An efficient naive generalisation of the Hill estimator—discrepancy between asymptotic and finite sample behaviour

Helena Penalva

ESCE - IPS and CEAUL, Universidade de Lisboa, Portugal, e-mail: helena.penalva@esce.ips.pt

Frederico Caeiro

CMA and DM, Universidade Nova de Lisboa, Portugal, e-mail: fac@fct.unl.pt

M. Ivette Gomes

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

M. Manuela Neves

ISA and CEAUL, Universidade de Lisboa, Portugal, e-mail: manela@isa.ulisboa.pt

Abstract

The Lehmer mean of order p of k positive numbers $(a_1, \ldots, a_k)$ is defined by $\sum_{i=1}^k a_i^p / \sum_{i=1}^k a_i^{p-1}$, generalizing both the arithmetic mean ($p = 1$) and the harmonic mean ($p = 0$). Given a random sample $(X_1, \ldots, X_n)$ and the associated sample of ascending order statistics $(X_{1:n} \leq \cdots \leq X_{n:n})$, the classical Hill estimator of a positive extreme value index (EVI), the primary parameter of extreme events, can thus be considered as the Lehmer mean of order 1 of the k log-excesses $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \leq i \leq k < n$. We now more generally consider the Lehmer mean of order p of the log-excesses and an associated Lehmer EVI-estimator. Apart from the derivation of the asymptotic behaviour of this class of EVI-estimators, an asymptotic comparison, at optimal levels, of the members of such a class reveals that for the optimal p they are able to overall outperform a recent and promising generalization of the Hill EVI-estimator. A large-scale Monte-Carlo simulation study is developed, giving emphasis to the discrepancies between asymptotic and finite sample behaviour of the estimators. A bootstrap algorithm for an adaptive estimation of the tuning parameters under play is also put forward.

Keywords and phrases. *Bootstrap methodology; Heavy tails; Monte-Carlo simulation; Optimal levels; Semi-parametric estimation; Statistics of extremes.*

1 The new class of estimators and scope of the article

On the basis of a sample of size n of independent, identically distributed *random variables* (RVs), $X_1, \ldots, X_n$, from a *cumulative distribution function* (CDF) F, let us consider the notation, $X_{1:n} \leq \cdots \leq X_{n:n}$, for the associated ascending *order statistics* (OSs). As usual in a framework of statistical *extreme value theory* (EVT) let us 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, we use the notation $F \in \mathcal{D}_{\mathcal{M}}(\mathrm{EV}_\xi)$, and the parameter ξ is the *extreme value index* (EVI), the primary parameter of extreme events.

Let us denote by $\mathcal{R}_a$ 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 Seneta, 1976, and Bingham *et al.*, 1987, among others). For weakly dependent and stationary sequences from $F(\cdot)$, the EVI measures the heaviness of the right-tail function

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

and the heavier the right tail, the larger ξ is.

In this article we work with Pareto-type underlying CDFs, with a positive EVI, or equivalently, models such that $\overline{F}(x) = x^{-1/\xi}L(x)$, $\xi > 0$, with $L \in \mathcal{R}_0$, a slowly varying function at infinity, i.e. a regularly varying function with an index of regular variation equal to zero. These heavy-tailed models are quite common in a large variety of fields of application, like bibliometrics, biostatistics, computer science, insurance, finance, social sciences and telecommunications, among others.

For Pareto-type models, the classical EVI-estimators are the *Hill* (H) estimators (Hill, 1975), which are the averages of the log-excesses,

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \le i \le k < n. \tag{1.3}$$

We thus have

$$\hat{\xi}^{\rm H}(k) \equiv {\rm H}(k) := \frac{1}{k}\sum_{i=1}^{k} V_{ik}, \quad 1 \le k < n. \tag{1.4}$$

One of the interesting facts concerning the H EVI-estimator is that various asymptotically equivalent versions of H(k) can be derived through essentially different methods, such as the *maximum likelihood* (ML) method or the mean excess function approach, showing that the Hill estimator is very natural. Details can be found in Beirlant *et al.* (2004), among others. We merely note that from a quantile point of view, i.e. with $F^{\leftarrow}(x) := \inf\{y : F(y) \ge x\}$ denoting the generalised inverse function of F, and

$$U(t) := F^{\leftarrow}(1 - 1/t), \quad t \ge 1, \tag{1.5}$$

the reciprocal tail quantile function, we can write the distributional identity $X \stackrel{d}{=} U(Y)$, with Y a unit Pareto RV, i.e. a RV with a CDF $F_{_Y}(y) = 1 - 1/y,\ y \ge 1$. For the OSs associated with a random unit Pareto sample $(Y_1, \ldots, Y_n)$, we have the distributional identity

$$\frac{Y_{n-i+1:n}}{Y_{n-k:n}} \stackrel{d}{=} Y_{k-i+1:k}, \quad 1 \le i \le k.$$

Moreover, $kY_{n-k:n}/n \underset{n\to\infty}{\overset{p}{\longrightarrow}} 1$, i.e. $Y_{n-k:n} \overset{p}{\sim} n/k$. Consequently, and provided that $k = k_n$, $1 \le k < n$, is an intermediate sequence of integers, i.e. if

$$k = k_n \to \infty \qquad \text{and} \qquad k_n = o(n), \quad \text{as } n \to \infty, \tag{1.6}$$

we get

$$V_{ik} := \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} = \ln \frac{U(Y_{n-i+1:n})}{U(Y_{n-k:n})} = \ln \frac{U(Y_{n-k:n}\ Y_{k-i+1:k})}{U(Y_{n-k:n})}$$
$$= \xi \ln Y_{k-i+1:k}\ (1 + o_p(1)) = \xi E_{k-i+1:k}(1 + o_p(1)), \quad (1.7)$$

with E denoting a standard exponential RV and the $o_p(1)$-term uniform in i, $1 \le i \le k$. Hence, we have the approximation

$$V_{ik} \approx \xi \ln Y_{k-i+1:k} = \xi E_{k-i+1:k}, \quad 1 \le i \le k.$$

The log-excesses, V_{ik}, $1 \le i \le k$, in (1.3), are thus approximately the k top OSs of a sample of size k from an exponential parent with mean value ξ. This justifies the H EVI-estimator in (1.4), the average of the k log-excesses in (1.3).

Apart from the average of the log-excesses, the p-moments of log-excesses, i.e.

$$M_{k,n}^{(p)} := \frac{1}{k}\sum_{i=1}^{k} \left\{ \ln X_{n-i+1:n} - \ln X_{n-k:n} \right\}^p, \quad p \ge 1, \quad (1.8)$$

introduced in Dekkers *at al.* (1989) $\left[M_{k,n}^{(1)} \equiv \mathrm{H}(k)\right]$ have also played a relevant role in the estimation of the EVI, and can more generally be parameterized in $p \in \mathbb{R} \setminus \{0\}$. And a simple generalization of the mean is Lehmer's mean of order p (see Havil, 2003). Given a set of positive numbers $\underline{\mathbf{a}} = (a_1, \ldots, a_k)$, such a mean generalizes both the arithmetic mean ($p = 1$) and the harmonic mean ($p = 0$), being defined as

$$L_p(\underline{\mathbf{a}}) := \frac{\sum_{i=1}^{k} a_i^p}{\sum_{i=1}^{k} a_i^{p-1}}, \quad p \in \mathbb{R}.$$

Further note that $\lim_{p\to-\infty} L_p(\underline{\mathbf{a}}) = \min_{1\le i\le k} a_i$ and $\lim_{p\to+\infty} L_p(\underline{\mathbf{a}}) = \max_{1\le i\le k} a_i$.

The H EVI-estimator can thus be considered as the Lehmer mean of order 1 of the k log-excesses $\underline{\mathbf{V}} := (V_{ik},\ 1 \le i \le k)$, in (1.3). We now more generally consider the Lehmer mean of order p of those statistics. Since from (1.7),

$$V_{ik}^p \ \approx\ \xi^p E_{k-i+1:k}^p, \quad 1 \le i \le k,$$

and $\mathbb{E}(E^p) = \Gamma(p+1)$ for any real $p > -1$, with $\Gamma(\cdot)$ denoting the Gamma function, the law of large numbers enables us to say that

$$\frac{1}{k}\sum_{i=1}^{k} V_{ik}^p \underset{n\to\infty}{\overset{p}{\longrightarrow}} \Gamma(p+1)\xi^p.$$

Hence the reason for the class of *Lehmer* (L) EVI-estimators,

$$\hat{\xi}^{\mathrm{L}_p}(k) \equiv \mathrm{L}_p(k) := \frac{L_p(\underline{\mathbf{V}})}{p} = \frac{1}{p}\,\frac{\sum_{i=1}^{k} V_{ik}^p}{\sum_{i=1}^{k} V_{ik}^{p-1}} = \frac{M_{k,n}^{(p)}}{pM_{k,n}^{(p-1)}} \qquad \Big[\mathrm{L}_1(k) \equiv \mathrm{H}(k)\Big], \tag{1.9}$$

consistent for all $\xi \geq 0$ and real $p > 0$, and where $M_{k,n}^{(p)}$ is given in (1.8).

As a possible competitive class of EVI-estimators, we further refer the one recently studied in Brilhante *et al.* (2013) and Gomes and Caeiro (2014) (see also Paulaskas and Vaičiulis, 2013, 2015; Brilhante *et al.*, 2014; Caeiro *et al.*, 2015; and Gomes *et al.*, 2015a, 2016a, among others.) On the basis of the fact that the Hill EVI-estimator in (1.4) is the logarithm of the *geometric mean* of the statistics $U_{ik} := X_{n-i+1:n}/X_{n-k:n}$, the consideration of the Hölder's *mean of order-p* (MO_p) of those same statistics led to the so-called MO_p EVI-estimators

$$\hat{\xi}^{\mathrm{H}_p}(k) \equiv \mathrm{H}_p(k) := \begin{cases} \left(1 - \left(\frac{1}{k}\sum_{i=1}^{k} U_{ik}^p\right)^{-1}\right)/p, & \text{if} \quad p \neq 0 < 1/\xi, \\ \mathrm{H}(k), & \text{if} \quad p = 0. \end{cases} \tag{1.10}$$

As shown in Brilhante *et al.* (2013), this is a very flexible class of EVI-estimators, which is even able to overpass, for finite sample size n and a wide variety of underlying parents F, one of the simplest and one of the most efficient EVI-estimators in the literature, the *reduced-bias* (RB) *corrected-Hill* (CH) EVI-estimators in Caeiro *et al.* (2005), to be introduced in Section 2.2.

In this article, after the introduction, in Section 2, of a few technical details in the field of EVT, we deal in Section 3 with a few details on the asymptotic behaviour of the class of L_p EVI-estimators, in (1.9). In Section 4, after showing that at optimal k-levels and for the optimal p, the members of such a class are able to overall outperform the optimal EVI-estimators in (1.10), which on its turn had been shown in Brilhante *et al.* (2013) to have a similar behavior comparatively with the optimal Hill EVI-estimator, we compare asymptotically, at optimal levels, a large set of alternative classes of EVI-estimators, drawing some concluding remarks. Section 5 is dedicated to the finite sample properties of the new class of L_p EVI-estimators, comparatively with the behaviour of some of the aforementioned EVI-estimators, done through a large-scale Monte-Carlo simulation study, and giving emphasis to a few peculiar discrepancies between the asymptotic and the finite sample behaviour of these EVI-estimators. In Section 6, we put forward a method similar to the one in Brilhante *et al.* (2013), for the adaptive choice of the vector (k, p) of tuning parameters, based on the bootstrap methodology, but implementation of the method and application to real and simulated data is beyond the scope of this article. Finally, in the Appendix, we provide the proof of the asymptotic normal behavior of the L_p EVI-estimators for all finite $p \geq 1$.

2 Preliminary results in the area of EVT

After a reference, in Section 2.1, to the most common first and second-order frameworks for heavy-tailed models, we briefly review, in Sections 2.2 and 2.3, some extra popular EVI-estimators. Recent reviews on the topic of statistical univariate EVT can be found in Beirlant *et al.* (2012) and Gomes and Guillou (2015).

2.1 A brief review of first and second-order conditions

In the area of statistical EVT and whenever working with large values, a model F is commonly said to be heavy-tailed whenever the right tail function $\overline{F}$, in (1.2), is a regularly varying function with a negative index of regular variation equal to $-1/\xi$, $\xi > 0$, or equivalently, the reciprocal quantile function U, in (1.5), is of regular variation with an index ξ, i.e.

$$F \in \mathcal{D}^+_{\mathcal{M}} := \mathcal{D}_{\mathcal{M}} \left(EV_\xi\right)_{\xi>0} \quad \Longleftrightarrow \quad \overline{F} \in \mathrm{RV}_{-1/\xi} \quad \Longleftrightarrow \quad U \in \mathrm{RV}_\xi \tag{2.1}$$

for all $x > 0$ (Gnedenko, 1943; de Haan, 1984).

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

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

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). This condition has been widely accepted as an appropriate condition to specify the right-tail of a Pareto-type distribution in a semi-parametric way. For technical simplicity, we often assume that we are working in Hall-Welsh class of models (Hall and Welsh, 1985), with a right tail function,

$$\overline{F}(x) = \big(x/C\big)^{-1/\xi}\Big(1 + \beta(x/C)^{\rho/\xi}/\rho + o(x^{\rho/\xi})\Big), \quad \text{as } x \to \infty,$$

$C > 0$, $\beta \neq 0$ and $\rho < 0$. Equivalently, we can say that, with (β, ρ) the vector of second-order parameters, the general second-order condition in (2.2) holds with $A(t) = \xi\beta t^\rho,\ \rho < 0$. Also, and equivalently,

$$U(t) = C\ t^\xi\Big(1 + \xi\ \beta\ t^\rho/\rho + o(t^\rho)\Big), \quad \text{as } t \to \infty. \tag{2.3}$$

Models like the log-Gamma ($\rho = 0$) 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, $F_\xi(x) = \exp(-x^{-1/\xi}),\ x \geq 0$, and the Student's t_ν CDFs, among others, belong to Hall-Welsh class. Further details on first and second-order conditions can be found in Beirlant *et al.* (2004), de Haan and Ferreira (2006) and Fraga Alves *et al.* (2007), among others.

2.2 Explicit EVI-estimators under consideration

Due to its simplicity, the most popular EVI-estimator, valid only for $\xi \geq 0$, is the Hill estimator in (1.4). Apart from the Hill and the aforementioned MO_p EVI-estimators estimators in (1.10), we shall consider the now well-known *moment* (M) EVI-estimators, studied in Dekkers *et al.* (1989), based on $\big(M_{k,n}^{(1)}, M_{k,n}^{(2)}\big)$, with $M_{k,n}^{(p)}$ defined in (1.8). They are given by

$$\hat{\xi}^{\text{M}}(k) \equiv \text{M}(k) := M_{k,n}^{(1)} + \tfrac{1}{2}\Big\{1 - \Big(M_{k,n}^{(2)}/\big(M_{k,n}^{(1)}\big)^2 - 1\Big)^{-1}\Big\}, \tag{2.4}$$

consistent for all $\xi \in \mathbb{R}$. We also mention the possibly RB EVI-estimators introduced and studied in Gomes and Martins (2001), consistent for $\xi \geq 0$ and $p > -1$,

$$\hat{\xi}^{\mathrm{GM}_p}(k) \equiv \mathrm{GM}_p(k) := \frac{M_{k,n}^{(p)}}{\Gamma(p+1)\left[M_{k,n}^{(1)}\right]^{p-1}}$$

$$\Big[\mathrm{GM}_1(k) \equiv \mathrm{H}(k) \equiv \mathrm{L}_1(k), \quad \mathrm{GM}_2(k) \equiv \mathrm{L}_2(k)\Big]. \quad (2.5)$$

This class is a particular case of the also possibly RB class of EVI-estimators in Caeiro and Gomes (2002b) (see also, Caeiro and Gomes, 2002a, 2014b),

$$\hat{\xi}^{\mathrm{CG}_{p,\delta}}(k) \equiv \mathrm{CG}_{p,\delta}(k) := \frac{\Gamma(p)}{M_{k,n}^{(p-1)}}\left(\frac{M_{k,n}^{(\delta p)}}{\Gamma(\delta p+1)}\right)^{1/\delta}, \quad \delta > 0,\ p > 0 \quad \Big[\mathrm{CG}_{1,1}(k) \equiv \mathrm{H}(k)\Big]. \quad (2.6)$$

For $\delta = 2$ in (2.6), we obtain a class studied in Caeiro and Gomes (2002a), which generalizes the estimator $\mathrm{CG}_{1,2}(k) = \sqrt{M_{k,n}^{(2)}/2}$, studied in Gomes *et al.* (2000). And for the L EVI-estimators in (1.9), $\mathrm{L}_p(k) \equiv \mathrm{CG}_{p,1}(k)$.

With the additional notation

$$L_{k,n}^{(j)} := \frac{1}{k}\sum_{i=1}^{k}\left(1 - \frac{X_{n-k:n}}{X_{n-i+1:n}}\right)^j, \quad j \geq 1, \quad (2.7)$$

we also consider in the asymptotic comparison at optimal levels performed in Section 4, the following classes of EVI-estimators:

- The *generalised Hill* (GH) estimator (Beirlant *et al.*, 1996), based on the Hill estimator in (1.4) and with the functional form

  $$\hat{\xi}^{\mathrm{GH}}(k) \equiv \mathrm{GH}(k) := \hat{\xi}^{\mathrm{H}}(k) + \frac{1}{k}\sum_{i=1}^{k}\left\{\ln \hat{\xi}^{\mathrm{H}}(i) - \ln \hat{\xi}^{\mathrm{H}}(k)\right\}, \quad (2.8)$$

  further studied in Beirlant *et al.* (2005).

- The *mixed moment* (MM) estimator (Fraga Alves *et al.*, 2009), based on the statistics $M_{k,n}^{(1)}$ and $L_{k,n}^{(1)}$, respectively given in (1.8) and (2.7), and defined by

  $$\hat{\xi}^{\mathrm{MM}}(k) \equiv \mathrm{MM}(k) := \frac{\hat{\varphi}_{k,n} - 1}{1 + 2\min(\hat{\varphi}_{k,n} - 1, 0)}, \quad \text{with} \quad \hat{\varphi}_{k,n} := \frac{M_{k,n}^{(1)} - L_{k,n}^{(1)}}{\left(L_{k,n}^{(1)}\right)^2}. \quad (2.9)$$

- The *Pareto* (P) *probability weighted moment* (PPWM) class of EVI-estimators introduced in Caeiro and Gomes (2011) (see also Caeiro *et al.*, 2014, 2016), dependent on the statistics,

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

  defined by

$$\hat{\xi}^{\mathrm{PPWM}}(k) \equiv \mathrm{PPWM}(k) := 1 - \frac{\hat{a}_1(k)}{\hat{a}_0(k) - \hat{a}_1(k)}, \tag{2.10}$$

  with $k = 1, 2, \cdots, n-1$, and consistent for $\xi < 1$.

