Competitive estimation of the extreme value index*

M. Ivette Gomes
Universidade de Lisboa, Faculdade de Ciências (DEIO) and CEAUL

and

Lígia Henriques-Rodrigues
Universidade de São Paulo, IME, Brasil, and CEAUL

May 7, 2016

Abstract

The *mean-of-order-p* (MO_p) *extreme value index* (EVI) estimators are based on Hölder's mean of an adequate set of statistics, and generalize the classical Hill EVI-estimators, associated with $p = 0$. Such a class of estimators, dependent on the tuning parameter $p \in \mathbb{R}$, has revealed to be highly flexible, but it is not invariant for changes in location. To make the MO_p location-invariant, it is thus sensible to use the *peaks over a random threshold* (PORT) methodology, based upon the excesses over an adequate ascending order statistic. In this article, apart from an asymptotic comparison at optimal levels of the optimal MO_p class and some competitive EVI-estimators, like a Pareto probability weighted moments EVI-estimator, a few details on PORT classes of EVI-estimators are provided, enhancing their high efficiency both asymptotically and for finite samples.

Keywords. Heavy tails, PORT methodology, Statistical extreme value theory

1 Introduction

Given a sample of n *random variables* (RVs), $\underline{\mathbf{X}}_n := (X_1, \ldots, X_n)$, either independent, identically distributed or possible weakly dependent and stationary from a *cumulative distribution function* (CDF)

*Research partially supported by FCT—Fundação para a Ciência e a Tecnologia, project UID/MAT/00006/2013 (CEA/UL)

F, let us denote by $X_{1:n} \le \cdots \le X_{n:n}$ the associated ascending order statistics. 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 normalized, i.e. $(X_{n:n} - b_n)/a_n$, converges in distribution to a non-degenerate RV. Then, the limiting CDF is compulsory of the type of an *extreme value* (EV) CDF, dependent upon a generalized shape parameter $\xi \in \mathbb{R}$, and with the functional form

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

It is then said that F belongs to the max-domain of attration of EV_ξ, in (1.1), and the notation $F \in \mathcal{D}_{\mathcal{M}}(\mathrm{EV}_\xi)$ is commonly used in the field of *extreme value theory* (EVT). The parameter ξ, the so-called *extreme value index* (EVI), is the primary parameter of extreme events. The EVI measures the heaviness of the *right tail-function* (RTF),

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

and the heavier the right-tail the larger the EVI is. For Paretian-type RTFs ($\xi > 0$), i.e. the often called heavy RTFs, the most popular EVI-estimator is the Hill (H) EVI-estimator (Hill, 1975), the average of the k log-excesses,

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

being thus the logarithm of the geometric mean (or mean-of-order-0) of

$$U_{ik} := X_{n-i+1:n}/X_{n-k:n},\ 1 \le i \le k < n, \tag{1.3}$$

i.e.

$$\hat{\xi}_k^{\mathrm{H}} = \hat{\xi}_k^{\mathrm{H}}(\underline{\mathbf{X}}_n) := \frac{1}{k}\sum_{i=1}^{k} V_{ik} = \ln\left(\prod_{i=1}^{k} \frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^{1/k}. \tag{1.4}$$

Brilhante *et al.* (2013) considered as basic statistics the *mean-of-order-p* (MO_p) of U_{ik}, in (1.3), for $p \ge 0$. More generally, Gomes and Caeiro (2014) considered the same statistics for $p \in \mathbb{R}$, i.e.

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

and, with $\hat{\xi}_k^{\mathrm{H}_0} \equiv \hat{\xi}_k^{\mathrm{H}}$, defined in (1.4), the class of MO_p EVI-estimators,

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

The class of MO_p EVI-estimators in (1.5) depends on the tuning parameter $p \in \mathbb{R}$, it is highly flexible, it is scale-invariant but it is not invariant for changes in location, a property enjoyed by the EVI itself. To make the MO_p EVI-estimators in (1.5) location-invariant, it is sensible to use the *peaks over a random threshold* (PORT) methodology, introduced in Araújo-Santos *et al.* (2006), and further computationally studied in Gomes *et al.* (2008a). The PORT methodology is based on the sample of excesses over the random threshold $X_{n_q:n}$, $0 \le q < 1$, $n_q := \lfloor nq \rfloor + 1$, where $\lfloor x \rfloor$ denotes the integer part of x, i.e. it is based on the sample of size $n^{(q)} = n - n_q$,

