MEAN-OF-ORDER-p LOCATION-INVARIANT EXTREME VALUE INDEX ESTIMATION

M. Ivette Gomes

CEAUL and DEIO, FCUL, Universidade de Lisboa, Portugal, e-mail: ivette.gomes@fc.ul.pt

Lígia Henriques-Rodrigues

Universidade de São Paulo, IME, São Paulo, Brasil, and CEAUL, Portugal, e-mail: ligiahr@ime.usp.br

B.G. Manjunath

CEAUL, Universidade de Lisboa, Portugal, e-mail: bgmanjunath@gmail.com

January 5, 2015

Abstract

A simple generalisation of the classical Hill estimator of a positive extreme value index (EVI) has been recently introduced in the literature. Indeed, the Hill estimator can be regarded as the logarithm of the mean of order $p = 0$ of a certain set of statistics. Instead of such a geometric mean, we can more generally consider the mean of order p (MOP) of those statistics, with p real, and even an optimal MOP (OMOP) class of EVI-estimators. These estimators are scale invariant but not location invariant. With PORT standing for peaks over random threshold, new classes of PORT-MOP and PORT-OMOP EVI-estimators are now introduced. These classes are dependent on an extra tuning parameter q, $0 \leq q < 1$, and they are both location and scale invariant, a property also played by the EVI. The asymptotic normal behaviour of those PORT classes is derived. These EVI-estimators are further studied for finite samples, through a Monte-Carlo simulation study. An adequate choice of the *tuning* parameters under play is put forward, and some concluding remarks are provided.

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

Keywords and phrases. Bootstrap and/or heuristic threshold selection, Heavy tails, Location/scale invariant semi-parametric estimation, Monte-Carlo simulation, Optimal levels, Statistics of extremes.

1 Introduction

Given a sample of size n of *independent, identically distributed* (IID) random variables (RVs), $\underline{\mathbf{X}}_n := (X_1, \ldots, X_n)$, with a common *cumulative distribution function* (CDF) F, let us denote by $X_{1:n} \leq \cdots \leq X_{n:n}$ the associated ascending order statistics. As usual in a framework of *extreme value theory* (EVT), let us further assume that there exist sequences of real constants $\{a_n > 0\}$ and $\{b_n \in \mathbb{R}\}$ such that the maximum, linearly normalised, i.e. $(X_{n:n} - b_n)/a_n$, converges in distribution to a non-degenerate RV. Then, the limit distribution is necessarily of the type of the general *extreme value* (EV) CDF, given by

$$\mathrm{EV}_\xi(x) = \begin{cases} \exp(-(1+\xi x)^{-1/\xi}), \ 1+\xi x > 0, & \text{if} \quad \xi \neq 0, \\ \exp(-\exp(-x)), \ x \in \mathbb{R}, & \text{if} \quad \xi = 0. \end{cases} \tag{1.1}$$

The CDF F is said to belong to the max-domain of attraction of EV_ξ, and we consider the common notation $F \in \mathcal{D}_\mathcal{M}(\mathrm{EV}_\xi)$. The parameter ξ is the *extreme value index* (EVI), the primary parameter of extreme events.

The EVI measures the heaviness of the survival function or right tail-function

$$\overline{F}(x) := 1 - F(x), \tag{1.2}$$

and the heavier the right tail, the larger ξ is. Let us further use the notation $\mathcal{R}_a$ for the class of regularly varying functions at infinity, with an index of regular variation equal to $a \in \mathbb{R}$, i.e. positive measurable functions $g(\cdot)$ such that for all $x > 0$, $g(tx)/g(t) \to x^a$, as $t \to \infty$ (see Bingham *et al.*, 1987, among others, for details on the theory of regular variation). In this paper we work with Pareto-type underlying models, i.e. with a positive EVI, a quite common assumption in many areas of application, like bibliometrics, biostatistics, computer science, insurance, finance, social sciences and telecommunications, among others. The right-tail function $\overline{F}$, in (1.2), belongs then to $\mathcal{R}_{-1/\xi}$. Indeed, and more generally,

$$\overline{F} \in \mathcal{D}_\mathcal{M}(\mathrm{EV}_{\xi>0}) =: {\mathcal{D}_\mathcal{M}}^+ \qquad \Longleftrightarrow \qquad \overline{F} \in \mathcal{R}_{-1/\xi}, \tag{1.3}$$

a result due to Gnedenko (1943).

For the class of Pareto-type models in (1.3), the most well-known EVI-estimators are the Hill (H) estimators (Hill, 1975), which are the averages of the log-excesses,

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \le i \le k < n.$$

We can thus define the H-class of EVI-estimators as:

$$\mathrm{H}(k) := \mathrm{H}(k; \underline{\mathbf{X}}_n) := \frac{1}{k} \sum_{i=1}^{k} V_{ik}, \quad 1 \le k < n. \tag{1.4}$$

We can further write

$$\mathrm{H}(k) = \sum_{i=1}^{k} \ln \left(\frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^{1/k} = \ln \left(\prod_{i=1}^{k} \frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^{1/k}, \quad 1 \le k < n.$$

The Hill estimator is thus the logarithm of the *geometric mean* (or *mean of order* 0) of

$$U_{ik} := X_{n-i+1:n}/X_{n-k:n}, \quad 1 \le i \le k < n.$$

Brilhante *et al.* (2013) considered as basic statistics, the *mean of order p* (MOP) of U_{ik}, $1 \le i \le k$, for $p \ge 0$. More generally, Gomes and Caeiro (2014) considered those same statistics for any $p \in \mathbb{R}$, i.e. the class of statistics

$$\mathrm{M}_p(k) = \begin{cases} \left(\frac{1}{k} \sum_{i=1}^{k} U_{ik}^p \right)^{1/p}, & \text{if} \quad p \neq 0, \\ \\ \left(\prod_{i=1}^{k} U_{ik} \right)^{1/k}, & \text{if} \quad p = 0, \end{cases}$$

and the following associated class of MOP EVI-estimators:

$$\mathrm{H}_p(k) = \mathrm{H}_p(k; \underline{\mathbf{X}}_n) \equiv \mathrm{MOP}(k) := \begin{cases} \Big(1 - \mathrm{M}_p^{-p}(k) \Big)/p, & \text{if} \quad p < 1/\xi, \\ \\ \ln \mathrm{M}_0(k) = \mathrm{H}(k), & \text{if} \quad p = 0, \end{cases} \tag{1.5}$$

with $\mathrm{H}_0(k) \equiv \mathrm{H}(k)$, given in (1.4). This class of MOP EVI-estimators depends now on this *tuning* parameter $p \in \mathbb{R}$, it is highly flexible, but, as often desired, it is not location-invariant, depending strongly on possible shifts in the model underlying the data. To make the EVI-estimators $\mathrm{H}_p(k)$, in (1.5), location-invariant, it is thus sensible to use the *peaks over a random threshold* (PORT) technique now applied to the MOP EVI-estimation. The PORT

methodology, introduced in Araújo Santos *et al.* (2006) and further studied in Gomes *et al.* (2008a), is based on a *sample of excesses* over a random threshold $X_{n_q:n}$, $n_q := \lfloor nq \rfloor + 1$, where $\lfloor x \rfloor$ denotes the integer part of x, i.e. it is based on the sample of size $n^{(q)} = n - n_q$, defined by

$$\underline{\mathbf{X}}_n^{(q)} := \big(X_{n:n} - X_{n_q:n}, \ldots, X_{n_q+1:n} - X_{n_q:n}\big). \tag{1.6}$$

After the introduction, in Section 2, of a few technical details in the field of EVT and a brief reference to the most simple *minimum-variance reduced-bias* (MVRB) EVI-estimators, the *corrected-Hill* (CH) EVI-estimators introduced and studied in Caeiro *et al.* (2005), we refer a class of *optimal* MOP (OMOP) EVI-estimators recently studied in Brilhante *et al.* (2014). We further introduce the new classes of PORT-MOP and PORT-OMOP EVI-estimators. Section 3 is essentially dedicated to consistency and asymptotic normal behaviour of these new classes of EVI-estimators, with a brief reference to the known asymptotic behaviour of the CH and MOP EVI-estimators. Section 4 is dedicated to the finite sample properties of the new classes of estimators, comparatively to the behaviour of the aforementioned MVRB and even PORT-MVRB EVI-estimators, done through a small-scale simulation study. In Section 5, we refer possible methods for the adaptive choice of the tuning parameters (k, p, q), either based on the bootstrap or on heuristic methodologies, and provide some concluding remarks.