- Just as in de Haan and Ferreira (2006), we further consider, also for $\xi < 1$, the *generalized Pareto* (GP) PWM (GPPWM) EVI-estimators of ξ, based on the sample of exceedances over the high random level $X_{n-k:n}$ and defined by

$$\hat{\xi}^{\mathrm{GPPWM}}(k) \equiv \mathrm{GPPWM}(k) := 1 - \frac{2\hat{a}_1^\star(k)}{\hat{a}_0^\star(k) - 2\hat{a}_1^\star(k)}, \tag{2.11}$$

  with $k = 1, 2, \ldots, n-1$, and

$$\hat{a}_s^\star(k) := \frac{1}{k}\sum_{i=1}^{k} \left(\frac{i-1}{k-1}\right)^s (X_{n-i+1:n} - X_{n-k:n}), \quad s = 0, 1.$$

The estimators in (2.4), (2.8) and (2.9) are consistent for any $\xi \in \mathbb{R}$.

In the simulation study we consider the simplest class of CH EVI-estimators, the one introduced in Caeiro *et al.* (2005),

$$\hat{\xi}^{\mathrm{CH}}(k) \equiv \hat{\xi}^{\mathrm{CH}}_{\hat{\alpha},\hat{\beta},\hat{\rho}}(k) \equiv \mathrm{CH}(k) := \hat{\xi}^{\mathrm{H}}(k)\Big(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})\Big). \tag{2.12}$$

The estimators in (2.12) can be second-order *minimum-variance reduced-bias* (MVRB) estimators, for adequate levels k and an adequate external estimation of the vector of second-order parameters, (β, ρ), in (2.3), i.e., the use of $\hat{\xi}^{\mathrm{CH}}(k)$, with an adequate estimation of (β, ρ), enables us to eliminate the dominant component of bias of the H EVI-estimator, $\hat{\xi}^{\mathrm{H}}(k)$, keeping its asymptotic variance. For details on algorithms for the (β, ρ)-estimation, see Gomes and Pestana (2007a,b) and Gomes *et al.* (2008b). We have so far suggested the use of the ρ-estimators in Fraga Alves *et al.* (2003) and the β-estimators in Gomes and Martins (2002). However, recent classes of β-estimators (Caeiro and Gomes, 2006; Gomes *et al.*, 2010; Henriques-Rodrigues *et al.*, 2015) and ρ-estimators (Goegebeur *et al.*, 2008, 2010; Ciuperca and Mercadier, 2010; de Wet *et al.*, 2012; Worms and Worms, 2012; Deme *et al.*, 2013; Caeiro and Gomes, 2014a, 2015a; Henriques-Rodrigues *et al.*, 2014) are potential candidates for the (β, ρ)-estimation. Overviews on reduced-bias estimation can be found in Chapter 6 of Reiss and Thomas (2007), Beirlant *et al.* (2012) and Gomes and Guillou (2015).

2.3 ML EVI-estimators

As mentioned in de Haan and Ferreira (2006), the class of CDFs $F \in \mathcal{D}_{\mathcal{M}}(\mathrm{EV}_\xi)$, either for $\xi \geq 0$ or, more generally, for $\xi \in \mathbb{R}$, cannot be parameterised with a finite number of parameters, and consequently, there does not exist an ML estimator for ξ in such a wide class of models. There exists however an estimator, introduced by Smith (1987), usually denoted as the ML EVI-estimator. Such an estimator is based on the excesses over a deterministic high level u, but can be easily rephrased on the basis of the excesses over the high random threshold $X_{n-k:n}$,

$$W_{ik} := X_{n-i+1:n} - X_{n-k:n}, \quad 1 \leq i \leq k < n. \tag{2.13}$$

For models in (2.1), these k excesses are approximately distributed as the whole set of the k top OSs associated with a sample of size k from a GP CDF, $\mathrm{GP}(x; \xi, \alpha) = 1 - (1 + \alpha\ x)^{-1/\xi}$, $x > 0$, $(\alpha,\ \xi > 0)$, a quite relevant re-parametrization due to Davison (1984). Indeed, αW_{ik} is well approximated by $Y^{\xi}_{k-i+1:k} - 1$, with Y a unit Pareto RV. The solution of the associated ML equations gives then rise to an explicit expression for the ML EVI-estimator, a function of the ML implicit estimator $\hat{\alpha}_{\text{ML}}$ of α and the sample of excesses, given by

$$\hat{\xi}^{\text{ML}}(k) \equiv \hat{\xi}^{\text{ML}}(k, \hat{\alpha}_{\text{ML}}) \equiv \text{ML}(k) := \frac{1}{k} \sum_{i=1}^{k} \ln(1 + \hat{\alpha}_{\text{ML}}\ W_{ik}). \tag{2.14}$$

A comprehensive study of the asymptotic properties of the ML estimator in (2.14) has been undertaken in Drees *et al.* (2004). As recently shown by Zhou (2009, 2010), such an EVI–estimator is also consistent for $\xi > -1$. We can also consider the random threshold $X_{n-k:n}$ replaced by a deterministic threshold u, working then under the POT methodology, introduced in Smith (1987).

Remark 1. *A simple heuristic estimator of* α *is* $1/X_{n-k:n}$. *If we consider* $\hat\alpha = 1/X_{n-k:n}$ *and the excesses* W_{ik}, $1 \le i \le k$, *in* (2.13), $1 + \hat\alpha\ W_{ik} = X_{n-i+1:n}/X_{n-k:n}$. *Then,* $\hat\xi^{\rm ML}(k, \hat\alpha) = \frac{1}{k}\sum_{i=1}^{k}\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}$ *is the average of the log-excesses* V_{ik}, $1 \le i \le k$, *i.e. it is the classical* H EVI-*estimator in* (1.4).

Remark 2. *Note that all the aforementioned* EVI-*estimators are scale invariant. The* GPPWM *estimators, in* (2.11) *and the* ML *estimators in* (2.14) *are also location invariant, and can be regarded as classes of peaks over random threshold* (PORT) EVI-*estimators, the acronymous introduced in* Araújo Santos *et al.* (2006) *for estimators of parameters of rare events based on excesses over a central empirical quantile and even over the minimum of the available sample whenever possible, i.e. when the underlying parent* F *has a finite left endpoint* (*see* Gomes *et al.*, 2008a, *for details on the topic*). *Further* PORT EVI-*estimation can be found in* Gomes *et al.* (2011a; 2012a; 2013a; 2015c), Gomes and Henriques-Rodrigues (2013) *and* Caeiro *et al.* (2016).

Remark 3. *Further note that the* MM EVI-*estimators, in* (2.9), *are very close to the* ML EVI-*estimators for a wide class of models with* $\xi \ge 0$ (*see* Fraga Alves *et al.*, 2009).

3 Asymptotic behaviour of the EVI-estimators

Under the validity of the second-order condition in (2.2), trivial adaptations of the results in de Haan and Peng (1998), Drees *et al.* (2004), Beirlant *et al.* (2005), Caeiro *et al.* (2005), de Haan and Ferreira (2006), Fraga Alves *et al.* (2009) and Caeiro and Gomes (2011), respectively for the H, ML, GH, CH, GPPWM, MM and PPWM EVI-estimators, enable us to state the following theorem.

Theorem 1. *Assume that condition* (2.2) *holds. Let* $k = k_n$ *be such that* (1.6) *holds, and let us 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* $\xi > 0$, *the* H, M, GH, MM, PPWM, GPPWM *and* ML EVI-*estimators, generally denoted* $\hat\xi^\bullet(k)$, *and respectively defined in* (1.4), (2.4), (2.8), (2.9), (2.10), (2.11) *and* (2.14), *are asymptotically normal, i.e.*

$$\sqrt{k}\left(\hat\xi^\bullet(k) - \xi\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{N}\left(\lambda_A b_\bullet, \sigma_\bullet^2\right), \tag{3.1}$$

where $\mathcal{N}(\mu, \sigma^2)$ *stands for a normal* RV *with mean value* μ *and variance* σ^2,

$$b_{\rm H} = \frac{1}{1-\rho}, \quad \sigma^2_{\rm H} = \xi^2, \quad b_{\rm M} = b_{\rm GH} = \frac{\xi - \xi\rho + \rho}{\xi(1-\rho)^2}, \quad \sigma^2_{\rm M} = \sigma^2_{\rm GH} = 1 + \xi^2,$$

$$b_{\rm MM} = b_{\rm ML} = \frac{(1+\xi)(\xi+\rho)}{\xi(1-\rho)(1+\xi-\rho)}, \quad \sigma^2_{\rm MM} = \sigma^2_{\rm ML} = (1+\xi)^2,$$

and for $\xi < 1/2$,

$$b_{\rm PPWM} = \frac{(1-\xi)(2-\xi)}{(1-\xi-\rho)(2-\xi-\rho)}, \qquad \sigma^2_{\rm PPWM} = \frac{\xi^2(1-\xi)(2-\xi)^2}{(1-2\xi)(3-2\xi)},$$

$$b_{\rm GPPWM} = \frac{(\xi+\rho)\ b_{\rm PPWM}}{\xi} \qquad \text{and} \qquad \sigma^2_{\rm GPPWM} = \frac{(1-\xi+2\xi^2)(1-\xi)(2-\xi)^2}{(1-2\xi)(3-2\xi)}.$$

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* $b_{\rm CH} = 0$ *and* $\sigma^2_{\rm CH} = \sigma^2_{\rm H} = \xi^2$*, for the* RB EVI-*estimator in* (2.12).

3.1 Further details on the asymptotic behaviour of the EVI-estimators under consideration

For the EVI-estimators dependent on a tuning parameter p, trivial adaptations of the results in Gomes and Martins (2001), Caeiro and Gomes (2002b) and Brilhante *et al.* (2013), respectively for the GM_p, $\text{CG}_{p,\delta}$ and H_p classes of EVI-estimators, enable us to state the following theorem.

Theorem 2. *Under the validity of the first-order condition, in* (2.1)*, and for intermediate sequences* $k = k_n$*, i.e. if* (1.6) *holds, the classes* H_p*,* GM_p *and* $\text{CG}_{p,\delta}$*, respectively defined in* (1.10)*,* (2.5) *and* (2.6)*, are consistent for the estimation of* $\xi \ge 0$*, provided that* $p \in \mathcal{R}_\bullet$*, where* $\mathcal{R}_{\rm H} = \{p : p \le 1/\xi\}$*,* $\mathcal{R}_{\rm GM} = \{p : p > -1\}$ *and* $\mathcal{R}_{\rm CG} = \{(p,\delta) : p > 0,\ \delta > -1/p\}$.

Under the conditions of **Theorem 1***, for any of the estimators in* (1.10)*,* (2.5) *and* (2.6)*, also generally denoted* $\hat\xi^\bullet(k)$*, and for adequate regions of the tuning parameters* p *and* δ*,* (3.1) *holds, with*

$$b_{\text{GM}_p} = \frac{1-(1-\rho)^p}{\rho(1-\rho)^p} - \frac{p-1}{1-\rho}, \quad \sigma^2_{\text{GM}_p} = \xi^2 \left\{ \frac{\Gamma(2p+1)}{\Gamma^2(p+1)} - p^2 \right\} \quad (p > -1/2), \tag{3.2}$$

$$
b_{\mathrm{CG}_{p,\delta}} = \frac{(1-\rho)^{-\delta p} - \delta(1-\rho)^{-p+1} + \delta - 1}{\delta\rho},
$$

$$
\sigma^2_{\mathrm{CG}_{p,\delta}} = \frac{\xi^2}{\delta^2}\left\{\frac{2\Gamma(2\delta p)}{\delta p \Gamma^2(\delta p)} + \frac{\delta^2\Gamma(2p-1)}{\Gamma^2(p)} - \frac{2\Gamma((\delta+1)p)}{p\Gamma(p)\Gamma(\delta p)} - (\delta-1)^2\right\} \quad (p \geq 1,\ \delta > 0), \quad (3.3)
$$

$$
b_{\mathrm{H}_p} = \frac{1-p\xi}{1-p\xi-\rho} \quad \sigma^2_{\mathrm{H}_p} = \frac{\xi^2(1-p\xi)^2}{1-2p\xi} \quad (p < 1/(2\xi). \quad (3.4)
$$

For the particular case $\delta = 1$, in (2.6), i.e. for the L_p EVI-estimator in (1.9), we can state:

Corollary 1. *Under the validity of the initial first-order conditions in* **Theorem 2**, *the class of* L_p EVI-*estimator sin* (1.9), *is consistent for the estimation of* $\xi \geq 0$, *provided that* $p \in \mathcal{R}_\mathrm{L} = \{p : p > 0\}$. *Under the conditions of* **Theorem 1**, (3.1) *holds, with*

$$
b_{\mathrm{L}_p} = \frac{1}{(1-\rho)^p} \quad \text{and} \quad \sigma^2_{\mathrm{L}_p} = \frac{\xi^2\ \Gamma(2p-1)}{\Gamma^2(p)} \quad (p \geq 1). \quad (3.5)
$$

For an isolated proof of the **Corollary 1** related to the L_p EVI-estimators, in (1.9), see the **Appendix**.

Remark 4. *Note that regarding the* L_p EVI-*estimators, in* (1.9), **Corollary1** *is a particular case of* **Theorem 1** *in* Caeiro and Gomes (2002b), *but trivially generalizing consistency for* $p > 0$ *rather than* $p \geq 1$. *Further note that there is a full agreement between* (3.5) *and* (3.3), *the results provided in* **Theorem 1** *of* Caeiro and Gomes (2002b), *whenever* $\delta = 1$.

Remark 5. *More specifically than in* **Corollary 1**, *note that the validity of the second-order condition in* (2.2) *and* $p \geq 1$, *enables us to write for all* $\rho \leq 0$ *the asymptotic distributional representation*

$$
\mathrm{L}_p(k) \overset{d}{=} \xi + \frac{\sigma_{\mathrm{L}_p}\ Z_k^{(p)}}{\sqrt{k}} + b_{\mathrm{L}_p}\ A(n/k) + o_p(A(n/k)) \quad (3.6)
$$

with $(b_{\mathrm{L}_p}, \sigma^2_{\mathrm{L}_p})$ *given in* (3.5), *and where* $Z_k^{(p)}$ *is an asymptotically standard normal* RV. *A similar result for* $p = 1$, *i.e. for the* H EVI-*estimator was derived in* de Haan and Peng (1998).

Remark 6. *Note that for* $p = 2$, *in* (1.8), *and a heavy tail* (*see* Gomes *et al.*, 2000), *we get*

$$
M_{k,n}^{(2)} \overset{d}{=} 2\xi^2 + \frac{\xi^2\ P_n^{(2)}}{\sqrt{k}} + \frac{2\xi(2-\rho)A(n/k)}{(1-\rho)^2} + o_p(A(n/k),
$$

where $P_n^{(2)} \stackrel{d}{=} \sqrt{k}\{\sum_{i=1}^k E_i^2/k - 2\}$, *and* $(P_n^{(1)}, P_n^{(2)})$ *is asymptotically bivariate normal with null mean and covariance matrix* $\Sigma_2 = \begin{bmatrix} 1 & 4 \\ 4 & 20 \end{bmatrix}$ (Dekkers *et al.*, 1989). *Further note that an alternative to the* H EVI-*estimator, and related to the* EVI-*estimator in* (1.9), *but with* $p = 2$, *was considered in* de Haan and Peng (1998), *and was there attributed to* Casper de Vries. *For such estimator,* $\mathrm{L}_2(k) = M_{k,n}^2/\big(2M_{k,n}^{(1)}\big)$, Gomes *et al.* (2000) *derived the asymptotic distributional representation*

$$\mathrm{L}_2(k) \stackrel{d}{=} \xi + \frac{\xi\sqrt{2}Z_n^{(2)}}{\sqrt{k}} + \frac{A(n/k)}{(1-\rho)^2} + o_p(A(n/k)) + o_p(1/\sqrt{k}),$$

where $Z_n^{(2)} = \big(P_n^{(2)}/2 - P_n^{(1)}\big)/\sqrt{2}$ *is asymptotically* Normal(0,1). *For an asymptotic comparison at optimal levels of* L_2 *and* L_1, *see* Gomes *et al.* (2000).

Remark 7. Gomes and Martins (2001) *considered* $M_{k,n}^{(p)}$, *in* (1.8), *and got the asymptotic distributional representation*

$$M_{k,n}^{(p)} \stackrel{d}{=} \xi^p \left\{ \Gamma(p+1) + \frac{P_n^{(p)}}{\sqrt{k}} + \frac{\Gamma(p+1)\,(1-(1-\rho)^p)\,A(n/k)}{\xi\rho(1-\rho)^p} \right\} + o_p(A(n/k)), \qquad (3.7)$$

where $(P_n^{(1)}, P_n^{(p)})$ *is asymptotically a bivariate normal vector, with zero mean and covariance matrix*

$$\Sigma_p = \begin{bmatrix} 1 & p\Gamma(p+1) \\ p\Gamma(p+1) & \Gamma(2p+1) - \Gamma^2(p+1) \end{bmatrix}.$$

In the above mentioned article, the class of EVI-*estimators in* (2.5) *was introduced and studied both asymptotically and for finite samples. Then, an asymptotic distributional representation of the type of the one in* (3.6) *holds for the* EVI-*estimator in* (2.5) *and* $p \geq 1$, *with* $\big(\sigma^2_{\mathrm{L}_p},\ b_{\mathrm{L}_p}\big)$ *replaced by* $\big(\sigma^2_{\mathrm{GM}_p},\ b_{\mathrm{GM}_p}\big)$, *in* (3.2). *This class of* EVI-*estimators can be second-order* RB, *i.e. there exists a non-explicit value of* p_0 *such that* $b_{\mathrm{GM}_{p_0}} = 0$.

Remark 8. *Further note that for the* MO_p EVI-*estimators, now denoted* H_p *and defined in* (1.10), *a distributional representation of the type of the one in* (3.6) *holds for* $p < 1/(2\xi)$, *with* $\big(\sigma^2_{\mathrm{H}_p}, b_{\mathrm{H}_p}\big)$ *given in* (3.4).

For any $\xi > 0$, the asymptotic variance $\sigma^2_{\mathrm{L}_p}(\xi)$, in (3.5), has a minimum at $p = 1$. In Figure 1 (*left*), we present the normalized standard deviation, $\sigma_{\mathrm{L}_p}(\xi)/\xi$, independent of ξ, as a function of p. On the other side, the asymptotic bias ruler, $b_{\mathrm{L}_p}(\rho)$, also in (3.5), is independent of ξ and always decreasing in p. Such a performance is shown in Figure 1 (*right*).