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

The PORT-MO_p EVI-estimators are thus estimators with the same functional form of the EVI-estimators in (1.5), but with the original sample $\underline{\mathbf{X}}_n$ replaced by the sample of excesses $\underline{\mathbf{X}}_n^{(q)}$, in (1.6). Consequently, the PORT-MO_p EVI-estimators are given by

$$\hat{\xi}_k^{\mathrm{H}_p^{(q)}} := \hat{\xi}_k^{\mathrm{H}_p}(\underline{\mathbf{X}}_n^{(q)}).$$

Remark 1.1. *We can have $q = 0$, in the case $F(\cdot)$ has a finite left endpoint, $x_{_F} := \inf\{x : F(x) > 0\}$, (the random level can then be the minimum). Note that the choice $q = 0$ is appealing in practice, but should be used with care. Such a random threshold can indeed lead to under-estimation and even inconsistency (see, for instance* Gomes *et al.*, 2008a*). Generally, we can have $0 < q < 1$ (the random level is then an empirical quantile).*

As a measure of comparison, the recent and promising *Pareto probability weighted moments* (PPWM) (see Caeiro and Gomes, 2011; Caeiro *et al.*, 2014) will be considered. The PPWM EVI-estimators are consistent for $\xi < 1$, depend on the statistics,

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

and are defined by

$$\hat{\xi}_k^{\rm PPWM} = 1 - \frac{\hat{a}_1(k)}{\hat{a}_0(k) - \hat{a}_1(k)}, \quad 1 \le k < n. \tag{1.7}$$

PORT-PPWM EVI-estimators (see Caeiro *et al.*, 2016) will also be included in the comparative studies to be developed in this article. In Section 2 a few details on second-order frameworks in EVT, reduced-bias estimation and asymptotic behavior of the estimators will be provided. Section 3 is dedicated to the finite sample properties of the EVI-estimators under study as well as their PORT-versions, done through a large-scale simulation study. Section 4 is devoted to a few final comments on the advantages of PORT EVI-estimators and on possible choices of the vector of tuning parameters.

2 Second-order frameworks for heavy RTFs, reduced-bias and PORT EVI-estimation

Let us consider for the reciprocal *right tail quantile function* (RTQF) associated with F, the notation

$$U(t) := F^{\leftarrow}(1 - 1/t), \quad \text{with} \quad F^{\leftarrow}(y) := \inf\{x : F(x) \ge y\}. \tag{2.1}$$

For heavy right-tails, it is usual to work under the validity of a first-order condition expressed either in terms of the RTF, in (1.2) or the RTQF, in (2.1),

$$F \in \mathcal{D}_{\mathcal{M}}^{+} := \mathcal{D}_{\mathcal{M}}(\mathrm{EV}_{\xi>0}) \quad \Longleftrightarrow \quad \overline{F} \in \mathcal{R}_{-1/\xi} \quad \Longleftrightarrow \quad U \in \mathcal{R}_{\xi}, \tag{2.2}$$

where $\mathcal{R}_\alpha$ stands for the class of regularly varying functions with an index of regular variation $\alpha \in \mathbb{R}$, i.e. positive measurable functions $g(\cdot)$ such that for all $x > 0$, $g(tx)/g(t) \to t^\alpha$, as $t \to \infty$. Details on regular variation theory can be found in Seneta (1976) and Bingham *et al.* (1987), among others.

Consistency of most estimators of parameters of extreme events, like the EVI, can be easily obtained under the first-order framework, i.e. provided that (2.2) holds. To obtain adequate information on the non-degenerate behavior of such estimators it is usual to assume the validity of a second-order condition. Such a second-order condition rules the rate of convergence in any of the first-order conditions in (2.2), and can be written as,

$$\lim_{t\to\infty} \Big(\frac{U(tx)}{U(t)} - x^\xi\Big)/A(t) = x^\xi \begin{cases} \frac{x^\rho - 1}{\rho} & \text{if} \quad \rho < 0, \\ \ln x & \text{if} \quad \rho = 0, \end{cases} \tag{2.3}$$

$\forall x > 0$, where $\rho \le 0$ is a generalized shape second-order parameter.

Whenever working with bias-reduction and also for asymptotic comparison of estimators at optimal levels in the sense of minimal *mean square error* (MSE), it is common to work in a slightly less general class than the one in (2.3), the Hall-Welsh class (Hall and Welsh, 1985), $\mathcal{H} \subset \mathcal{D}_{\mathcal{M}}^{+}$, where

$$U(t) = C\ t^{\xi}\big(1 + \xi\ \beta\ t^{\rho}/\rho + o(t^{\rho})\big), \tag{2.4}$$

as $t \to \infty$, with $\rho < 0$ and $\beta \neq 0$. The vector (β, ρ) of second-order parameters can then be adequately and consistently estimated, so that we can work with *minimum-variance reduced-bias* (MVRB) EVI-estimators. We refer one of the simplest classes of MVRB EVI-estimators, introduced in Caeiro *et al.* (2005), with the functional form,

$$\hat{\xi}_k^{\rm CH} = \hat{\xi}_k^{\rm CH}(\underline{\mathbf{X}}_n) := \hat{\xi}_k^{\rm H}\left(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})\right),\ 1 \le k < n. \tag{2.5}$$

Other MVRB EVI-estimators can be found in Gomes *et al.* (2007a; 2008b), among others. See also the recent overviews in Beirlant *et al.* (2012) and Gomes and Guillou (2015).