2 Preliminary results in the area of EVT

n the area of EVT and whenever working with large values, i.e. with the right tail of the model F underlying the available sample, the model F is usually said to be *heavy-tailed* whenever (1.3) holds. Moreover, with the notation $F^{\leftarrow}(t) := \inf\{x : F(x) \geq t\}$ for the generalised inverse function of F, the condition $F \in \mathcal{D}_{\mathcal{M}}^{+}$ is equivalent to say that the tail quantile function $U(t) := F^{\leftarrow}(1 - 1/t)$ is of regular variation with index ξ (de Haan, 1984). We thus assume the validity of any of the following first-order conditions:

$$F \in \mathcal{D}_{\mathcal{M}}^{+} \qquad \Longleftrightarrow \qquad \overline{F} \in \mathcal{R}_{-1/\xi} \qquad \Longleftrightarrow \qquad U \in \mathcal{R}_{\xi}. \tag{2.1}$$

The *second-order parameter* ρ (≤ 0) rules the rate of convergence in the first-order condition, in (2.1), and can be defined as the non-positive parameter appearing in the limiting relation

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \psi_\rho(x) := \begin{cases} \frac{x^\rho - 1}{\rho}, & \text{if} \quad \rho < 0, \\ \\ \ln x, & \text{if} \quad \rho = 0, \end{cases} \tag{2.2}$$

which is assumed to hold for every $x > 0$, and where $|A|$ must then be of regular variation with index ρ (Geluk and de Haan, 1987). For related details on the topic, see Beirlant *et al.* (2004) and de Haan and Ferreira (2006).

Whenever dealing with bias reduction in the field of extremes, it is usual to consider a slightly more restrict class than $\mathcal{D}^+_{\mathcal{M}}$, the class of models

$$U(t) = C\, t^\xi \big\{1 + A(t)/\rho + o(t^\rho)\big\}, \qquad A(t) = \xi\beta t^\rho, \tag{2.3}$$

as $t \to \infty$, where $C > 0$, $\xi > 0$, $\rho < 0$ and $\beta \neq 0$ (Hall and Welsh, 1985). This means that the slowly varying function $L(t)$ in $U(t) = t^\xi L(t)$ is assumed to behave asymptotically as a constant. To assume (2.3) is equivalent to choose $A(t) = \xi\beta t^\rho$, $\rho < 0$, in the more general second-order condition in (2.2). Models like the log-Gamma and the log-Pareto ($\rho = 0$) are thus excluded from this class. The standard Pareto ($\rho = -\infty$) is also excluded. But most heavy-tailed models used in applications, like the EV_ξ, in (1.1), the Fréchet, $F(x) = \exp(-x^{-1/\xi})$, $x \geq 0$, both for $\xi > 0$, and the well-known Student's t CDFs, among others, belong to Hall-Welsh class.

2.1 The CH class of EVI-estimators

Due to its simplicity and just as mentioned above, the most popular EVI-estimators, consistent only for non-negative values of ξ, are Hill estimators in (1.4). We further consider the simplest class of CH EVI-estimators, the one introduced in Caeiro *et al.* (2005),

$$\text{CH}(k) = \text{CH}(k; \underline{\mathbf{X}}_n) := \text{H}(k) \left(1 - \frac{\hat\beta (n/k)^{\hat\rho}}{1 - \hat\rho}\right). \tag{2.4}$$

The estimators in (2.4) can be second-order MVRB EVI-estimators, for adequate levels k and an adequate external estimation of the vector of second-order parameters, (β, ρ), in (2.3), algorithmically given in Gomes and Pestana (2007), among others, i.e. the use of $\text{CH}(k)$, and

an adequate estimation of (β, ρ), enables us to eliminate the dominant component of the bias of the Hill estimator, $\mathrm{H}(k)$, keeping its asymptotic variance. Like that, and theoretically, $\mathrm{CH}(k)$ outperforms $\mathrm{H}(k)$ for all k.

We again suggest the use of the class of β-estimators in Gomes and Martins (2002) and the simplest class of ρ-estimators in Fraga Alves *et al.* (2003). In the simulations, we have considered only models with $|\rho| \leq 1$. Indeed, this is the case where alternatives to the H-class of EVI-estimators are welcome due to the high bias of H EVI-estimators for moderate up to large values of k, including the optimal k in the sense of minimal *root mean square error* (RMSE). In such cases, we suggest the use of the *tuning* parameter $\tau = 0$ in the simplest class of ρ-estimators in Fraga Alves *et al.* (2003), given by

$$\hat{\rho}_\tau(k) \equiv \hat{\rho}_\tau(k; \underline{\mathbf{X}}_n) := \min\left(0,\ \frac{3(R_n^{(\tau)}(k; \underline{\mathbf{X}}_n) - 1)}{R_n^{(\tau)}(k; \underline{\mathbf{X}}_n) - 3}\right), \tag{2.5}$$

and dependent on the statistics

$$R_n^{(\tau)}(k; \underline{\mathbf{X}}_n) := \frac{\left(M_n^{(1)}(k; \underline{\mathbf{X}}_n)\right)^\tau - \left(M_n^{(2)}(k; \underline{\mathbf{X}}_n)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k; \underline{\mathbf{X}}_n)/2\right)^{\tau/2} - \left(M_n^{(3)}(k; \underline{\mathbf{X}}_n)/6\right)^{\tau/3}}, \quad \tau \in \mathbb{R},$$

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

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

As already suggested in previous papers, we have here decided for the computation of $\hat{\rho}_\tau(k)$ at $k = k_1$, given by $k_1 = \lfloor n^{1-\epsilon} \rfloor$, $\epsilon = 0.001$, the threshold used in Caeiro *et al.* (2005) and Gomes and Pestana (2007).

For the estimation of the scale second-order parameter β, in (2.3), and again on the basis of a sample $\underline{\mathbf{X}}_n$, we consider

$$\hat{\beta}_{\hat{\rho}}(k) \equiv \hat{\beta}_{\hat{\rho}}(k; \underline{\mathbf{X}}_n) := \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{d_{\hat{\rho}}(k)\ D_0(k) - D_{\hat{\rho}}(k)}{d_{\hat{\rho}}(k)\ D_{\hat{\rho}}(k) - D_{2\hat{\rho}}(k)}, \tag{2.6}$$

dependent on the estimator $\hat{\rho} = \hat{\rho}_0(k_1; \underline{\mathbf{X}}_n)$, with $\hat{\rho}_\tau(k)$ defined in (2.5), and where, for any $\alpha \leq 0$,

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

with $U_i,\ 1 \leq i \leq k$, the *scaled log-spacings* associated with $\underline{\mathbf{X}}_n$. Details on the distributional behaviour of the estimator in (2.6) can be found in Gomes and Martins (2002) and more

recently in Gomes *et al.* (2008b) and Caeiro *et al.* (2009). Interesting alternative classes of estimators of the 'shape' and 'scale' second-order parameters have recently been introduced. References to those classes can be found in recent overviews on reduced-bias estimation (Chapter 6 of Reiss and Thomas, 2007; Beirlant *et al.*, 2012; Gomes and Guillou, 2014).