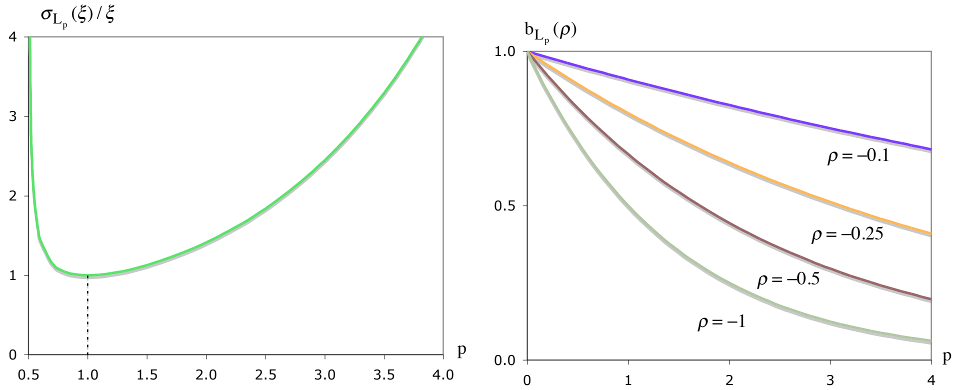

Figure 1: Graph of $\sigma_{\mathrm{L}_p}(\xi)/\xi$, as a function of $p > 0.5$ (*left*) and of the asymptotic bias ruler $b_{\mathrm{L}_p}(\rho)$, for $\rho = -0.1, -0.25, -0.5$ and -1, as a function of $p \geq 0$

These aforementioned results claim for an asymptotic comparison, at the optimal k and (k, p)-values, of the class of EVI-estimators in (1.10), a topic to be dealt with in Section 4. In such a comparison we shall consider all EVI-estimators non-dependent of p, together with the L_p and H_p EVI-estimators, respectively given in (1.9) and in (1.10), computed at the optimal p. The classes GM_p and $\mathrm{CG}_{p,\delta},\ \delta > 1$, at optimal tuning parameters' values, will be excluded due to the fact that such classes can be second-order RB in the whole (ξ, ρ)-plane.

4 Asymptotic comparison at optimal levels

We next proceed to the comparison of 'classical' EVI-estimators at their optimal levels. This is again done in a way similar to the one used in de Haan and Peng (1998), Gomes and Martins (2001), Gomes *et al.* (2005, 2007, 2013b, 2015b), Gomes and Neves (2008), Gomes and Henriques-Rodrigues (2010) and Brilhante *et al.* (2013). Let us assume that $\hat{\xi}^{\bullet}(k)$ denotes any arbitrary semi-parametric EVI-estimator, for which we have the asymptotic distributional representation

$$\hat{\xi}^{\bullet}(k) = \xi + \frac{\sigma_{\bullet}\ Z_k^{\bullet}}{\sqrt{k}} + b_{\bullet}\ A(n/k) + o_p(A(n/k)), \tag{4.1}$$

for any intermediate sequence of integers $k = k_n$, and where $Z_k^{\bullet}$ is asymptotically standard normal. Then, $\sqrt{k}\big(\hat{\xi}^{\bullet}(k) - \xi\big) \overset{d}{\to} N(\lambda_{_A} b_{\bullet}, \sigma_{\bullet}^2)$ provided that k is such that $\sqrt{k}\ A(n/k) \to \lambda_{_A}$, finite, as $n \to \infty$. We then write $\mathrm{Bias}_{\infty}\big(\hat{\xi}^{\bullet}(k)\big) := b_{\bullet}\ A(n/k)$ and $\mathrm{Var}_{\infty}\big(\hat{\xi}^{\bullet}(k)\big) := \sigma_{\bullet}^2/k$. The so-called *asymptotic mean square error* (AMSE) is then given by

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

Regular variation theory (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

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

where $\hat{\xi}_{n0}^{\bullet} := \hat{\xi}^{\bullet}(k_{0|\bullet}(n))$ and $k_{0|\bullet}(n) := \arg\min_k \mathrm{MSE}\left(\hat{\xi}^{\bullet}(k)\right)$. Moreover, if we slightly restrict the second-order condition in (2.2), assuming that $A(t) = \xi\beta t^{\rho}$, $\rho < 0$, just as happens for the class in (2.3), we can write

$$k_{0|\bullet}(n) := \arg\min_k \mathrm{MSE}\left(\hat{\xi}^{\bullet}(k)\right) = \left(\frac{\sigma_{\bullet}^2\ n^{-2\rho}}{b_{\bullet}^2 \xi^2 \beta^2(-2\rho)}\right)^{1/(1-2\rho)} (1 + o(1)).$$

We again consider the following:

Definition 1. *Given two biased estimators* $\hat{\xi}^{(1)}(k)$ *and* $\hat{\xi}^{(2)}(k)$*, for which a distributional representation of the type of the one in* (4.1) *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 root efficiency* (AREFF) *of* $\hat{\xi}_{n0}^{(1)}$ *relatively to* $\hat{\xi}_{n0}^{(2)}$ *is*

$$\mathrm{AREFF}_{1|2} \equiv \mathrm{AREFF}_{\hat{\xi}_{n0}^{(1)}|\hat{\xi}_{n0}^{(2)}} := \sqrt{\mathrm{LMSE}\big(\hat{\xi}_{n0}^{(2)}\big)/\mathrm{LMSE}\big(\hat{\xi}_{n0}^{(1)}\big)} = \left(\left(\frac{\sigma_2}{\sigma_1}\right)^{-2\rho}\left|\frac{b_2}{b_1}\right|\right)^{\frac{1}{1-2\rho}}. \tag{4.2}$$

Remark 9. *Note that the* AREFF*-indicator, in* (4.2)*, has been conceived so that the highest the* AREFF *indicator is, the better is the first estimator.*

Remark 10. *Further note that some of the aforementioned 'classical'* EVI-*estimators can be second-order* RB *in some regions of the* (ξ,ρ)*-plane. We can thus not apply* **Definition 1**, *because* $b_2 = 0$. *This happens with the* MM, *the* GPPWM *and the* ML EVI-*estimators, in* (2.9), (2.11) *and* (2.14), *respectively, all second-order* RB EVI-*estimators in the region* $\xi+\rho=0$ $\big($*where* $b_{\rm MM} = b_{\rm GPPWM} = b_{\rm ML} = 0\big)$. *Consequently, and for* $\xi+\rho=0$, *they are expected to asymptotically outperform at optimal levels any of the other* EVI-*estimators. Despite of the fact that* $\sigma_{\rm GPPWM} > \sigma_{\rm MM} = \sigma_{\rm ML} > \sigma_{\rm H} = \sigma_{\rm CH}$, *this does not mean too much. All depends on the dominant component of bias ... and it is without doubt a challenge for further research, out of the scope of this paper, already partially dealt with in* Caeiro *et al.* (2009). *A similar comment applies to the behaviour of the* M *and the* GH EVI-*estimators in the region* $\xi = -\rho/(1-\rho)$ $\big($*where* $b_{\rm M} = b_{\rm GH} = 0\big)$. *Again, despite of the fact that the* M *and the* GH EVI-*estimators have an asymptotic variance equal to* $1+\xi^2 > \xi^2$, *the asymptotic variance of* H *and* CH, *all depends on the comparative behaviour of the mean square errors. At the optimal* p *in the sense of minimal* RMSE, *the classes of* GM_p *and* $\mathrm{CG}_{p,\delta}$, $\delta>1$, EVI-*estimators, respectively given in* (2.5) *and* (2.6) *are second-order* RB EVI-*estimators in the whole* (ξ,ρ)*-plane, and just as the* CH *class, in* (2.12), *will be excluded from the asymptotic comparison at optimal levels. However the* L_p *and* H_p EVI-*estimators, respectively given in* (1.9) *and* (1.10), *are never second-order* RB EVI-*estimators, and will be crucially included in the asymptotic comparison in* Section 4.1.

4.1 Asymptotic comparison of L_p and H_p EVI-estimators at optimal levels

Let us now turn back to the L_p EVI-estimators $\mathrm{L}_p(k)$ in (1.9). We have

$$\mathrm{LMSE}(\mathrm{L}_{0|p}) = \big(\xi^2\Gamma(2p-1)/\Gamma^2(p)\big)^{-\frac{2\rho}{1-2\rho}}\,\big((1-\rho)^{-2p}\big)^{\frac{1}{1-2\rho}}$$

and

$$\mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1} = \left(\left(\Gamma(p)/\sqrt{\Gamma(2p-1)}\right)^{-2\rho}(1-\rho)^{p-1}\right)^{\frac{1}{1-2\rho}}. \qquad (4.3)$$

Remark 11. *In* Gomes *et al.* (2000) *was shown that the asymptotic relative efficiency* (AREFF) *of* $\mathrm{L}_2(k)$ *comparatively to* $\mathrm{L}_1(k)$ *is given by* $\mathrm{AREFF}_{\mathrm{L}_2|\mathrm{L}_1} = [2^\rho(1-\rho)]^{1/(1-2\rho)}$, *in agreement with* (4.3). *As noticed in the aforementioned article,* $\mathrm{AREFF}_{\mathrm{L}_2|\mathrm{L}_1} > 1 \iff -1<\rho<0$.

To measure the performance of $\mathrm{H}_{0|p}$, with H_p the MO_p EVI-estimator in (1.10), Brilhante *et al.* (2013) computed a similar AREFF-indicator, given by

$$\mathrm{AREFF}_{\mathrm{H}_p|\mathrm{H}_0} = \left(\left(\frac{\sqrt{1-2p\xi}}{1-p\xi} \right)^{-2\rho} \left| \frac{1-p\xi-\rho}{(1-\rho)(1-p\xi)} \right| \right)^{\frac{1}{1-2\rho}}, \tag{4.4}$$

reparameterised in $(\rho,\ a = p\xi < 1/2)$, and denoted $\mathrm{AREFF}^*_{a|0}$. In Figure 2, we picture the contour plots of $\mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1}$ in (4.3) (*left*) and of $\mathrm{AREFF}^*_{a|0}$ (*right*), in (4.4).

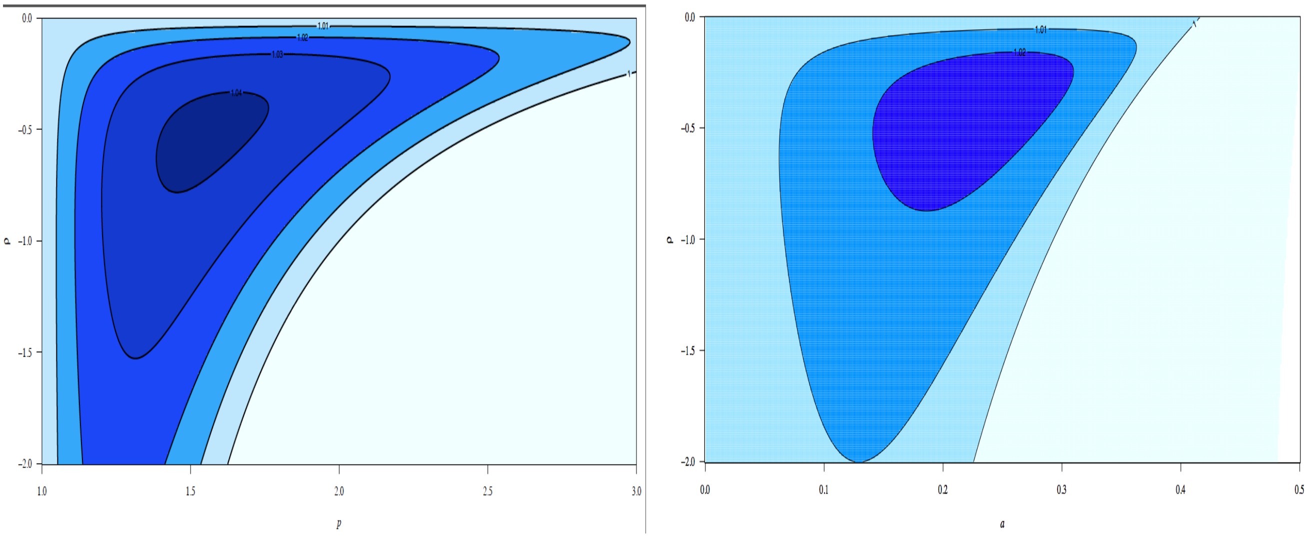

Figure 2: Contour plots of $\mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1}$, in (4.3) (*left*) and of $\mathrm{ARREF}^*_{a|0}$, in (4.4) (*right*)

The gain in efficiency is not terribly high, but, at optimal levels, there is a wide region of the (p,ρ)-plane where the new class of L_p EVI-estimators performs better than the Hill estimators, with efficiencies slightly higher than the ones associated with the comparison of H_p and the Hill, in the (a,ρ)-plane. This result together with the fact that as far as we know, the EVI-estimator in (1.10) computed at the optimal p in the sense of minimal $\mathrm{AREFF}_{\mathrm{H}_p|\mathrm{H}_0}$, i.e. at $p_{\mathrm{M|H}} \equiv p_{\mathrm{M|H}}(\rho) := \arg\max_p \mathrm{AREFF}_{\mathrm{H}_p|\mathrm{H}_0}$, explicitly given by

$$p_{\mathrm{M|H}} = \varphi_\rho/\xi, \quad \text{with} \quad \varphi_\rho := 1-\rho/2-\sqrt{\rho^2-4\rho+2}\Big/2$$

and $b_{p_{\mathrm{M|H}}} \neq 0$, is the unique non-RB EVI-estimator which is able to beat the Hill EVI-estimator in the whole (ξ,ρ)-plane, immediately leads us to think on what happens for the optimal value of p associated with the L_p EVI-estimation.

In Figure 3 (*left*), we picture the indicator $\mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1}$, as a function of p for $|\rho| = 0(0.1)2$. In the same Figure (*right*), the value of $p_{_{\mathrm{M|L}}} = p_{_{\mathrm{M|L}}}(\rho) := \arg\max_p \mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1}$ is pictured as a function of $|\rho|$.

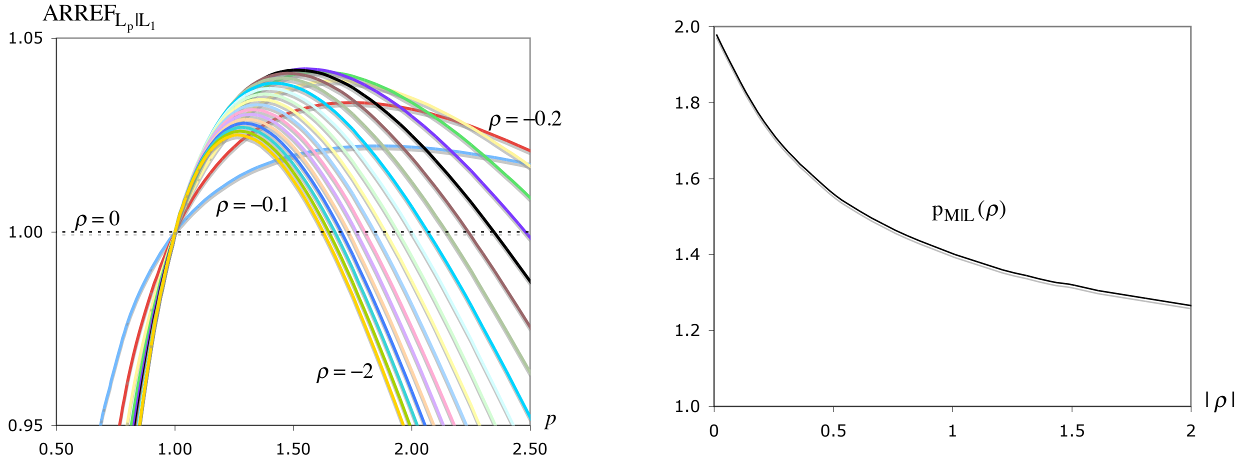

Figure 3: $\mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1}$, as a function of p, for $|\rho| = 0(0.1)2$ (*left*) and the value of $p_{_{\mathrm{M|L}}} = p_{_{\mathrm{M|L}}}(\rho)$, as a function of $|\rho|$ (*right*)

Indeed, just as $\mathrm{AREFF}_{\mathrm{H}_{p_{\mathrm{M|H}}}|\mathrm{H}_0} > 1$, for any $\rho < 0$ and $\xi > 0$, also, at $p_{_{\mathrm{M|L}}} \equiv p_{_{\mathrm{M|L}}}(\rho) := \arg\max_p \mathrm{AREFF}_{\mathrm{L}_p|\mathrm{L}_1}$,

$$\mathrm{AREFF}_{\mathrm{L}_{p_{\mathrm{M|L}}}|\mathrm{L}_1} > 1,$$

for any $\rho < 0$ and $\xi > 0$. Moreover,

$$\mathrm{AREFF}_{\mathrm{L}_{p_{\mathrm{M|L}}}|\mathrm{L}_1} > \mathrm{AREFF}_{\mathrm{H}_{p_{\mathrm{M|H}}}|\mathrm{H}_0},$$

as illustrated in Figure 4.

4.2 An overall comparison of EVI-estimators at optimal levels

As mentioned above and first detected by Brilhante *et al.* (2013), the optimal MO_p EVI-estimator can beat the optimal Hill EVI-estimator in the whole (ξ, ρ)-plane. But it is now beaten by the optimal Lehmer EVI-estimator also in the whole (ξ, ρ), an atypical behaviour among other classical EVI-estimators.

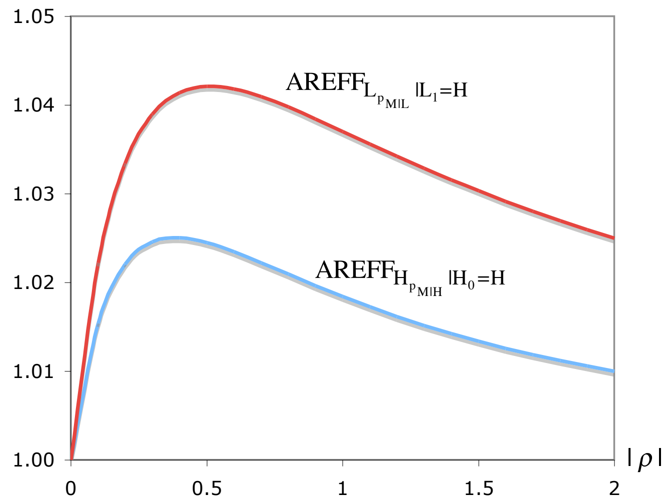

Figure 4: $\text{AREFF}_{\text{L}_{p_{\text{M}|\text{L}}}|\text{L}_1}$ and $\text{AREFF}_{\text{H}_{p_{\text{M}|\text{H}}}|\text{H}_0}$ as a function of $|\rho| = 0(0.1)2$

But again, as happened before with the optimal MO_p EVI-estimator, the optimal Lehmer EVI-estimator can be beaten by the M EVI-estimator in a region close to $\xi = -\rho/(1-\rho)$, where $b_{\text{M}} = 0$. The MM-estimator in (2.9), asymptotically equivalent to the ML-estimator, unless $\xi + \rho = 0$ and $(\xi, \rho) \neq (0, 0)$, outperforms the M EVI-estimator at optimal levels, in a region around $\xi + \rho = 0$, and can even outperform the optimal Lehmer EVI-estimator, as can be seen in Figure 5, where we exhibit the comparative behaviour of all 'classical' EVI-estimators under consideration, including both the L and the H classes (*bottom*), after including only the H class (*top*), as done in Brilhante *et al.* (2013). The GPPWM EVI-estimator is RB for $\xi + \rho = 0$, and can beat the MM EVI-estimator in a short region of the (ξ, ρ)-plane. The PPWM can beat even the optimal Lehmer for a few values of ξ around 0.1.

