Mean-of-order p reduced-bias extreme value index estimation under a third-order framework

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

Jan Beirlant

Catholic University of Leuven, Belgium, e-mail: Jan.Beirlant@wis.kuleuven.be

Tertius de Wet

Stellenbosch University, South Africa, e-mail: tdewet@sun.ac.za.

April 26, 2015

Abstract

Recent reduced-bias versions of a very simple generalization of the 'classical' Hill estimator of a positive extreme value index (EVI) are put forward. The Hill estimator can be regarded as the logarithm of the mean-of-order-0 of a certain set of statistics. Instead of such a geometric mean, it is sensible to consider the *mean-of-order-p* (MOP) of those statistics, with $p \in \mathbb{R}$. Under a third-order framework, the asymptotic behaviour of the MOP, optimal MOP and associated reduced-bias classes of EVI-estimators is derived. Information on the dominant non-null asymptotic bias is also provided so that we can deal with an asymptotic comparison at optimal levels of some of those classes. Large-scale Monte-Carlo simulation experiments are undertaken to provide finite sample comparisons.

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

Keywords and phrases. *Bias estimation*; *heavy tails*; *optimal levels; semi-parametric reduced-bias estimation; statistics of extremes.*

1 Introduction

Given a sample of size n of *independent and identically distributed* (IID) or possibly weakly dependent *random variables* (RVs), $X_1, \ldots, X_n$, with a common *cumulative distribution function* (CDF) F, let us denote the associated ascending *order statistics* (OSs) by $X_{1:n} \leq \cdots \leq X_{n:n}$. Let us further assume that there exist sequences of real constants $\{a_n > 0\}$ and $\{b_n \in \mathbb{R}\}$ such that $(X_{n:n} - b_n)/a_n$ converges in distribution to a non-degenerate RV. Then, as first proved in Gnedenko (1943), the limiting CDF 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 then said to belong to the max-domain of attraction of EV_ξ and, as usual, we use the notation $F \in \mathcal{D}_\mathcal{M}(\mathrm{EV}_\xi)$. The parameter ξ in (1.1) is the *extreme value index* (EVI) for maxima, the primary parameter of large extreme events. This EVI measures the heaviness of the *right tail-function* (RTF)

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

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

Let us further use the notation $\mathcal{RV}_a$ for the class of regularly varying functions at infinity, with an index of regular variation equal to $a \in \mathbb{R}$, i.e. positive measurable functions $g(\cdot)$ such that for all $x > 0$, $g(tx)/g(t) \to x^a$, as $t \to \infty$ (see Bingham *et al.*, 1987, for details on regular variation). In this article we work with Pareto-type underlying models, with a positive EVI, or equivalently, RTFs such that $\overline{F}(x) = x^{-\alpha}L(x)$, $\alpha = 1/\xi > 0$, with $L \in \mathcal{RV}_0$, a regularly varying function with an index of regular variation equal to zero, i.e. a slowly varying function at infinity. These heavy-tailed models are quite common in many areas of application, such as biostatistics, computer science, finance, insurance and social sciences, among others.

For Pareto-type models, the 'classical' EVI-estimators are the Hill estimators (Hill, 1975), which are the averages of the log-excesses, i.e.

$$\mathrm{H}(k) := \frac{1}{k}\sum_{i=1}^{k} V_{ik}, \quad V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \le i \le k < n. \tag{1.3}$$

Note that we can write

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

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

$$\underline{\mathbf{U}} := \{U_{ik} := X_{n-i+1:n}/X_{n-k:n},\ 1 \le i \le k < n\}\,. \tag{1.4}$$

Brilhante *et al.* (2013) considered the *mean-of-order-p* (MOP) of $\underline{\mathbf{U}}$, in (1.4), with $p \ge 0$. Here, we more generally consider $p \in \mathbb{R}$, i.e., the class of statistics

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

and the class of MOP EVI-estimators,

$$\mathrm{H}_p(k) := \begin{cases} \left(1 - \mathrm{M}_p^{-p}(k)\right)/p, & \text{if} \quad p < 1/\xi,\ p \neq 0, \\ \\ \ln \mathrm{M}_0(k) = \mathrm{H}(k), & \text{if} \quad p = 0, \end{cases} \tag{1.5}$$

with $\mathrm{H}_0(k) \equiv \mathrm{H}(k)$, given in (1.3). This class of MOP EVI-estimators depends on this tuning parameter $p \in \mathbb{R}$, and it is consistent for any real $p < 1/\xi$. It is indeed a highly flexible class of EVI-estimators, but it is not asymptotically unbiased for the moderate k-values leading to minimum *mean square error* (MSE).

For technical simplicity, consider Hall-Welsh class of models (Hall and Welsh, 1985), dependent on a vector (β, ρ) of second-order parameters and detailed in Section 2.1. Further considering values of p such that the asymptotic normality of the estimators in (1.5) was known to hold at the time, i.e. $0 \le p < 1/(2\xi)$, Brilhante *et al.* (2014) noticed that there is an optimal value

$$p \equiv p_{\mathrm{M}} = \varphi_\rho/\xi, \quad \text{with} \quad 0 < \varphi_\rho = 1 - \rho/2 - \sqrt{\rho^2 - 4\rho + 2}\Big/2 \le 1 - \sqrt{2}/2 < 1/2, \tag{1.6}$$

which maximises the asymptotic efficiency of the class of estimators in (1.5). Indeed, this result holds also for any real $p < 1/(2\xi)$. Then, an optimal MOP RV,

$$\overline{\mathrm{H}}(k) := \mathrm{H}_{p_{\mathrm{M}}}(k), \tag{1.7}$$

with $\mathrm{H}_p(k)$ and p_{M} given in (1.5) and (1.6), respectively, was considered and its asymptotic behaviour was further derived. In the aforementioned Hall-Welsh class of models, the dominant component of the bias of $\overline{\mathrm{H}}(k)$, given in (1.7), is $\xi\beta(1-\varphi_\rho)(n/k)^\rho/(1-\rho-\varphi_\rho)$. Such a behaviour has led Gomes *et al.* (2013) to directly remove the bias of $\overline{\mathrm{H}}(k)$, introducing an associated class of reduced-bias optimal MOP RVs. With the notation RB for *reduced-bias*, and on the basis of

$$\mathrm{RB}(k; p, \beta, \rho, \phi) := \mathrm{H}_p(k)\left(1 - \frac{\beta(1-\phi)}{1-\rho-\phi}\left(\frac{n}{k}\right)^\rho\right), \tag{1.8}$$

they defined

$$\overline{\mathrm{RB}}(k) := \overline{\mathrm{RB}}_{\mathrm{p_M}}(k) = \mathrm{RB}(k; p_{\mathrm{M}}, \beta, \rho, \varphi_\rho) = \overline{\mathrm{H}}(k)\Big(1 - \frac{\beta\big(1-\varphi_\rho\big)}{1-\rho-\varphi_\rho}\Big(\frac{n}{k}\Big)^\rho\Big), \tag{1.9}$$

with φ_ρ and $\overline{\mathrm{H}}(k)$ given in (1.6) and (1.7), respectively. The class of RVs in (1.9) is similar in spirit to the *minimum-variance reduced-bias* (MVRB) *corrected-Hill* (CH) EVI-estimators in Caeiro *et al.* (2005), and the main reasons for such a consideration are also similar to the ones presented in the aforementioned article.

But we cannot forget that $p_{\mathrm{M}} = \varphi_\rho/\xi$ depends on ξ and ρ, and *a priori*, we have to estimate both ξ and ρ in order to have EVI-estimators, based in (1.7) and (1.9), respectively. It is thus sensible to consider, just as done in Gomes *et al.* (2015), the partially RB MOP class of EVI-estimators,

$$\overline{\mathrm{RB}}_p(k) := \mathrm{RB}\left(k; p, \hat\beta, \hat\rho, \varphi_{\hat\rho}\right) = \mathrm{H}_p(k)\Big(1 - \frac{\hat\beta(1-\varphi_{\hat\rho})}{1-\hat\rho-\varphi_{\hat\rho}}\left(\frac{n}{k}\right)^{\hat\rho}\Big), \tag{1.10}$$

still dependent on a tuning parameter p. Moreover, we can further consider the class of RB MOP EVI-estimators,

$$\overline{\overline{\mathrm{RB}}}_p(k) := \mathrm{RB}\left(k; p, \hat\beta, \hat\rho, p\mathrm{H}_p(k)\right) = \mathrm{H}_p(k)\Big(1 - \frac{\hat\beta(1-p\mathrm{H}_p(k))}{1-\hat\rho-p\mathrm{H}_p(k)}\left(\frac{n}{k}\right)^{\hat\rho}\Big), \tag{1.11}$$

also dependent on a tuning parameter p, and the class of RVs,

$$\begin{aligned}\overline{\overline{\mathrm{RB}}}(k) &:= \overline{\overline{\mathrm{RB}}}_{\mathrm{p_M}}(k) = \mathrm{RB}\left(k; \mathrm{p_M}, \beta, \rho, \varphi_\rho\overline{\mathrm{H}}(k)/\xi\right)\\ &= \overline{\mathrm{H}}(k)\Big(1 - \frac{\beta\big(1-\varphi_\rho\overline{\mathrm{H}}(k)/\xi\big)}{1-\rho-\varphi_\rho\overline{\mathrm{H}}(k)/\xi}\Big(\frac{n}{k}\Big)^\rho\Big),\end{aligned} \tag{1.12}$$

with p_{M}, $\overline{\mathrm{H}}(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$ given in (1.6), (1.7) and (1.11), respectively. Further note that for $p = 0$ in (1.11) we get the simplest class of CH EVI-estimators provided in Caeiro *et al.* (2005), the pioneering article on MVRB EVI-estimation, i.e. the class in (1.11) generalizes the CH class of EVI-estimators,

$$\mathrm{CH}(k) \equiv \mathrm{CH}(k; \hat{\beta}, \hat{\rho}) := \mathrm{RB}\left(k; 0, \hat{\beta}, \hat{\rho}, 0\right) = \overline{\overline{\mathrm{RB}}}_0(k) = \mathrm{H}(k)\left(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})\right), \quad (1.13)$$

with $\mathrm{RB}(k; p, \beta, \rho, \phi)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$ given in (1.8) and (1.11), respectively.

In Section 2, next to a few technical details in the field of *extreme value theory* (EVT), we make a brief reference to the estimation of second-order parameters. In Section 3, and under a third-order framework, we deal with the asymptotic behaviour of the class of MOP EVI-estimators in (1.5), for any real $p < 1/\xi$, and consequently of the particular case in (1.7). We further proceed with the study of the RB classes, in (1.9), (1.10), (1.11) and (1.12), providing information on the dominant non-null asymptotic bias, so that we can deal, in Section 4.1, with the asymptotic comparison, at optimal levels, of the CH and the optimal RB MOP classes of EVI-estimators. In Section 4.2 the asymptotic comparison at optimal levels of the MOP EVI-estimators, for a real p, is developed. Section 5 is dedicated to the finite sample properties of the classes of RB MOP EVI-estimators, in comparison to to the behaviour of CH EVI-estimators, in (1.13), done through a large-scale simulation study. Some overall conclusions are drawn in Section 6. Finally, in Section 7, the proofs of the theorems in Section 3 are given.

2 Preliminary results in the area of EVT

In statistics of univariate extremes, whenever working with large values, i.e. with the RTF of the model F underlying the data, F is usually said to be heavy-tailed whenever the RTF, in (1.2), is a regularly varying function with a negative index of regular variation equal to $-1/\xi$, $\xi > 0$. Then (Gnedenko, 1943), if $\overline{F} \in \mathcal{RV}_{-1/\xi}$, with $\xi > 0$, F is in the max-domain of attraction of $\mathrm{EV}_\xi(\cdot)$, in (1.1), $\xi > 0$, and we use the notation $F \in \mathcal{D}_{\mathcal{M}}\left(EV_\xi\right)_{\xi>0} =: \mathcal{D}_{\mathcal{M}}^+$. Conversely, as also proved by Gnedenko (1943), if $F \in \mathcal{D}_{\mathcal{M}}^+$ we necessarily have $\overline{F} \in \mathcal{RV}_{-1/\xi}$.