2.2 The OMOP class of EVI-estimators

Working in the class of models in (2.3) for technical simplicity, Brilhante *et al.* (2014) noticed that there is an optimal value $p \equiv p_{\text{M}} = \varphi_\rho/\xi$, with

$$\varphi_\rho = 1 - \rho/2 - \sqrt{\rho^2 - 4\rho + 2}\Big/2 \in \big(0, 1 - \sqrt{2}/2\big), \tag{2.7}$$

which maximises the asymptotic efficiency of the class of estimators in (1.5). They then considered the MOP EVI-estimator associated with the optimal $p \equiv p_{\text{M}}$ estimated through $\hat{p}_{\text{M}}$, based on any initial consistent estimator of ξ and ρ, i.e. an *optimal* MOP (OMOP) class of EVI-estimators. Here, we estimate the optimal k-value for the H EVI-estimation, $k_{0|0} := \arg\min_k \text{RMSE}\big(\text{H}_0(k)\big)$, computing, as given in Hall (1982),

$$\hat{k}_{0|0} \equiv \hat{k}_{0|\text{H}_0} = \Big((1-\hat{\rho})n^{-\hat{\rho}}/\big(\hat{\beta}\,\sqrt{-2\hat{\rho}}\big)\Big)^{2/(1-2\hat{\rho})},$$

the associated observed value of the EVI-estimator $\text{H}_{00} := \text{H}(\hat{k}_{0|0})$, and, with φ_ρ given in (2.7), the OMOP EVI-estimators

$$\text{H}^*(k) \equiv \text{H}^*(k; \underline{\mathbf{X}}_n) := \text{H}_{\hat{p}_{\text{M}}}(k; \underline{\mathbf{X}}_n), \quad 1 \le k < n, \quad \hat{p}_{\text{M}} = \varphi_{\hat{\rho}}/\text{H}_{00}. \tag{2.8}$$

Neither the H nor the CH nor the MOP EVI–estimators are invariant for changes in location, but they can easily be made location-invariant with the technique introduced in Araújo Santos *et al.* (2006), briefly discribed in the following Section.

2.3 The PORT methodology

The EVI-estimators in (1.4), (1.5), (2.4) and (2.8) are scale-invariant, but not location-invariant, as often desired, due to the fact that the EVI itself enjoys such a property, i.e. it is location and scale invariant. Indeed, note that a general first-order condition to have $F \in \mathcal{D}_{\mathcal{M}}\,(\text{EV}_\xi)$, given in de Haan (1984), can be written as

$$F \in \mathcal{D}_{\mathcal{M}}\,(\text{EV}_\xi) \iff \lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \psi_\xi(x), \tag{2.9}$$

for an adequate function $a(\cdot)$, with an absolute value necessarily in $\mathcal{R}_\xi$, and where $\psi_\rho(\cdot)$ is the Box-Cox function, already defined in (2.2). If a shift s is induced in data associated with the RV X, i.e. if we consider $Y = X + s$, the relationship between the tail quantile functions of Y and X is given by $U_{_Y}(t) = s + U_{_X}(t)$. Consequently, $U_{_Y}(tx) - U_{_Y}(t) = U_{_X}(tx) - U_{_X}(t)$ and from (2.9), the EVI, ξ, is the same for X and $Y = X + s$, for any shift $s \in \mathbb{R}$.

Just as mentioned above, the class of PORT-Hill estimators is based on the *sample of excesses* in (1.6). In this article, we shall work with PORT-MOP and PORT-OMOP EVI-estimators, generally denoted E. They have the same functional form of the associated EVI-estimators in (1.5) and (2.8) but with the original sample $\underline{\mathbf{X}}_n$ replaced everywhere by the sample of excesses $\underline{\mathbf{X}}_n^{(q)}$, in (1.6). Consequently, they are given by the functional equations,

$$\mathrm{E}^{(q)}(k) := \mathrm{E}\big(k; \underline{\mathbf{X}}_n^{(q)}\big), \qquad \text{with } \mathrm{E} \equiv \mathrm{H}_p \text{ and } \mathrm{E} \equiv \mathrm{H}^*. \tag{2.10}$$

These estimators are now invariant for both changes of location and scale, and depend on the extra *tuning parameter* q, which only influences the asymptotic bias, making them highly flexible and even able to compare favourably with the MVRB EVI-estimators in (2.4), for a large variety of underlying models in the domain of attraction for maxima of the EV_ξ CDF, in (1.1). In the simulation procedure, we further include the PORT-MVRB EVI-estimators,

$$\mathrm{CH}^{(q)}(k) = \mathrm{CH}\big(k; \underline{\mathbf{X}}_n^{(q)}\big), \tag{2.11}$$

studied by simulation in Gomes *et al.* (2011a, 2013), with $\underline{\mathbf{X}}_n^{(q)}$ and $\mathrm{CH}\big(k; \underline{\mathbf{X}}_n\big)$ respectively given in (1.6) and (2.4).

3 Asymptotic behaviourof EVI-estimators

Consistency of the Hill EVI-estimators, $\mathrm{H} \equiv \mathrm{H}_0$, written both in (1.4) and (1.5), is achieved in the whole $\mathcal{D}_{\mathcal{M}}^+$ whenever we work with intermediate values of k, i.e.

$$k = k_n \to \infty,\ 1 \le k < n, \quad \text{and} \quad k_n = o(n),\ \text{as } n \to \infty. \tag{3.1}$$

3.1 Asymptotic normal behaviour of MOP and OMOP EVI-estimators

Let us consider the notation $\mathcal{N}(\mu, \sigma^2)$ for a normal RV with mean value μ and variance σ^2. Under the aforementioned second-order framework, in (2.2), and as a generalization

of the results in de Haan and Peng (1998), Brilhante *et al.* (2013) derived, for the MOP EVI-estimators in (1.5) and $0 \le p \le 1/(2\xi)$, the asymptotic distributional representation,

$$\sqrt{k}\Big(\mathrm{H}_p(k) - \xi\Big) \stackrel{d}{=} \mathcal{N}\left(0, \tfrac{\xi^2(1-p\xi)^2}{1-2p\xi}\right) + \frac{(1-p\xi)\sqrt{k}A(n/k)}{1-\rho-p\xi}\big(1+o_p(1)\big),$$

more generally valid for $p \in \mathbb{R}$ (Gomes and Caeiro, 2014). For the OMOP EVI-estimators, in (2.8), Brilhante *et al.* (2014) got the obvious validity of a similar asymptotic distributional representation, but with $p\xi$ replaced by φ_ρ, in (2.7), i.e.

$$\sqrt{k}\Big(\mathrm{H}^*(k) - \xi\Big) \stackrel{d}{=} \mathcal{N}\left(0, \tfrac{\xi^2(1-\varphi_\rho)^2}{1-2\varphi_\rho}\right) + \frac{(1-\varphi_\rho)\,\sqrt{k}A(n/k)}{1-\varphi_\rho-\rho}\big(1+o_p(1)\big).$$

The asymptotic variance increases when p moves away from $p = 0$, but the bias decreases and, at optimal levels in the sense of minimal RMSE, the OMOP EVI-estimators outperform the H EVI-estimators.

Under the same conditions as before, but with $\mathrm{CH}(k)$ given in (2.4) and assuming that (2.3) holds, an adequate estimation of the second-order parameters, (β, ρ), enables to guarantee that $\sqrt{k}\big(\mathrm{CH}(k) - \xi\big)$ can be asymptotically normal with variance also equal to ξ^2 but with a null mean value. Indeed, from the results in Caeiro *et al.* (2005), we know that it is possible to get

$$\sqrt{k}\Big(\mathrm{CH}(k) - \xi\Big) \stackrel{d}{=} \mathcal{N}\left(0, \xi^2\right) + o_p\Big(\sqrt{k}A(n/k)\Big).$$

On the basis of the results in the aforementioned papers, and generally denoting by $\mathrm{E}(k)$ any of the EVI-estimators in (1.5) and (2.8), we can state the following theorem.

Theorem 1. (de Haan and Peng, 1998; Caeiro *et al.*, 2005; Brilhante *et al.*, 2013, 2014) *Under the validity of the first-order condition, in* (2.1)*, and for intermediate sequences* $k = k_n$*, i.e. if* (3.1) *holds, the classes of* EVI*-estimators* $\mathrm{H}_p(k)$*, in* (1.5)*, for* $p < 1/\xi$*, and the* EVI*-estimators in* (2.4) *and* (2.8) *are consistent for the estimation of* ξ*. If we assume the validity of the second-order condition in* (2.2) *and additionally assume that we are working with values of* k *such that* $\lambda_{_A} := \lim_{n\to\infty} \sqrt{k}\, A(n/k)$ *is finite, we can then guarantee that for* $p < 1/(2\xi)$ *whenever dealing with* $\mathrm{H}_p(k)$*,*

$$\sqrt{k}\,(\mathrm{E}(k) - \xi) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(\lambda_{_A} b_\bullet, \sigma^2_\bullet\right),$$

where

$$b_{_{\mathrm{H}_p}} = \frac{1-p\xi}{1-\rho-p\xi}, \quad b_{_{\mathrm{H}^*}} = \frac{1-\varphi_\rho}{1-\rho-\varphi_\rho}, \quad \sigma^2_{_{\mathrm{H}_p}} = \frac{\xi^2(1-p\xi)^2}{1-2p\xi}, \quad \sigma^2_{_{\mathrm{H}^*}} = \frac{\xi^2(1-\varphi_\rho)^2}{1-2\varphi_\rho}.$$

If we further assume to be working in Hall-Welsh *class of models in* (2.3)*, and estimate* β *and* ρ *consistently through* $\hat{\beta}$ *and* $\hat{\rho}$*, with* $\hat{\rho} - \rho = o_p(1/\ln n)$*, we get the aforementioned normal behaviour also for* $\mathrm{E} = \mathrm{CH}$*, in* (2.4)*, but now with* $b_{\scriptscriptstyle \mathrm{CH}} = 0$ *and* $\sigma^2_{\scriptscriptstyle \mathrm{CH}} = \sigma^2_{\scriptscriptstyle \mathrm{H}} = \xi^2$.