As expected, none of the estimators can always dominate the alternatives, but the L_p EVI-estimators have a nice performance, being unexpectedly able to beat the $\text{MO}_p \equiv \text{H}_p$ EVI-estimators at optimal levels in the whole (ξ, ρ)-plane.

Remark 12. *As already mentioned in* Brilhante *et al.* (2013), note *that in the region* $\xi + \rho \neq 0$ *and* $\xi \neq -\rho/(1-\rho)$*, the* CH-*estimators, in* (2.12)*, overpass at optimal levels all other classical estimators under consideration. They were thus not included in* Figure 5*, so that we can see the*

ρ \ ξ	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	1.90	2.00
0.00	MM	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*	H*
-0.10	H*	ML	MM	GP	GP	MM	MM	MM	MM	M	M	M	M	M	M	M	M	M	M	M	M
-0.20	H*	P	ML	GP	MM	MM	MM	MM	MM	MM	MM	M	M	M	M	M	M	M	M	M	M
-0.30	H*	P	P	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	M	M	M	M	M	M	M	M
-0.40	H*	P	P	M	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	M	M	M	M	M	M
-0.50	H*	P	P	M	M	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	M	M	M	M
-0.60	H*	P	P	H*	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	M	M
-0.70	H*	P	P	H*	M	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-0.80	H*	P	P	H*	M	M	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-0.90	H*	P	P	H*	H*	M	M	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.00	H*	P	P	H*	H*	M	M	M	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.10	H*	P	H*	H*	H*	M	M	M	H*	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.20	H*	P	H*	H*	H*	M	M	M	H*	H*	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM
-1.30	H*	P	H*	H*	H*	H*	M	M	H*	H*	H*	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM
-1.40	H*	P	H*	H*	H*	H*	M	M	H*	H*	H*	H*	MM	MM	ML	MM	MM	MM	MM	MM	MM
-1.50	H*	P	H*	H*	H*	H*	M	M	H*	H*	H*	H*	MM	MM	MM	ML	MM	MM	MM	MM	MM
-1.60	H*	P	H*	H*	H*	H*	M	M	H*	H*	H*	H*	H*	MM	MM	MM	ML	MM	MM	MM	MM
-1.70	H*	P	H*	H*	H*	H*	M	M	H*	H*	H*	H*	H*	H*	MM	MM	MM	ML	MM	MM	MM
-1.80	H*	P	H*	H*	H*	H*	M	M	H*	H*	H*	H*	H*	H*	H*	MM	MM	MM	ML	MM	MM
-1.90	H*	P	H*	H*	H*	H*	H*	M	H*	H*	H*	H*	H*	H*	H*	H*	MM	MM	MM	ML	MM
-2.00	H*	P	H*	H*	H*	H*	H*	M	H*	H*	H*	H*	H*	H*	H*	H*	MM	MM	MM	MM	ML

ρ \ ξ	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	1.90	2.00
0.00	MM	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*
-0.10	L*	ML	MM	GP	GP	MM	MM	MM	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*
-0.20	L*	L*	ML	GP	MM	MM	MM	MM	MM	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*
-0.30	L*	L*	L*	ML	MM	MM	MM	MM	MM	MM	MM	L*	L*	L*	L*	L*	L*	L*	L*	L*	L*
-0.40	L*	P	L*	M	ML	MM	MM	MM	MM	MM	MM	MM	MM	L*	L*	L*	L*	L*	L*	L*	L*
-0.50	L*	P	L*	M	M	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	L*	L*	L*	L*	L*	L*
-0.60	L*	P	L*	L*	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	L*	L*	L*	L*
-0.70	L*	P	L*	L*	M	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	L*	L*	L*
-0.80	L*	P	L*	L*	M	M	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	L*	L*
-0.90	L*	L*	L*	L*	L*	M	M	M	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.00	L*	L*	L*	L*	L*	M	M	M	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.10	L*	L*	L*	L*	L*	M	M	M	L*	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.20	L*	L*	L*	L*	L*	M	M	M	L*	L*	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM
-1.30	L*	L*	L*	L*	L*	L*	M	M	L*	L*	L*	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM
-1.40	L*	L*	L*	L*	L*	L*	M	M	L*	L*	L*	L*	MM	MM	ML	MM	MM	MM	MM	MM	MM
-1.50	L*	L*	L*	L*	L*	L*	M	M	L*	L*	L*	L*	MM	MM	MM	ML	MM	MM	MM	MM	MM
-1.60	L*	L*	L*	L*	L*	L*	M	M	L*	L*	L*	L*	L*	MM	MM	MM	ML	MM	MM	MM	MM
-1.70	L*	L*	L*	L*	L*	L*	M	M	L*	L*	L*	L*	L*	L*	MM	MM	MM	ML	MM	MM	MM
-1.80	L*	L*	L*	L*	L*	L*	M	M	L*	L*	L*	L*	L*	L*	L*	MM	MM	MM	ML	MM	MM
-1.90	L*	L*	L*	L*	L*	L*	L*	M	L*	L*	L*	L*	L*	L*	L*	L*	MM	MM	MM	ML	MM
-2.00	L*	L*	L*	L*	L*	L*	L*	M	L*	L*	L*	L*	L*	L*	L*	L*	MM	MM	MM	MM	ML

Figure 5: Comparative overall behaviour of the classical EVI-estimators under consideration, considering only the optimal H_p, denoted H^* (*top*) and including both the optimal H_p and L_p, denoted L^* (*bottom*)

comparative behaviour of the non reduced-bias EVI-*estimators. A similar comment applies to the optimal* GM *and* CG EVI-*estimators, as mentioned above.*

5 Finite sample properties of the EVI-estimators

We have implemented multi-sample Monte Carlo simulation experiments of size 5000×10 for the class of L_p EVI-estimators, in (1.9), comparatively with the MVRB EVI-estimators, in (2.12), for sample sizes $n = 100,\ 200,\ 500,$ and 1000, from the following underlying models:

(2) the *extreme value* model, with CDF $F(x) = \mathrm{EV}_\xi(x)$, with $\mathrm{EV}_\xi(x)$ given in (1.1), $\xi = 0.1, 0.25, 0.5$ and 1 $(\rho = -\xi)$;

(1) the *Fréchet* model, with CDF $F(x) = \exp(-x^{-1/\xi}),\ x \geq 0$, for the same values of ξ $(\rho = -1)$;

(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 $\nu = 2, 3, 4$ $(\xi = 1/\nu, \rho = -2/\nu)$.

For details on multi-sample simulation, see Gomes and Oliveira (2001), among others.

5.1 Mean values and mean square error patterns

For each value of n and for each of the above-mentioned models, we have first simulated the mean value (E) and the root mean square error (RMSE) of the estimators $\mathrm{L}_p(k)$, in (1.9), as functions of the number of top order statistics k involved in the estimation and for a few values of $p \geq 1$. We have first implemented a simulation for $\mathrm{EV}_{0.25}$ parents and values of p from 1 until 3, with step 0.5. The results obtained were slight astonishing from a theoretical point of view, because the highest efficiency was obtained for $p = 3$, being increasing with p. But as can be seen in Figure 6, based on the first replicate with a size 5000, and at optimal levels, in the sense of minimal RMSE, even $\mathrm{L}_{1.5}$ beats the MVRB EVI-estimators, CH, in (2.12), also pictured in Figure 6, being $p = 1.5$ a value close to $p_{_{\mathrm{M|L}}} = 1.7$.

And for $\xi = 0.25$ an increase in the values of p provided even better results, as can be seen in Figure 7, where we present the simulated mean values and RMSEs of L_p for $p = 3, 6(1)$.

However, as $p \to +\infty$, $\mathrm{L}_p(k) \to 0$, being no longer consistent for the estimation of $\xi > 0$ (look at Figure 8, where apart from $p = 4$ and 6 we also represent $p = 8, 10$ and 15. In these cases

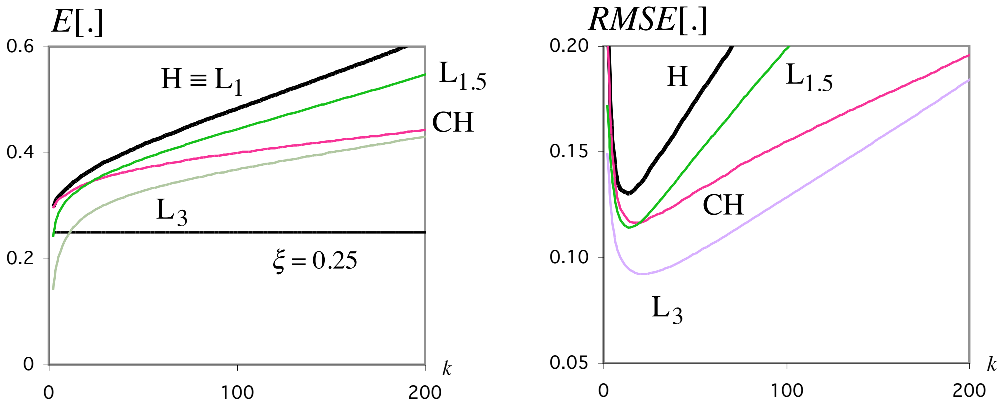

Figure 6: Mean values (*left*) and RMSEs (*right*) of the CH and L_p EVI-estimators under study for an EV_ξ CDF with $\xi = 0.25$, and $p = 1.5$ and 3

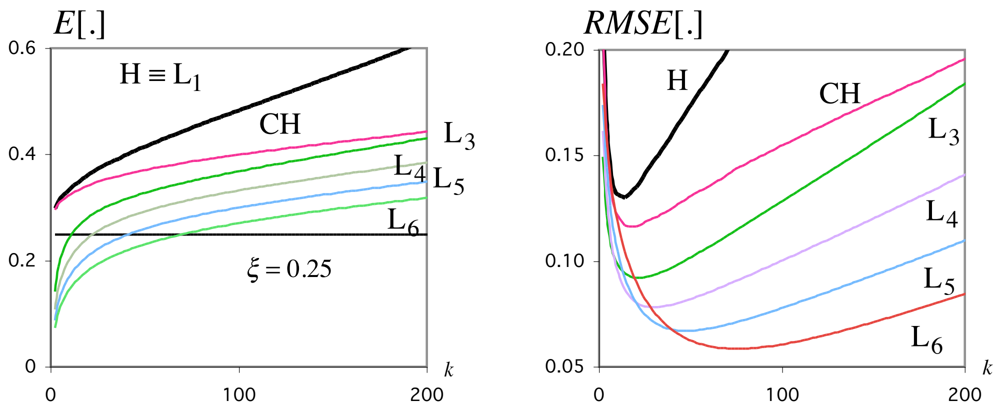

Figure 7: Mean values (*left*) and RMSEs (*right*) of the CH and L_p EVI-estimators under study for an EV_ξ CDF with $\xi = 0.25$, and $p = 3, 4, 5$ and 6

we still have a minimum RMSE at $k < n$, but such a minimum RMSE is attained at $k = n - 1$ for larger values of p, as can be seen in Figure 9, similar to Figure 8, but for $p = 30, 40$ and 50. In any adaptive choice of (k, p) we should thus avoid a minimum RMSE estimate attained at $k = n - 1$. We indeed believe that the aforementioned behaviour for small ξ is due to the the fact that $L_p(k)$ goes to zero as $p \to +\infty$. However, even for very small values of ξ, if we go on increasing p, we finally detect a decreasing in efficiency. An adaptive choice of (k, p) associated to RMSE minimization, avoiding values of k close to $n - 1$, is thus advisable.

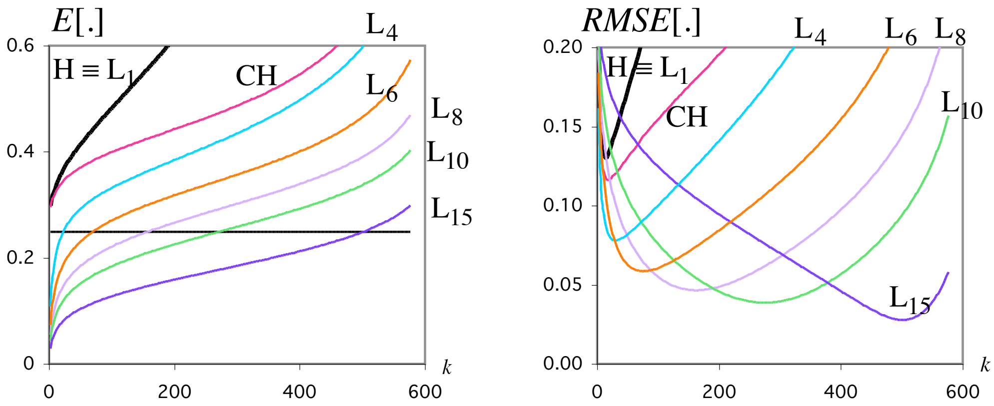

Figure 8: Mean values (*left*) and RMSEs (*right*) of the CH and L_p EVI-estimators under study for an EV_ξ CDF with $\xi = 0.25$, and $p = 30, 40$ and 50

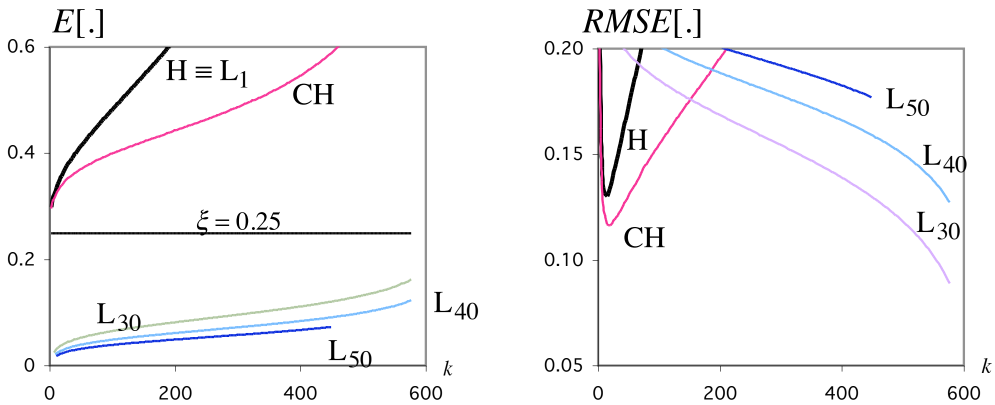

Figure 9: Mean values (*left*) and RMSEs (*right*) of the CH and L_p EVI-estimators under study for an EV_ξ CDF with $\xi = 0.25$, and $p = 30, 40$ and 50

Note however that the type of pattern mentioned above, i.e. an increasing efficiency as p increases and an outperformance comparatively with the CH EVI-estimator, has been obtained only for values of ξ not far away from zero, as can be seen in Figure 10, associated with an EV_1 underlying parent. We still include in Figure 11 a similar picture for a Student t_4 parent ($\xi = 1/\nu = 0.25, \rho = -2/\nu = -0.5$).

Similar results have been obtained for all other simulated models. We can always find an

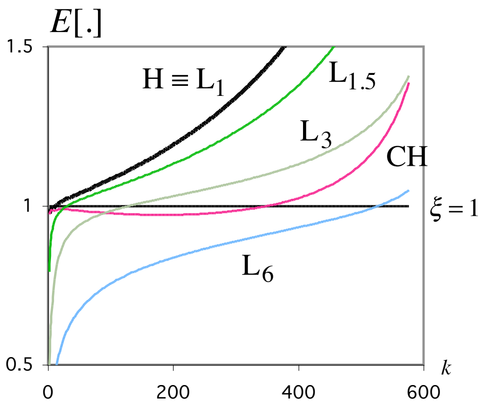

Figure 10: Mean values (*left*) and RMSEs (*right*) of the EVI-estimators under study for an EV_ξ CDF with $\xi = 1$, and $p = 1.5, 3$ and 6

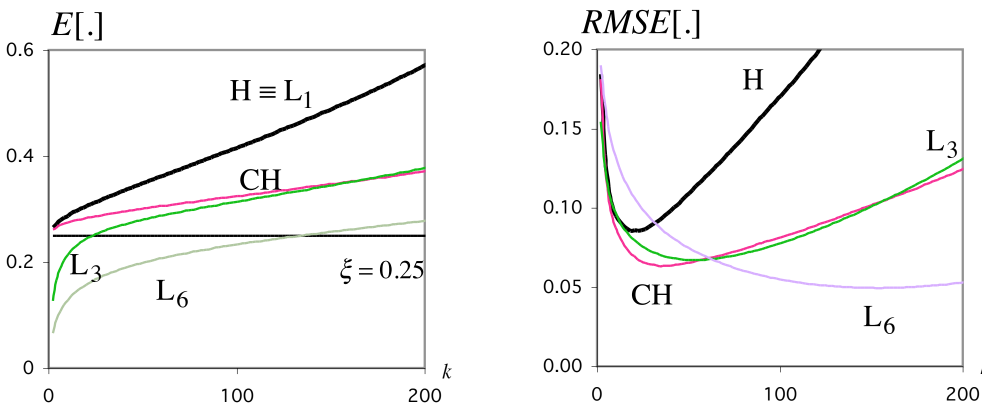

Figure 11: Mean values (*left*) and RMSEs (*right*) of the EVI-estimators under study for a Student t_4 CDF ($\xi = 0.25$), and $p = 3$ and 6

optimal value for p, clear from these pictures in what concerns RMSE, but often also valid for mean values at optimal levels, in the sense of minimal RMSE, as we shall see next, in Section 5.1.1.

5.1.1 Mean values of the EVI-estimators at optimal levels

Tables 1 and 2 are respectively related to the simulated EV_ξ and GP_ξ parents with $\xi < 1$. We there present, for $n = 100,\ 200,\ 500$ and 1000, the simulated mean values at optimal levels (levels where RMSE are minima as functions of k) of the EVI-estimators CH, in (2.12) and $\mathrm{L}_p(k)$, in (1.9), for a few values of p, including $p = 1$ (H). Information on 95% confidence intervals, computed on the basis of the 10 replicates with 5000 runs each, is also provided. Among the estimators considered, the one providing the smallest squared bias is underlined, and written in **bold**.