2.1 A brief review of first and higher-order conditions for a heavy RTF

If $F \in \mathcal{D}_{\mathcal{M}}^{+}$, and with the notation $F^{\leftarrow}(t) := \inf\{x : F(x) \geq t\}$ for the generalised inverse function of F, the tail quantile function,

$$U(t) := F^{\leftarrow}(1 - 1/t), \quad t > 1, \tag{2.1}$$

is of regular variation with index ξ (de Haan, 1984). We thus assume the validity of any of the following equivalent first-order conditions (FOCs):

$$F \in \mathcal{D}_{\mathcal{M}}^{+} \quad \Longleftrightarrow \quad \overline{F} \in \mathcal{RV}_{-1/\xi} \quad \Longleftrightarrow \quad U \in \mathcal{RV}_{\xi}. \tag{2.2}$$

The *second-order parameter* ρ (≤ 0) rules the rate of convergence in the FOC, in (2.2), and can be defined, for $U(\cdot)$ in (2.1), as the non-positive parameter appearing in the limiting relation

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

which is assumed to hold for every $x > 0$, with A ultimately decreasing and where $|A|$ must then be of regular variation with index ρ (Geluk and de Haan, 1987). This second-order condition (SOC) has been widely accepted as an appropriate condition to specify a Pareto-type distribution and enables easily the derivation of the non-degenerate bias of EVI-estimators, under a semi-parametric framework. For further details on the topic, see Beirlant *et al.* (2004) and de Haan and Ferreira (2006).

Whenever dealing with bias reduction, it is usual to consider a slightly more restricted class than $\mathcal{D}_{\mathcal{M}}^{+}$, where

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

as $t \to \infty$, $C > 0$, $\xi > 0$, $\rho < 0$ and $\beta \neq 0$ (Hall and Welsh, 1985). To assume (2.4) is equivalent to assume that (2.3) holds with $A(t) = \xi\beta t^{\rho}$, $\rho < 0$. Models like the log-gamma are thus excluded from this class. The standard Pareto ($\rho = -\infty$) is also excluded. But most heavy-tailed models used in applications, such as the $\mathrm{EV}_\xi(\cdot)$, in (1.1), the *generalized Pareto*, $\mathrm{GP}_\xi(x) = 1 + \ln \mathrm{EV}_\xi(x),\ x \geq 0$, the Fréchet, $F_\xi(x) = \exp(-x^{-1/\xi}),\ x > 0$, and the Student's t CDFs, among others, belong to Hall-Welsh class.

To obtain information on the asymptotic bias of RB EVI-estimators, it is sensible to further assume a third-order condition (TOC), ruling the rate of convergence in (2.3), and which guarantees that

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx)-\ln U(t)-\xi \ln x}{A(t)} - \psi_\rho(x)}{B(t)} = \psi_{\rho+\rho'}(x), \tag{2.5}$$

again with $U(\cdot)$ given in (2.1), and where $|B| \in RV_{\rho'}$. It is often assumed that (2.5) holds with $\rho,\ \rho' < 0$. More specifically, we assume the slightly more restrictive TOC,

$$U(t) = C\ t^\xi \Big(1 + \xi\beta t^\rho/\rho + O\big(t^{2\rho}\big)\Big). \tag{2.6}$$

To assume that (2.6) holds is equivalent to saying that (2.5) holds with $\rho = \rho' < 0$ and

$$A(t) = \xi\ \beta\ t^\rho, \quad B(t) = \beta'\ t^\rho = \frac{\beta' A(t)}{\beta\xi} =: \frac{\zeta A(t)}{\xi}, \quad \zeta = \beta'/\beta \neq 0, \tag{2.7}$$

where β and β' can possibly be arbitrary slowly varying functions.

2.2 Asymptotic behaviour of EVI-estimators under a SOC

In order to have consistency of the Hill EVI-estimators, over the whole $\mathcal{D}^+_{\mathcal{M}}$, we need to work with intermediate values of k, i.e., a sequence of integers $k = k_n$, $1 \le k < n$, such that

$$k = k_n \to \infty \quad \text{and} \quad k_n = o(n),\ \text{as } n \to \infty. \tag{2.8}$$

Under the aforementioned SOC, in (2.3), the asymptotic distributional representation

$$\mathrm{H}(k) \ \stackrel{d}{=}\ \xi + \frac{\xi\ Z_k}{\sqrt{k}} + \frac{A(n/k)}{1-\rho}(1 + o_p(1)) \tag{2.9}$$

holds (de Haan and Peng, 1998), where, with $\{E_i\}_{i\ge1}$ a sequence of IID standard exponential RVs,

$$Z_k := \sqrt{k}\Big(\sum_{i=1}^{k} E_i/k - 1\Big) \tag{2.10}$$

is an asymptotically standard normal RV.

The Hill EVI–estimators usually reveal a high asymptotic bias. Indeed, and with the notation $\mathcal{N}(\mu, \sigma^2)$ indicating a normal RV with mean value μ and variance σ^2, it follows from (2.9) that under the general SOC in (2.3),

$$\sqrt{k}\,(\mathrm{H}(k) - \xi) \ \stackrel{d}{=}\ \mathcal{N}\left(0, \xi^2\right) + b^{(1)}_{\mathrm{H}} \sqrt{k} A(n/k) + o_p\Big(\sqrt{k} A(n/k)\Big),$$

where the bias $b^{(1)}_{\mathrm{H}}\sqrt{k}A(n/k) = \sqrt{k}A(n/k)/(1-\rho)$ can be very large, moderate or small, i.e. go to infinity, constant or zero as $n \to \infty$. Under the same conditions as before, and with CH given in (1.13), $\sqrt{k}\,(\mathrm{CH}(k) - \xi)$ is asymptotically normal with variance also equal to ξ^2 but with a null mean value. Indeed, from the results in Caeiro *et al.* (2005), we know that it is possible to adequately estimate the second-order parameters β and ρ, so that we get

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

More specifically, under the Pareto-type TOC in (2.6), we can adequately estimate the vector of second-order parameters, (β, ρ), and write (Caeiro *et al.*, 2009; Caeiro and Gomes, 2011),

$$\sqrt{k}\,(\mathrm{CH}(k) - \xi) \;\stackrel{d}{=}\; \mathcal{N}\left(0, \xi^2\right) + b^{(2)}_{\mathrm{CH}}\sqrt{k}\; A^2(n/k) + O_p\big(A(n/k)\big) + o_p\Big(\sqrt{k}A^2(n/k)\Big),$$
$$b^{(2)}_{\mathrm{CH}} = \frac{1}{\xi}\Big(\frac{\zeta}{1-2\rho} - \frac{1}{(1-\rho)^2}\Big),$$

with ζ given in (2.7). Consequently, $\mathrm{CH}(k)$ outperforms $\mathrm{H}(k)$ for all integer k.

2.3 Comments on the estimation of second-order parameters

We shall consider the class of estimators of the second-order parameter ρ proposed by Fraga Alves *et al.* (2003). Under adequate general conditions, they are semi-parametric asymptotically normal estimators of ρ, whenever $\rho < 0$, which show, for a large variety of models and for a wide range of large k-values, highly stable sample paths as functions of k, the number of top OSs used. Such a class of estimators has been first parameterised by a tuning parameter $\tau > 0$, but more generally τ can be considered as a real number (Caeiro and Gomes, 2006). It is defined as

$$\hat\rho(k;\tau) \equiv \hat\rho_\tau(k) := -\Big|3\big(\mathrm{V}_n(k;\tau) - 1\big)/\big(\mathrm{V}_n(k;\tau) - 3\big)\Big|, \tag{2.11}$$

where, with V_{ik} given in (1.3), $\mathrm{M}^{(\ell)}_{k,n} := \frac{1}{k}\sum_{i=1}^k V^{\ell}_{ik}$, $\ell \geq 1$, and the notation $a^{b\tau} = b\ln a$, whenever $\tau = 0$,

$$\mathrm{V}_n(k;\tau) := \frac{\big(\mathrm{M}^{(1)}_{k,n}\big)^{\tau} - \big(\mathrm{M}^{(2)}_{k,n}/2\big)^{\tau/2}}{\big(\mathrm{M}^{(2)}_{k,n}/2\big)^{\tau/2} - \big(\mathrm{M}^{(3)}_{k,n}/6\big)^{\tau/3}}, \qquad \tau \in \mathbb{R}.$$

Consistency and asymptotic normality of the estimators in (2.11) were proved in Fraga Alves *et al.* (2003). Under the SOC in (2.3), with $\rho < 0$, if (2.8) holds and $\sqrt{k}\;A(n/k) \to \infty$, as $n \to \infty$, the statistics $\hat\rho(k;\tau)$ in (2.11) converge in probability to ρ, as $n \to \infty$, for any real τ. Moreover, under the TOC, in (2.5), if $\sqrt{k}A^2(n/k) \to \lambda_{_A}$, finite, and $\sqrt{k}A(n/k)B(n/k) \to \lambda_{_B}$, also finite, $\sqrt{k}A(n/k)\,(\hat\rho(k;\tau) - \rho)$ is asymptotically normal with asymptotic variance $\sigma^2_{_R}$ and a possibly non-null asymptotic bias. Caeiro *et al.* (2009) made explicit both the asymptotic bias and variance of the ρ-estimators in (2.11).

Remark 1 (Adequate choice of k_1 for the ρ-estimation)**.** *As stated in* Caeiro and Gomes (2008) *the ideal situation would be the choice of an 'optimal' level* $k_1^{\rm opt}$ *for the estimation of* ρ*, in the sense of minimal* MSE. *Denoting* $\hat\rho = \hat\rho(k_1^{\rm opt};\tau)$ *any of the* ρ*-estimators in* (2.11) *computed at such a* $k_1^{\rm opt}$*,*

$$\hat\rho - \rho = o_p(1/\ln n), \quad \textit{as } n \to \infty, \tag{2.12}$$

a condition needed for an MVRB EVI*-estimation. We stress that in practice, such a* $k_1^{\rm opt}$*, being of a high theoretical interest, has only a 'limited' interest, at the current state-of-the-art. However, if we consider a level* k_1 *of the order of* $n^{1-\epsilon}$*, for some small* $\epsilon > 0$*, we can also guarantee* (2.12) *for a large class of models* (*see* Caeiro *et al.*, 2009, *among others*)*. This is the reason why, such as done in* Caeiro *et al.* (2005)*,* Gomes and Pestana (2007) *and* Gomes *et al.* (2007; 2008)*, the pioneering articles in* MVRB*-estimation, we advise in practice, as a compromise between theoretical and practical considerations, the use of any intermediate level such as* $k_1 = \lfloor n^{1-\epsilon} \rfloor$ *for some* $\epsilon > 0$*, small, with* $\lfloor x \rfloor$ *denoting the integer part of* x*. The choice of* ϵ *is not crucial. Further considerations on the choice of* k_1 *can be found in* Caeiro *et al.* (2009)*.*

Remark 2 (Choice of the tuning parameter τ in the ρ-estimation)**.** *The theoretical and simulated results in* Fraga Alves *et al.* (2003)*, together with their use in* RB *estimation, lead us again to advise in practice the use of* $\tau = 0$ *for* $\rho \in [-1, 0)$ *and* $\tau = 1$ *for* $\rho \in (-\infty, -1)$*. However, practitioners should not choose blindly the value of* τ *in* (2.11)*. It is sensible to draw a few sample paths of* $\hat\rho_\tau(k) = \hat\rho(k;\tau)$*, as functions of* k*, electing the value of* τ *which provides the highest stability for large* k*, by means of any stability criterion, such as the one suggested in* Gomes and Pestana (2007)*. Note that in practice, the choice of* τ *is much more crucial than the choice of* k_1*. In the simulation study in* Section 5 *we are interested in models with* $\rho \in [-1, 0)$*, the region where bias reduction is indeed strongly needed, and we thus always consider* $\tau = 0$*.*