2.1 A brief reference to the asymptotic behavior of the aforementioned classes of EVI-estimators

Consistency of any of the aforementioned EVI-estimators is achieved in the whole $\mathcal{D}_{\mathcal{M}}^{+}$ only if we work with an intermediate sequence of values of k, i.e. a sequence of integers $k = k_n$, $1 \le k < n$, such that

$$k = k_n \to \infty \quad \text{and} \quad k_n = o(n),\ \text{as}\ n \to \infty. \tag{2.6}$$

Under the validity of the second-order condition in (2.3), and denoting by $\hat{\xi}_k^{\bullet}$ any of the estimators H, H_p, PPWM and CH, respectively defined in (1.4), (1.5), (1.7) and (2.5), the asymptotic distributional identity

$$\hat{\xi}_k^{\bullet} \stackrel{d}{=} \xi + \frac{\sigma_{\bullet}}{\sqrt{k}}\ Z_k^{\bullet} + b_{\bullet} A(n/k)(1 + o_p(1)) \tag{2.7}$$

holds, where $Z_k^{\bullet}$ is an asymptotically standard normal RV. The distributional representations in (2.7) for $\mathrm{H} \equiv \mathrm{H}_0$, $\mathrm{H}_p,\ p \ge 0$, $\mathrm{H}_p,\ p \in \mathbb{R}$, PPWM and CH were respectively derived in de Haan and Peng (1998), Brilhante *et al.* (2013), Gomes and Caeiro (2014), Caeiro and Gomes (2011) and Caeiro *et al.*

(2005). Under the validity of (2.3) and (2.6), (2.7) holds, with

$$\sigma^2_{\mathrm{H}_p} := \frac{\xi^2(1-p\xi)^2}{1-2p\xi}, \qquad b_{\mathrm{H}_p} := \frac{1-p\xi}{1-p\xi-\rho}, \quad \forall p < 1/(2\xi),$$

and $\forall \xi < 1/2$,

$$\sigma^2_{\mathrm{PPWM}} := \frac{\xi^2(1-\xi)(2-\xi)^2}{(1-2\xi)(3-2\xi)}, \quad b_{\mathrm{PPWM}} := \frac{(1-\xi)(2-\xi)}{(1-\xi-\rho)(2-\xi-\rho)}.$$

For an adequate estimation of (β,ρ), in (2.4), we get $b_{\mathrm{CH}} = 0$ e $\sigma^2_{\mathrm{CH}} = \xi^2$ (see Caeiro *et al.*, 2005).

For the MO_p EVI-estimation, the optimal value of p, in the sense of minimal MSE is given by

$$p_{\mathrm{M}} = \left(1-\rho/2-\sqrt{\rho^2-4\rho+2}/2\right)/\xi =: \varphi(\rho)/\xi. \tag{2.8}$$

Such a value of p leads to a class of *optimal* MO_p EVI-estimators, the so-called OMO_p class. We use the notation

$$\hat{\xi}^{\mathrm{H}^*}_k := \hat{\xi}^{\mathrm{H}_{p_\mathrm{M}}}_k. \tag{2.9}$$

Consequently, and with $\varphi(\rho)$ defined in (2.8), we get for $\mathrm{H}^*(k)$ a distributional representation of the type of the one given in (2.7), but with

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

Remark 2.1. *At optimal levels k, again in the sense of minimal asymptotic* MSE*, the class* H^**, in* (2.9)*, outperforms the classs* H*, in* (1.4)*, in the whole* (ξ,ρ)*-plane* (*see* Brilhante *et al.*, 2013)*. This behavior, together with the fact that the* PPWM EVI-*estimators have revealed to be an interesting alternative to the* H EVI-*estimators in a wide area of the* (ξ,ρ)*-plane, leads us to consider in* Section 2.3 *the asymptotic comparison at optimal levels of* H^* *and* PPWM EVI-*estimators, after a brief reference to classic* PORT EVI-*estimation, in* Section 2.2.