Remark 1. *Note again that* $\sigma^2_{\scriptscriptstyle \mathrm{H}} < \sigma^2_{\scriptscriptstyle \mathrm{H}_p}$ *for all* $\xi > 0$ *and* $0 \neq p < 1/\xi$*. The other way round,* $b_{\scriptscriptstyle \mathrm{H}} \geq b_{\scriptscriptstyle \mathrm{H}_p}$ *for all* ξ*. And as can be seen in* Brilhante *et al.* (2013; 2014)*, at the optimal* p*,* $\mathrm{H}_p(k)$ *can asymptotically outperform* $\mathrm{H}(k)$ *at optimal levels in the sense of minimal* RMSE*, in the whole* (ξ, ρ)*-plane. As far as we know, such a property is so far achieved only by this class of EVI-estimators. Se also* Paulauskas and Vaiciulis (2013).

3.2 Asymptotic behaviour of PORT-MOP EVI-estimators

Note first that if there is a possible shift s in the model, i.e. if the CDF $F(x) \equiv F_s(x) = F(x; s)$ depends on (x, s) through the difference $x - s$, the parameter ξ does not change, as mentioned above in Section 2.3, but the parameter ρ, as well as the A-function, in (2.2), depend on such a shift s, i.e. $\rho = \rho_s$, $A = A_s$, and

$$(A_s(t), \rho_s) := \begin{cases} \big(-\xi s/U_0(t), -\xi \big), & \text{if} \quad \xi + \rho_0 < 0 \ \wedge \ s \neq 0, \\ \big(A_0(t) - \xi s/U_0(t), \rho_0 \big), & \text{if} \quad \xi + \rho_0 = 0 \ \wedge \ s \neq 0, \\ \big(A_0(t), \rho_0 \big), & \text{otherwise.} \end{cases}$$

Further details on the influence of such a shift in $\big(\beta, \rho, A(\cdot)\big)$ and on the estimation of 'shape' and 'scale' second-order parameters can be found in Henriques-Rodrigues *et al.* (2014, 2015).

To study the asymptotic properties of the PORT-MOP (and PORT-OMOP) EVI-estimators for $p \neq 0$, it is convenient to study first the behaviour of the statistics,

$$W_p(k; q) := \frac{1}{k} \sum_{i=1}^{k} \left(\frac{X_{n-i+1:n} - X_{n_q:n}}{X_{n-k:n} - X_{n_q:n}} \right)^p, \qquad p \neq 0, \tag{3.2}$$

for $X = X_0 \frown F_0$. Indeed,

$$\mathrm{H}_p\Big(k; \underline{\mathbf{X}}^{(q)}_n\Big) = \frac{1 - W_p^{-1}(k; q)}{p} \quad \text{if} \quad p \neq 0. \tag{3.3}$$

Remark 2. *Note that with*

$$Q_r(k; q) = \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k} \right)^r \frac{X_{n-i+1:n} - X_{n_q:n}}{X_{n-k:n} - X_{n_q:n}},$$

the statistics studied in Caeiro *et al.* (2014)*, we get, with* $W_p(k;q)$ *given in* (3.2)*,* $W_1(k;q) = Q_0(k;q)$*.*

Remark 3. *It is also worth noting that, as already detected in* Fraga Alves *et al.* (2009)*, for invariant versions of the mixed moment, and in* Caeiro *et al.* (2014)*, for invariant versions of the* Pareto *probability weighted moment* EVI*-estimators, due to the fact that*

$$X_{\lfloor nq \rfloor + 1:n} - U_0(1/(1-q)) = O_p\Big(1/\sqrt{n}\Big),$$

$X_{n_q:n}$ *can be replaced by the* q*-quantile*

$$\chi_q := U_0(1/(1-q)). \tag{3.4}$$

The asymptotic behaviour of the statistics $W_p(k;q)$, in (3.2), comes then straightforwardly from the behaviour of the non-shifted statistics, as stated in the following proposition.

Theorem 2. *Under the second order framework in* (2.2)*, and for intermediate* k*, i.e. whenever* (3.1) *holds, we can guarantee, under general broad conditions, the asymptotic normality of* $W_p(k;q)$*, in* (3.2)*. Indeed, we can write, for* $p\xi < 1/2$*,*

$$W_p(k;q) \stackrel{d}{=} \frac{1}{1-p\xi} + \frac{\sigma_p(\xi)\mathcal{N}(0,1)}{\sqrt{k}} + \frac{pA_0(n/k)(1+o_p(1))}{(1-p\xi)(1-p\xi-\rho_0)} + \frac{p\xi\chi_q(1+o_p(1))}{(1-p\xi)(1-(p-1)\xi)U_0(n/k)}, \tag{3.5}$$

where

$$\sigma_p^2(\xi) := \frac{(p\xi)^2}{(1-p\xi)^2(1-2p\xi)}. \tag{3.6}$$

Proof. It is well-known that $U_0(X_{i:n}) \stackrel{d}{=} Y_{i:n}$, where Y is a standard unit Pareto RV, with CDF $F_Y(y) = 1 - 1/y$, $y > 1$. Moreover, $Y_{n-i+1:n}/Y_{n-k:n} \stackrel{d}{=} Y_{k-i+1:k}$, $1 \le i \le k$. Under the second order framework in (2.2), and thinking on the fact that we are now working with $s = 0$ due to the location invariance property of the statistics in (3.2), we can write

$$\frac{X_{n-i+1:n}}{X_{n-k:n}} \stackrel{d}{=} \frac{U_0\big(\frac{Y_{n-i+1:n}}{Y_{n-k:n}} Y_{n-k:n}\big)}{U_0(Y_{n-k:n})} \stackrel{d}{=} Y_{k-i+1:k}^{\xi}\left(1 + \tfrac{Y_{k-i+1:k}^{\rho}-1}{\rho} A_0(Y_{n-k:n})(1+o_p(1))\right).$$

Next, with the notation $\chi_q = U_0(1/(1-q))$, already introduced in (3.4),

$$\frac{X_{n-i+1:n}-\chi_q}{X_{n-k:n}-\chi_q} = \frac{X_{n-i+1:n}}{X_{n-k:n}}\left(\frac{1-\chi_q/X_{n-i+1:n}}{1-\chi_q/X_{n-k:n}}\right) = \frac{X_{n-i+1:n}}{X_{n-k:n}}\left(1 + \frac{\chi_q}{X_{n-k:n}}\left(1 - \frac{X_{n-k:n}}{X_{n-i+1:n}}\right)(1+o_p(1))\right).$$

Consequently,

$$W_p(k;q) := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{X_{n-i+1:n}-X_{n_q:n}}{X_{n-k:n}-X_{n_q:n}}\right)^p$$
$$= \frac{1}{k}\sum_{i=1}^{k}\left(\frac{X_{n-i+1:n}}{X_{n-k:n}}\left(1+\frac{\chi_q}{X_{n-k:n}}\left(1-\frac{X_{n-k:n}}{X_{n-i+1:n}}\right)(1+o_p(1))\right)\right)^p,$$

and we can write

$$W_p(k;q) \stackrel{d}{=} \frac{1}{k}\sum_{i=1}^{k} Y_{i:k}^{p\xi} + \frac{p\xi\chi_q}{U_0(n/k)}\frac{1}{k}\sum_{i=1}^{k} Y_{i:k}^{p\xi}\,\frac{Y_{i:k}^{-\xi}-1}{-\xi}(1+o_p(1))$$
$$+ \frac{p}{k}\sum_{i=1}^{k} Y_{i:k}^{p\xi}\,\frac{Y_{i:k}^{\rho}-1}{\rho}A_0(n/k)(1+o_p(1)).$$