Due to the aforementioned difference found for $\xi = 1$, we present Tables 3 and 4, respectively associated with EV_1 and GP_1 underlying parents and a slightly larger region of values of p between 1 and 2. Table 5 is related with Fréchet underlying models. Note that in all cases, the mean value of $\mathrm{L}_p(k)$ is decreasing in p, and we thus can always find a value of p associated with minimal squared bias, often not visible in the tables. Finally, we still present Table 6, related to the simulated Student underlying parents.

Remark 13. *We may draw the following specific comments:*

- *As intuitively expected,* $\mathrm{L}_{0|p}$ *are decreasing in* p *until a value* p_{min}*, approaching the true value of* ξ*, for all simulated models.*
- *But we cannot forget that as* p *increases to* $+\infty$*,* $\mathrm{L}_{0|p}$ *approaches zero, being no longer consistent. We thus need to have estimated reliable values of the* RMSE*.*
- *For* $\xi < 1$*, the* L_p EVI–*estimators outperform the* MVRB EVI-*estimators.*

Table 1: Simulated mean values, at optimal levels, of $\mathrm{H}(k) \equiv \mathrm{L}_1(k)$, $\mathrm{CH}(k)$ and $\mathrm{L}_p(k)$, $\mathrm{L}_p(k)$, $p = 1, 2(2)10$, and 15, for EV_ξ underlying parents with $\xi = 0.1, 0.25$ and 0.5, together with 95% confidence intervals

n	100	200	500	1000
		EV_ξ **parent,** $\xi = 0.1$		
CH	0.276 ± 0.0016	0.258 ± 0.0014	0.234 ± 0.0012	0.221 ± 0.0013
$p = 1$ (H)	0.334 ± 0.0009	0.284 ± 0.0007	0.243 ± 0.0005	0.223 ± 0.0016
$p = 2$	0.253 ± 0.0006	0.220 ± 0.0006	0.191 ± 0.0005	0.175 ± 0.0006
$p = 4$	0.159 ± 0.0004	0.138 ± 0.0004	0.121 ± 0.0003	0.111 ± 0.0004
$p = 6$	0.112 ± 0.0003	0.097 ± 0.0003	0.097 ± 0.0003	0.098 ± 0.0004
$p = 8$	**0.101** ± 0.0007	0.098 ± 0.0008	0.099 ± 0.0005	0.099 ± 0.0003
$p = 10$	0.097 ± 0.0002	0.099 ± 0.0004	0.099 ± 0.0002	0.099 ± 0.0002
$p = 15$	0.099 ± 0.0002	**0.099** ± 0.0002	**0.100** ± 0.0001	**0.100** ± 0.0001
		EV_ξ **parent,** $\xi = 0.25$		
CH	0.382 ± 0.0027	0.372 ± 0.0021	0.353 ± 0.0014	0.342 ± 0.0017
$p = 1$ (H)	0.427 ± 0.0012	0.391 ± 0.0026	0.365 ± 0.0019	0.348 ± 0.0012
$p = 2$	0.332 ± 0.0009	0.314 ± 0.0026	0.320 ± 0.0018	0.317 ± 0.0011
$p = 4$	0.252 ± 0.0007	0.256 ± 0.0012	0.260 ± 0.0008	0.265 ± 0.0004
$p = 6$	0.249 ± 0.0010	0.252 ± 0.0006	0.253 ± 0.0004	0.254 ± 0.0003
$p = 8$	**0.249** ± 0.0009	**0.250** ± 0.0005	0.251 ± 0.0003	0.251 ± 0.0003
$p = 10$	0.246 ± 0.0011	0.249 ± 0.0005	**0.250** ± 0.0002	**0.250** ± 0.0001
$p = 15$	0.190 ± 0.0040	0.226 ± 0.0032	0.248 ± 0.0002	0.249 ± 0.0002
		EV_ξ **parent,** $\xi = 0.5$		
CH	0.554 ± 0.0053	0.573 ± 0.0016	0.564 ± 0.0014	0.558 ± 0.0010
$p = 1$ (H)	0.654 ± 0.0032	0.624 ± 0.0033	0.596 ± 0.0011	0.579 ± 0.0016
$p = 2$	0.591 ± 0.0026	0.585 ± 0.0020	0.575 ± 0.0009	0.565 ± 0.0009
$p = 4$	0.523 ± 0.0011	0.535 ± 0.0010	0.546 ± 0.0008	0.549 ± 0.0011
$p = 6$	**0.499** ± 0.0019	**0.511** ± 0.0007	0.519 ± 0.0005	0.525 ± 0.0006
$p = 8$	0.422 ± 0.0068	0.484 ± 0.0045	**0.507** ± 0.0005	0.512 ± 0.0003
$p = 10$	0.350 ± 0.0059	0.410 ± 0.0047	0.479 ± 0.0037	**0.503** ± 0.0004
$p = 15$	0.242 ± 0.0043	0.289 ± 0.0036	0.348 ± 0.0031	0.387 ± 0.0023

Table 2: Simulated mean values, at optimal levels, of $\mathrm{H}(k) \equiv \mathrm{L}_1(k)$, $\mathrm{CH}(k)$ and $\mathrm{L}_p(k)$, $p = 1, 2(2)10$, and 15, for GP_ξ underlying parents with $\xi = 0.1, 0.25$ and 0.5, together with 95% confidence intervals

n	100	200	500	1000
		GP_ξ **parent,** $\xi = 0.1$		
CH	0.303 ± 0.0012	0.270 ± 0.0011	0.238 ± 0.0022	0.221 ± 0.0028
$p = 1$ (H)	0.325 ± 0.0013	0.282 ± 0.0012	0.242 ± 0.0012	0.224 ± 0.0027
$p = 2$	0.249 ± 0.0015	0.219 ± 0.0011	0.190 ± 0.0008	0.175 ± 0.0014
$p = 4$	0.156 ± 0.0010	0.138 ± 0.0007	0.120 ± 0.0006	0.111 ± 0.0009
$p = 6$	0.110 ± 0.0007	0.097 ± 0.0005	0.097 ± 0.0004	0.098 ± 0.0006
$p = 8$	0.099 ± 0.0006	0.099 ± 0.0013	0.099 ± 0.0008	0.099 ± 0.0005
$p = 10$	$\mathbf{0.099} \pm 0.0004$	0.099 ± 0.0003	0.099 ± 0.0003	0.099 ± 0.0003
$p = 15$	0.099 ± 0.0006	$\mathbf{0.100} \pm 0.0003$	$\mathbf{0.099} \pm 0.0001$	$\mathbf{0.100} \pm 0.0002$
		GP_ξ **parent,** $\xi = 0.25$		
CH	0.404 ± 0.0044	0.382 ± 0.0027	0.358 ± 0.0018	0.344 ± 0.0030
$p = 1$ (H)	0.418 ± 0.0030	0.389 ± 0.0037	0.364 ± 0.0028	0.348 ± 0.0028
$p = 2$	0.326 ± 0.0020	0.314 ± 0.0039	0.320 ± 0.0022	0.318 ± 0.0024
$p = 4$	0.251 ± 0.0033	0.256 ± 0.0016	0.261 ± 0.0010	0.266 ± 0.0008
$p = 6$	0.249 ± 0.0015	0.251 ± 0.0007	0.253 ± 0.0007	0.254 ± 0.0006
$p = 8$	$\mathbf{0.250} \pm 0.0011$	0.251 ± 0.0007	0.251 ± 0.0004	0.251 ± 0.0003
$p = 10$	0.249 ± 0.0004	0.249 ± 0.0006	0.250 ± 0.0003	0.251 ± 0.0003
$p = 15$	0.248 ± 0.0003	$\mathbf{0.250} \pm 0.0004$	$\mathbf{0.250} \pm 0.0002$	$\mathbf{0.250} \pm 0.0002$
		GP_ξ **parent,** $\xi = 0.5$		
CH	0.612 ± 0.0043	0.593 ± 0.0032	0.576 ± 0.0022	0.565 ± 0.0013
$p = 1$ (H)	0.646 ± 0.0082	0.621 ± 0.0037	0.593 ± 0.0040	0.577 ± 0.0021
$p = 2$	0.594 ± 0.0031	0.586 ± 0.0016	0.573 ± 0.0015	0.564 ± 0.0017
$p = 4$	0.530 ± 0.0014	0.539 ± 0.0012	0.547 ± 0.0016	0.550 ± 0.0018
$p = 6$	0.509 ± 0.0011	0.514 ± 0.0007	0.521 ± 0.0007	0.527 ± 0.0010
$p = 8$	0.503 ± 0.0015	0.506 ± 0.0009	0.510 ± 0.0004	0.514 ± 0.0005
$p = 10$	$\mathbf{0.500} \pm 0.0007$	$\mathbf{0.502} \pm 0.0011$	0.505 ± 0.0005	0.508 ± 0.0001
$p = 15$	0.458 ± 0.0009	0.485 ± 0.0008	$\mathbf{0.499} \pm 0.0016$	$\mathbf{0.501} \pm 0.0024$

Table 3: Simulated mean values, at optimal levels, of $\mathrm{H}(k) \equiv \mathrm{L}_1(k)$, $\mathrm{CH}(k)$ and $\mathrm{L}_p(k)$, $p = 1(0.2)1.6, 2$ and 6, for EV_1 underlying parent, together with 95% confidence intervals

n	100	200	500	1000
CH	**0.894** $\pm$ 0.0099	**0.975** $\pm$ 0.0046	**1.003** $\pm$ 0.0024	**1.004** $\pm$ 0.0013
$p = 1$ (H)	1.159 $\pm$ 0.0049	1.124 $\pm$ 0.0032	1.091 $\pm$ 0.0030	1.072 $\pm$ 0.0020
$p = 1.2$	1.138 $\pm$ 0.0078	1.113 $\pm$ 0.0048	1.085 $\pm$ 0.0024	1.068 $\pm$ 0.0017
$p = 1.4$	1.132 $\pm$ 0.0052	1.109 $\pm$ 0.0039	1.081 $\pm$ 0.0020	1.066 $\pm$ 0.0016
$p = 1.6$	1.125 $\pm$ 0.0047	1.105 $\pm$ 0.0021	1.079 $\pm$ 0.0021	1.065 $\pm$ 0.0013
$p = 2$	1.116 $\pm$ 0.0029	1.100 $\pm$ 0.0022	1.079 $\pm$ 0.0010	1.065 $\pm$ 0.0008
$p = 6$	0.798 $\pm$ 0.0088	0.901 $\pm$ 0.0062	**0.999** $\pm$ 0.0045	1.044 $\pm$ 0.0021

Table 4: Simulated mean values, at optimal levels, of $\mathrm{H}(k) \equiv \mathrm{L}_1(k)$, $\mathrm{CH}(k)$ and $\mathrm{L}_p(k)$, $p = 1(0.2)1.6, 2$ and 6, for GP_1 underlying parent, together with 95% confidence intervals

n	100	200	500	1000
CH	**1.008** $\pm$ 0.0036	**1.006** $\pm$ 0.0027	**1.002** $\pm$ 0.0017	**1.001** $\pm$ 0.0025
$p = 1$ (H)	1.136 $\pm$ 0.0072	1.110 $\pm$ 0.0063	1.078 $\pm$ 0.0041	1.063 $\pm$ 0.0040
$p = 1.2$	1.124 $\pm$ 0.0050	1.101 $\pm$ 0.0041	1.073 $\pm$ 0.0019	1.060 $\pm$ 0.0017
$p = 1.4$	1.118 $\pm$ 0.0038	1.097 $\pm$ 0.0021	1.072 $\pm$ 0.0019	1.058 $\pm$ 0.0019
$p = 1.6$	1.115 $\pm$ 0.0033	1.093 $\pm$ 0.0023	1.071 $\pm$ 0.0016	1.058 $\pm$ 0.0009
$p = 2$	1.106 $\pm$ 0.0023	1.091 $\pm$ 0.0017	1.071 $\pm$ 0.0013	1.057 $\pm$ 0.0021
$p = 6$	1.026 $\pm$ 0.0057	1.042 $\pm$ 0.0014	1.051 $\pm$ 0.0021	1.053 $\pm$ 0.0032

Table 5: Simulated mean values, at optimal levels, of $\mathrm{H}(k)/\xi \equiv \mathrm{L}_1(k)/\xi$, $\mathrm{CH}(k)/\xi$ and $\mathrm{L}_p(k)/\xi$, $p = 1(0.2)1.6, 2$ and 6, for Fréchet underlying parent, together with 95% confidence intervals

n	100	200	500	1000
CH	**0.982** $\pm$ 0.0030	**0.986** $\pm$ 0.0395	0.995 $\pm$ 0.0016	**0.998** $\pm$ 0.0026
$p = 1$ (H)	1.109 $\pm$ 0.0027	1.085 $\pm$ 0.0028	1.063 $\pm$ 0.0013	1.049 $\pm$ 0.0035
$p = 1.2$	1.099 $\pm$ 0.0033	1.081 $\pm$ 0.0026	1.059 $\pm$ 0.0012	1.047 $\pm$ 0.0014
$p = 1.4$	1.095 $\pm$ 0.0025	1.078 $\pm$ 0.0034	1.057 $\pm$ 0.0010	1.046 $\pm$ 0.0009
$p = 1.6$	1.094 $\pm$ 0.0020	1.076 $\pm$ 0.0020	1.057 $\pm$ 0.0012	1.046 $\pm$ 0.0006
$p = 2$	1.090 $\pm$ 0.0014	1.075 $\pm$ 0.0011	1.057 $\pm$ 0.0007	1.046 $\pm$ 0.0014
$p = 6$	0.904 $\pm$ 0.0015	0.954 $\pm$ 0.0012	**1.000** $\pm$ 0.0008	1.020 $\pm$ 0.0017

Table 6: Simulated mean values, at optimal levels, of $\mathrm{H}(k) \equiv \mathrm{L}_1(k)$, $\mathrm{CH}(k)$ and $\mathrm{L}_p(k)$, $p = 2(1)6$, 8, 10 and 15, for Student t_ν underlying parents with $\nu = 4, 3$ and 2, together with 95% confidence intervals

n	100	200	500	1000
Student t_4 parent ($\xi = 0.25, \rho = -0.5$)				
$p = 1$ (H)	0.360 ± 0.0014	0.339 ± 0.0042	0.315 ± 0.0027	0.306 ± 0.0022
CH	0.312 ± 0.0027	0.310 ± 0.0014	0.300 ± 0.0020	0.295 ± 0.0011
$p = 2$	0.297 ± 0.0034	0.301 ± 0.0016	0.297 ± 0.0016	0.292 ± 0.0011
$p = 3$	0.265 ± 0.0038	0.275 ± 0.0012	0.282 ± 0.0016	0.283 ± 0.0013
$p = 4$	0.253 ± 0.0026	0.261 ± 0.0013	0.269 ± 0.0008	0.272 ± 0.0009
$p = 5$	**0.251** ± 0.0022	0.255 ± 0.0009	0.260 ± 0.0005	0.264 ± 0.0006
$p = 6$	0.247 ± 0.0004	0.252 ± 0.0009	0.256 ± 0.0004	0.259 ± 0.0003
$p = 8$	0.240 ± 0.0049	**0.249** ± 0.0003	0.252 ± 0.0004	0.254 ± 0.0002
$p = 10$	0.204 ± 0.0068	0.242 ± 0.0023	**0.250** ± 0.0003	**0.252** ± 0.0004
$p = 15$	0.141 ± 0.0049	0.176 ± 0.0048	0.220 ± 0.0018	0.242 ± 0.0018
Student t_3 parent, ($\xi = 1/3, \rho = -2/3$)				
$p = 1$ (H)	0.441 ± 0.0046	0.416 ± 0.0035	0.394 ± 0.0026	0.385 ± 0.0018
CH	0.361 ± 0.0058	0.377 ± 0.0027	0.369 ± 0.0009	0.364 ± 0.0011
$p = 2$	0.390 ± 0.0050	0.387 ± 0.0027	0.380 ± 0.0010	0.375 ± 0.0012
$p = 3$	0.360 ± 0.0022	0.367 ± 0.0014	0.371 ± 0.0010	0.370 ± 0.0013
$p = 4$	0.344 ± 0.0019	0.353 ± 0.0012	0.360 ± 0.0011	0.363 ± 0.0011
$p = 5$	**0.338** ± 0.0018	0.343 ± 0.0010	0.350 ± 0.0005	0.355 ± 0.0008
$p = 6$	0.327 ± 0.0041	**0.338** ± 0.0005	0.344 ± 0.0005	0.348 ± 0.0005
$p = 8$	0.268 ± 0.0064	0.324 ± 0.0039	**0.337** ± 0.0006	0.340 ± 0.0003
$p = 10$	0.222 ± 0.0056	0.277 ± 0.0065	0.328 ± 0.0011	**0.335** ± 0.0004
$p = 15$	0.152 ± 0.0040	0.1945 ± 0.0049	0.241 ± 0.0035	0.272 ± 0.0028
Student t_2 parent, ($\xi = 0.5, \rho = -1$)				
$p = 1$ (H)	0.602 ± 0.0069	0.578 ± 0.0040	0.556 ± 0.0019	0.544 ± 0.0015
CH	0.471 ± 0.0095	**0.505** ± 0.0034	**0.512** ± 0.0019	0.507 ± 0.0009
$p = 2$	0.563 ± 0.0012	0.558 ± 0.0020	0.546 ± 0.0008	0.537 ± 0.0005
$p = 3$	0.538 ± 0.0019	0.544 ± 0.0012	0.540 ± 0.0008	0.534 ± 0.0007
$p = 4$	**0.515** ± 0.0011	0.530 ± 0.0015	0.533 ± 0.0011	0.529 ± 0.0009
$p = 5$	0.467 ± 0.0077	0.514 ± 0.0013	0.524 ± 0.0013	0.524 ± 0.0008
$p = 6$	0.412 ± 0.0071	0.479 ± 0.0059	0.516 ± 0.0008	0.519 ± 0.0006
$p = 8$	0.329 ± 0.0061	0.394 ± 0.0052	0.467 ± 0.0050	**0.501** ± 0.0033
$p = 10$	0.271 ± 0.0053	0.330 ± 0.0047	0.402 ± 0.0050	0.442 ± 0.0046
$p = 15$	0.186 ± 0.0038	0.230 ± 0.0035	0.288 ± 0.0041	0.325 ± 0.0040

5.1.2 RMSEs and relative efficiency indicators at optimal levels

We have computed the Hill estimator, in (1.9) whenever $p = 1$, at the simulated value of $k_{0|1} := \arg\min_k \mathrm{RMSE}\big(\mathrm{L}_1(k)\big)$, the simulated optimal k in the sense of minimum RMSE. Such an estimator is denoted $\mathrm{L}_{0|1}$. We have also compute $\mathrm{L}_{0|p}$, the estimator L_p computed at the simulated value of $k_{0|p} := \arg\min_k \mathrm{RMSE}\big(\mathrm{L}_p(k)\big)$. The simulated indicators are

$$\mathrm{REFF}_{p|1} := \frac{\mathrm{RMSE}\left(\mathrm{L}_{0|1}\right)}{\mathrm{RMSE}\left(\mathrm{L}_{0|p}\right)}. \tag{5.1}$$

A similar indicator has also been computed for the CH EVI-estimator, and as mentioned in Remark 9, the higher these indicators are, the better the associated EVI-estimators perform, comparatively to $\mathrm{H}_0 = \mathrm{L}_{0|1}$.