2.2 Asymptotic behavior of classic PORT EVI-estimators

Since $X_{\lfloor nq\rfloor+1:n} - \chi_q = X_{\lfloor nq\rfloor+1:n} - U(1/(1-q)) = O_p(1/\sqrt{n})$, the asymptotic behavior of the PORT-$\bullet$ EVI-estimators coincides with the one of the $\bullet$ EVI-estimators, given the replacement of $X_{n-i+1:n}$ by $X_{n-i+1:n} - \chi_q = X_{n-i+1:n} - U(1/(1-q))$, $1 \le i \le n$. Generally speaking, and denoting by $\hat{\xi}^{\bullet(q)}_k$ the PORT estimator associated with $\hat{\xi}^{\bullet}_k$, for which (2.7) holds with $b_\bullet \neq 0$,we get

$$\hat{\xi}^{\bullet(q)}_k \overset{d}{=} \xi + \frac{\sigma_\bullet}{\sqrt{k}}\, Z^\bullet_k + \Big(b_\bullet A_0(n/k) + \frac{\xi\chi_q}{U_0(n/k)}\Big)(1+o_p(1)),$$

where the functions $U_0(\cdot)$ and $A_0(\cdot)$ are related to the standard parent F_0, with location at zero and unit scale, associated with F. We thus get only a change in bias. Such a bias is ruled by the function

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

Further details on these classic PORT EVI-estimators can be found in Araújo-Santos *et al.* (2006), Gomes *et al.* (2008a; 2016b), Fraga Alves *et al.* (2009) and Caeiro *et al.* (2016).

2.3 Asymptotic comparison at optimal levels

We next proceed to the comparison at optimal levels of the estimators H_p and PPWM. This will be done in a way similar to the one found in several articles like de Haan and Peng (1998), Gomes and Martins (2001), Gomes *et al.* (2005; 2007b; 2013b; 2015), Gomes and Neves (2008), Gomes and Henriques-Rodrigues (2010) and Brilhante *et al.* (2013).

Let us assume the validity of (2.7). Then, if (2.6) holds, i.e. for all intermediate sequence of integers, $k=k_n$,

$$\sqrt{k}\big(\hat{\xi}_k^\bullet - \xi\big) \ \overset{d}{\to}\ N(\lambda_{_A} b_\bullet, \sigma_\bullet^2)$$

provided that $\sqrt{k}\ A(n/k) \to \lambda_{_A}$, finite, as $n\to\infty$. It is then common to write

$$\mathrm{Bias}_\infty\big(\hat{\xi}_k^\bullet\big) := b_\bullet\ A(n/k) \qquad \text{and} \qquad \mathrm{Var}_\infty\big(\hat{\xi}_k^\bullet\big) := \sigma_\bullet^2/k.$$

The so-called *asymptotic* MSE (AMSE) is then given by

$$\mathrm{AMSE}\big(\hat{\xi}_k^\bullet\big) := \sigma_\bullet^2/k + b_\bullet^2\ A^2(n/k).$$

Regular variation theory (see Bingham *et al.*, 1987) enabled Dekkers and de Haan (1993) to show that whenever $b_\bullet \neq 0$ there exists a function $\varphi(n)=\varphi(n,\xi,\rho)$, such that

$$\lim_{n\to\infty} \varphi(n)\ \mathrm{AMSE}\big(\hat{\xi}_{n0}^\bullet\big) = \big(\sigma_\bullet^2\big)^{-\frac{2\rho}{1-2\rho}}\ \big(b_\bullet^2\big)^{\frac{1}{1-2\rho}} =: \mathrm{LMSE}\big(\hat{\xi}_{n0}^\bullet\big),$$

where $\hat{\xi}_{n0}^\bullet := \hat{\xi}_{k_{0|\bullet}(n)}^\bullet$ and $k_{0|\bullet}(n) := \arg\min\limits_k \mathrm{MSE}\big(\hat{\xi}_k^\bullet\big)$. Under the validity of (2.3), with $A(t)=\xi\beta t^\rho$, $\rho<0$, or equivalently, under the validity of (2.4), we can guarantee that

$$k_{0|\bullet}(n) = \big(\sigma_\bullet^2\ n^{-2\rho}/(b_\bullet^2\xi^2\beta^2(-2\rho))\big)^{1/(1-2\rho)}\,(1+o(1)),$$

and consider the following:

Definition 2.1. *Given two first-order biased estimators, $\hat{\xi}_k^{(1)}$ and $\hat{\xi}_k^{(2)}$, for which* (2.7) *holds with constants* (σ_1, b_1) *and* (σ_2, b_2)*,* $b_1, b_2 \neq 0$*, respectively, both computed at their optimal levels, the asymptotic relative efficiency* (AREFF) *of $\hat{\xi}_{n0}^{(1)}$ relatively to $\hat{\xi}_{n0}^{(2)}$ is*

$$\mathrm{AREFF}_{1|2} := \sqrt{\mathrm{LMSE}(\hat{\xi}_{n0}^{(2)})/\mathrm{LMSE}(\hat{\xi}_{n0}^{(1)})} = \left(\left(\frac{\sigma_2}{\sigma_1}\right)^{-2\rho}\left|\frac{b_2}{b_1}\right|\right)^{\frac{1}{1-2\rho}}. \qquad (2.10)$$