Since, for $p\xi < 1$

$$\frac{1}{k}\sum_{i=1}^{k} Y_{i:k}^{p\xi} \xrightarrow{\mathbb{P}} \frac{1}{1-p\xi}$$

and if we further assume that $\rho < 0$,

$$\frac{1}{k}\sum_{i=1}^{k} Y_{i:k}^{p\xi}\left(\frac{Y_{i:k}^{\rho}-1}{\rho}\right) \xrightarrow{\mathbb{P}} \frac{1}{(1-p\xi)(1-p\xi-\rho)},$$

equation (3.5) follows. Moreover, $\sigma_p^2(\xi)$, given in (3.6), is merely the variance of $\sum_{i=1}^{k} Y_{i:k}^{p\xi}/k = \sum_{i=1}^{k} Y_i^{p\xi}/k$. ■

We next state the main theoretical result in this article, related to the shift invariant versions of the EVI-estimators in (1.5) and (2.8), i.e. the shift-invariant EVI-estimators, generally denoted $\mathrm{E}^{(q)}(k)$ in (2.10). Again, the asymptotic variance is kept at the same level of the unshifted EVI-estimators, but the dominant component of bias changes only in a few cases.

Theorem 3. *Under the second order framework in* (2.2)*, with* $p\xi < 1/2$*, and for intermediate* k*, i.e. if* (3.1) *holds, the asymptotic bias of the* PORT-MOP *and* PORT-OMOP EVI-*estimators, in* (2.10)*, is going to be ruled by*

$$B(t) = \begin{cases} \xi\chi_q/U_0(t), & \text{if } \xi+\rho_0 < 0 \wedge \chi_q \neq 0, \\ A_0(t) + \xi\chi_q/U_0(t), & \text{if } \xi+\rho_0 = 0 \wedge \chi_q \neq 0, \\ A_0(t), & \text{otherwise,} \end{cases}$$

with χ_q *defined in* (3.4). *If we assume that* $\sqrt{k}\ A_0(n/k) \to \lambda_{_A}$ *and/or* $\sqrt{k}/U_0(n/k) \to \lambda_{_U}$, *finite, as* $n \to \infty$, *and with* E *denoting either* H_p *or* H^*, *as given in* (2.10),

$$\sqrt{k}\left(\mathrm{E}^{(q)}(k) - \xi\right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(b_{\mathrm{E}|q}, \sigma^2_{\mathrm{E}}\right),$$

where

$$b_{\mathrm{E}|q} = \begin{cases} \frac{\xi(1-p\xi)\chi_q}{1-(p-1)\xi}\lambda_{_U}, & \text{if } \ \xi + \rho_0 < 0 \wedge \chi_q \neq 0, \\ \frac{1-p\xi}{1-(p-1)\xi}\ \lambda_{_A} + \frac{\xi(1-p\xi)\chi_q}{1-(p-1)\xi}\lambda_{_U}, & \text{if } \ \xi + \rho_0 = 0 \wedge \chi_q \neq 0, \\ \frac{1-p\xi}{1-p\xi-\rho_0}\ \lambda_{_A}, & \text{otherwise.} \end{cases}$$

Proof. For $p \neq 0$, (3.3) and the use of Taylor's expansion $(1+x)^{-1} = 1 - x + o(x)$, as $x \to 0$, enables us to get

$$\mathrm{H}_p^{(q)}(k) \ \stackrel{d}{=}\ \xi + \frac{\sigma_p(\xi)(1-p\xi)^2\mathcal{N}(0,1)(1+o_p(1))}{|p|\sqrt{k}} + \frac{(1-p\xi)A_0(n/k)(1+o_p(1))}{(1-p\xi-\rho_0)} + \frac{\xi(1-p\xi)\chi_q(1+o_p(1))}{(1-(p-1)\xi)U_0(n/k)}.$$

Consequently, the result in the theorem follows. ■

4 Finite sample properties of the EVI-estimators

We have implemented multi-sample Monte-Carlo simulation experiments of size 5000×20, i.e. 20 independent replicates with 5000 runs each, for the classes of MOP and PORT-MOP EVI-estimators associated with $p = p_\ell = 2\ell/(5\xi)$, $\ell = 0,\ 1,\ 2$, and also for the OMOP and PORT-OMOP EVI-estimators. The values $q = 0$ and $q = 0.25$ were considered. We further proceeded to the comparison with the MVRB and the PORT MVRB EVI-estimators, for the same values of q as mentioned above. Sample sizes from $n = 100$ until $n = 5000$ were simulated from a set of underlying models that include the ones shown here as an illustration, the EV model, with CDF $F(x) = \mathrm{EV}_\xi(x)$, with $\mathrm{EV}_\xi(x)$ given in (1.1), $\xi = 0.1, 0.25$, and the *Student-*t_ν, with $\nu = 4, 2$ degrees-of-freedom ($\xi = 1/\nu = 0.25, 0.5$). For details on multi-sample simulation, see Gomes and Oliveira (2001), among others. For the EV parents, results are presented essentially for $q = 0$, the value of q associated with the best performance of

the PORT methodology for these models. For Student parents we consider $q = 0.25$. This is due to the fact that for the Student model the left endpoint is infinite and we cannot thus consider $q = 0$ (see Araújo Santos *et al.*, 2006, and Gomes *et al.*, 2008a, for further details related to the topic).

Remark 4. *Note that, as already stated in the aforementioned articles dealing with a PORT framework, if there are only positive observed values in the sample, we gain nothing with the use of the PORT methodology. The other way round, if there are negative elements in the sample, as happens with EV and Student models and, in practice, with log-returns in financial data, among other types of data, the gain is quite high, as we shall see in the following. This is the main reason for the choice of the aforementioned parents.*

4.1 Mean values and mean square error patterns as k-functionals

For each value of n and for each of the above-mentioned models, we have first simulated the mean value (E) and the RMSE of the estimators under consideration, as functions of the number of top order statistics k involved in the estimation. Apart from the MOP, H_p, in (1.5), $p = 0$ $\big(\mathrm{H}_0 \equiv \mathrm{H}\big)$ and $p = p_\ell = 2\ell/(5\xi)$, $\ell = 1$ (for which asymptotic normality holds), and $\ell = 2$ (where only consistency was proved), the OMOP (H^*), in (2.8), and the MVRB (CH) EVI-estimators, in (2.4), we have also included their PORT versions, respectively given in (2.10) and (2.11), for the above mentioned values of q.

The results are illustrated in Figure 1, for an EV_ξ underlying parent, with $\xi = 0.25$ and $q = 0$. In this case, and for all k, there is a clear reduction in RMSE, as well as in bias, with the obtention of estimates closer to the target value ξ, particularly when we consider H_{p_2} and the associated PORT-version. However, at optimal levels, even the PORT-H^* and PORT-H_{p_1} versions beat the MVRB EVI-estimators. Indeed, the PORT-H_{p_1} can even beat the PORT-MVRB EVI-estimators, as happens in this illustration.

Similar patterns have been obtained for all other simulated models, with the PORT-MVRB outperforming the PORT-MOP only in a few cases and for large sample sizes n.