Regarding the β-estimation, we shall consider the β-estimators first obtained in Gomes and Martins (2002) and based on the scaled log-spacings $U_i := i\{\ln X_{n-i+1:n} - \ln X_{n-i:n}\}$, $1 \le i < n$. On the basis of any consistent estimator $\hat\rho$ of the second-order parameter ρ, we shall consider the β-estimator, $\hat\beta(k;\hat\rho)$, where, with $\rho < 0$,

$$\hat\beta(k;\rho) \equiv \hat\beta_\rho(k) := \frac{\left(\frac{k}{n}\right)^\rho \left\{\left(\frac1k \sum_{i=1}^k \left(\frac ik\right)^{-\rho}\right)\left(\frac1k \sum_{i=1}^k U_i\right) - \left(\frac1k \sum_{i=1}^k \left(\frac ik\right)^{-\rho} U_i\right)\right\}}{\left(\frac1k \sum_{i=1}^k \left(\frac ik\right)^{-\rho}\right)\left(\frac1k \sum_{i=1}^k \left(\frac ik\right)^{-\rho} U_i\right) - \left(\frac1k \sum_{i=1}^k \left(\frac ik\right)^{-2\rho} U_i\right)}. \tag{2.13}$$

Gomes and Martins (2002), under the second-order framework in (2.3), obtained the asymptotic behaviour of $\hat\beta(k;\rho)$ in (2.13). The full derivation of the asymptotic behaviour of $\hat\beta(k;\hat\rho)$ under a third-order framework, was derived in Caeiro *et al.* (2009), as a generalization of a result in Gomes *et al.* (2008). If we replace, in (2.13), ρ by $\hat\rho(k;\tau)$ given in (2.11), the rate of convergence of $\hat\beta(k;\hat\rho(k;\tau))$ is no longer of the order of $1/\big(\sqrt{k}A(n/k)\big)$ but of the order of $\ln(n/k)/(\sqrt{k}\ A(n/k))$, which must converge to zero, so that $\hat\beta(k;\hat\rho(k;\tau))$ is consistent for the estimation of β, and

$$\hat\beta(k;\hat\rho(k;\tau)) - \beta \overset{p}{\sim} -\beta \ln(n/k)\big(\hat\rho(k;\tau) - \rho\big). \tag{2.14}$$

If apart from $\sqrt{k}\ A(n/k)/\ln(n/k) \to \infty$, we further assume that $\sqrt{k}\ A^2(n/k) \to \lambda_{_A}$, finite, then $\sqrt{k}\ A(n/k)\big(\hat\beta(k;\hat\rho(k;\tau)) - \beta\big)/(\beta\ln(n/k))$ is asymptotically normal.

Remark 3. *If we consider* $\hat\beta \equiv \hat\beta(k_1;\hat\rho)$*, with* $\hat\rho$ *any of the estimators in* (2.11)*, computed also at the level* k_1*,* $\hat\beta - \beta$ *is thus, from* (2.14)*, of the order of* $(\hat\rho - \rho)\ln(n/k_1)$*. Consequently, the validity of* (2.12)*, enables us to guarantee the consistency of* $\hat\beta \equiv \hat\beta(k_1;\hat\rho)$*.*

Remark 4. *We have so far advised the use of the* ρ*-estimators in* Fraga Alves *et al.* (2003) *and the* β*-estimators in* Gomes and Martins (2002)*. Note however that several recent classes of* β *and* ρ*–estimators are potential candidates for the* (β,ρ)*-estimation. For information on the topic, see the recent reviews by* Beirlant *et al.* (2012) *and* Gomes and Guillou (2014)*.*

3 Asymptotic behaviour under a TOC

Trivial adaptations of the results in de Haan and Peng (1998), Caeiro *et al.* (2005), Brilhante *et al.* (2013, 2014) and Gomes *et al.* (2013; 2015) enable us to state the following theorem, still under FOC and SOC only.

Theorem 1. *Under the validity of the* FOC, *in* (2.2), *and for intermediate sequences* $k = k_n$, *i.e. if* (2.8) *holds, the classes of* EVI-*estimators* $\mathrm{H}_p(k)$, $\overline{\mathrm{RB}}_p(k)$, $\overline{\overline{\mathrm{RB}}}_p(k)$ *and* $\mathrm{CH}(k) = \overline{\overline{\mathrm{RB}}}_0(k)$, *respectively in* (1.5), (1.10), (1.11), *and* (1.13), *with* $p < 1/\xi$, *as well as the classes of* RVs *in* (1.7), (1.9) *and* (1.12), *generally denoted* $\mathrm{E}^\bullet(k)$, *are consistent for the estimation of* ξ. *If we assume the validity of the* SOC *in* (2.3) *and additionally assume that we are working with values of* k *such that* $\lambda := \lim_{n\to\infty} \sqrt{k}\, A(n/k)$ *is finite, we can then guarantee that, with* $p < 1/(2\xi)$,

$$\sqrt{k}\,(\mathrm{E}^\bullet(k) - \xi) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{N}\left(\lambda b_\bullet^{(1)}, \sigma_\bullet^2\right),$$

where, with φ_ρ *defined in* (1.6),

$$b_{\mathrm{H}_p}^{(1)} = \frac{1-p\xi}{1-\rho-p\xi}, \qquad b_{\overline{\mathrm{H}}}^{(1)} = \frac{1-\varphi_\rho}{1-\rho-\varphi_\rho}, \quad \sigma_{\mathrm{H}_p}^2 = \frac{\xi^2(1-p\xi)^2}{1-2p\xi}, \quad \sigma_{\overline{\mathrm{H}}}^2 = \frac{\xi^2(1-\varphi_\rho)^2}{1-2\varphi_\rho}.$$

If we further assume to be working in Hall-Welsh *class of models in* (2.6), *and estimate* β *and* ρ *consistently through* $\hat\beta$ *and* $\hat\rho$, *with* $\hat\rho - \rho = o_p(1/\ln n)$, *we get*

$$b_{\mathrm{CH}}^{(1)} = b_{\overline{\overline{\mathrm{RB}}}_p}^{(1)} = b_{\overline{\mathrm{RB}}}^{(1)} = b_{\overline{\overline{\mathrm{RB}}}}^{(1)} = 0, \quad b_{\overline{\mathrm{RB}}_p}^{(1)} = \frac{\rho(p\xi-\varphi_\rho)}{(1-p\xi-\rho)(1-\rho-\varphi_\rho)}, \tag{3.1}$$

$$\sigma_{\mathrm{CH}}^2 = \sigma_{\mathrm{H}}^2 = \xi^2, \qquad \sigma_{\overline{\mathrm{RB}}_p}^2 = \sigma_{\overline{\overline{\mathrm{RB}}}_p}^2 = \sigma_{\mathrm{H}_p}^2 = \frac{\xi^2(1-p\xi)^2}{1-2p\xi},$$
$$\sigma_{\overline{\mathrm{RB}}}^2 = \sigma_{\overline{\overline{\mathrm{RB}}}}^2 = \sigma_{\overline{\mathrm{H}}}^2 = \frac{\xi^2(1-\varphi_\rho)^2}{1-2\varphi_\rho}. \tag{3.2}$$

Remark 5. *On the basis of* Theorem 1 *we can say that in* Hall-Welsh *class of models in* (2.6) *both* $\overline{\mathrm{RB}}(k)$ *and* $\overline{\overline{\mathrm{RB}}}(k)$ *outperform* $\overline{\mathrm{H}}(k)$ *for all* k, *just as* $\mathrm{CH}(k)$ *outperforms* $\mathrm{H}(k)$.

A generalisation of Theorem 2 in Brilhante *et al.* (2013) enables us to further state:

Theorem 2. *Under the validity of the* TOC *in* (2.6), *or equivalently, if we assume that both* (2.5) *and* (2.7) *hold, we have for any real number* $p < 1/(2\xi)$ *the validity of the following asymptotic distributional representation,*

$$\mathrm{H}_p(k) \overset{d}{=} \xi + \frac{\xi(1-p\xi)V_k^{(p)}}{\sqrt{k(1-2p\xi)}} + \frac{(1-p\xi)A(n/k)}{1-p\xi-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)$$
$$+ \frac{(1-p\xi)\big(\zeta(1-p\xi-\rho)^2 + p\xi\rho\big)A^2(n/k)}{\xi(1-p\xi-\rho)^2(1-p\xi-2\rho)} + o_p(A^2(n/k)), \tag{3.3}$$

with $V_k^{(p)}$ *asymptotically standard normal. Consequently, if we consider* $\overline{\mathrm{H}}$, *in* (1.7), *we get* (3.3), *with* $p\xi$ *replaced by* φ_ρ, *the function defined in* (1.6).

Remark 6. *If we consider* $p = 0$ *in* (3.3) *and with* $V_k^{(0)} = Z_k$*, as defined in* (2.10)*, we get*

$$\mathrm{H}_0(k) \equiv \mathrm{H}(k) \;\stackrel{d}{=}\; \xi + \frac{\xi Z_k}{\sqrt{k}} + \frac{A(n/k)}{1-\rho} + \frac{\zeta A^2(n/k)}{\xi(1-2\rho)} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) + o_p(A^2(n/k)),$$

in agreement with the results in Caeiro *et al.* (2009)*.*

Generalizing Theorem 2 in Gomes *et al.* (2013) and Theorems 2 and 3 in Gomes *et al.* (2015), we next provide the behaviour of the RB classes in (1.9), (1.10), (1.11) and (1.12). We consider first, in Theorem 3, the behaviour of the RVs, denoted $\mathrm{RB}(k;\beta,\rho)$, for sake of simplicity. In Theorem 4 we state the asymptotic behaviour of $\mathrm{RB}(k;\hat\beta,\hat\rho)$, for a suitable estimation of the vector (β,ρ).

Theorem 3. *Under the* TOC*, in* (2.6)*, further assuming that* $A(\cdot)$ *and* $B(\cdot)$ *can be chosen as in* (2.7)*, for levels* k *such that* (2.8) *holds, and with* Z_k^{RB} *asymptotically standard normal* RVs*, we can write*

$$\mathrm{RB}(k;\beta,\rho) \;\stackrel{d}{=}\; \xi + \frac{\sigma_{\mathrm{RB}} Z_k^{\mathrm{RB}}}{\sqrt{k}} + b_{\mathrm{RB}}^{(1)} A(n/k) + \Big(b_{\mathrm{RB}}^{(2)} A^2(n/k) + O_p\big(A(n/k)/\sqrt{k}\big)\Big)\,(1+o_p(1)), \tag{3.4}$$

with $b_{\mathrm{RB}}^{(1)}$ *defined in* (3.1)*, and* $(\sigma_{\overline{\overline{\mathrm{RB}}}_p}, \sigma_{\overline{\overline{\mathrm{RB}}}}) = (\sigma_{\overline{\mathrm{RB}}_p}, \sigma_{\overline{\mathrm{RB}}}) = (\sigma_{\mathrm{H}_p}, \sigma_{\overline{\mathrm{H}}})$*, defined in* (3.2)*,*

$$b_{\overline{\mathrm{RB}}_p}^{(2)} = \frac{1-p\xi}{\xi(1-p\xi-\rho)}\left(\frac{\zeta(1-p\xi-\rho)^2 + p\xi\rho}{(1-p\xi-\rho)(1-p\xi-2\rho)} - \frac{1-\varphi_\rho}{1-\varphi_\rho-\rho}\right),$$

$$b_{\overline{\overline{\mathrm{RB}}}_p}^{(2)} = \frac{1-p\xi}{\xi(1-p\xi-\rho)^3}\left(\frac{\zeta(1-p\xi-\rho)^3 + p\xi\rho^2 - (1-p\xi)(1-p\xi-\rho)(1-p\xi-2\rho)}{1-p\xi-2\rho}\right) \tag{3.5}$$

$$\begin{aligned} b_{\overline{\mathrm{RB}}}^{(2)} = b_{\overline{\mathrm{RB}}_{\varphi_\rho/\xi}}^{(2)} &= \frac{1-\varphi_\rho}{\xi(1-\varphi_\rho-\rho)}\left(\frac{\zeta(1-\varphi_\rho-\rho)^2 + \varphi_\rho\rho}{(1-\varphi_\rho-\rho)(1-\varphi_\rho-2\rho)} - \frac{1-\varphi_\rho}{1-\varphi_\rho-\rho}\right) \\ &= \frac{1-\varphi_\rho}{\xi(1-\varphi_\rho-\rho)^2}\left(\frac{\zeta(1-\varphi_\rho-\rho)^2 + \varphi_\rho\rho - (1-\varphi_\rho)(1-\varphi_\rho-2\rho)}{1-\varphi_\rho-2\rho}\right), \end{aligned}$$