Remark 2.2. *Note that the* AREFF*-indicator, in* (2.10)*, was conceived in such a way that the higher is* $\mathrm{AREFF}_{1|2}$ *the better is the first estimator.*

Regarding the PPWM EVI-estimators, it is obvious that for all ξ, $0 < \xi < 1/2$, $\sigma^2_{\mathrm{H}} = \xi^2 < \sigma^2_{\mathrm{PPWM}}$. However, $b_{\mathrm{PPWM}} < b_{\mathrm{H}} = 1/(1-\rho)$, if $\rho \neq 0$. An asymptotic comparison of H and PPWM (denoted P) at optimal levels leads us to the results presented in Figure 1. We can then see that the gain in efficiency of the PPWM comparatively to the H EVI-estimator happens in a wide region of the (ξ, ρ)-plane.

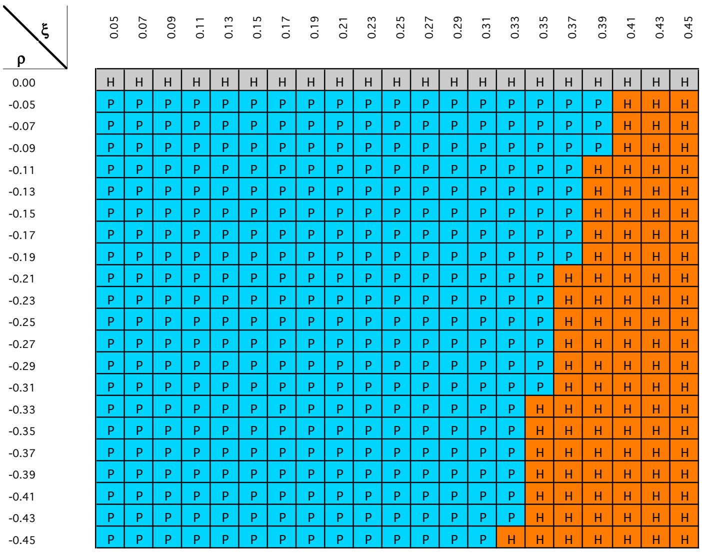

Figure 1: Best EVI-estimator asymptotically and at optimal levels, for a choice between H and PPWM

On the other hand, the asymptotic comparison of PPWM and H* at optimal levels leads us to the results presented in Figure 2. This figure reveals again the importance of the OMO_p EVI-estimation, despite of the competitiveness of the PPWM methodology.

ρ \ ξ	0.05	0.07	0.09	0.11	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45
0.00	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
-0.05	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	H*	H*	H*	H*
-0.07	H*	H*	P	P	P	P	P	P	P	P	P	P	P	P	P	P	H*	H*	H*	H*	H*
-0.09	H*	H*	H*	H*	P	P	P	P	P	P	P	P	P	P	P	H*	H*	H*	H*	H*	H*
-0.11	H*	H*	H*	H*	H*	H*	H*	P	P	P	P	P	P	H*	H*	H*	H*	H*	H*	H*	H*
-0.13	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.15	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.17	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.19	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.21	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.23	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.25	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.27	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.29	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.31	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.33	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.35	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.37	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.39	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.41	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.43	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.45	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*

Figure 2: Best EVI-estimator asymptotically and at optimal levels, for a choice between H* and PPWM

3 Monte-Carlo simulation

Large-scale Monte-Carlo experiments were run for H, H_p, PPWM, CH and H^*, respectively defined in (1.4), (1.5), (1.7), (2.5) and (2.9), as well as their PORT versions associated with $q = 0, 0.1, 0.25$, for the following underlying parents:

(1) the *Fréchet* model, with CDF $F(x) = \exp(-x^{-1/\xi}),\ x \geq 0,\ \xi = 0.1, 0.25, 0.5$ and 1 ($\rho = -1$);

(2) the *extreme value* model, with CDF $F(x) = \mathrm{EV}_\xi(x)$, with $\mathrm{EV}_\xi(x)$ given in (1.1), for the same values of ξ ($\rho = -\xi$);

(3) the *generalised Pareto* model, with CDF $F(x) = 1 + \ln \mathrm{EV}_\xi(x) = 1 - \big(1 + \xi x\big)^{-1/\xi}$, $0 \leq x < -1/\xi$, $\mathrm{EV}_\xi(x)$ given in (1.1), also for the same values of ξ ($\rho = -\xi$);

(4) the *Student*-t_ν, with $\nu = 2, 3, 4$ ($\xi = 1/\nu, \rho = -2/\nu$).