Again as an illustration of the results obtained, we present Tables 7–12. In the first row, we provide the RMSE of $\mathrm{L}_{0|1}$, so that we can easily recover the RMSE of all other estimators $\mathrm{L}_{0|p}$. The following rows provide the REFF indicators of CH and $\mathrm{REFF}_{_{p|1}}$ in (5.1), for the same L_p EVI-estimators considered in the equivalent tables of Section 5.1.1. A similar mark (underlined and **bold**) is used for the highest REFF indicator.

Remark 14. *We now provide a few comments related with the* REFF*-indicators:*

- *For all simulated models but the* Fréchet (*independently of* ξ) *and all the simulated models with* $\xi = 1$, *the new* L_p EVI-*estimators are clearly able to outperform the* MVRB EVI-*estimators.*
- *The values of p associated with the highest efficiency are higher than expected, and not in agreement with the theoretical results, certainly due to the non-consistency of the* L_p-*statistics as* $p = +\infty$.

Table 7: Simulated RMSE of H (first row) and REFF-indicators of CH(k) and $L_p(k)$, $p = 1.5, 2(2)10$ and 15, for EV_ξ underlying parents, $\xi = 0.1, 0.25$ and 0.5, together with 95% confidence intervals

n	100	200	500	1000
		EV_ξ **parent,** $\xi = 0.1$		
$RMSE_0$(H)	0.268 ± 0.2186	0.216 ± 0.1757	0.174 ± 0.1392	0.151 ± 0.1180
CH	1.245 ± 0.0050	1.140 ± 0.0027	1.070 ± 0.0019	1.045 ± 0.0015
$p = 1.5$	1.221 ± 0.0019	1.209 ± 0.0022	1.196 ± 0.0024	1.188 ± 0.0034
$p = 2$	1.445 ± 0.0037	1.422 ± 0.0040	1.399 ± 0.0050	1.383 ± 0.0062
$p = 4$	2.948 ± 0.0134	2.949 ± 0.0135	2.899 ± 0.0150	2.804 ± 0.0165
$p = 6$	5.457 ± 0.0404	5.005 ± 0.0302	4.384 ± 0.0229	4.029 ± 0.0350
$p = 8$	7.206 ± 0.0411	6.362 ± 0.0327	5.599 ± 0.0273	5.158 ± 0.0277
$p = 10$	8.737 ± 0.0514	7.744 ± 0.0406	6.837 ± 0.0349	6.302 ± 0.0343
$p = 15$	$\underline{\mathbf{12.312}} \pm 0.0740$	$\underline{\mathbf{11.116}} \pm 0.0474$	$\underline{\mathbf{9.933}} \pm 0.0513$	$\underline{\mathbf{9.163}} \pm 0.0510$
		EV_ξ **parent,** $\xi = 0.25$		
$RMSE_0$(H)	0.246 ± 0.2154	0.200 ± 0.1710	0.157 ± 0.1385	0.133 ± 0.1200
CH	1.328 ± 0.0108	1.237 ± 0.0056	1.171 ± 0.0042	1.130 ± 0.0031
$p = 1.5$	1.245 ± 0.0039	1.202 ± 0.0048	1.158 ± 0.0028	1.135 ± 0.0034
$p = 2$	1.473 ± 0.0077	1.369 ± 0.0087	1.268 ± 0.0057	1.220 ± 0.0063
$p = 4$	2.532 ± 0.0183	2.198 ± 0.0149	1.863 ± 0.0133	1.663 ± 0.0146
$p = 6$	3.506 ± 0.0331	3.036 ± 0.0208	2.542 ± 0.0226	2.236 ± 0.0295
$p = 8$	4.382 ± 0.0285	3.825 ± 0.0227	3.197 ± 0.0222	2.798 ± 0.0233
$p = 10$	$\underline{\mathbf{5.093}} \pm 0.0347$	4.569 ± 0.0267	3.850 ± 0.0269	3.365 ± 0.0268
$p = 15$	3.560 ± 0.1549	$\underline{\mathbf{4.734}} \pm 0.1575$	$\underline{\mathbf{5.172}} \pm 0.0347$	$\underline{\mathbf{4.680}} \pm 0.0364$
		EV_ξ **parent,** $\xi = 0.5$		
$RMSE_0$(H)	0.256 ± 0.2059	0.202 ± 0.1667	0.151 ± 0.1419	0.122 ± 0.1293
CH	1.492 ± 0.0258	1.501 ± 0.0097	1.476 ± 0.0059	1.452 ± 0.0059
$p = 1.5$	1.183 ± 0.0029	1.149 ± 0.0033	1.117 ± 0.0029	1.099 ± 0.0026
$p = 2$	1.295 ± 0.0055	1.226 ± 0.0057	1.163 ± 0.0045	1.132 ± 0.0044
$p = 4$	1.787 ± 0.0141	1.539 ± 0.0125	1.308 ± 0.0096	1.193 ± 0.0105
$p = 6$	$\underline{\mathbf{2.324}} \pm 0.0269$	1.974 ± 0.0177	1.582 ± 0.0157	1.366 ± 0.0202
$p = 8$	2.119 ± 0.0696	$\underline{\mathbf{2.264}} \pm 0.0187$	1.888 ± 0.0149	1.602 ± 0.0155
$p = 10$	1.531 ± 0.0458	1.728 ± 0.0488	$\underline{\mathbf{2.012}} \pm 0.0322$	$\underline{\mathbf{1.831}} \pm 0.0159$
$p = 15$	0.976 ± 0.0165	0.931 ± 0.0152	0.937 ± 0.0176	0.981 ± 0.0156

Table 8: Simulated RMSE of H (first row) and REFF-indicators of CH(k) and $L_p(k)$, $p = 2(2)10$ and 15, for GP_ξ underlying parents, $\xi = 0.1, 0.25$ and 0.5, together with 95% confidence intervals

n	100	200	500	1000
		GP_ξ **parent,** $\xi = 0.1$		
$RMSE_0(H)$	0.259 ± 0.1606	0.213 ± 0.1258	0.172 ± 0.0950	0.150 ± 0.0764
CH	1.100 ± 0.0011	1.061 ± 0.0008	1.032 ± 0.0006	1.023 ± 0.0012
$p = 2$	1.433 ± 0.0068	1.415 ± 0.0054	1.394 ± 0.0109	1.375 ± 0.0095
$p = 4$	2.919 ± 0.0226	2.936 ± 0.0156	2.889 ± 0.0288	2.786 ± 0.0270
$p = 6$	5.338 ± 0.0446	4.946 ± 0.0278	4.347 ± 0.0451	3.991 ± 0.0337
$p = 8$	7.038 ± 0.0564	6.308 ± 0.0425	5.564 ± 0.0488	5.126 ± 0.0436
$p = 10$	8.516 ± 0.0521	7.693 ± 0.0545	6.793 ± 0.0594	6.263 ± 0.0506
$p = 15$	$\underline{\mathbf{12.205}} \pm 0.0701$	$\underline{\mathbf{11.110}} \pm 0.0938$	$\underline{\mathbf{9.883}} \pm 0.0767$	$\underline{\mathbf{9.104}} \pm 0.0721$
		GP_ξ **parent,** $\xi = 0.25$		
$RMSE_0(H)$	0.237 ± 0.1724	0.195 ± 0.1380	0.154 ± 0.1074	0.131 ± 0.0887
CH	1.149 ± 0.0057	1.117 ± 0.0042	1.088 ± 0.0049	1.069 ± 0.0030
$p = 2$	1.447 ± 0.0103	1.350 ± 0.0114	1.262 ± 0.0117	1.214 ± 0.0103
$p = 4$	2.445 ± 0.0182	2.157 ± 0.0171	1.840 ± 0.0218	1.649 ± 0.0182
$p = 6$	3.379 ± 0.0266	2.965 ± 0.0253	2.498 ± 0.0256	2.205 ± 0.0213
$p = 8$	4.298 ± 0.0340	3.768 ± 0.0354	3.154 ± 0.0306	2.768 ± 0.0248
$p = 10$	5.186 ± 0.0508	4.564 ± 0.0422	3.806 ± 0.0344	3.330 ± 0.0291
$p = 15$	$\underline{\mathbf{7.091}} \pm 0.0721$	$\underline{\mathbf{6.415}} \pm 0.0595$	$\underline{\mathbf{5.405}} \pm 0.0425$	$\underline{\mathbf{4.724}} \pm 0.0356$
		GP_ξ **parent,** $\xi = 0.5$		
$RMSE_0(H)$	0.239 ± 0.1914	0.190 ± 0.1648	0.1444 ± 0.1408	0.118 ± 0.1259
CH	1.423 ± 0.0087	1.380 ± 0.0082	1.339 ± 0.0069	1.302 ± 0.0105
$p = 2$	1.271 ± 0.0083	1.209 ± 0.0062	1.160 ± 0.0081	1.129 ± 0.0076
$p = 4$	1.703 ± 0.0185	1.484 ± 0.0195	1.288 ± 0.0169	1.182 ± 0.0152
$p = 6$	2.232 ± 0.0254	1.883 ± 0.0267	1.534 ± 0.0222	1.333 ± 0.0198
$p = 8$	2.754 ± 0.0310	2.309 ± 0.0331	1.842 ± 0.0250	1.566 ± 0.0416
$p = 10$	$\underline{\mathbf{3.187}} \pm 0.0360$	2.720 ± 0.0371	2.163 ± 0.0274	1.820 ± 0.0471
$p = 15$	3.021 ± 0.0237	$\underline{\mathbf{3.108}} \pm 0.0385$	$\underline{\mathbf{2.798}} \pm 0.0291$	$\underline{\mathbf{2.424}} \pm 0.0618$

Table 9: Simulated RMSE of H (first row) and REFF-indicators of CH(k) and $L_p(k)$, $p = 1.2(0.2)1.6, 2$ and 6, for EV_1 underlying parents, together with 95% confidence intervals

n	100	200	500	1000
$RMSE_0(H)$	0.314 ± 0.2292	0.239 ± 0.2077	0.170 ± 0.1967	0.132 ± 0.1938
CH	0.814 ± 0.1168	**1.182** ± 0.0230	**1.410** ± 0.0212	**1.678** ± 0.0198
$p = 1.2$	1.071 ± 0.0037	1.060 ± 0.0036	1.052 ± 0.0014	1.045 ± 0.0019
$p = 1.4$	1.114 ± 0.0061	1.093 ± 0.0060	1.078 ± 0.0030	1.066 ± 0.0030
$p = 1.6$	1.142 ± 0.0082	1.112 ± 0.0077	1.089 ± 0.0039	1.073 ± 0.0041
$p = 2$	**1.176** ± 0.0059	1.127 ± 0.0062	1.084 ± 0.0035	1.062 ± 0.0042
$p = 6$	1.074 ± 0.0230	1.073 ± 0.0156	1.008 ± 0.0100	0.914 ± 0.0096

Table 10: Simulated RMSE of H (first row) and REFF-indicators of CH(k) and $L_p(k)$, $p = 1.2(0.2)1.6, 2$ and 6, for GP_1 underlying parents, together with 95% confidence intervals

n	100	200	500	1000
$RMSE_0(H)$	0.266 ± 0.2783	0.205 ± 0.2673	0.147 ± 0.2604	0.115 ± 0.4072
CH	**1.982** ± 0.0164	**2.126** ± 0.0175	**2.427** ± 0.0177	**2.685** ± 0.0439
$p = 1.2$	1.064 ± 0.0032	1.057 ± 0.0022	1.047 ± 0.0020	1.043 ± 0.0026
$p = 1.4$	1.101 ± 0.0042	1.087 ± 0.0029	1.069 ± 0.0038	1.061 ± 0.0036
$p = 1.6$	1.124 ± 0.0059	1.102 ± 0.0038	1.078 ± 0.0046	1.065 ± 0.0046
$p = 2$	1.147 ± 0.0079	1.110 ± 0.0058	1.073 ± 0.0059	1.055 ± 0.0102
$p = 6$	1.449 ± 0.0212	1.219 ± 0.0198	0.990 ± 0.0153	0.872 ± 0.0164

Table 11: Simulated RMSE of H/ξ (first row) and REFF-indicators of CH(k) and $L_p(k)$ (independent on ξ), $p = 1.2(0.2)1.6, 2$ and 6, for Fréchet underlying parents, together with 95% confidence intervals

n	100	200	500	1000
$RMSE_0(H)$	0.212 ± 0.2373	0.163 ± 0.2272	0.117 ± 0.2189	0.091 ± 0.4484
CH	**1.257** ± 0.0072	**1.237** ± 0.1591	**1.337** ± 0.0080	**1.456** ± 0.0101
$p = 1.2$	1.057 ± 0.0022	1.048 ± 0.0017	1.043 ± 0.0016	1.041 ± 0.0017
$p = 1.4$	1.086 ± 0.0030	1.071 ± 0.0037	1.061 ± 0.0026	1.056 ± 0.0025
$p = 1.6$	1.099 ± 0.0044	1.080 ± 0.0048	1.065 ± 0.0039	1.058 ± 0.0024
$p = 2$	1.103 ± 0.0046	1.074 ± 0.0056	1.049 ± 0.0033	1.038 ± 0.0041
$p = 6$	0.966 ± 0.0071	0.876 ± 0.0100	0.776 ± 0.0065	0.722 ± 0.0080

Table 12: Simulated RMSE of H (first row) and REFF-indicators of CH(k) and $L_p(k)$, $p = 2(1)6$, 8, 10 and 15, for Student-t_ν underlying parents, $\nu = 4, 3, 2$ ($\xi = 1/\nu$), together with 95% confidence intervals

n	100	200	500	1000
Student t_4 parent, ($\xi = 0.25, \rho = -0.5$)				
$RMSE_0(H)$	0.182 ± 0.5069	0.143 ± 0.4256	0.106 ± 0.3561	0.086 ± 0.3174
CH	1.395 ± 0.0951	1.400 ± 0.0172	1.360 ± 0.0097	1.325 ± 0.0108
$p = 2$	1.423 ± 0.0137	1.316 ± 0.0111	1.229 ± 0.0112	1.185 ± 0.0091
$p = 3$	1.814 ± 0.0191	1.581 ± 0.0170	1.37 ± 0.0237	1.268 ± 0.0178
$p = 4$	2.211 ± 0.0208	1.887 ± 0.0191	1.557 ± 0.0312	1.385 ± 0.0240
$p = 5$	2.593 ± 0.0246	2.200 ± 0.0223	1.777 ± 0.0359	1.540 ± 0.0274
$p = 6$	2.939 ± 0.0265	2.510 ± 0.0271	2.009 ± 0.0383	1.717 ± 0.0302
$p = 8$	$\underline{\mathbf{3.365}} \pm 0.0501$	3.054 ± 0.0324	2.473 ± 0.0423	2.090 ± 0.0344
$p = 10$	2.824 ± 0.1637	$\underline{\mathbf{3.327}} \pm 0.0452$	$\underline{\mathbf{2.883}} \pm 0.0425$	2.459 ± 0.0383
$p = 15$	1.605 ± 0.0568	1.780 ± 0.0818	2.438 ± 0.0559	$\underline{\mathbf{2.837}} \pm 0.0440$
Student t_3 parent, ($\xi = 1/3, \rho = -2/3$)				
$RMSE_0(H)$	0.189 ± 0.4864	0.145 ± 0.4081	0.105 ± 0.3414	0.084 ± 0.3027
CH	1.431 ± 0.0802	1.511 ± 0.0133	1.551 ± 0.0126	1.569 ± 0.0115
$p = 2$	1.347 ± 0.0084	1.259 ± 0.0082	1.183 ± 0.0090	1.136 ± 0.0083
$p = 3$	1.626 ± 0.0148	1.429 ± 0.0170	1.257 ± 0.0161	1.162 ± 0.0147
$p = 4$	1.925 ± 0.0159	1.636 ± 0.0230	1.358 ± 0.0203	1.205 ± 0.0193
$p = 5$	2.208 ± 0.0142	1.862 ± 0.0259	1.495 ± 0.0227	1.283 ± 0.0224
$p = 6$	$\underline{\mathbf{2.420}} \pm 0.0220$	2.083 ± 0.0280	1.652 ± 0.0251	1.389 ± 0.0244
$p = 8$	2.077 ± 0.0906	$\underline{\mathbf{2.364}} \pm 0.0372$	1.965 ± 0.0284	1.634 ± 0.0264
$p = 10$	1.534 ± 0.0555	1.891 ± 0.1020	$\underline{\mathbf{2.13}} \pm 0.0279$	$\underline{\mathbf{1.860}} \pm 0.0282$
$p = 15$	1.026 ± 0.0209	1.013 ± 0.0308	1.059 ± 0.0332	1.166 ± 0.0359
Student t_2 parent, ($\xi = 0.5, \rho = -1$)				
$RMSE_0(H)$	0.204 ± 0.4496	0.153 ± 0.3788	0.108 ± 0.3183	0.084 ± 0.2830
CH	1.045 ± 0.1172	1.397 ± 0.0126	$\underline{\mathbf{1.694}} \pm 0.00247$	$\underline{\mathbf{1.951}} \pm 0.0214$
$p = 1.2$	1.089 ± 0.0020	1.072 ± 0.0025	1.060 ± 0.0021	1.054 ± 0.0016
$p = 1.4$	1.146 ± 0.0032	1.11 ± 0.0035	1.095 ± 0.0032	1.085 ± 0.0023
$p = 1.6$	1.189 ± 0.0041	1.148 ± 0.0040	1.115 ± 0.0047	1.102 ± 0.0031
$p = 2$	1.257 ± 0.0069	1.188 ± 0.0062	1.136 ± 0.0065	1.117 ± 0.0045
$p = 3$	1.412 ± 0.0127	1.266 ± 0.0118	1.162 ± 0.0115	1.125 ± 0.0081
$p = 4$	1.576 ± 0.0171	1.368 ± 0.0167	1.204 ± 0.0155	1.144 ± 0.0114
$p = 5$	$\underline{\mathbf{1.636}} \pm 0.0371$	1.482 ± 0.0186	1.270 ± 0.0190	1.185 ± 0.0140
$p = 6$	1.463 ± 0.0513	$\underline{\mathbf{1.513}} \pm 0.0227$	1.344 ± 0.0206	1.240 ± 0.0159
$p = 8$	1.067 ± 0.0337	1.124 ± 0.0308	1.288 ± 0.0376	1.319 ± 0.0138
$p = 10$	0.853 ± 0.0211	0.830 ± 0.0188	0.897 ± 0.0307	0.978 ± 0.0333
$p = 15$	0.643 ± 0.0104	0.557 ± 0.0078	0.495 ± 0.0093	0.463 ± 0.0090

5.1.3 Visualization of a few tables above

For a better visualization of tables in Section 5.1.1, we present Figures 12-13, respectively associated with $EV_{0.25}$ and Student-t_4 underlying parents.