$$b_{\overline{\overline{\mathrm{RB}}}}^{(2)} = b_{\overline{\overline{\mathrm{RB}}}_{\varphi_\rho/\xi}}^{(2)} = \frac{1-\varphi_\rho}{\xi(1-\rho-\varphi_\rho)^3}\left(\frac{\zeta(1-\rho-\varphi_\rho)^3 + \rho^2\varphi_\rho - (1-\varphi_\rho)(1-\rho-\varphi_\rho)(1-2\rho-\varphi_\rho)}{1-\varphi_\rho-2\rho}\right),$$

with ζ *given in* (2.7)*. Consequently, if* $\sqrt{k}\,A(n/k) \to \lambda$*, finite,*

$$\sqrt{k}\,(\mathrm{RB}(k;\beta,\rho) - \xi) \;\underset{n\to\infty}{\stackrel{d}{\longrightarrow}}\; N\Big(\lambda b_{\mathrm{RB}}^{(1)}, \sigma_{\mathrm{RB}}^2\Big), \tag{3.6}$$

with a null mean value unless we are dealing with $\overline{\mathrm{RB}}_p$*. Moreover, for* $\mathrm{RB} \neq \overline{\mathrm{RB}}_p$*, if* $\sqrt{k}\;A^2(n/k) \to \lambda_A$*, finite,*

$$\sqrt{k}\,(\mathrm{RB}(k;\beta,\rho) - \xi) \;\underset{n\to\infty}{\stackrel{d}{\longrightarrow}}\; N\Big(\lambda_A b_{\mathrm{RB}}^{(2)}, \sigma_{\mathrm{RB}}^2\Big). \tag{3.7}$$

Remark 7. *Note that if we consider $p = 0$ in* (3.5)*, we get*

$$b^{(2)}_{\overline{\overline{\mathrm{RB}}}_0} \equiv b^{(2)}_{\mathrm{CH}} = \frac{\zeta}{\xi(1-2\rho)} - \frac{1}{\xi(1-\rho)^2},$$

the value already derived in Caeiro *et al.* (2009).

Theorem 4. *Under the same conditions of* Theorem 3*, let us again generally denote all the aforementioned* RB*-classes by* $\mathrm{RB}(k;\hat\beta,\hat\rho)$*, for any of the estimators $\hat\rho$ and $\hat\beta$ in* (2.11) *and in* (2.13)*, respectively, both computed at a level k_1 such that* (2.12) *holds. Then, $\sqrt{k}\big\{\mathrm{RB}(k;\hat\beta,\hat\rho)-\xi\big\}$ are asymptotically normal with a null mean value and a variance $(\sigma_{\overline{\overline{\mathrm{RB}}}_p},\sigma_{\overline{\overline{\mathrm{RB}}}}) = (\sigma_{\overline{\mathrm{RB}}_p},\sigma_{\overline{\mathrm{RB}}}) = (\sigma_{\mathrm{H}_p},\sigma_{\overline{\mathrm{H}}})$, defined in* (3.2)*, not only when $\sqrt{k}\ A(n/k)\to 0$, but also when $\sqrt{k}\ A(n/k)\to\lambda$, finite, i.e.* (3.4) *and* (3.6) *hold with* $\mathrm{RB}(k;\beta,\rho)$ *replaced by* $\mathrm{RB}(k;\hat\beta,\hat\rho)$*.*

We can further get the same limiting result in (3.7) *for levels k such that $\sqrt{k}\ A(n/k)\to\infty$, provided that $k = o(k_1)$, as $n\to\infty$, and we choose k_1 optimal for the estimation of ρ, i.e., such that $\sqrt{k_1}A^2(n/k_1)\to\lambda_{_{A_1}}$, finite. If for k such that $\sqrt{k}A^2(n/k)\to\lambda_{_A}$, finite, we assume that $\hat\rho-\rho = o_p\big(A(n/k)/\ln(n/k)\big)$, the result in* (3.7) *also holds for* $\mathrm{RB}(k;\hat\beta,\hat\rho)$*.*

4 Asymptotic comparison at optimal levels

We next proceed to the comparison of some of the aforementioned EVI-estimators at their optimal levels. This is done in a way similar to the one used, for classical EVI-estimation, in de Haan and Peng (1998), among others, and in Gomes *et al.* (2007) and Caeiro and Gomes (2011), for specific sets of RB EVI-estimators. Let us assume that $\hat\xi^\bullet_n(k)$ denotes any arbitrary semi-parametric EVI-estimator, for which we have the asymptotic distributional representation

$$\hat\xi^\bullet_n(k) = \xi + \frac{\sigma_\bullet Z^\bullet_k}{\sqrt{k}} + b^{(1)}_\bullet A(n/k) + b^{(2)}_\bullet A^2(n/k) + o_p(A^2(n/k)), \tag{4.1}$$

for any intermediate sequence of integers $k = k_n$, and where $Z^\bullet_k$ is asymptotically standard normal. Then, $\sqrt{k}\big(\hat\xi^\bullet_n(k)-\xi\big) \overset{d}{\to} \mathcal{N}(\lambda^{(1)}_{_A} b^{(1)}_\bullet, \sigma^2_\bullet)$ provided that k is such that $\sqrt{k}\ A(n/k)\to\lambda^{(1)}_{_A}$, finite, as $n\to\infty$. If $b^{(1)}_\bullet = 0$, and we consider levels k such that

$\sqrt{k}\ A^2(n/k) \to \lambda_{_A}^{(2)}$, finite, as $n \to \infty$, $\sqrt{k}\big(\hat{\xi}_n^{\bullet}(k) - \xi\big) \xrightarrow{d} \mathcal{N}(\lambda_{_A}^{(2)} b_{\bullet}^{(2)}, \sigma_{\bullet}^2)$. We then write

$$\mathrm{Bias}_\infty\big(\hat{\xi}_n^{\bullet}(k)\big) := \begin{cases} b_{\bullet}^{(1)} A(n/k), & \text{if} \quad b_{\bullet}^{(1)} \neq 0, \\ b_{\bullet}^{(2)} A^2(n/k), & \text{if} \quad b_{\bullet}^{(1)} = 0,\ b_{\bullet}^{(2)} \neq 0, \end{cases}$$

and $\mathrm{Var}_\infty\big(\hat{\xi}_n^{\bullet}(k)\big) := \sigma_{\bullet}^2/k$. The so-called *asymptotic mean square error* (AMSE) is then given by

$$\mathrm{AMSE}\left(\hat{\xi}_n^{\bullet}(k)\right) := \begin{cases} \sigma_{\bullet}^2/k + \left(b_{\bullet}^{(1)}\right)^2\ A^2(n/k), & \text{if} \quad b_{\bullet}^{(1)} \neq 0, \\ \sigma_{\bullet}^2/k + \left(b_{\bullet}^{(2)}\right)^2\ A^4(n/k), & \text{if} \quad b_{\bullet}^{(1)} = 0,\ b_{\bullet}^{(2)} \neq 0. \end{cases}$$

Regular variation theory (Bingham *et al.*, 1987), enables us to show that there exists a function $\eta(n) = \eta(n, \xi, \rho)$, such that

$$\lim_{n\to\infty} \eta(n)\ \mathrm{AMSE}\left(\hat{\xi}_{n0}^{\bullet}\right) = \left\{ \begin{array}{ll} (\sigma_{\bullet}^2)^{-\frac{2\rho}{1-2\rho}} \left(b_{\bullet}^{(1)}\right)^{\frac{2}{1-2\rho}}, & \text{if} \quad b_{\bullet}^{(1)} \neq 0, \\ (\sigma_{\bullet}^2)^{-\frac{4\rho}{1-4\rho}} \left(b_{\bullet}^{(2)}\right)^{\frac{2}{1-4\rho}}, & \text{if} \quad b_{\bullet}^{(1)} = 0,\ b_{\bullet}^{(2)} \neq 0 \end{array} \right\} =: \mathrm{LMSE}\left(\hat{\xi}_{n0}^{\bullet}\right),$$

where $\hat{\xi}_{n0}^{\bullet} := \hat{\xi}_n^{\bullet}(k_{0|\bullet}(n))$ and $k_{0|\bullet}(n) := \arg\min_k \mathrm{MSE}\Big(\hat{\xi}_n^{\bullet}(k)\Big)$.

We again consider the following:

Definition 1. *Given two biased estimators $\hat{\xi}_n^{(1)}(k)$ and $\hat{\xi}_n^{(2)}(k)$, for each of which a distributional representation of the type in* (4.1) *holds, with constants $(\sigma_1, b_1^{(1)}, b_1^{(2)})$ and $(\sigma_2, b_2^{(1)}, b_2^{(2)})$, respectively, both computed at their optimal levels, the asymptotic root efficiency* (AREFF) *of $\hat{\xi}_{n0}^{(1)}$ relatively to $\hat{\xi}_{n0}^{(2)}$ is*

$$\mathrm{AREFF}_{1|2} \equiv \mathrm{AREFF}_{\hat{\xi}_{n0}^{(1)}|\hat{\xi}_{n0}^{(2)}} := \sqrt{\frac{\mathrm{LMSE}\big(\hat{\xi}_{n0}^{(2)}\big)}{\mathrm{LMSE}\big(\hat{\xi}_{n0}^{(1)}\big)}}$$
$$= \begin{cases} \left(\left(\frac{\sigma_2}{\sigma_1}\right)^{-2\rho} \left|\frac{b_2^{(1)}}{b_1^{(1)}}\right|\right)^{\frac{1}{1-2\rho}}, & \textit{if} \quad b_1^{(1)} \neq 0 \wedge b_2^{(1)} \neq 0, \\ \\ \left(\left(\frac{\sigma_2}{\sigma_1}\right)^{-4\rho} \left|\frac{b_2^{(2)}}{b_1^{(2)}}\right|\right)^{\frac{1}{1-4\rho}}, & \textit{if} \quad b_1^{(1)} = b_2^{(1)} = 0 \wedge b_1^{(2)}, b_2^{(2)} \neq 0. \end{cases} \tag{4.2}$$

Remark 8. *Note that the* AREFF *indicator, in* (4.2)*, is defined in such a way that the higher the* AREFF *indicator is, the better is the first estimator.*

4.1 Asymptotic comparison at optimal levels of CH and optimal RB MOP classes

We first proceed to the comparison of CH, $\overline{\mathrm{RB}}$ and $\overline{\overline{\mathrm{RB}}}$, at their optimal levels. In Figures 1 and 2, we respectively present the patterns of the AREFF indicators of $\overline{\overline{\mathrm{RB}}}|\mathrm{CH}$ and of $\overline{\overline{\mathrm{RB}}}|\overline{\mathrm{RB}}$ in the whole (ζ, ρ)–plane, with ζ the parameter given in (2.7). We further present in Figure 3 the AREFF indicator of $\overline{\mathrm{RB}}|\mathrm{CH}$.

