiv \hat{\xi}^{\bullet,*}(n; n_1) = \hat{\xi}_{\hat{k}_{0|T}^{\bullet,*}(n;n_1),n},$$

already provided in (13).

Such an algorithm needs to be computationally validated, a topic we deal with in the next section. Further, note that bootstrap *confidence intervals* (CIs) are easily associated with the estimates presented through the replication of this algorithm r times.

3 A Small-Scale Simulation Study

In this section, we have implemented a multi-sample Monte Carlo simulation experiment of size 1000, to obtain the distributional behaviour of the EVI adaptive bootstrap estimates $\hat{\xi}^{H,*}$ and $\hat{\xi}^{PPWM,*}$ in (6) and (7), respectively. We have considered a resample of size $n_1 = [n^{0.955}]$ for samples of size $n = 100, 200, 500, 750, 1000, 2000$ and 5000 from the following models:

- the Fréchet model, with d.f.

$$F(x) = \exp(-x^{-1/\xi}), \quad x > 0, \quad \xi > 0,$$

  with $\xi = 0.25$ ($\rho = -1$);
- the Burr model, with d.f.

$$F(x) = 1 - (1 + x^{-\rho/\xi})^{1/\rho}, \quad x > 0,$$

  with $(\xi, \rho) = (0.25, -0.75)$;
- the Half-t_4 model, i.e., the absolute value of a Student's t with $\nu = 4$ degrees of freedom ($\xi = 0.25$, $\rho = -0.5$).

In Table 2 we present, for the above mentioned models, the multi-sample simulated (mean) double bootstrap optimal sample fraction (OSF), the mean (E) and median (med) of the EVI-estimates and the simulated RMSE for both EVI-estimators, as a function of the sample size n. The less biased EVI-estimate and the smallest RMSE is presented in **bold**. Although both estimators over-estimate the EVI, the consideration of the PPWM EVI-estimator leads to a less biased EVI-estimate, as expected. The PPWM estimation can also lead to a smaller RMSE, for models with $|\rho| < 1$.

4 A Case Study

Here, the performance of the adaptive double bootstrap procedure is illustrated through the analysis of a real dataset. The analysis was made in R software with the computer code developed in Caeiro and Gomes [14]. We used the dataset Auto-Claims from a motor insurance portfolio. The data is available in the R package `insuranceData` [15]. The variable of interest is the amount paid on a closed claim, in dollars. There are $n = 6773$ claims available. Since large claims are a topic of great concern in the Insurance Industry, accurate modelling of the right tail of the underlying distribution is extremely important. The Histogram and the Pareto Quantile-Quantile (QQ) Plot, in Fig. 2, are compatible with a Pareto-type underlying distribution.

In Fig. 3, we present the EVI-estimates provided by the Hill and the PPWM EVI-estimators in (6) and (7), respectively. Both estimators are upward biased for large k.

Table 2 Simulated OSF, EVI-estimates (provided by the mean and median) and RMSE estimates obtained through the Hill and PPWM estimators and the double bootstrap methodology