4.2 Mean values and relative efficiency indicators at optimal levels

Table 1 is also related to the EV_ξ model, with $\xi = 0.25$. We there present, for different sample sizes n, the simulated mean values at optimal levels (levels where RMSEs are minima

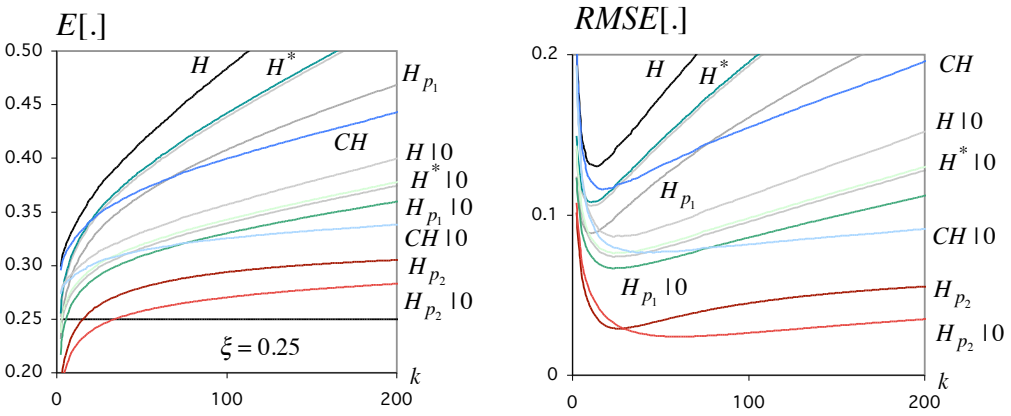

Figure 1: Mean values (*left*) and root mean square errors (*right*) of H, H* (OMOP), CH, and H$_p$, $p = p_\ell = 2\ell/(5\xi)$, $\ell = 1, 2$ (MOP), together with their PORT versions, associated with $q = 0$ and generally denoted $\bullet|0$, for $\text{EV}_{0.25}$ underlying parents and sample size $n = 1000$

as functions of k) of the EVI-estimators under consideration in this study. Information on standard errors, computed on the basis of the 20 replicates with 5000 runs each, are available from the authors, upon request. Among the estimators considered, and distinguishing 3 regions, a first one with (H, CH, H*, H$_{p_1}$), a second one with the associated PORT versions, (H|0, CH|0, H*|0, H$_{p_1}$|0), and a third one with (H$_{p_2}$, H$_{p_2}$|0), for which an asymptotic normal behaviour is not available, the one providing the smallest squared bias is underlined and written in **bold** whenever there is an out-performance of the behaviour achieved in the previous region.

Table 1: Simulated mean values of the semi-parametric EVI-estimators under consideration, at their simulated optimal levels for underlying $\text{EV}_{0.25}$ parents.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
H	0.4202	0.3915	0.3646	0.3482	0.3348	0.3212
CH	0.3816	0.3716	0.3533	0.3416	0.3295	0.3174
H*	0.3398	0.3351	0.3303	0.3226	0.3167	0.3082
H$_{p_1}$	**0.3059**	**0.3077**	**0.3034**	**0.3013**	**0.2998**	**0.2940**
H|0	0.3663	0.3464	0.3261	0.3154	0.3053	0.2957
CH|0	0.3510	0.3369	0.3210	0.3114	0.3033	0.2945
H*|0	0.3292	0.3208	0.3106	0.3046	0.2980	0.2904
H$_{p_1}$|0	**0.3052**	**0.3001**	**0.2963**	**0.2928**	**0.2895**	**0.2848**
H$_{p_2}$	0.2723	0.2698	0.2669	0.2651	0.2638	0.2620
H$_{p_2}$|0	**0.2669**	**0.2650**	**0.2625**	**0.2614**	**0.2603**	**0.2590**

We have further computed the Hill estimator, given in (1.5) when $p = 0$, at the simulated value of $k_{0|0} = \arg\min_k \mathrm{RMSE}\big(\mathrm{H}_0(k)\big)$, the simulated optimal k in the sense of minimum RMSE, not relevant in practice, but providing an indication of the best possible performance of Hill's estimator. Such an estimator is denoted by $\widetilde{\mathrm{H}}_{00}$. For any of the estimators under study, generally denoted $\mathrm{E}(k)$, we have also computed E_0, the estimator $\mathrm{E}(k)$ computed at the simulated value of $k_{0|E} := \arg\min_k \mathrm{RMSE}\big(\mathrm{E}(k)\big)$. The simulated indicators are

$$\mathrm{REFF}_{\mathrm{E}|0} := \frac{\mathrm{RMSE}\left(\widetilde{\mathrm{H}}_{00}\right)}{\mathrm{RMSE}\left(\mathrm{E}_0\right)}. \tag{4.1}$$

Remark 5. *Note that, as usual, an indicator higher than one means a better performance than the Hill estimator. Consequently, the higher these indicators are, the better the associated EVI-estimators perform, comparatively to* $\widetilde{\mathrm{H}}_{00}$.

Again as an illustration of the results obtained, we present Table 2. In the first row, we provide RMSE_0, the RMSE of $\widetilde{\mathrm{H}}_{00}$, so that we can easily recover the RMSE of all other estimators. The following rows provide the REFF-indicators for the different EVI-estimators under study. A similar mark (underlined and **bold**) is used for the highest REFF indicator, again considering the aforementioned three regions.

Table 2: Simulated values of RMSE_0 (first row) and of $\mathrm{REFF}_{\bullet|0}$ indicators, for underlying $\mathrm{EV}_{0.25}$ parents.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
RMSE_0	0.246	0.200	0.157	0.133	0.113	0.092
CH	1.3256	1.2374	1.1711	1.1304	1.1008	1.0716
H^*	1.4391	1.3384	1.2491	1.2021	1.1653	1.1333
H_{p_1}	**1.9307**	**1.7443**	**1.5646**	**1.4633**	**1.3785**	**1.2999**
H|0	1.4875	1.4991	1.5169	1.5309	1.5405	1.5542
CH|0	1.9212	1.8505	1.7790	1.7366	1.6958	1.6633
$\mathrm{H}^*\|0$	1.8966	1.8156	1.7511	1.7217	1.6995	1.6868
$\mathrm{H}_{p_1}\|0$	**2.3988**	**2.2171**	**2.0478**	**1.9564**	**1.8828**	**1.8230**
H_{p_2}	6.4033	5.6755	4.9396	4.4849	4.0943	3.6784
$\mathrm{H}_{p_2}\|0$	**7.5643**	**6.7594**	**5.9369**	**5.4315**	**4.9769**	**4.4991**

For a better visualization of the results presented in Table 1 and Table 2, we further present Figure 2. Due to the high REFF-indicators of H_{p_2} and associated PORT estimators,

we present them in a different scale, at the top of Figure 2, *right*, the one related to the REFF-indicators.

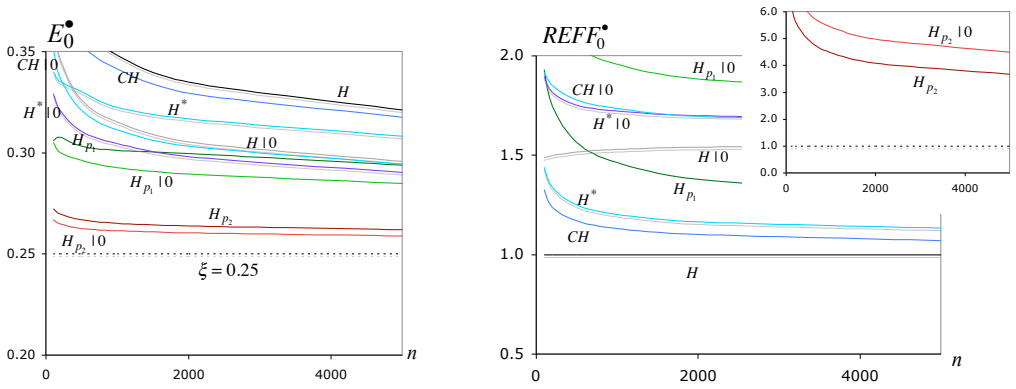

Figure 2: Mean values (*left*) and REFF-indicators (*right*) at optimal levels of the different estimators under study, for an underlying $\mathrm{EV}_{0.25}$ parent and sample sizes $n = 100(100)500$ and $500(500)5000$

Tables 3–4, 5–6 and 7–8 are similar to Tables 1–2, respectively for $\mathrm{EV}_{0.1}$, Student-t_4 and Student-t_2 underlying parents.