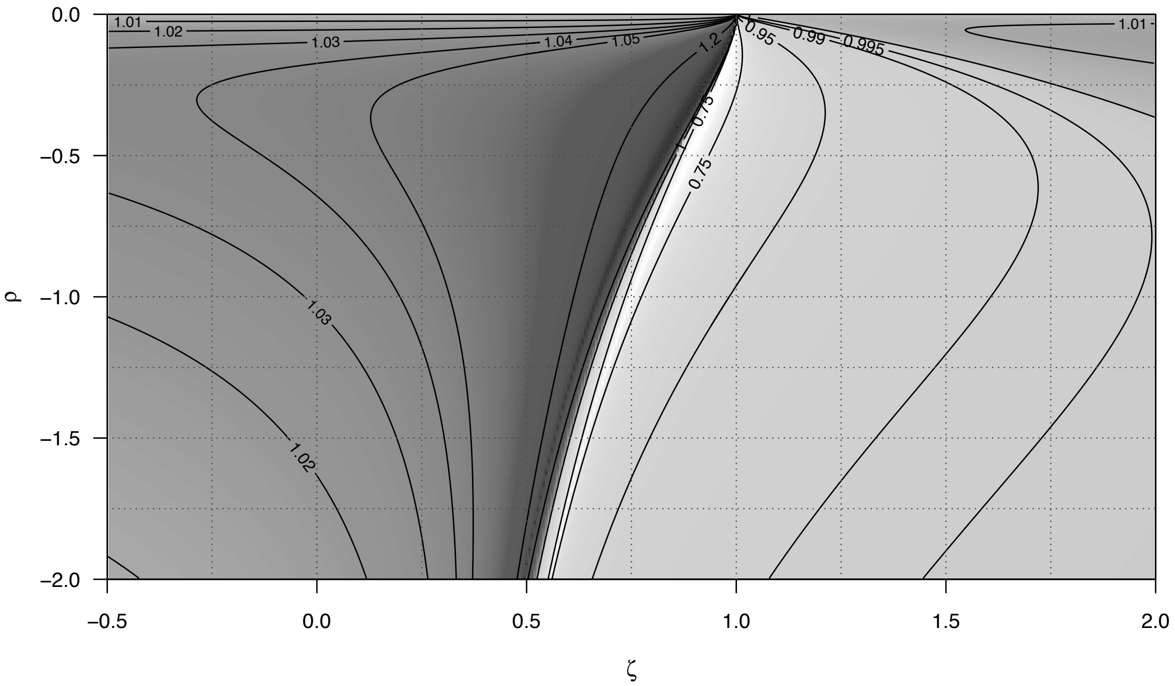

Figure 1: The indicator $\mathrm{ARREF}_{\overline{\overline{\mathrm{RB}}}|\mathrm{CH}}$, as a function of (ζ, ρ)

As was detected by simulation in Gomes *et al.* (2013), and despite of the fact that asymptotically at optimal levels $\overline{\mathrm{H}}$ beats H in the whole (ξ, ρ)-plane, neither $\overline{\mathrm{RB}}$ nor $\overline{\overline{\mathrm{RB}}}$ beats CH in the whole (ζ, ρ)-plane, but both $\overline{\mathrm{RB}}$ and $\overline{\overline{\mathrm{RB}}}$ beat CH in a wide, interesting region of the (ζ, ρ)-plane. Also, asymptotically $\overline{\overline{\mathrm{RB}}}$ beats $\overline{\mathrm{RB}}$ in a wide and relevant region of the (ζ, ρ)-plane.

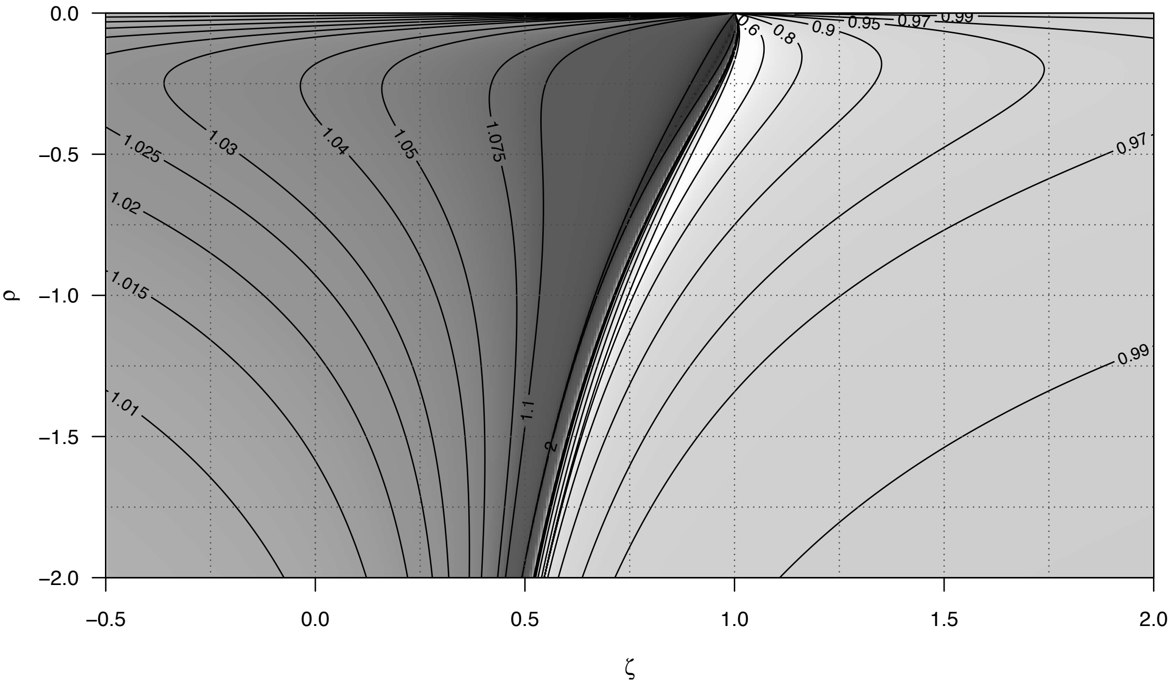

Figure 2: The indicator $\mathrm{ARREF}_{\overline{\overline{\mathrm{RB}}}|\overline{\mathrm{RB}}}$, as a function of (ζ, ρ)

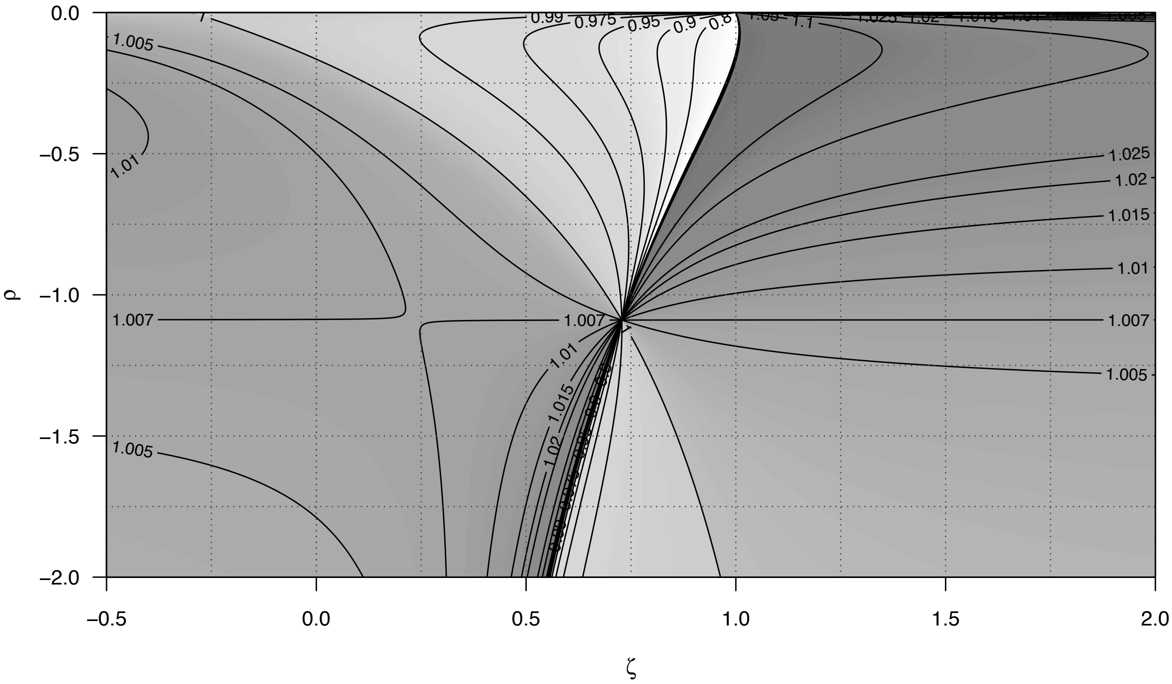

Figure 3: The indicator $\mathrm{ARREF}_{\overline{\mathrm{RB}}|\mathrm{CH}}$, as a function of (ζ, ρ)

4.2 Asymptotic comparison of optimal MOP EVI-estimators

Let us next consider the MOP EVI-estimators $\mathrm{H}_p(k)$ in (1.5). We have

$$\mathrm{LMSE}(H_{p0}) = \left(\frac{\xi^2(1-p\xi)^2}{1-2p\xi}\right)^{-\frac{2\rho}{1-2\rho}} \left(\frac{1-p\xi}{1-p\xi-\rho}\right)^{\frac{2}{1-2\rho}}.$$

To measure the performance of H_{p0}, we have computed the AREFF-indicator, in (4.2), now denoted as follows:

$$\mathrm{AREFF}_{p|0} = \left(\left(\frac{\sqrt{1-2p\xi}}{1-p\xi}\right)^{-2\rho} \left|\frac{1-p\xi-\rho}{(1-\rho)(1-p\xi)}\right|\right)^{\frac{1}{1-2\rho}}.$$

As was done in Brilhante *et al.* (2013) for $p \geq 0$, we can reparameterise $\mathrm{AREFF}_{p|0}$, so that we have a dependence on two parameters only, the second-order parameter ρ and the parameter $a = p\xi < 1/2$, possibly negative. In Figure 4, we picture the values of

$$\mathrm{AREFF}^*_{a|0} = \left(\left(\frac{\sqrt{1-2a}}{1-a}\right)^{-2\rho} \left|\frac{1-a-\rho}{(1-\rho)(1-a)}\right|\right)^{\frac{1}{1-2\rho}}. \tag{4.3}$$

We always loose efficiency when $p < 0$, and for $p \geq 0$, the gain in efficiency is not terribly high, as already detected in Brilhante *et al.* (2013). But, at optimal levels, there is a wide region of the (a, ρ)-plane where the new class of MOP EVI-estimators performs better than the Hill estimator.

Remark 9. *For an asymptotic comparison at optimal levels of the* MOP EVI-*estimators, with* $p \geq 0$, *and the classical moment* (Dekkers *et al.*, 1989), *generalized* Hill (Beirlant *et al.*, 1996, 2005) *and mixed-moment* (Fraga Alves *et al.*, 2009) EVI-*estimators, see* Brilhante *et al.* (2013).

Remark 10. *The lack of efficiency of the* MOP EVI-*estimators with* $p < 0$ *together with the results in* Stehlík *et al.* (2010) *and* Beran *et al.* (2014), *related to the robustness of the* MOP EVI-*estimators associated with* $p = -1$ $(a = -\xi < 0)$, *requires a further discussion of the topic, 'robustness versus efficiency', as initiated in* Del'Aquilla and Embrechts (2006), *or for the finding of indicators that simultaneously take both concepts into account. This is, however, a relevant topic beyond the scope of this article.*

Remark 11. *Note further that as detected in* Brilhante *et al.* (2013), *at optimal levels, the optimal class,* $\overline{\mathrm{H}}$, *in* (1.7), *can beat the optimal* Hill EVI-*estimator,* H_{00}, *in the whole* (ξ, ρ)*-plane.*