The simulations were multi-sample simulations of size 5000×20. We have considered sample sizes $n = 100, 500(100)$, and $1000, 5000(1000)$. For each value of n, on the basis of a run of size 5000 and for each of the aforementioned models, we have first simulated the mean value (E) and root MSE (RMSE), as functions of the sample fraction, k/n.

As an illustration, we consider the $EV_{0.1}$ underlying parent. In Figure 3, we represent, at the *left*, the simulated mean values of H_p, $p = 0$ $\big(H_0 \equiv H\big)$ and $p = p_\ell = 2\ell/(5\xi)$, $\ell = 1$ (where an asymptotic normal behavior holds), and $\ell = 2$ (where we can only guarantee consistency), H^*, PPWM (again denoted P) and CH. Due to the specificity of the EV_ξ model in (1.1), with a finite left endpoint, $q = 0$ was considered and simulated mean values of associated PORT EVI-estimators are represented also in Figure 3 (*right*). The notation $\bullet|0$ is used for the PORT-$\bullet$ EVI-estimators.

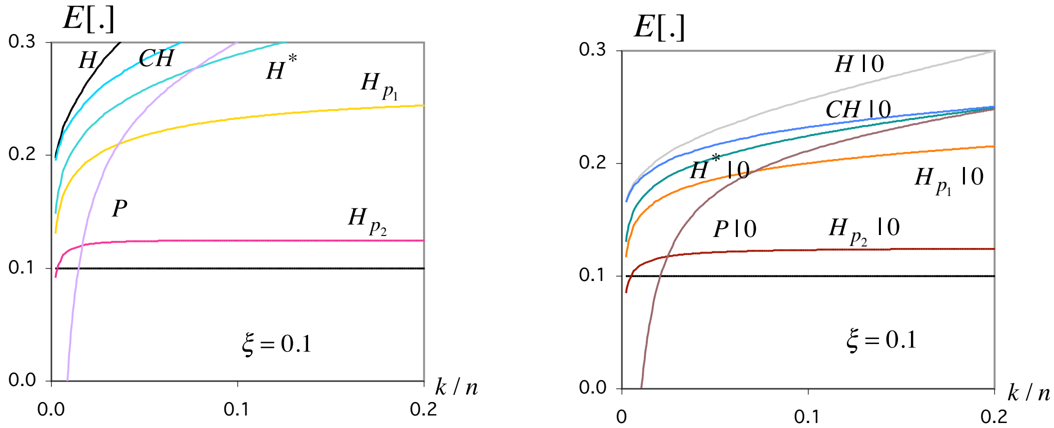

Figure 3: Mean values of the EVI-estimators (*left*) and their PORT versions associated with $q = 0$ (*right*)

Figure 4 is similar to Figure 3, but we now represent the RMSE of the EVI-estimators.

We have further computed the Hill EVI-estimator at the simulated value of $k_{0|0} = \arg\min_k \text{RMSE}\big(\hat\xi_k^{H_0}\big)$, a value that provides the best possible estimation provided by the Hill EVI-estimator. For such estimator the notation $\tilde\xi_0^{H_0}$ is used. We have next computed $\tilde\xi_0^\bullet$, the $\hat\xi_k^\bullet$ EVI-estimator, computed at the simulated value of $k_{0|\bullet} := \arg\min_k \text{RMSE}\big(\hat\xi_k^\bullet\big)$. The simulated mean values of $\tilde\xi_0^\bullet$, denoted by $E_0^\bullet$, are represented in Figure 5. The simulated values of the indicators $\text{REFF}_0^\bullet := \text{RMSE}\big(\tilde\xi_0^{H_0}\big)/\text{RMSE}\big(\tilde\xi_0^\bullet\big)$, are shown in Figure 6.