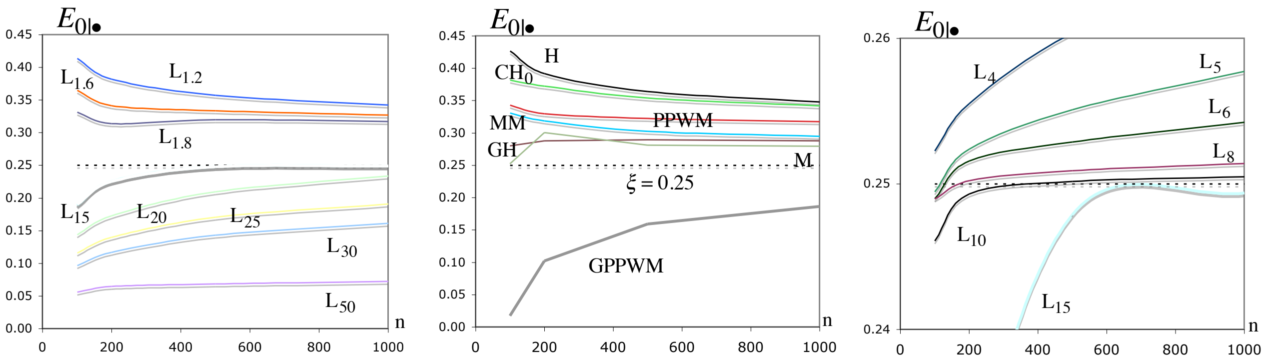

Figure 12: Mean values at optimal levels, as a function of the sample size n, and for an $EV_{0.25}$ underlying parent

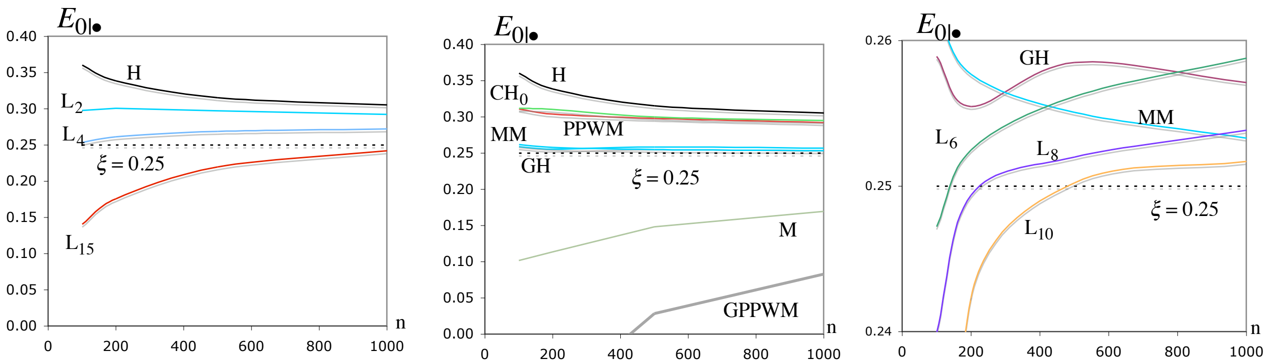

Figure 13: Mean values at optimal levels, as a function of the sample size n, and for a Student t_4 underlying parent

Figures 14–15 are associated with the results in Section 5.1.2, and for the same parents as Figures 12–13.

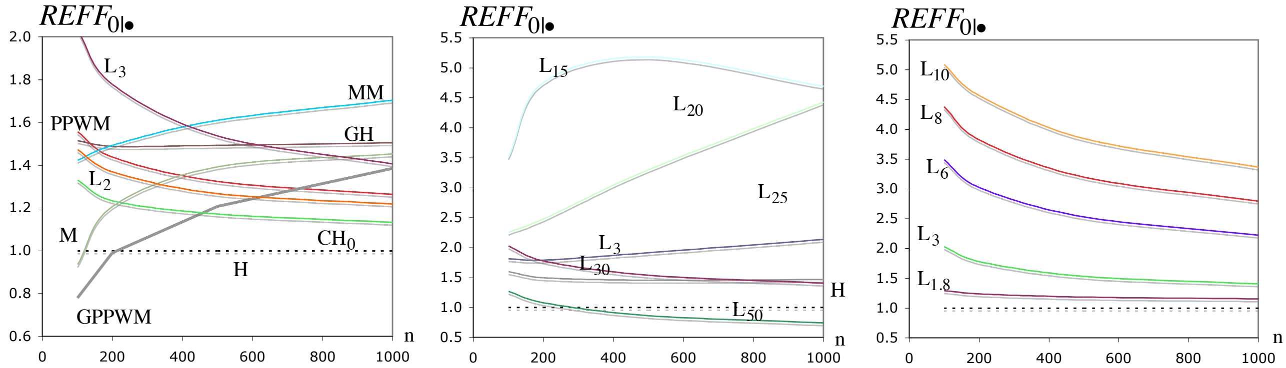

Figure 14: REFF-indicators at optimal levels, as a function of the sample size n, and for an $\mathrm{EV}_{0.25}$ underlying parent

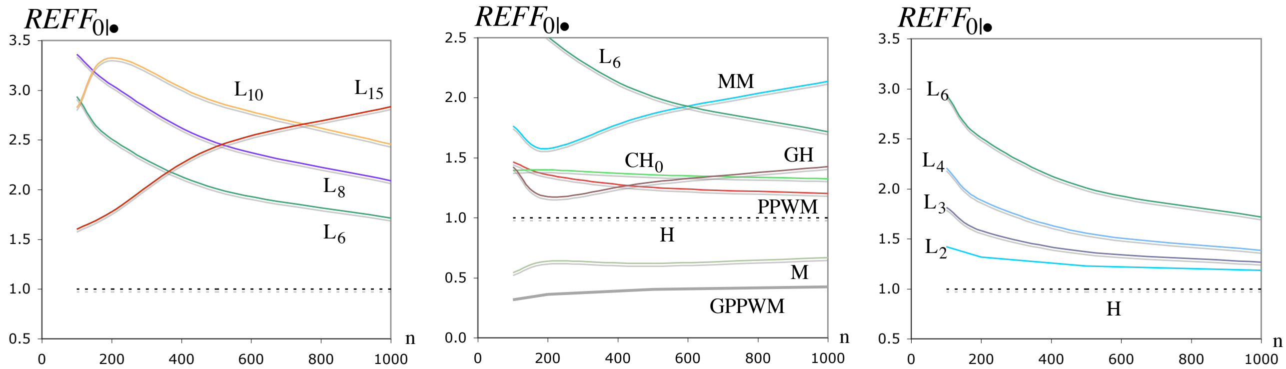

Figure 15: REFF-indicators at optimal levels, as a function of the sample size n, and for a Student t_4 underlying parent

6 An adaptive choice of (p, k)

A reasonably sophisticated algorithm, that has proved to work properly in many situations, including the choice of (p, k) in the MO_p EVI-estimation, is the double-bootstrap algorithm. The basic framework for such algorithm is next provided, but such an algorithm needs modifications to lead to large values of p, the ones that have revealed to be the most adequate ones for an efficient estimation of values of ξ close to zero.

For the new class of L_p EVI-estimators $L_p(k)$, in (1.9),

$$k_{0|p}(n) = \arg\min_k \mathrm{MSE}(\mathrm{L}_p(k)) = k_{A|p}(n)(1 + o(1)), \tag{6.1}$$

with

$$k_{A|p}(n) := \arg\min_k \mathrm{AMSE}\big(\mathrm{L}_p(k)\big). \tag{6.2}$$

For any admissible p, and provided that we can assure the asymptotic normality of the estimator under play, i.e. if $p \geq 1$, the bootstrap methodology can thus enable us to consistently estimate the optimal sample fraction (OSF), $k_{0|p}(n)/n$, with $k_{0|p}(n)$ given in (6.1), on the basis of a consistent estimator of $k_{A|p}(n)$, in (6.2), 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, in Brilhante *et al.* (2013), for the MO_p EVI-estimator, and in Gomes *et al.* (2011b, 2012b, 2016b), for second-order reduced-bias estimation. See also Caeiro and Gomes (2015b) for a review on OSF-estimation. With the notation $\lfloor x \rfloor$ for the integer part of x, we use again the auxiliary statistics

$$T_{k,n} \equiv T(k|\mathrm{L}_p) \equiv T_{k,n|p} := \mathrm{L}_p(\lfloor k/2 \rfloor) - \mathrm{L}_p(k), \quad k = 2, \ldots, n-1, \tag{6.3}$$

which converge in probability to zero, for any intermediate k, and have an asymptotic behaviour strongly related with the asymptotic behaviour of $\mathrm{L}_p(k)$. Indeed, under the aforementioned second-order framework in (2.2), we get, for all $p \geq 1$,

$$T(k|\mathrm{L}_p) \stackrel{d}{=} \frac{\sigma_{\mathrm{L}_p}\ Z_k^{(p)}}{\sqrt{k}} + b_{\mathrm{L}_p}(2^\rho - 1)\ A(n/k)(1 + o_p(1)),$$

with $Z_k^{(p)}$ asymptotically standard normal, and $\big(\sigma_{\mathrm{L}_p}, b_{\mathrm{L}_p}\big) = \big(\sigma_{\mathrm{L}_p}(\xi), b_{\mathrm{L}_p}(\rho)\big)$ given in (3.5).

Consequently, denoting $k_{0|T}(n) := \arg\min_k \mathrm{MSE}(T_{k,n})$, we have

$$k_{0|p}(n) = k_{0|T}(n) \times (1 - 2^\rho)^{\frac{2}{1-2\rho}}\ (1 + o(1)). \tag{6.4}$$

Given the random sample $\underline{X}_n = (X_1, \ldots, X_n)$ from any unknown model F, and the functional in (6.3), $T_{k,n} =: \phi_k(\underline{X}_n)$, $1 < k < n$, consider for any $n_1 = O(n^{1-\epsilon})$, $0 < \epsilon < 1$, 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 available sample, $\underline{X}_n$. Next, associate to the bootstrap sample, the corresponding bootstrap auxiliary statistic, $T^*_{k_1,n_1} := \phi_{k_1}(\underline{X}^*_{n_1})$, $1 < k_1 < n_1$. Then, with $k^*_{0|T}(n_1) = \arg\min_{k_1} \text{MSE}\big(T^*_{k_1,n_1}\big)$,

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

Consequently, for another sample size, $n_2 = \lfloor n_1^2/n \rfloor + 1$, we have

$$\big(k^*_{0|T}(n_1)\big)^2 / k^*_{0|T}(n_2) = k_{0|T}(n)(1 + o(1)), \quad \text{as } n \to \infty. \tag{6.5}$$

On the basis of (6.5), we are now able to first consistently estimate $k_{0|T}$, and next $k_{0|p}$ through (6.4), on the basis of any estimate $\hat{\rho}$ of the second-order parameter ρ. With $\hat{k}^*_{0|T}$ denoting the sample counterpart of $k^*_{0|T}$, $\hat{\rho}$ an adequate ρ-estimate, and $c_\rho = (1 - 2^\rho)^{\frac{2}{1-2\rho}}$, we thus have the k_0-estimate

$$\hat{k}^*_{0|p} \equiv \hat{k}^*_{0|p}(n; n_1) := \min\Big(n - 1,\ \big\lfloor c_{\hat{\rho}}\ (\hat{k}^*_{0|T}(n_1))^2 / \hat{k}^*_{0|T}(\lfloor n_1^2/n \rfloor + 1)\big\rfloor + 1\Big). \tag{6.6}$$

The adaptive estimate of ξ is then given by

$$\text{L}^*_p \equiv \text{L}^*_{p,n,n_1|T} := \text{L}_p(\hat{k}^*_{0|p}(n; n_1)). \tag{6.7}$$

6.1 A double-bootstrap algorithm for an adaptive L_p EVI-estimation

We now proceed with the description of an algorithm for the adaptive estimation of ξ. In **Steps 2**, **3** and **4**, we reproduce the algorithm provided in Gomes and Pestana (2007b) for the estimation of the second-order parameters β and ρ, already tested in several articles on a related subject.

Algorithm 6.1.

Step 1 *Given an observed sample* $(x_1, \ldots, x_n)$*, compute, for the tuning parameters* $\tau = 0$ *and* $\tau = 1$*, the observed values of* $\hat{\rho}_\tau(k)$*, the most simple class of estimators in* Fraga Alves *et al.* (2003). *Such estimators have the functional form*

$$\hat{\rho}_\tau(k) := \min\left(0,\ 3(W_{k,n}^{(\tau)} - 1)/(W_{k,n}^{(\tau)} - 3)\right), \tag{6.8}$$

dependent on the statistics

$$W_{k,n}^{(0)} := \frac{\ln\left(M_{k,n}^{(1)}\right) - \frac{1}{2}\ln\left(M_{k,n}^{(2)}/2\right)}{\frac{1}{2}\ln\left(M_{k,n}^{(2)}/2\right) - \frac{1}{3}\ln\left(M_{k,n}^{(3)}/6\right)}, \quad W_{k,n}^{(1)} := \frac{M_{k,n}^{(1)} - \left(M_{k,n}^{(2)}/2\right)^{1/2}}{\left(M_{k,n}^{(2)}/2\right)^{1/2} - \left(M_{k,n}^{(3)}/6\right)^{1/3}},$$

where $M_{k,n}^{(j)}$*,* $j = 1, 2, 3$*, are given in* (2.7).

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

Step 3 *Work with* $\hat{\rho} \equiv \hat{\rho}_{\tau^*} = \hat{\rho}_{\tau^*}(k_1)$ *and* $\hat{\beta} \equiv \hat{\beta}_{\tau^*} := \hat{\beta}_{\hat{\rho}_{\tau^*}}(k_1)$*, with* $k_1 = \lfloor n^{0.999} \rfloor$*, being* $\hat{\beta}_{\hat{\rho}}(k)$ *the estimator in* Gomes and Martins (2002)*, given by*

$$\hat{\beta}_{\hat{\rho}}(k) := \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{d_k(\hat{\rho})\ D_k(0) - D_k(\hat{\rho})}{d_k(\hat{\rho})\ D_k(\hat{\rho}) - D_k(2\hat{\rho})},$$

dependent on the estimator $\hat{\rho} = \hat{\rho}_{\tau^*}(k_1)$*, and where, for any* $\alpha \le 0$*,*

$$d_k(\alpha) := \frac{1}{k}\sum_{i=1}^{k} (i/k)^{-\alpha} \quad \text{and} \quad D_k(\alpha) := \frac{1}{k}\sum_{i=1}^{k} (i/k)^{-\alpha}\, U_i,$$

with $U_i = i\ \left(\ln X_{n-i+1:n} - \ln X_{n-i:n}\right)$*,* $1 \le i \le k < n$*, the scaled log-spacings.*

Step 4 *For* $p = 1(0.1)10$*, compute* $\mathrm{L}_p(k)$*,* $k = 1, 2, \ldots, n-1$.

Step 5 *Next, consider sub-sample sizes* $n_1 = \lfloor n^b \rfloor$*,* $b = 0.925(0.001)0.999$*,* $n_2 = \lfloor n_1^2/n \rfloor + 1$.

Step 6 *For* l *from* 1 *until* $B = 250$ (*number of bootstrap iterations*)*, generate independently, from the empirical* CDF $F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}}$ *associated with the observed sample, the 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.*

Step 7 *Denoting* $T^*_{k,n}$ *the bootstrap counterpart of* $T_{k,n}$, *in* (6.3), *obtain, for* $1 \le l \le B$, $t^*_{k,n_1,l}$, $1 < k < n_1$, $t^*_{k,n_2,l}$, $1 < k < n_2$, *the observed values of the statistic* T^*_{k,n_i}, $i = 1,2$, *and compute, for* $i = 1,2$ *and* $k = 2,\ldots,n_i - 1$, $\mathrm{MSE}^*(n_i,k) = \sum_{l=1}^{B} \big(t^*_{k,n_i,l}\big)^2/B$.

Step 8 *Obtain* $\hat{k}^*_{0|T}(n_i) := \arg\min_{1<k<n_i} \mathrm{MSE}^*(n_i,k)$, $i = 1,2$, *and return to* **Step 6** *if* $\hat{k}^*_{0|T}(n_2) > \hat{k}^*_{0|T}(n_1)$.

Step 9 *Compute* $\hat{k}^*_{0|p} \equiv \hat{k}^*_{0|p}(n;n_1)$, *given in* (6.6).

Step 10 *Compute* $\mathrm{L}^*_p \equiv \mathrm{L}^*_{p,n,n_1|T}$, *given in* (6.7), *and the* MSE-*estimate*

$$\widehat{\mathrm{MSE}}^*_p \equiv \widehat{\mathrm{MSE}}^*_p(n_1) \equiv \hat{\mathrm{MSE}}(\hat{k}^*_{0|p}|\mathrm{L}^*_p)$$
$$:= \frac{\sigma^2_{\mathrm{L}_p}(\mathrm{L}^*_p)}{\hat{k}^*_{0|p}} + \left(\frac{\mathrm{L}^*_p\hat{\beta}\sqrt{\Gamma(2p-1)}\big(n/\hat{k}^*_{0|p}\big)^{\hat{\rho}}}{\Gamma(p)}\right)^2 =: (\hat{\sigma}^*_{0p})^2 + (\hat{b}^*_{0p})^2, \quad (6.9)$$

where $\sigma_{\mathrm{L}_p}(\xi)$ *has been defined in* (3.5).

Step 11 *Compute* $n^*_1(p) := \arg\min_{n_1} \widehat{\mathrm{MSE}}^*_p(n_1)$, *with* $\widehat{\mathrm{MSE}}^*_p(n_1)$ *obtained in* **Step 10**, *and* $p^*_{min} := \arg\min n_1(p)$, *with the values* $n_1 = n_1(p)$ *given in* **Step 5**.