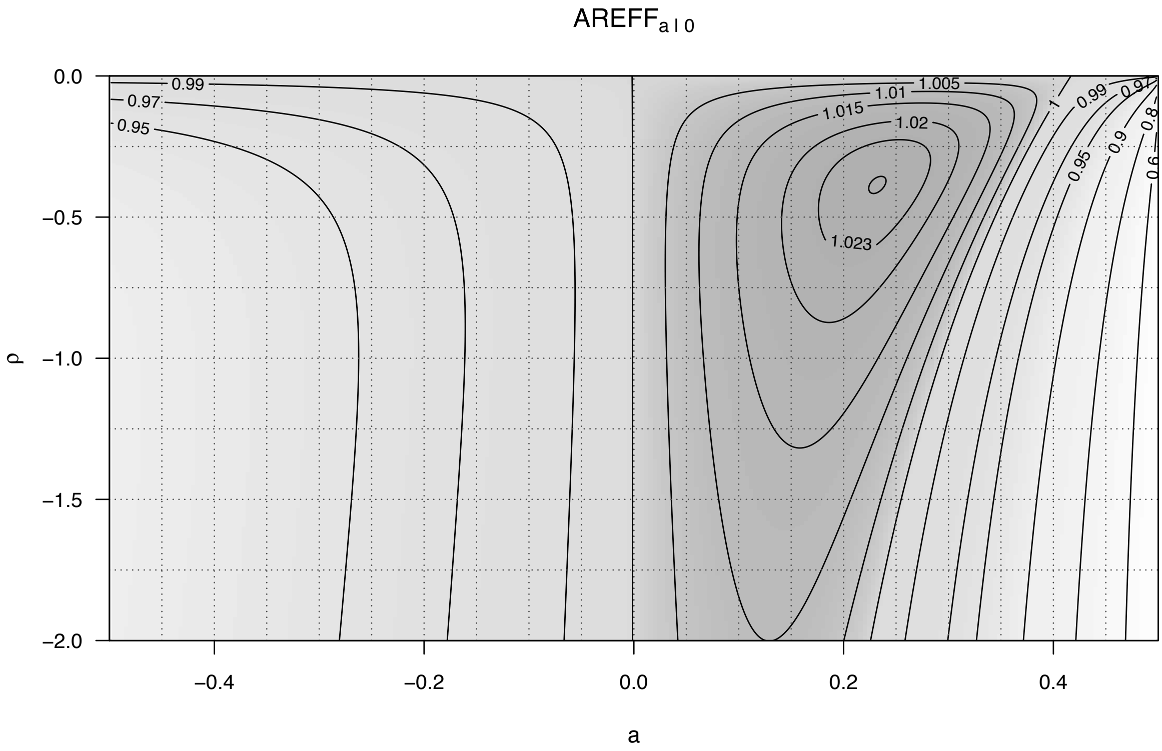

Figure 4: The indicator $\text{ARREF}^*_{a|0}$, in (4.3), as a function of (a, ρ)

5 Finite sample properties of the EVI-estimators

In Brilhante *et al.* (2013) they implemented multi-sample Monte-Carlo simulation experiments of size 5000×20 for the class of MOP EVI-estimators, given in(1.5), compared to the MVRB CH EVI-estimators, given in (1.13), for a large set of ξ-values and for several sample sizes n from the following underlying models:

(1) Fréchet$_\xi$ model, $F(x) = \exp(-x^{-1/\xi}),\ x \geq 0\ (\rho = -1)$;

(2) Extreme value model, $F(x) = \text{EV}_\xi(x)$, given in (1.1) $(\rho = -\xi)$;

(3) Generalised Pareto model, $F(x) = 1 + \ln \text{EV}_\xi(x) = 1 - \big(1 + \xi x\big)^{-1/\xi}$, $x \geq 0$ $(\rho = -\xi)$;

(4) the *Student*-t_ν, with $\nu = 1, 2, 4$, i.e. for values of $\xi = 1, 0.5, 0.25$ $(\xi = 1/\nu,\ \rho = -2/\nu)$.

For the same models as before, further including, for a large set of (ξ, ρ)-values, the

(5) Burr$_{\xi,\rho}$ model, $F(x) = 1 - (1 + x^{-\rho/\xi})^{1/\rho}$, $x \geq 0$,

a similar comparative analysis of $\overline{\text{H}}$ and $\overline{\text{RB}}$, given in (1.7) and (1.9), respectively, was carried out in Gomes *et al.* (2013), where the notation H^* and CH^* is used. Similar, but smaller scale simulation experiments have been performed in Gomes *et al.* (2015) for the class $\overline{\mathrm{RB}}_p(k)$ given in (1.10). For all the aforementioned models we have now further included in the simulation experiments $\overline{\overline{\mathrm{RB}}}_p$ and $\overline{\overline{\mathrm{RB}}}$, given in (1.11) and (1.12), respectively.

We shall now also estimate the optimal k-value for the H EVI-estimation, as given in Hall (1982), computing $\hat{k}_{0|\mathrm{H}_0} = \big((1-\hat\rho)n^{-\hat\rho}/\big(\hat\beta\sqrt{-2\hat\rho}\big)\big)^{2/(1-2\hat\rho)}$, $\hat{\mathrm{H}}_{00} := \mathrm{H}(\hat{k}_{0|\mathrm{H}_0})$ and, with φ_ρ given in (1.6), we further consider the EVI-estimators

$$\mathrm{H}^*(k) := \mathrm{H}_{\hat p_{\mathrm{M}}}(k),\quad \overline{\mathrm{RB}}^*(k) := \overline{\mathrm{RB}}_{\hat p_{\mathrm{M}}}(k),\quad \overline{\overline{\mathrm{RB}}}^*(k) := \overline{\overline{\mathrm{RB}}}_{\hat p_{\mathrm{M}}}(k),\quad \hat p_{\mathrm{M}} = \varphi_{\hat\rho}/\hat{\mathrm{H}}_{00}. \tag{5.1}$$

For details on multi-sample simulation, we refer Gomes and Oliveira (2001).

5.1 Mean values and mean square error patterns

For each value of n and for each of the aforementioned models, we first simulated the mean values (E) and root MSEs (RMSEs) of the aforementioned EVI-estimators, as functions of the k top OSs involved in the estimation, and on the basis of the first run of size 5000. As an illustration, we present Figure 5, associated with an $\mathrm{EV}_{0.25}$ parent. Regarding the EVI-estimators $\overline{\mathrm{RB}}^*$ and $\overline{\overline{\mathrm{RB}}}^*$, given in (5.1), and due to its proximity, we picture only the most efficient one. We further represent H, H^*, CH, and among the values of $p = \ell/(10\xi)$, $\ell = 1,2,3,4$, the more efficient $(\overline{\mathrm{RB}}_p, \overline{\overline{\mathrm{RB}}}_p)$ EVI-estimator, with $\mathrm{H}(k)$, $\overline{\mathrm{RB}}_p(k)$, $\overline{\overline{\mathrm{RB}}}_p(k)$, $\mathrm{CH}(k)$ and $\mathrm{H}^*(k)$, respectively defined in (1.3), (1.10), (1.11), (1.13) and (5.1).

A few comments:

- There is always a reduction in RMSE, as well as in bias. We get estimates closer to the target value ξ, when we move from H to H^*, next to CH and finally to either $\overline{\mathrm{RB}}^*$ or to $\overline{\overline{\mathrm{RB}}}^*$.
- Such a reduction is particularly high for values of ρ approaching zero. This happens even when we work with models out of the scope of the stated theorems, such as the log-gamma ($\rho = 0$).
- However, there is always a value of p that enables both $\overline{\mathrm{RB}}_p(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$ to outperform the corresponding estimated RB EVI-estimator, given in (5.1).

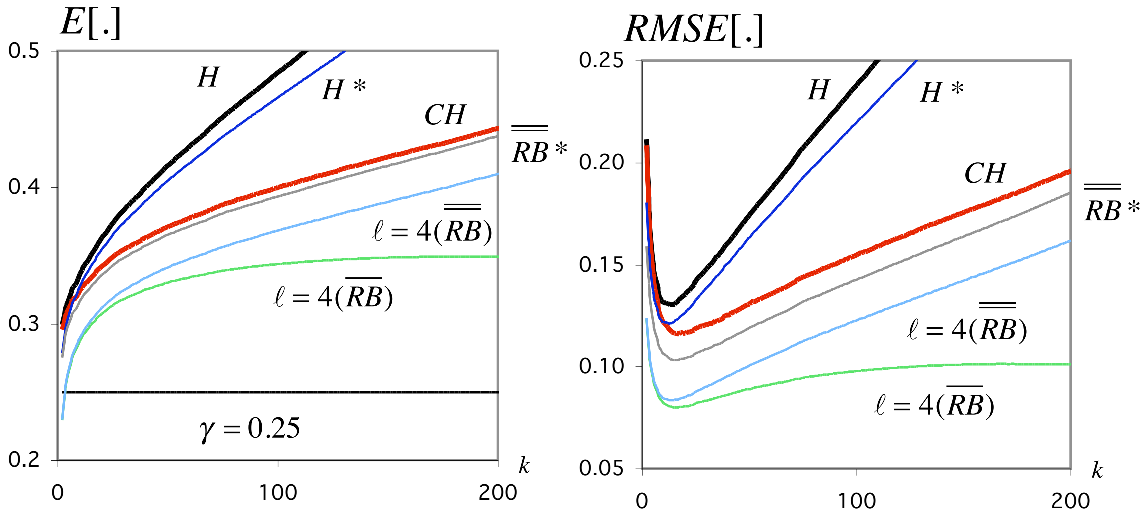

Figure 5: Mean values (*left*) and RMSEs (*right*) for an $\mathrm{EV}_{0.25}$ underlying parent and $n = 1000$

5.1.1 Mean values of the EVI-estimators at optimal levels

As an illustration of the bias reduction achieved with the RB MOP EVI-estimators in (1.10) and (1.11) at simulated optimal levels (levels where RMSE are minimal as functions of k), see Tables 1 and 2, respectively related to $\mathrm{EV}_{0.25}$ and Fréchet models. We there present, for $n = 100,\ 200,\ 500,\ 1000,\ 2000$ and 5000, the simulated mean values at optimal levels of H, H*, CH, ($\overline{\mathrm{RB}}^*$, $\overline{\overline{\mathrm{RB}}}^*$), and ($\overline{\mathrm{RB}}_p, \overline{\overline{\mathrm{RB}}}_p$), $p = \ell/(10\xi)$, $\ell = 1, 2, 3, 4$, with $\mathrm{H}(k)$, $\overline{\mathrm{RB}}_p(k)$, $\overline{\overline{\mathrm{RB}}}_p(k)$, $\mathrm{CH}(k)$ and ($\mathrm{H}^*(\mathrm{k})$, $\overline{\mathrm{RB}}^*(k)$, $\overline{\overline{\mathrm{RB}}}^*(k)$), respectively defined in (1.3), (1.10), (1.11), (1.13) and (5.1). Information on 95% confidence intervals, computed on the basis of the 20 replicates with 5000 runs each, is also provided. The indicator is written in **bold** when its performance regarding squared bias exceeds that of the estimator in the previous class. If the estimate in a class does not improve on the one in the immediately previous class, we write it in *italics*. Generally denoting by T any of the aforementioned EVI-estimators, note that for Fréchet underlying parents, T/ξ does not depend on ξ. For a proof, see Gomes *et al.* (2013).

5.1.2 Mean square errors and relative efficiency indicators at optimal levels

We have computed the Hill estimator at the simulated value of $k_{0|\mathrm{H}} := \arg\min_k \mathrm{RMSE}\big(\mathrm{H}(k)\big)$, the simulated optimal k in the sense of minimum RMSE,

Table 1: Simulated mean values, at simulated optimal levels, of $\mathrm{H}(k) \equiv \mathrm{H}_0(k)$, $\mathrm{H}^*(k)$, $\mathrm{CH}(k)$, $\overline{\mathrm{RB}}^*(k)$, $\overline{\overline{\mathrm{RB}}}^*(k)$, $\overline{\mathrm{RB}}_p(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$, $p = \ell/(10\xi)$, $\ell = 1, 2, 4$, for EV_ξ underlying parents, $\xi = 0.25$, together with 95% confidence intervals