Table 3: Simulated mean values of the semi-parametric EVI-estimators under consideration, at their simulated optimal levels for underlying $\mathrm{EV}_{0.1}$ parents.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
H	0.2918	0.2644	0.2403	0.2225	0.2089	0.1952
CH	0.2714	0.2544	0.2341	0.2214	0.2076	0.1946
H*	0.1895	0.1745	0.1605	0.1516	0.1442	0.1464
H_{p_1}	**0.1601**	**0.1496**	**0.1396**	**0.1330**	**0.1274**	**0.1315**
H|0	0.2404	0.2191	0.2009	0.1895	0.1801	0.1688
CH|0	0.2346	0.2176	0.1989	0.1887	0.1793	0.1689
H*|0	0.1611	0.1499	0.1435	0.1441	0.1458	0.14440
$\mathrm{H}_{p_1}\|0$	**0.1400**	**0.1317**	**0.1278**	**0.1290**	**0.1271**	**0.1291**
H_{p_2}	0.1159	0.1149	0.1133	0.1127	0.1114	0.1105
$\mathrm{H}_{p_2}\|0$	**0.1131**	**0.1124**	**0.1110**	**0.1104**	**0.1098**	**0.1090**

Remark 6. *As intuitively expected,* $\mathrm{H}_{p|\bullet}$ *are decreasing in* p*, approaching the true value of* ξ*, or all simulated models.*

Table 4: Simulated values of RMSE_0 (first row) and of $\mathrm{REFF}_{\bullet|0}$ indicators, for underlying $\mathrm{EV}_{0.1}$ parents.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
RMSE_0	0.2524	0.2109	0.1732	0.1511	0.1329	0.1136
CH	1.1778	1.1141	1.0684	1.0450	1.0293	1.0186
H^*	2.0954	1.9436	1.7846	1.6708	1.5618	1.4483
H_{p_1}	**3.0221**	**2.7527**	**2.4758**	**2.2837**	**2.1044**	**1.9174**
H|0	1.4292	1.4185	1.4153	1.4093	1.4006	1.3967
CH|0	1.5680	1.5140	1.4760	1.4509	1.4290	1.4134
H^*|0	2.5865	2.3621	2.1291	1.9935	1.8775	1.7709
H_{p_1}|0	**3.5906**	**3.2188**	**2.8408**	**2.6229**	**2.4277**	**2.2369**
H_{p_2}	12.1731	10.5862	9.1739	8.3307	7.6068	6.8415
H_{p_2}|0	**13.3178**	**11.6827**	**10.1972**	**9.2846**	**8.5188**	**7.6951**

Table 5: Simulated mean values of the semi-parametric EVI-estimators under consideration, at their simulated optimal levels for underlying Student-t_4 parents ($\xi = 0.25$).

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
H	0.3607	0.3392	0.3167	0.3055	0.2959	0.2862
CH	0.3109	0.3104	0.3005	0.2939	0.2879	0.2805
H^*	0.3236	0.3135	0.3028	0.2959	0.2891	0.2818
H_{p_1}	**0.2964**	**0.2914**	**0.2881**	**0.2844**	**0.2810**	**0.2765**
H|0.25	0.3078	0.2935	0.2806	0.2728	0.2672	0.2613
CH|0.25	0.2869	0.2783	**0.2686**	**0.2641**	**0.2599**	**0.2561**
H^*|0.25	0.2923	0.2861	0.2764	0.2699	0.2658	0.2607
H_{p_1}|025	**0.2797**	**0.2762**	0.2709	0.2671	0.2640	0.2599
H_{p_2}	0.2662	0.2646	0.2616	0.2604	0.2589	0.2575
H_{p_2}|0.25	**0.2613**	**0.2591**	**0.2570**	**0.2558**	**0.2550**	**0.2539**

Remark 7. *For adequate values of q and p, the* PORT-MOP EVI*–estimators are able to outperform the MVRB and even the PORT-MVRB, in some cases.*

Table 6: Simulated values of RMSE_0 (first row) and of $\text{REFF}_{\bullet|0}$ indicators, for underlying Student-t_4 parents.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
RMSE_0	0.1830	0.1431	0.1059	0.0854	0.0696	0.0535
CH	1.4349	1.3982	1.3615	1.3223	**1.2834**	**1.2358**
H*	1.2984	1.2280	1.1625	1.1297	1.1046	1.0822
H_{p_1}	**1.7501**	**1.5845**	**1.4200**	**1.3285**	1.2554	1.1819
H|0.25	1.6242	1.6823	1.7745	1.8702	1.9850	2.1777
CH|0.25	2.4005	**2.5115**	**2.7219**	**2.8846**	**3.1153**	**3.5054**
H*|0.25	1.9459	1.9360	1.9712	2.0386	2.1329	2.3108
H_{p_1}|025	**2.4223**	2.3048	2.2245	2.2166	2.2410	2.3346
H_{p_2}	5.3556	4.7308	4.0399	3.5993	3.2243	2.7827
H_{p_2}|0.25	**6.6674**	**6.0186**	**5.2884**	**4.8145**	**4.3920**	**3.8883**

Table 7: Simulated mean values of the semi-parametric EVI-estimators under consideration, at their simulated optimal levels for underlying Student-t_2 parents ($\xi = 0.5$).

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
H	0.6015	0.5769	0.5560	0.5439	0.5355	0.5257
CH	**0.4644**	**0.5059**	**0.5117**	**0.5073**	**0.5041**	**0.5019**
H*	0.5823	0.5671	0.5510	0.5404	0.5324	0.5233
H_{p_1}	0.5553	0.5486	0.5393	0.5325	0.5261	0.5182
H|0.25	0.5203	0.5139	0.5063	0.5037	0.5020	0.5009
CH|0.25	**0.4885**	**0.4940**	**0.4974**	**0.4988**	**0.4995**	**0.4997**
H*|0.25	0.5194	0.5142	0.5070	0.5035	0.5018	0.5009
H_{p_1}|025	0.5186	0.5130	0.5078	0.5048	0.5023	0.5011
H_{p_2}	0.5206	0.5168	0.5137	0.5111	0.5086	0.5053
H_{p_2}|0.25	0.5120	0.5096	0.5072	0.5051	0.5036	0.5018

5 AN ADAPTIVE CHOICE OF (k, p, q) AND CONCLUDING REMARKS

Apart from heuristic choices based on sample path stability, similar to the ones in Neves *et al.* (2015), we suggest the use of the double-bootstrap methodology, briefly described in the following Section.

Table 8: Simulated values of RMSE_0 (first row) and of $\mathrm{REFF}_{\bullet|0}$ indicators, for underlying Student-t_2 parents.

	$n=100$	$n=200$	$n=500$	$n=1000$	$n=2000$	$n=5000$
RMSE_0	0.2028	0.1528	0.1078	0.0835	0.0652	0.0470
CH	0.9803	**1.4180**	**1.7059**	**1.9437**	**2.2267**	**2.6414**
H^*	1.1363	1.1047	1.0811	1.0695	1.0666	1.0644
H_{p_1}	**1.4333**	1.3224	1.2344	1.1957	1.1841	1.1844
H|0.25	1.8476	1.9699	2.2126	2.4120	2.6709	3.0481
CH|0.25	**2.4870**	**2.6495**	**2.9310**	**3.1988**	**3.5307**	**4.0413**
H^*|0.25	1.9814	2.0820	2.3071	2.5030	2.7652	3.1490
H_{p_1}|025	2.2140	2.2306	2.3726	2.5269	2.7644	3.1234
H_{p_2}	3.7572	3.2811	2.7464	2.4304	2.2496	2.1766
H_{p_2}|0.25	**4.5942**	**4.1347**	**3.6354**	**3.3598**	3.2719	3.3502

5.1 Bootstrap adaptive PORT-MOP EVI-estimation

A reasonably sophisticated and time-consuming algorithm, that has proved to work properly in many situations, is the double-bootstrap algorithm. The basic framework for such algorithm is related to the fact that for any class of EVI-estimators, generally denoted $\mathrm{E}(k)$,

$$k_{0|\mathrm{E}}(n) = \arg\min_k \mathrm{RMSE}\big(\mathrm{E}(k)\big) = k_{A|\mathrm{E}}(n)\big(1+o(1)\big), \tag{5.1}$$

with $k_{A|\mathrm{E}}(n) := \arg\min_k \mathrm{ARMSE}\big(\mathrm{E}(k)\big)$ and ARMSE standing for *asymptotic root mean square error*. The bootstrap methodology can then enable us to consistently estimate the optimal sample fraction, $k_{0|\mathrm{E}}(n)/n$, with $k_{0|\mathrm{E}}(n)$ given in (5.1), on the basis of a consistent estimator of $k_{A|\mathrm{E}}(n)$, in a way similar to the one used in Draisma *et al.* (1999), Danielson *et al.* (2001) and Gomes and Oliveira (2001), for the classical adaptive Hill EVI-estimation, performed through $\mathrm{H}(k) \equiv \mathrm{H}_0(k)$, in (1.4), in Brilhante *et al.* (2013), for the MOP EVI-estimation throught $\mathrm{H}_p(k)$, in (1.5), in Gomes *et al.* (2011b, 2012), for second-order reduced-bias estimation, and in Gomes *et al.* (2015) for the CH and PORT-CH EVI-estimation.