Step 12 *Consider the adaptive threshold estimate* $\hat{k}^{**}_0 := \hat{k}_{0|p^*_{min}}(n;n^*_1)$, $n^*_1 := n_1(p^*_{min})$, *and the final* EVI-*estimate* $\mathrm{L}^{**} := \mathrm{L}^*_{p^*_{min}} = \mathrm{L}^*_{p,n,n^*_1|T}$.

Remark 15. *For small values of* ξ, *this 'parametric' method, based on asymptotic variance and asymptotic dominant component of bias of the* L_p EVI-*estimators, is not leading to the large values of* p *associated with minimum simulated* RMSE. *It is instead leading to values close to the optimal asymptotic* p, *denoted* $p_{\mathrm{M|L}}$ *in* Section 4.1. *We thus think sensible to try any kind of non-parametric approach, out of the scope of this article. Note however that we have still tried the replacement of* **Steps 7–10** *by:*

Step 7' *Denoting* $T^*_{k,n}$ *the bootstrap counterpart of* $T_{k,n}$, *in* (6.3), *obtain, for* $1 \le l \le B$, $t^*_{k,n_1,l}$, $1 < k < n_1$, $t^*_{k,n_2,l}$, $1 < k < n_2$, *the observed values of the statistic* T^*_{k,n_i}, $i = 1,2$, *and compute, for* $i = 1,2$ *and* $k = 2,\ldots,n_i - 1$,

$$\mathrm{B}^*_{i,k} = \frac{1}{B}\sum_{l=1}^{B} t^*_{k,n_i,l}, \qquad \mathrm{M}^*_{i,k} = \frac{1}{B}\sum_{l=1}^{B} \big(t^*_{k,n_i,l}\big)^2.$$

Compute the bootstrap MSE-*estimate,*

$$\widehat{\mathrm{MSE}}_p^* \equiv \widehat{\mathrm{MSE}}_p^*(n_1) = \frac{M_{1,k}^* + M_{2,k}^* - \left((B_{1,k}^*)^2 + (B_{2,k}^*)^2\right)}{2} + \left(\frac{\left(B_{1,k}\right)^2}{\left(2^{\hat\rho}-1\right)B_{2,k}^*}\right)^2.$$

However, the results obtained were not far away from the ones obtained with **Algorithm 6.1**

Remark 16. *For any* $p \geq 1$*, and with* $\hat{k}_{0|p}^*$ *and* $(\sigma_{0p}^*, b_{0p}^*)$ *given in* (6.6) *and* (6.9)*, respectively, the* RV $\left(\mathrm{L}_p(\hat{k}_{0|p}^*) - \xi - b_{0p}^*\right)/\sigma_{0p}^*$ *is approximately* $\mathcal{N}(0,1)$*. We can then get approximate* $100(1-\alpha)\%$ *confidence intervals* (CIs) *for* ξ*, given by*

$$\left(\mathrm{L}_p(\hat{k}_{0|p}^*) - b_{0p}^* - \chi_{1-\alpha/2}\sigma_{0p}^*,\ \mathrm{L}_p(\hat{k}_{0|p}^*) - b_{0p}^* + \chi_{1-\alpha/2}\sigma_{0p}^*\right),$$

where χ_p *denotes the quantile of probability* p *of a standard normal* CDF.

Remark 17. *We further make the following general comments:*

(i) *The value of* n *must be replaced by* $n_0 := \sum_{i=1}^n I_{[X_i>0]}$*, the number of positive elements in the sample, whenever there are negative elements in the sample. A similar comment applies to the bootstrap sample sizes* n_1 *and* n_2*.*

(ii) *As already mentioned in several papers essentially related with bias reduction, in* **Step 2** *of the algorithm we are led in almost all situations to the tuning parameter* $\tau^* = 0$ *whenever* $-1 \leq \rho < 0$ *and* $\tau^* = 1$*, otherwise. We thus claim again for the relevance of the choice* $\tau = 0$*, in* (6.8)*, due to the importance of the region* $|\rho| \geq 1$*.*

(iv) *In* **Algorithm 6.1** *above, we have also dealt with the choice of the tuning parameter* n_1 *associated with the bootstrap methodology, but again, the method is only moderately dependent on the choice of the nuisance parameter* n_1*. This enhances the practical value of the method. Moreover, although aware of the need of* $n_1 = o(n)$*, it seems that, once again, we get good results up till* n*.*

(v) *The Monte-Carlo procedure in the* **Steps 6–12** *of* **Algorithm 6.1** *can be replicated, if we want to associate standard bootstrap errors to the* OSF *and to the* EVI-*estimates. The value of* B *can also be adequately chosen.*

(vi) *We would like to stress again that the use of the random sample of size* n_2*,* $(x_1^*, \ldots, x_{n_2}^*)$*, and of the extended sample of size* n_1*,* $(x_1^*, \ldots, x_{n_2}^*, x_{n_2+1}^*, \ldots, x_{n_1}^*)$*, leads us to increase the* *precision of the result with a smaller B. Indeed, if we had generated the sample of size n_1 independently of the sample of size n_2, we would have got a wider confidence interval for the* EVI*, should we have kept the same value for B. This is quite similar to the use of the simulation technique of "Common Random Numbers" in comparison algorithms, when we want to decrease the variance of a final answer to $z = y_1 - y_2$, inducing a positive dependence between y_1 and y_2.*

7 Appendix

Proof. (Theorem 5). As we have seen before in Section 1, the law of large numbers enables us to say the statistics in (1.9) are consistent for the estimation of $\xi \geq 0$ for all $p > 0$. With Y denoting again a unit Pareto RV, and working under the second-order framework in (2.2), we can write

$$
\begin{aligned}
\frac{1}{k}\sum_{i=1}^{k}\left(\ln\frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^p &= \frac{1}{k}\sum_{i=1}^{k}\left(\ln\frac{U(Y_{n-i+1:n})}{U(Y_{n-k:n})}\right)^p \\
&= \frac{1}{k}\sum_{i=1}^{k}\Big(\xi\ln Y_i + A(n/k)\ (Y_i^\rho - 1)/\rho + o_p(A(n/k))\Big)^p \\
&= \xi^p\ \frac{1}{k}\sum_{i=1}^{k}(\ln Y_i)^p + \frac{p\xi^{p-1}A(n/k)}{\rho}\ \frac{1}{k}\sum_{i=1}^{k}(\ln Y_i)^{p-1}\ (Y_i^\rho - 1)\ (1 + o_p(1)).
\end{aligned}
$$

Consequently, since

$$
\frac{1}{k}\sum_{i=1}^{k}(\ln Y_i)^{p-1}\ (Y_i^\rho - 1) \ \underset{n\to\infty}{\overset{p}{\longrightarrow}}\ \Gamma(p)\ \big[(1-\rho)^{-p} - 1\big],
$$

we can write

$$
M_n^{(p)}(k) \ \stackrel{d}{=}\ \xi^p\ \frac{1}{k}\sum_{i=1}^{k}E_i^p + \frac{\xi^{p-1}\Gamma(p+1)\big[(1-\rho)^{-p} - 1\big]A(n/k)}{\rho} + o_p(A(n/k)).
$$

Let $P_n^{(p)} := \sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}E_i^p - \Gamma(p+1)\right)$. Then (3.7), and the second-order structure of $\{P_n^{(p)}\}$ follows, with $\mathbb{V}\mathrm{ar}\big(P_n^{(p)}\big) = \Gamma(2p+1) - \Gamma^2(p+1)$. We can thus write

$$
M_n^{(p)}(k) \ \stackrel{d}{=}\ \xi^p\Gamma(p+1)\left\{1 + \frac{P_n^{(p)}}{\Gamma(p+1)\sqrt{k}} + \frac{(1-\rho)^{-p} - 1}{\xi\rho}A(n/k)(1 + o_p(1))\right\},
$$

and, for $p \geq 1$,

$$\frac{M_n^{(p)}(k)}{pM_n^{(p-1)}(k)} \stackrel{d}{=} \xi \left\{ 1 + \frac{P_n^{(p)}}{\Gamma(p+1)\sqrt{k}} - \frac{P_n^{(p-1)}}{\Gamma(p)\sqrt{k}} + \frac{(1-\rho)^{-p}}{\xi} A(n/k)(1+o_p(1)) \right\}.$$

Next note that

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\big(P_n^{(p-1)}, P_n^{(p)}\big) &= \mathbb{C}\mathrm{ov}\left(\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k} E_i^{p-1} - \Gamma(p) \right), \sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k} E_i^{p} - \Gamma(p+1) \right) \right) \\
&= k\ \mathbb{C}\mathrm{ov}\left(\frac{1}{k}\sum_{i=1}^{k} E_i^{p-1}, \frac{1}{k}\sum_{i=1}^{k} E_i^{p} \right) = \frac{\sum_{i=1}^{k} \mathbb{C}\mathrm{ov}\left(E_i^{p-1},\ E_i^{p}\right)}{k} \\
&= \mathbb{E}\big(E^{2p-1}\big) - \mathbb{E}\big(E^{p-1}\big)\ \mathbb{E}\big(E^{p}\big) = \Gamma(2p) - \Gamma(p)\Gamma(p+1).
\end{aligned}$$

We thus have

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}\left(\frac{P_n^{(p)}}{\Gamma(p+1)} - \frac{P_n^{(p-1)}}{\Gamma(p)} \right) \\
&\qquad = \frac{\Gamma(2p+1) - \Gamma^2(p+1)}{\Gamma^2(p+1)} + \frac{\Gamma(2p-1) - \Gamma^2(p)}{\Gamma^2(p)} - \frac{2\left(\Gamma(2p) - \Gamma(p)\Gamma(p+1)\right)}{\Gamma(p)\Gamma(p+1)} \\
&= \frac{\Gamma(2p-1)}{\Gamma^2(p)} = \frac{\Gamma(2p)}{(2p-1)\Gamma^2(p)} = \frac{(2\pi)^{-1/2}\ 2^{2p-\frac{1}{2}}\Gamma(p)\Gamma(p+\frac{1}{2})}{(2p-1)\Gamma^2(p)} = \frac{(2\pi)^{-1/2}\ 2^{2p-\frac{1}{2}}\Gamma(p+\frac{1}{2})}{(2p-1)\Gamma(p)}
\end{aligned}$$

Consequently, we obtain the validity of the asymptotic distributional representation in (3.6), with $\big(b_{\mathrm{L}_p}, \sigma_{\mathrm{L}_p}\big)$ given in (3.5). ■

Acknowledgements. Research partially supported by National Funds through **FCT** – Fundação para a Ciência e a Tecnologia, projects UID/MAT/00006/2013 (CEA/UL), PEst-OE/MAT/UI0297/2013 (CMA/UNL), and COST Action IC1408.

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., Vynckier, P. and Teugels, J. (1996). Excess functions and estimation of the extreme-value index. *Bernoulli* **2**, 293–318.

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

[4] Beirlant, J., Dierckx, G. and Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots. *Bernoulli* **11**:6, 949–970.

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

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

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

[8] Brilhante, M.F., Gomes, M.I. and Pestana, D. (2014). Brilhante, M.F., Gomes, M.I. and Pestana, D. (2014). The mean-of-order p extreme value index estimator revisited. In A. Pacheco *et al.* (Eds.), *New Advances in Statistical Modeling and Application*, Springer-Verlag, Berlin, Heidelberg, 163–175.

[9] Caeiro, F. and Gomes, M.I. (2002a). A class of asymptotically unbiased' semi-parametric estimators of the tail index. *Test* **11**:2, 345–364.

[10] Caeiro, F. and Gomes , M.I. (2002b). Bias reduction in the estimation of parameters of rare events. *Theory of Stochastic Processes* **8**:24, 67–76.

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

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

[13] Caeiro, F. and Gomes, M.I. (2014a). A semi-parametric estimator of a shape second order parameter. In A. Pacheco *et al.* (Eds.), *New Advances in Statistical Modeling and Application*, Springer-Verlag, Berlin, Heidelberg, 163–175.

[14] Caeiro, F. and Gomes, M.I. (2014b). Comparison of asymptotically unbiased extreme value index estimators: a Monte Carlo simulation study. *AIP Conference Proceedings* **1618**, 551–554.

[15] Caeiro, F. and Gomes, M.I. (2015a). Bias Reduction in the Estimation of a Shape Second-Order parameter of a Heavy Tailed Model. *J. Statist. Comput. and Simul.* **85**:17, 3405–3419.

[16] Caeiro, F. and Gomes, M.I. (2015b). Adaptive Choice of Optimal Threshold in EVA. Chapter in: Dipak Dey and Jun Yan, *Extreme Value Modeling and Risk Analysis*, Chapman-Hall/CRC, ISBN 9781498701297, pp. 69–87.

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

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

[19] 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.

[20] Caeiro, F., Gomes, M.I., Beirlant, J. and de Wet, T. (2015). *Mean-of-order p reduced-bias extreme value index estimation under a third-order framework.* Notas e Comunicações CEAUL 02/2015.

[21] 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.

[22] Ciuperca, G. and Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes* **13**:1, 55–87.

[23] Danielsson, J., de Haan, L., Peng, L. and de Vries, C.G. (2001). Using a bootstrap method to choose the sample fraction in the tail index estimation. *J. Multivariate Analysis* **76**, 226–248.

[24] Davison, A. (1984). Modeling excesses over high threshold with an application. In J. Tiago de Oliveira (Ed.), *Statistical Extremes and Applications*, D. Reidel, 461–482.

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

[26] Dekkers, A., Einmahl, J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* **17**, 1833–1855.

[27] Deme, E.H., Gardes, L. and Girard, S. (2013). On the estimation of the second order parameter for heavy-tailed distributions. *Revstat* **11**:3, 277–299.

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

[29] Drees, H., Ferreira, A. and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.* **14**, 1179–1201.

[30] Fraga Alves, M.I., Gomes, M.I. and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica* **60**:1, 193–213.

[31] Fraga Alves, M.I., Gomes, M.I., de Haan, L. and Neves, C. (2007). A note on second order conditions in extreme value theory: linking general and heavy tails conditions. *Revstat* **5**:3, 285–305.

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

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

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

[35] Goegebeur, Y., Beirlant, J. and de Wet, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation. *Revstat* **6**:1, 51–69.

[36] Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Planning and Inference* **140**:9, 2632–2654.

[37] 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 M. Gilli *et al.* (eds.). *Proceedings of COMPSTAT* 2014, The International Statistical Institute/International Association for Statistical Computing, pp. 289-298.

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

[39] 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.

[40] Gomes, M.I. and Henriques Rodrigues (2013). Adaptive PORT-MVRB Estimation of the Extreme Value Index. In Oliveira, P., Temido, M.G., Henriques, C and Vichi, M. (eds.) Studies in Theoretical and Applied Statistics: Subseries B: Recent Developments in Modeling and Applications in Statistics (SPE2010). Springer, pp. 117-125.

[41] 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.

[42] Gomes, M.I. and Martins, M.J. (2002). "Asymptotically unbiased" estimators of the extreme value index based on external estimation of the second order parameter. *Extremes* **5**:1, 5–31.

[43] 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.

[44] 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.

[45] Gomes, M.I. and Pestana, D.D. (2007a). A simple second order reduced-bias extreme value index estimator. *J. Statist. Computation and Simulation* **77**:6, 487–504.

[46] Gomes, M.I. and Pestana, D.D. (2007b). A sturdy reduced-bias extreme quantile (VaR) estimator. *J. American Statistical Association* **102**:477, 280–292.

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

[48] 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.

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

[50] Gomes, M.I., Fraga Alves, M.I. and Araújo-Santos, P. (2008a). PORT Hill and moment estimators for heavy-tailed models. *CommunṠtatist.—Simul. and Comput.* **37**, 1281–1306.

[51] 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.

[52] Gomes, M.I., Henriques-Rodrigues, L., Pereira, H. and Pestana, D. (2010). Tail index and second order parameters' semi-parametric estimation based on the log-excesses. *J. Statist. Comput. and Simul.* **80**:6, 653–666.

[53] Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2011a). Reduced-bias location-invariant extreme value index estimation: a simulation study. *Comm. Statist.—Simul. and Comput.* **40**:3, 424–447.

[54] Gomes, M.I., Mendonça, S. and Pestana, D.D. (2011b). Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology. *Comm. in Statistics—Theory and Methods* **40**:16, 2946–2968.

[55] Gomes, M.I., Caeiro, F. and Henriques-Rodrigues, L. (2012a). PORT-PPWM extreme value index estimation. In Colubi, A. *et al.* (eds.). *Proceedings of COMPSTAT* 2012, The International Statistical Institute/International Association for Statistical Computing, 259–270.

[56] Gomes, M.I., Figueiredo, F. and Neves, M.M. (2012b). Adaptive estimation of heavy right tails: the bootstrap methodology in action. *Extremes*, **15**, 463–489.

[57] 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.

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

[59] Gomes, M.I., Brilhante, M.F., Caeiro, F. and Pestana, D. (2015a). A new partially reduced-bias mean-of-order-p class of extreme value index estimators. *Computational Statistics and Data Analysis* **82**, 223–237.

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

[61] Gomes, M.I., Henriques-Rodrigues, L. and Manjunath, B.G. (2015c). Mean-of-order-p location-invariant extreme value index estimation. *Revstat*, Accepted.

[62] Gomes, M.I., Brilhante, M.F. and Pestana, D. (2016a). New reduced-bias estimators of a positive extreme value index. *Comm. in Statistics—Simulation and Computation* **45**, 1–30.

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

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

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

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

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

[68] Havil, J. (2003). *Gamma: Exploring Euler's Constant.* Princeton, NJ: Princeton Univ. Press.

[69] Henriques-Rodrigues, L., Gomes, M.I., Fraga Alves, M.I. and Neves, C. (2014). PORT-estimation of a shape second-order parameter. *Revstat* **12**:3, 299–328.

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

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

[72] Paulaskas, V. and Vaičiulis, M. (2013). On the improvement of Hill and some others estimators. *Lith. Math. J.* **53**, 336-355.

[73] Paulaskas, V. and Vaičiulis, M. (2015). A class of new tail index estimators. *Ann. Instit. Statist. Math.*, DOI: 10.1007/s10463-015-0548-3.

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

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

[76] Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15**:3, 1174–1207.

[77] de Wet, T., Goegebeur, Y. and Munch, M. (2012). Asymptotically unbiased estimation of the second order tail parameter in extreme value statistics. *Statistics and Probability Letters* **82**, 565–573.

[78] Worms, J. and Worms, R. (2012). Estimation of second order parameters using probability weighted moments. *ESAIM: Probability and Statistics* **16**, 97–113.

[79] Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *J. Multivariate Anal.*, **100**:4, 794–815.

[80] Zhou, C. (2010). The extent of the maximum likelihood estimator for the extreme value index. *J. Multivariate Anal.*, **101**:4, 971–983.