EV_ξ parent, $\xi = 0.25$						
n	100	200	500	1000	2000	5000
H	0.427 ± 0.0012	0.392 ± 0.0026	0.365 ± 0.0019	0.348 ± 0.0012	0.335 ± 0.0013	0.321 ± 0.0010
H^*	$\mathbf{0.382} \pm 0.0027$	$\mathbf{0.372} \pm 0.0021$	$\mathbf{0.353} \pm 0.0014$	$\mathbf{0.342} \pm 0.0017$	$\mathbf{0.330} \pm 0.0008$	$\mathbf{0.317} \pm 0.0008$
CH	$\mathit{0.407} \pm 0.0012$	$\mathbf{0.371} \pm 0.0029$	$\mathbf{0.351} \pm 0.0017$	$\mathbf{0.339} \pm 0.0012$	$\mathbf{0.327} \pm 0.0010$	$\mathbf{0.316} \pm 0.0009$
$\overline{\mathrm{RB}}^*$	$\mathbf{0.367} \pm 0.0022$	$\mathbf{0.359} \pm 0.0018$	$\mathbf{0.344} \pm 0.0014$	0.334 ± 0.0016	0.324 ± 0.0008	0.313 ± 0.0008
$\overline{\overline{\mathrm{RB}}}^*$	0.368 ± 0.0026	0.359 ± 0.0019	0.345 ± 0.0014	$\mathbf{0.334} \pm 0.0015$	$\mathbf{0.324} \pm 0.0007$	$\mathbf{0.313} \pm 0.0007$
$\overline{\mathrm{RB}}_p(\ell = 1)$	$\mathit{0.368} \pm 0.0025$	$\mathit{0.360} \pm 0.0018$	$\mathit{0.346} \pm 0.0014$	$\mathit{0.335} \pm 0.0014$	$\mathit{0.326} \pm 0.0007$	$\mathit{0.314} \pm 0.0008$
$\overline{\overline{\mathrm{RB}}}_p(\ell = 1)$	0.365 ± 0.0021	0.359 ± 0.0018	$\mathit{0.345} \pm 0.0014$	$\mathit{0.335} \pm 0.0015$	$\mathit{0.325} \pm 0.0007$	$\mathit{0.314} \pm 0.0008$
$\overline{\mathrm{RB}}_p(\ell = 2)$	0.353 ± 0.0024	0.345 ± 0.0017	0.335 ± 0.0013	0.327 ± 0.0014	0.319 ± 0.0007	0.310 ± 0.0009
$\overline{\overline{\mathrm{RB}}}_p(\ell = 2)$	0.346 ± 0.0029	0.343 ± 0.0017	0.336 ± 0.0014	0.327 ± 0.0012	0.320 ± 0.0008	0.310 ± 0.0009
$\overline{\mathrm{RB}}_p(\ell = 4)$	$\mathbf{0.283} \pm 0.0011$	$\mathbf{0.291} \pm 0.0009$	$\mathbf{0.291} \pm 0.0004$	$\mathbf{0.291} \pm 0.0007$	$\mathbf{0.290} \pm 0.0005$	$\mathbf{0.290} \pm 0.0002$
$\overline{\overline{\mathrm{RB}}}_p(\ell = 4)$	0.300 ± 0.0020	0.305 ± 0.0013	0.305 ± 0.0013	0.302 ± 0.0010	0.300 ± 0.0009	0.295 ± 0.0010

Table 2: Simulated mean values, at simulated optimal levels, of $\mathrm{H}(k) \equiv \mathrm{H}_0(k)/\xi$, $\mathrm{H}^*(k)/\xi$, $\mathrm{CH}(k)/\xi$, $\overline{\mathrm{RB}}^*(k)/\xi$, $\overline{\overline{\mathrm{RB}}}^*(k)/\xi$, $\overline{\mathrm{RB}}_p(k)/\xi$ and $\overline{\overline{\mathrm{RB}}}_p(k)/\xi$, $p = \ell/(10\xi)$, $\ell = 1, 2, 4$, for Fréchet underlying parents, together with 95% confidence intervals

Fréchet parents						
n	100	200	500	1000	2000	5000
H	1.109 ± 0.0027	1.085 ± 0.0028	1.063 ± 0.0013	1.049 ± 0.0014	1.039 ± 0.0009	1.029 ± 0.0006
H^*	$\mathbf{1.097} \pm 0.0021$	$\mathbf{1.076} \pm 0.0024$	$\mathbf{1.057} \pm 0.0010$	$\mathbf{1.047} \pm 0.0008$	$\mathbf{1.037} \pm 0.0011$	$\mathbf{1.029} \pm 0.0006$
CH	$\mathbf{0.982} \pm 0.0030$	$\mathbf{0.986} \pm 0.0395$	$\mathbf{0.995} \pm 0.0016$	$\mathbf{0.999} \pm 0.0008$	$\mathbf{1.000} \pm 0.0005$	$\mathbf{1.000} \pm 0.0004$
$\overline{\mathrm{RB}}^*$	$\mathbf{0.997} \pm 0.0028$	$\mathbf{1.001} \pm 0.0134$	$\mathbf{1.002} \pm 0.0009$	$\mathit{1.002} \pm 0.0007$	$\mathit{1.002} \pm 0.0004$	$\mathit{1.001} \pm 0.0004$
$\overline{\overline{\mathrm{RB}}}^*$	$\mathbf{0.997} \pm 0.0028$	0.997 ± 0.0062	1.003 ± 0.0013	$\mathit{1.003} \pm 0.0006$	$\mathit{1.003} \pm 0.0005$	$\mathit{1.001} \pm 0.0003$
$\overline{\mathrm{RB}}_p(\ell = 1)$	$\mathit{0.969} \pm 0.0018$	$\mathit{0.975} \pm 0.0008$	$\mathit{0.979} \pm 0.0007$	$\mathit{0.980} \pm 0.0005$	$\mathit{0.980} \pm 0.0005$	$\mathit{0.980} \pm 0.0002$
$\overline{\overline{\mathrm{RB}}}_p(\ell = 1)$	$\mathbf{1.002} \pm 0.0029$	$\mathit{1.006} \pm 0.0377$	$\mathit{1.010} \pm 0.0014$	$\mathit{1.010} \pm 0.0009$	$\mathit{1.010} \pm 0.0006$	$\mathit{1.009} \pm 0.0003$
$\overline{\mathrm{RB}}_p(\ell = 2)$	$\mathit{0.995} \pm 0.0030$	$\mathit{1.004} \pm 0.0303$	1.001 ± 0.0010	$\mathit{1.001} \pm 0.0008$	$\mathit{1.001} \pm 0.0006$	$\mathit{1.001} \pm 0.0004$
$\overline{\overline{\mathrm{RB}}}_p(\ell = 2)$	1.002 ± 0.0028	$\mathit{1.002} \pm 0.0186$	$\mathit{1.004} \pm 0.0008$	$\mathit{1.004} \pm 0.0007$	$\mathit{1.002} \pm 0.0006$	$\mathit{1.001} \pm 0.0002$
$\overline{\mathrm{RB}}_p(\ell = 4)$	$\mathit{0.989} \pm 0.0013$	$\mathbf{1.000} \pm 0.0164$	$\mathbf{1.000} \pm 0.0011$	$\mathbf{1.000} \pm 0.0005$	$\mathbf{1.000} \pm 0.0003$	$\mathbf{1.000} \pm 0.0003$
$\overline{\overline{\mathrm{RB}}}_p(\ell = 4)$	$\mathit{1.020} \pm 0.0018$	$\mathit{1.016} \pm 0.0295$	$\mathit{1.014} \pm 0.0009$	$\mathit{1.010} \pm 0.0006$	$\mathit{1.007} \pm 0.0004$	$\mathit{1.005} \pm 0.0003$

not relevant in practice, but providing an indication of the best possible performance of Hill's estimator. Such an estimator is denoted by H^s_{00}. With RB denoting $\overline{\mathrm{RB}}^*$ or $\overline{\overline{\mathrm{RB}}}^*$ or $\overline{\mathrm{RB}}_p$ or $\overline{\overline{\mathrm{RB}}}_p$ we have also computed RB^s_{00}, i.e. the EVI-estimator $\mathrm{RB}(k)$ computed at the simulated value of $k_{0|\mathrm{RB}} := \arg\min_k \mathrm{MSE}\big(\mathrm{RB}(k)\big)$. The simulated indicators are

$$\mathrm{REFF}_{\mathrm{RB}|\mathrm{H}} := \frac{\mathrm{RMSE}\,(\mathrm{H}^s_{00})}{\mathrm{RMSE}\,(\mathrm{RB}^s_{00})} = \sqrt{\frac{\mathrm{MSE}\,(\mathrm{H}^s_{00})}{\mathrm{MSE}\,(\mathrm{RB}^s_{00})}}. \tag{5.2}$$

Similar REFF-indicators, $\mathrm{REFF}_{\overline{\overline{\mathrm{H}}}^*|\mathrm{H}}$ and $\mathrm{REFF}_{\mathrm{CH}|\mathrm{H}}$, have also been computed for the $\overline{\overline{\mathrm{H}}}$ and CH EVI-estimators.

Remark 12. *An indicator higher than one means a better performance than the* Hill *estimator. Consequently, the higher these indicators are, the better the associated* EVI*-estimators perform, compared to* H^s_{00}.

Again as an illustration of the obtained results, we present Tables 3 and 4. In the first row, we provide the RMSE of H^s_{00}, denoted by RMSE_{00}, so that we can easily recover the RMSE of all other estimators. The following rows provide the REFF-indicators of the different EVI-estimators under study. As before, similar marks (**bold** and *italic*) are used for the REFF-indicators.

Table 3: Simulated RMSE of H^s_{00}/ξ (first row) and REFF-indicators of $\mathrm{H}^*(k)$, $\mathrm{CH}(k)$, $\overline{\mathrm{RB}}^*(k)$, $\overline{\overline{\mathrm{RB}}}^*(k)$, $\overline{\mathrm{RB}}_p(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$, $p = \ell/(10\xi)$, $\ell = 1, 2, 4$, for EV_ξ underlying parents, $\xi = 0.25$, together with 95% confidence intervals

EV_ξ parent, $\xi = 0.25$						
n	100	200	500	1000	2000	5000
RMSE_{00}	0.246 ± 0.1698	0.200 ± 0.1669	0.157 ± 0.1591	0.133 ± 0.1527	0.113 ± 0.1462	0.092 ± 0.1379
H^*	**1.103** ± 0.0011	**1.097** ± 0.0019	**1.082** ± 0.0012	**1.073** ± 0.0015	**1.065** ± 0.0010	**1.058** ± 0.0010
CH	**1.328** ± 0.0108	**1.237** ± 0.0056	**1.171** ± 0.0042	**1.130** ± 0.0021	**1.101** ± 0.0021	**1.072** ± 0.0020
$\overline{\mathrm{RB}}^*$	1.386 ± 0.0084	1.307 ± 0.0050	1.240 ± 0.0034	1.195 ± 0.0028	1.161 ± 0.0020	**1.127** ± 0.0013
$\overline{\overline{\mathrm{RB}}}^*$	**1.388** ± 0.0085	**1.309** ± 0.0051	**1.241** ± 0.0035	**1.196** ± 0.0028	**1.162** ± 0.0021	**1.127** ± 0.0013
$\overline{\mathrm{RB}}_p(\ell = 1)$	1.420 ± 0.0083	1.319 ± 0.0051	*1.236* ± 0.0034	*1.185* ± 0.0027	*1.148* ± 0.0190	*1.112* ± 0.0013
$\overline{\overline{\mathrm{RB}}}_p(\ell = 1)$	1.411 ± 0.0080	1.319 ± 0.0051	*1.238* ± 0.0035	*1.188* ± 0.0028	*1.151* ± 0.0020	*1.114* ± 0.0014
$\overline{\mathrm{RB}}_p(\ell = 2)$	1.624 ± 0.0091	1.478 ± 0.0057	1.351 ± 0.0034	1.277 ± 0.0033	1.222 ± 0.0027	1.171 ± 0.0011
$\overline{\overline{\mathrm{RB}}}_p(\ell = 2)$	1.543 ± 0.0054	1.435 ± 0.0052	1.327 ± 0.0031	1.263 ± 0.0032	1.213 ± 0.0026	1.166 ± 0.0011
$\overline{\mathrm{RB}}_p(\ell = 4)$	**3.707** ± 0.1013	**3.981** ± 0.0705	**3.637** ± 0.0362	**3.183** ± 0.0568	**2.782** ± 0.0349	**2.293** ± 0.0148
$\overline{\overline{\mathrm{RB}}}_p(\ell = 4)$	2.080 ± 0.0074	1.896 ± 0.0054	1.682 ± 0.0043	1.556 ± 0.0058	1.453 ± 0.0053	1.355 ± 0.0036

For all simulated models, excluding only the Fréchet, and among the values considered $p = \ell/(10\xi)$, $\ell = 1, 2, 3, 4$, the value of p leading to the best performance of the RB MOP EVI-estimators was the one associated with $\ell = 4$.