In this case, and for any k there is a clear reduction in the RMSE, as well as in the bias, with estimates closer to the target, ξ, particularly when we consider H_{p_2} and the associated PORT version. But at optimal levels, even the PORT-H^* and PORT-H_{p_1} EVI-estimators outperform the MVRB. Indeed, even the PORT-H_{p_1} EVI-estimator outperforms the PORT-MVRB EVI-estimator. The behavior of the PPWM and associated PORT version is also quite remarkable, being beaten only by the H_{p_2} EVI-

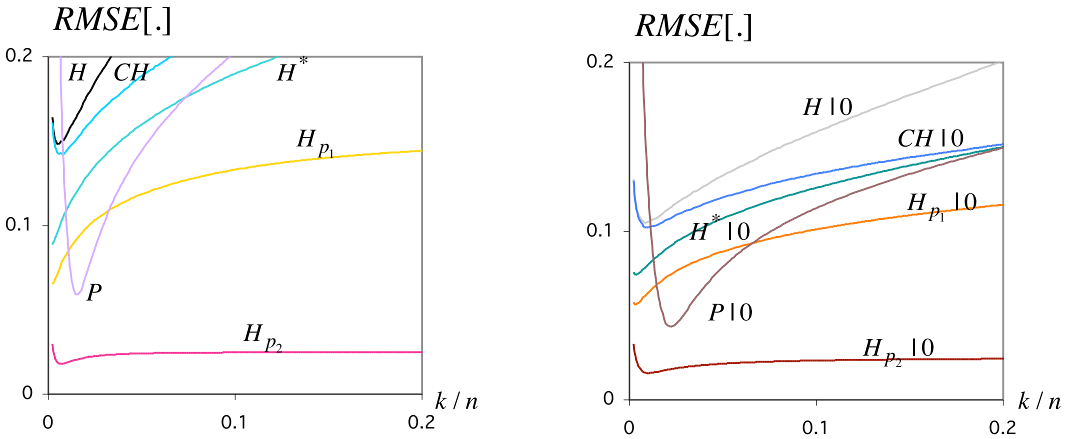

Figure 4: RMSE values of the EVI-estimators (*left*) and their PORT versions associated with $q = 0$ (*right*)

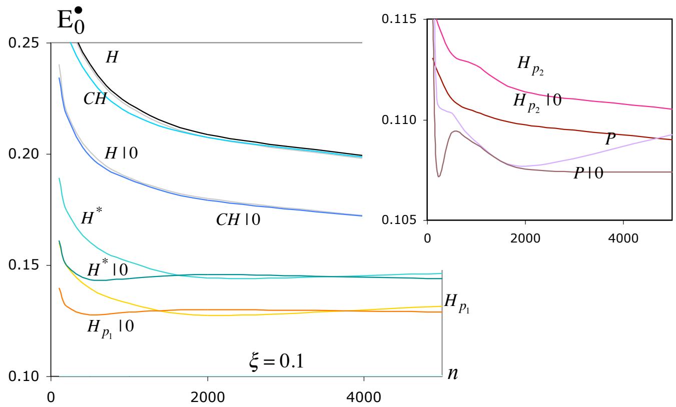

Figure 5: Simulated mean values of the EVI-estimators at optimal levels, as a function of n

estimator.

4 Final comments

The type of improvement described for the $\mathrm{EV}_{0.1}$ underlying parent happens for the great majority of simulated models, as can be seen in Gomes *et al.* (2016b), where some other models are used as an illustration of the behavior of PORT MO_p EVI-estimators only. For all k, the MVRB outperforms the Hill EVI-estimator, the OMO_p, i.e. H^*, outperforms the MVRB, regarding minimal RMSE, and

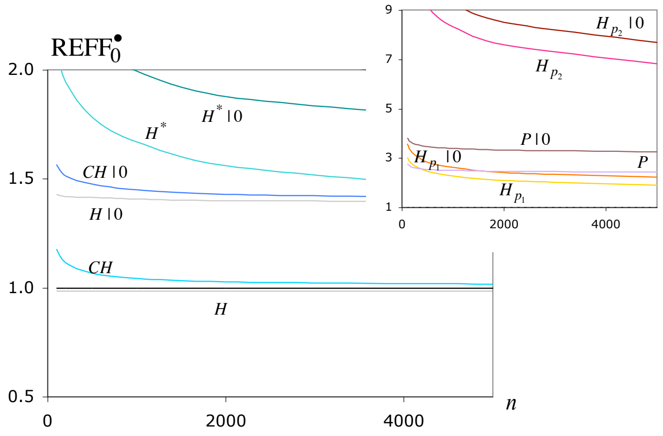

Figure 6: Simulated REFF-indicators for the different EVI-estimators under study, as a function of n

an adequate choice of p in the MO_p and PORT-MO_p enables them to outperform the PORT-MVRB EVI-estimator. However, and only regarding bias, the PORT-PPWM outperforms the best PORT-MO_p EVI-estimator.

Further note that the choice of the tuning parameters (k, p, q) can be done by algorithms like a bootstrap algorithm of the type of the one in Brilhante *et al.* (2013) (see also Caeiro and Gomes, 2015, and Gomes *et al.*, 2016a, where R-scripts are provided) or heuristic sample-path stability algorithms, like the ones in Gomes *et al.* (2013a).