|	Hill				PPWM			
n	OSF	$E(\hat{\xi}^*)$	med$(\hat{\xi}^*)$	RMSE$(\hat{\xi}^*)$	OSF	$E(\hat{\xi}^*)$	med$(\hat{\xi}^*)$	RMSE$(\hat{\xi}^*)$
	Fréchet with $\xi = 0.25$ $(\rho = -1)$							
100	0.3458	0.2665	0.2716	**0.0706**	0.3677	**0.2446**	**0.2533**	0.0728
200	0.3246	0.2675	0.2701	**0.0561**	0.3453	**0.2503**	**0.2559**	0.0573
500	0.2750	0.2656	0.2674	**0.0390**	0.3101	**0.2544**	**0.2587**	0.0395
750	0.2567	0.2646	0.2680	**0.0322**	0.2826	**0.2532**	**0.2571**	0.0353
1000	0.2420	0.2630	0.2640	**0.0282**	0.2632	**0.2528**	**0.2566**	0.0332
2000	0.2027	0.2614	0.2628	**0.0217**	0.2269	**0.2538**	**0.2569**	0.0244
5000	0.1602	0.2588	0.2601	**0.0163**	0.1832	**0.2539**	**0.2562**	0.0177
	Burr with $\xi = 0.25$, $\rho = -0.75$							
100	0.1756	0.2963	0.3020	0.1390	0.1676	**0.2611**	**0.2737**	**0.0969**
200	0.1503	0.2913	0.2988	0.0931	0.1389	**0.2614**	**0.2702**	**0.0761**
500	0.1202	0.2885	0.2927	0.0612	0.1133	**0.2614**	**0.2686**	**0.0590**
750	0.1021	0.2836	0.2882	0.0525	0.0979	**0.2602**	**0.2660**	**0.0500**
1000	0.0974	0.2833	0.2854	0.0503	0.0912	**0.2604**	**0.2654**	**0.0464**
2000	0.0768	0.2780	0.2813	0.0391	0.0733	**0.2595**	**0.2639**	**0.0372**
5000	0.0550	0.2703	0.2724	**0.0292**	0.0570	**0.2577**	**0.2623**	0.0306
	Half-t_4 ($\xi = 0.25$, $\rho = -0.5$)							
100	0.0986	0.3492	0.3520	0.2877	0.0951	**0.2922**	**0.3025**	**0.1318**
200	0.0843	0.3391	0.3463	0.1907	0.0761	**0.2887**	**0.3012**	**0.1105**
500	0.0628	0.3382	0.3371	0.2858	0.0562	**0.2862**	**0.2960**	**0.0874**
750	0.0550	0.3279	0.3361	0.1037	0.0492	**0.2849**	**0.2934**	**0.0794**
1000	0.0500	0.3243	0.3300	0.0973	0.0436	**0.2807**	**0.2906**	**0.0736**
2000	0.0392	0.3133	0.3179	0.0799	0.0340	**0.2755**	**0.2839**	**0.0618**
5000	0.0268	0.2993	0.3030	0.0604	0.0244	**0.2715**	**0.2786**	**0.0498**
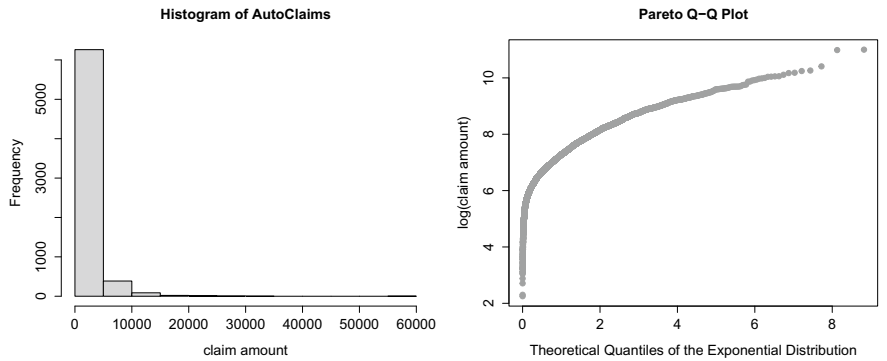
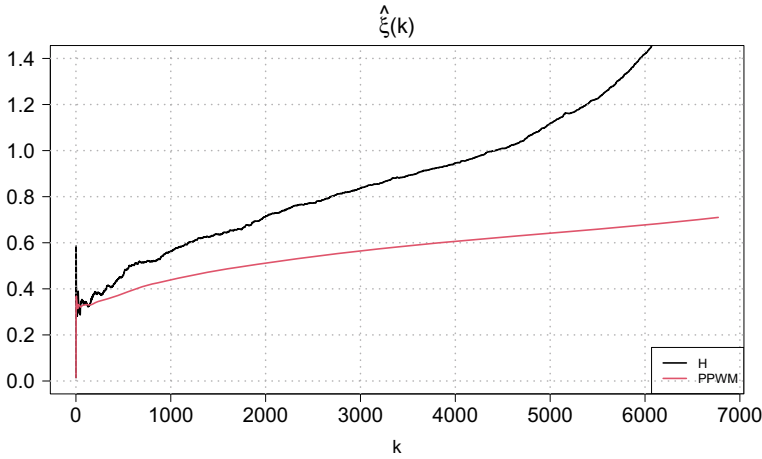**Fig. 2** Histogram and Pareto QQ plot for the AutoClaims dataset

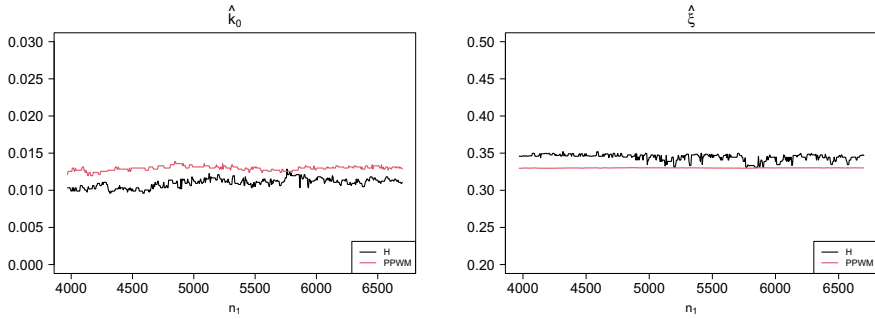**Fig. 3** Estimates of the EVI for the AutoClaims dataset