The bootstrap methodology is applied to sub-samples of size $m_1 = o(n)$ and $m_2 = m_1^2/n$, is practically independent on m_1 for an adequate PORT EVI-estimation and it is essentially based on the relationship between the optimal sample fraction of the EVI-estimator under consideration, and the one of the auxiliary statistics

$$T_{k,n} \equiv T(k|\mathrm{E}) := \mathrm{E}([k/2]) - \mathrm{E}(k), \quad k = 2, \ldots, n-1,$$

which converge in probability to the known value zero, for any intermediate k, and have an asymptotic behaviour strongly related with the asymptotic behaviour of $\mathrm{E}(k)$. For details, see Gomes *et al.* (2015), where an algorithm for the optimal choice of (k, q) is provided for the PORT-MVRB EVI-estimators, in (2.11). Indeed, for the adaptive choice of (k, p, q) based on minimal bootstrap RMSE, an algorithm of the type of the one in Gomes *et al.* (2015) can be conceived with the inclusion of the MOP and PORT-MOP together with the Hill, the PORT-Hill, the MVRB and the PORT-MVRB. This is however a topic out of the scope of this article.

5.2 Overall comments

A few concluding remarks:

- For both mean values and RMSEs at optimal levels, and for all simulated models, if we restrict ourselves to the region of values of p where we can guarantee asymptotic normality, i.e. $p < 1/(2\xi)$, the best results were obtained for the value of p closer to $1/(2\xi)$, i.e. $p = 2/(5\xi)$. The OMOP is not at all competitive with the MOP, regarding both bias and MSE.

- For the simulated models, the MOP can clearly beat the MVRB, being beaten by the MVRB only for Student-t_2 parents. A similar comment applies to the behaviour of the PORT-MOP comparatively to the PORT-MVRB EVI-estimators.

- The improvement achieved with the use of the PORT-MOP EVI-estimation can be highly significant, as illustrated. Indeed, the PORT-MOP can, for an adequate (p, q) beat the MVRB EVI-estimators for all k, being often able to beat the optimal PORT-MVRB. This is surely due to the small increase in the variance and the high reduction of bias of the PORT-MOP comparatively with the PORT-MVRB, a topic not yet investigated, due to the deep involvement of a third-order framework.

Acknowledgements. Research partially supported by National Funds through **FCT** – Fundação para a Ciência e a Tecnologia, project PEst-OE/MAT/UI0006/2014.

References

[1] Araújo Santos, P.; Fraga Alves, M.I. and Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation, *Revstat*, **4**:3, 227–247.

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

[3] Beirlant, J.; Caeiro, F. and Gomes, M.I. (2012). An overview and open research topics in the field of statistics of univariate extremes, *Revstat*, **10**:1, 1–31.

[4] Bingham, N.H.; Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.

[5] Brilhante, F.; Gomes, M.I. and Pestana, D. (2013). A simple generalization of the Hill estimator, *Computational Statistics and Data Analysis*, **57**:1, 518–535.

[6] Brilhante, M.F.; Gomes, M.I. and Pestana, D. (2014). The mean-of-order p extreme value index estimator revisited. In Pacheco, A., Santos, R., Oliveira, M.R., and Paulino, C.D. (eds.), *New Advances in Statistical Modeling and Application*, Studies in Theoretical and Applied Statistics, Selected Papers of the Statistical Societies, Springer-Verlag, Berlin and Heidelberg, 163–175.

[7] Caeiro, F.; Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat*, **3**:2, 113–136.

[8] Caeiro, F.; Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework, *Communications in Statistics—Theory & Methods*, **38**:7, 1019–1040.

[9] Caeiro, F.; Gomes, M.I. and Henriques-Rodrigues, L. (2014). A location invariant probability weighted moment EVI estimator, *International J. Computer Mathematics*, DOI: 10.1080/00207160.2014.975217

[10] 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.

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

[12] Fraga Alves, M.I.; Gomes, M.I. and Haan, L. de (2003). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Mathematica*, **60**:2, 193–213.

[13] Fraga Alves, M.I.; Gomes, M.I.; Haan, L. de and Neves, C. (2009). The mixed moment estimator and location invariant alternatives, *Extremes*, **12**, 149–185.

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

[15] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics*, **44**:6, 423–453.

[16] Gomes, M.I. and Caeiro, F. (2014). Efficiency of partially reduced-bias mean-of-order-p versus minimum-variance reduced-bias extreme value index estimation. In Gilli, M., Gonzalez-Rodriguez, G. and Nieto-Reyes, A. (eds.), *Proceedings of COMPSTAT 2014*, The International Statistical Institute/International Association for Statistical Computing, 289–298.

[17] Gomes, M.I. and Guillou, A. (2014). Extreme value theory and statistics of univariate extremes: A review, *International Statistical Review*, doi:10.1111/insr.12058.

[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.; Fraga Alves, M.I. and Araújo Santos, P. (2008a). PORT Hill and moment estimators for heavy-tailed models, *Communications in Statistics—Simulation & Computation*, **37**, 1281–1306.

[22] Gomes, M.I.; Haan, L. de and Henriques-Rodrigues, L. (2008b). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses, *J. Royal Statistical Society B*, **70**:1, 31–52.

[23] Gomes, M.I.; Henriques-Rodrigues, L. and Miranda, C. (2011a). Reduced-bias location-invariant extreme value index estimation: a simulation study, *Communications in Statistics—Simulation & Computation*, **40**:3, 424–447.

[24] Gomes, M.I.; Mendonça, S. and Pestana, D. (2011b). Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology, *Communications in Statistics—Theory & 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*, **15**, 463–489.

[26] Gomes, M.I.; Henriques-Rodrigues, L.; Fraga Alves, M.I. and Manjunath, B.G. (2013). Adaptive PORT-MVRB estimation: an empirical comparison of two heuristic algorithms, *J. Statistical Computation and Simulation*, **83**:6, 1129–1144.

[27] Gomes, M.I.; Caeiro, F.; Henriques-Rodrigues, L. and Manjunath, B.G. (2015). Bootstrap methods in statistics of extremes. In Longin, F. (ed.), *Handbook of Extreme Value Theory and Its Applications to Finance and Insurance*, Handbook Series in Financial Engineering and Econometrics (Ruey Tsay Adv. Ed.), John Wiley and Sons, in press.

[28] Haan, L. de (1984). Slow variation and characterization of domains of attraction, In Tiago de Oliveira, J. (ed.), *Statistical Extremes and Applications*, 31–48, D. Reidel, Dordrecht, Holland.

[29] Haan, L. de and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer, New York.

[30] Haan, L. de and Peng, L. (1998). Comparison of extreme value index estimators, *Statistica Neerlandica*, **52**, 60–70.

[31] Hall, P. (1982). On some simple estimates of an exponent of regular variation, *J. Royal Statistical Society B*, **44**, 37–42.

[32] Hall, P. and Welsh, A.H. (1985). Adaptative estimates of parameters of regular variation, *Annals of Statistics*, **13**, 331–341.

[33] Henriques-Rodrigues, L.; Gomes, M.I.; Fraga Alves, M.I. and Neves, C. (2014). PORT estimation of a shape second-order parameter, *Revstat* , **13**:3, 299–328.

[34] Henriques-Rodrigues, L.; Gomes, M.I. and Manjunath, B.G. (2015). Estimation of a scale second-order parameter related to the PORT methodology, *J. Statistical Theory and Practice*, **9**:3, 571–599.

[35] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1163–1174.

[36] Neves, M.M.; Gomes, M.I.; Figueiredo, F. and Prata-Gomes, D. (2015). Modeling extreme events: sample fraction adaptive choice in parameter estimation, *J. Statistical Theory and Practice*, **9**:1, 184–199.

[37] Paulauskas, V. and Vaiciulis, M. (2013). On the improvement of Hill and some others estimators, *Lithuanian Mathematical J.*, **53**, 336–355.

[38] Reiss, R.-D. and Thomas, M. (2001; 2007). *Statistical Analysis of Extreme Values, with Application to Insurance, Finance, Hydrology and Other Fields*, 2nd edition; 3rd edition, Birkhäuser Verlag.