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., Caeiro, F. and Gomes, M.I. (2012). An overview and open research topics in statistics of univariate extremes. *Revstat* **10**:1, 1–31.

[3] Bingham, N., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation.* Cambridge Univ. Press, Cambridge.

[4] 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.

[5] Caeiro, F. and Gomes, M.I. (2011). Semi-parametric tail inference through probability-weighted moments. *J. Statist. Planning and Inference* **141**:2, 937–950.

[6] Caeiro, F. and Gomes, M.I. (2015). Threshold Selection in Extreme Value Analysis. Chapter in: Dipak Dey and Jun Yan, *Extreme Value Modeling and Risk Analysis: Methods and Applications*, Chapman-Hall/CRC, ISBN 9781498701297, pp. 69–87.

[7] Caeiro, F., Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 111–136.

[8] Caeiro, F., Gomes, M.I. and Vandewalle, B. (2014). Semi-Parametric Probability-Weighted Moments Estimation Revisited. *Methodology and Computing in Applied Probability* **16**:1, 1–29.

[9] Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2016). A location invariant probability weighted moment EVI estimator. *International Journal of Computer Mathematics* **93**:4, 676–695.

[10] Dekkers, A.L.M. and de Haan, L. (1993). Optimal sample fraction in extreme value estimation. *J. Multivariate Analysis* **47**:2, 173–195.

[11] Fraga Alves, M.I., Gomes, M.I., de Haan, L. and Neves, C. (2009). Mixed moment estimators and location invariant alternatives. *Extremes* **12**, 149–185.

[12] 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.

[13] Gomes, M.I. and Guillou, A. (2015). Extreme Value Theory and Statistics of Univariate Extremes: A Review. *International Statistical Review* **83**:2, 263–292.

[14] Gomes, M.I. and Henriques-Rodrigues, L. (2010). Comparison at optimal levels of classical tail index estimators: a challenge for reduced-bias estimation? *Discussiones Mathematica: Probability and Statistics* **30**:1, 35–51.

[15] Gomes, M.I. and Martins, M.J. (2001). Generalizations of the Hill estimator — asymptotic versus finite sample behaviour. *J. Statist. Planning and Inference* **93**, 161–180.

[16] Gomes, M.I. and Neves, C. (2008). Asymptotic comparison of the mixed moment and classical extreme value index estimators. *Statistics and Probability Letters* **78**:6, 643–653.

[17] Gomes, M.I., Miranda, C. and Pereira, H. (2005). Revisiting the role of the Jackknife methodology in the estimation of a positive extreme value index. *Comm. in Statistics—Theory and Methods* **34**, 1–20.

[18] Gomes, M.I., Martins, M.J. and Neves, M.M. (2007a). Improving second order reduced bias extreme value index estimation. *Revstat* **5**:2, 177–207.

[19] Gomes, M.I., Miranda, C. and Viseu, C. (2007b). Reduced-bias tail index estimation and the Jackknife methodology. *Statistica Neerlandica* **61**:2, 243–270.

[20] Gomes, M.I., Fraga Alves, M.I. and Araújo-Santos, P. (2008a). PORT Hill and Moment Estimators for Heavy-Tailed Models. *Commun. Statist. – Simul. and Comput.* **37**, 1281–1306.

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

[22] Gomes, M.I., Henriques-Rodrigues, L., Fraga Alves, M.I. and Manjunath, B.G. (2013a). Adaptive PORT-MVRB estimation: an empirical comparison of two heuristic algorithms. *J. Statist. Comput. and Simul.* **83**:6, 1129–1144.

[23] Gomes, M.I., Martins, M.J. and Neves, M.M. (2013b). Generalized jackknife-based estimators for univariate extreme-value modeling. *Comm. in Statistics—Theory and Methods* **42**:7, 1227–1245.

[24] Gomes, M.I., Figueiredo, F., Martins, M.J. and Neves, M.M. (2015). Resampling methodologies and reliable tail estimation. *South African Statistical Journal* **49**, 1–20.

[25] Gomes, M.I., Caeiro, F., Henriques-Rodrigues, L. and Manjunath, B.G. (2016a). 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.

[26] Gomes, M.I., Henriques-Rodrigues, L. and Manjunath, B.L. (2016b). Mean-of-order-p Location-invariant Extreme Value Index Estimation. Accepted at *Revstat.* Available at: `https://www.ine.pt/revstat/pdf/PORT-MOP-REVSTAT.pdf`

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

[28] Hall, P. and A.H. Welsh (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13**, 331–341.

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

[30] Seneta, E. (1976). *Regularly Varying Functions*. Lecture Notes in Math. 508, Springer- Verlag, Berlin.