Fig. 4 Estimates of the OSF's $\hat{k}_0^{\bullet,*}/n$ (*left*) and the bootstrap adaptive extreme value index estimates $\hat{\xi}^{\bullet,*}$ (*right*), as functions of the sub-sample size n_1, for the amount paid on a closed claim

In Fig. 4, as a function of the sub-sample size n_1, ranging from $n_1 = 3990$ until $n_1 = 6700$, we picture, at the left, the estimates $\hat{k}_0^{\bullet,*}(n_1)/n$ of the optimal sample fraction, $k_0^{\bullet}/n$, for the adaptive double bootstrap estimation of ξ through the H and the PPWM estimators. Associated bootstrap EVI-estimates are pictured at the right. Contrarily to the bootstrap Hill, the bootstrap PPWM EVI-estimates are quite stable as a function of the sub-sample size n_1 (see Fig. 4, right).

For a re-sample size $n_1 = [n^{0.955}] = 4554$, and $B = 250$ bootstrap generations, we were led to $\hat{k}_0^{H,*} = 67$ and to $\hat{\xi}^{H,*} = 0.3463$. This same algorithm applied to the PPWM estimator provide the bootstrap estimates $\hat{k}_0^{PPWM,*} = 88$ and $\hat{\xi}^{PPWM,*} = 0.3301$.

5   Conclusions

In this paper we addressed the adaptive estimation of the EVI with the double boot-strap methodology associated to the Hill and the PPWM estimators. The presented simulation study shows that the adaptive PPWM EVI-estimator is usually less biased and provides a similar or a smaller RMSE than the adaptive Hill EVI-estimator. Moreover, the efficiency of the adaptive PPWM estimator relatively to the adaptive Hill estimator seems to improve as the asymptotic bias of estimators increases (as ρ increases). Further research concerning the sensitivity of the method on the choice of n_1 will be addressed in the future.

References

1. Jenkinson, A.F.: The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Q. J. R. Meteorol. Soc. **81**, 158–171 (1955)
2. Von Mises, R.: La distribution de la plus grande de *n* valuers. Rev. Math. Union Interbalcanique **1**, 141–160 (1936)
3. Hill, B.M.: A simple general approach to inference about the tail of a distribution. Ann. Stat. **3**(5), 1163–1174 (1975)
4. Caeiro, F., Gomes, M.I.: Semi-parametric tail inference through probability-weighted moments. J. Stat. Plann. Inference **141**(2), 937–950 (2011)
5. Beirlant, J., Herrmann, K., Teugels, J.: Estimation of the extreme value index. In: Extreme Events in Finance, pp. 97–115. Wiley (2016)
6. Fedotenkov, I.: A review of more than one hundred Pareto-tail index estimators. Statistica (Bologna) **80**(3), 245–299 (2020)
7. Beirlant, J., Caeiro, F., Gomes, M.I.: An overview and open research topics in statistics of univariate extremes. Revstat-Stat. J. **10**(1), 1–31 (2012)
8. Gomes, M.I., Oliveira, O.: The bootstrap methodology in statistics of extremes–choice of the optimal sample fraction. Extremes **4**(4), 331–358 (2001)
9. Gomes, M.I., Mendonça, S., Pestana, D.: Adaptive reduced-bias tail index and var estimation via the bootstrap methodology. Commun. Stat.-Theory and Methods **40**(16), 2946–2968 (2011)
10. Caeiro, F., Gomes, M.I., Vandewalle, B.: Semi-parametric probability-weighted moments estimation revisited. Methodol. Comput. Appl. Probab. **16**(1), 1–29 (2014)
11. Hall, P.: Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. J. Multivar. Anal. **32**(2), 177–203 (1990)
12. Fraga Alves, M.I., Gomes, M.I., De Haan, L.: A new class of semi-parametric estimators of the second order parameter. Portugaliae Math. **60**(2), 193–213 (2003)
13. Gomes, M.I., Martins, M.J.: Asymptotically unbiased estimators of the tail index based on external estimation of the second order parameter. Extremes **5**(1), 5–31 (2002)
14. Caeiro, F., Gomes, M.: Threshold selection in extreme value analysis. In: Extreme Value Modeling and Risk Analysis. Chapman and Hall/CRC, Boca Raton, FL
15. Wolny-Dominiak, A., Trzesiok, M.: `insuranceData`: A collection of insurance datasets useful in risk classification in non-life insurance (2014). R package version 1.0