6 Overall conclusions

- Regarding the MOP EVI-estimation, we always lose efficiency when $p < 0$, and for $p \geq 0$, the gain in efficiency is not terribly high, as already detected in Brilhante

Table 4: Simulated RMSE of H^{s}_{00}/ξ (first row) and REFF-indicators of $\mathrm{H}^*(k)$, $\mathrm{CH}(k)$, $\overline{\mathrm{RB}}^*(k)$, $\overline{\overline{\mathrm{RB}}}^*(k)$, $\overline{\mathrm{RB}}_p(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$, $p=\ell/(10\xi)$, $\ell=1,2,4$, for EV_ξ (independent on ξ), for Fréchet parents, together with 95% confidence intervals

Fréchet parents						
n	100	200	500	1000	2000	5000
RMSE_{00}	0.212 ± 0.1547	0.163 ± 0.1520	0.117 ± 0.1432	0.091 ± 0.1345	0.071 ± 0.1977	0.052 ± 0.1764
H^*	**1.059** ± 0.0015	**1.049** ± 0.0016	**1.040** ± 0.0013	**1.034** ± 0.0013	**1.028** ± 0.0016	**1.026** ± 0.0014
CH	**1.257** ± 0.0072	**1.237** ± 0.1591	**1.337** ± 0.0080	**1.460** ± 0.0123	**1.574** ± 0.0123	**1.795** ± 0.0097
$\overline{\mathrm{RB}}^*$	*1.230* ± 0.0068	*1.216* ± 0.1519	*1.311* ± 0.0076	1.425 ± 0.0072	1.534 ± 0.0117	1.741 ± 0.0092
$\overline{\overline{\mathrm{RB}}}^*$	**1.296** ± 0.0073	**1.277** ± 0.1609	**1.380** ± 0.0084	**1.502** ± 0.0078	**1.616** ± 0.0113	**1.839** ± 0.0095
$\overline{\mathrm{RB}}_p(\ell=1)$	**1.323** ± 0.0076	**1.306** ± 0.1659	**1.413** ± 0.0088	1.539 ± 0.0078	1.655 ± 0.0112	1.883 ± 0.0099
$\overline{\overline{\mathrm{RB}}}_p(\ell=1)$	*1.269* ± 0.0071	*1.251* ± 0.1554	*1.352* ± 0.0079	*1.471* ± 0.0079	*1.585* ± 0.0117	*1.804* ± 0.0095
$\overline{\mathrm{RB}}_p(\ell=2)$	1.308 ± 0.0086	1.290 ± 0.2260	1.405 ± 0.0090	**1.541** ± 0.0084	**1.666** ± 0.0122	**1.899** ± 0.0094
$\overline{\overline{\mathrm{RB}}}_p(\ell=2)$	*1.268* ± 0.0072	*1.249* ± 0.1611	*1.343* ± 0.0081	*1.455* ± 0.0073	*1.565* ± 0.0105	*1.778* ± 0.0094
$\overline{\mathrm{RB}}_p(\ell=4)$	1.306 ± 0.0104	1.278 ± 0.2822	1.386 ± 0.0118	1.526 ± 0.0154	1.655 ± 0.0137	1.884 ± 0.0106
$\overline{\overline{\mathrm{RB}}}_p(\ell=4)$	*1.265* ± 0.0083	*1.214* ± 0.2133	*1.251* ± 0.0085	*1.319* ± 0.0090	*1.394* ± 0.0118	*1.556* ± 0.0095

et al. (2013). But, at optimal levels, the optimal MOP class in (1.7) beats the H EVI-estimator in the whole (ξ,ρ)-plane.

- Despite this last comment, that asymptotically at optimal levels $\overline{\mathrm{H}}$ beats H in the whole (ξ,ρ)-plane, neither $\overline{\mathrm{RB}}$ nor $\overline{\overline{\mathrm{RB}}}$ beats CH in the whole (ζ,ρ)-plane. But, again asymptotically and at optimal levels, both $\overline{\mathrm{RB}}$ and $\overline{\overline{\mathrm{RB}}}$ beat CH, and also $\overline{\overline{\mathrm{RB}}}$ beats $\overline{\mathrm{RB}}$, in wide, interesting regions of the (ζ,ρ)-plane.
- For the large variety of simulated models considered, there is generally a reduction in RMSE, as well as in bias. We get estimates closer to the target value ξ, when we move from H to H^*, next to CH and finally to either $\overline{\mathrm{RB}}^*$ or to $\overline{\overline{\mathrm{RB}}}^*$. Such a reduction is particularly high for values of ρ approaching zero. This happens even when we work with models out of the scope of the stated theorems, such as the the log-gamma $(\rho=0)$.
- However, and as already mentioned above, there is always a value of p that enables both $\overline{\mathrm{RB}}_p(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$ to outperform the corresponding estimated RB EVI-estimator, in (5.1). This provides an indication for the use of an algorithm that adaptively chooses p among both $\overline{\mathrm{RB}}_p(k)$ and $\overline{\overline{\mathrm{RB}}}_p(k)$, respectively defined in (1.10) and (1.11). For the adaptive choice of the tuning parameters k and p, we suggest either heuristic choices or the use of simple and double-bootstrap methods, a topic out of the scope of this article.

7 Proofs

Proof. (Theorem 2) The consistency of the statistics in (1.5), for all $0 < p < 1/\xi$, has been proved in Brilhante *et al.* (2013), and it more generally holds for all real $p < 1/\xi$. Indeed, if (2.8) holds, and with Y_i, $i \geq 1$, IID unit Pareto RVs (with CDF $F_{_Y}(y) = 1 - 1/y,\ y \geq 1$), we get for any real p,

$$U_{ik}^p := \left(\frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^p = \left(\frac{U(Y_{n-i+1:n})}{U(Y_{n-k:n})}\right)^p = Y_{k-i+1:k}^{\xi p}(1 + o_p(1)).$$

Since $\mathbb{E}(Y^a) = 1/(1-a)$ if $a < 1$, the law of large numbers enables us to say that if $p < 1/\xi$, not necessarily positive,

$$\mathrm{H}_p(k) \ \stackrel{d}{=}\ \frac{1}{p}\Big(1 - \Big(\frac{1}{k}\sum_{i=1}^{k} Y_i^{\xi p}(1 + o_p(1))\Big)^{-1}\Big)$$

converges weakly to ξ, as $n \to \infty$. Let us now deal with the asymptotic non-degenerate behaviour of $\mathrm{H}_p(k)$, in (1.5), under the third-order framework in (2.5), but with $A(\cdot)$ and $B(\cdot)$ chosen as in (2.7). Note first that we can then write (2.5) as

$$\frac{U(tx)}{U(t)} = x^\xi\left(1 + A(t)\Big(\frac{x^\rho - 1}{\rho}\Big) + A^2(t)\Big(\frac{1}{2}\Big(\frac{x^\rho - 1}{\rho}\Big)^2 + \frac{\zeta}{\xi}\Big(\frac{x^{2\rho} - 1}{2\rho}\Big)\Big)(1 + o(1))\right).$$

But

$$\mathrm{M}_p^p(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^p = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{U(Y_{n-i+1:n})}{U(Y_{n-k:n})}\right)^p.$$

We can thus write

$$\mathrm{M}_p^p(k) = \frac{1}{k}\sum_{i=1}^{k} Y_i^{p\xi}\Big(1 + pA(n/k)\ (Y_i^\rho - 1)/\rho + \frac{p^2A^2(n/k)}{2}\big((Y_i^\rho - 1)/\rho\big)^2$$
$$+ p\zeta A^2(n/k))(Y_i^{2\rho} - 1)/(2\xi\rho) + o_p(A^2(n/k))\Big).$$

Consequently, with

$$U_k(\xi, p, \rho) = \frac{1}{k}\sum_{i=1}^{k} Y_i^{p\xi}\left(\frac{Y_i^\rho - 1}{\rho}\right),\ \ V_k(\xi, p, \rho) = \frac{1}{k}\sum_{i=1}^{k} Y_i^{p\xi}\left(\frac{Y_i^\rho - 1}{\rho}\right)^2,$$
$$W_k(\xi, p, \rho) = \frac{1}{k}\sum_{i=1}^{k} Y_i^{p\xi}\left(\frac{Y_i^{2\rho} - 1}{2\rho}\right),$$

$$\mathrm{M}_p^p(k) = \frac{1}{k}\sum_{i=1}^{k} Y_i^{p\xi} + pA(n/k)\ U_k(\xi,p,\rho)$$
$$+ pA^2(n/k)\Big(pV_k(\xi,p,\rho)/2 + \zeta\ W_k(\xi,p,\rho)/\xi\Big) + o_p(A^2(n/k)).$$

For $a < 1/2$, we have $\mathbb{V}ar(Y^a) = 1/(1-2a) - (1/(1-a))^2 = a^2(\big((1-a)^2(1-2a)\big)$. We thus know that for $p < 1/(2\xi)$,

$$\frac{\sqrt{k}(1-p\xi)\sqrt{1-2p\xi}\left(\frac{1}{k}\sum\limits_{i=1}^{k} Y_i^{p\xi} - \frac{1}{1-p\xi}\right)}{p\xi} =: V_k^{(p)}$$

is asymptotically standard normal, and we can write

$$\frac{1}{k}\sum_{i=1}^{k} Y_i^{p\xi} = \frac{1}{1-p\xi} + \frac{p\xi V_k^{(p)}}{\sqrt{k}(1-p\xi)\sqrt{1-2p\xi}}.$$

Also, and now for $p < 1/\xi$, and for any $a \le 0$,

$$\mathbb{E}\left(Y^{p\xi}(Y^a - 1)\right) = \frac{a}{(1-p\xi)(1-p\xi-a)}.$$

and

$$\mathbb{E}\left(Y^{p\xi}(Y^a - 1)^2\right) = \frac{2a^2}{(1-p\xi)(1-p\xi-a)(1-p\xi-2a)}.$$

We can thus write

$$\mathrm{M}_p^p(k) = \frac{1}{1-p\xi}\left(1 + \frac{p\xi V_k^{(p)}}{\sqrt{k(1-2p\xi)}} + \frac{pA(n/k)}{1-p\xi-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\right.$$
$$\left. + \frac{pA^2(n/k)}{1-p\xi-2\rho}\left(\frac{p\xi + \zeta(1-p\xi-\rho)}{\xi(1-p\xi-\rho)}\right) + o_p(A^2(n/k))\right).$$

Let's go back to the EVI-estimator in (1.5):

$$\mathrm{H}_p(k) = \frac{1 - \mathrm{M}_p^{-p}(k)}{p} = \frac{1}{p}\left(1 - (1-p\xi)\middle/\left(1 + \frac{p\xi V_k^{(p)}}{\sqrt{k(1-2p\xi)}} + \frac{pA(n/k)}{1-p\xi-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\right.\right.$$
$$\left.\left. + \frac{pA^2(n/k)}{1-p\xi-2\rho}\left(\frac{p\xi + \zeta(1-p\xi-\rho)}{\xi(1-p\xi-\rho)}\right) + o_p(A^2(n/k))\right)\right).$$

We can thus further write

$$\mathrm{H}_p(k) = \xi + \frac{\xi(1-p\xi)V_k^{(p)}}{\sqrt{k}\sqrt{1-2p\xi}} + \frac{(1-p\xi)A(n/k)}{1-p\xi-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)$$
$$- \frac{p(1-p\xi)A^2(n/k)}{(1-p\xi-\rho)^2} + \frac{(1-p\xi)A^2(n/k)}{1-p\xi-2\rho}\left(\frac{p\xi + \zeta(1-p\xi-\rho)}{\xi(1-p\xi-\rho)}\right) + o_p(A^2(n/k))\Big)\Big).$$

After some trivial computations, (3.3) follows. Indeed, since $\sqrt{k}\ O_p\big(A(n/k)/\sqrt{k}\big) = O_p(A(n/k)) \to 0$, as $n \to \infty$, the summands $O_p\big(A(n/k)/\sqrt{k}\big)$ are totally irrelevant for the asymptotic bias, that follows in a straight forward fashion from the representations above.

If we consider $\overline{\mathrm{H}}$, in (1.7), we obviously merely need to replace in (3.3) $p\xi$ by φ_ρ. ∎

Proof. (Theorem 3) Considering

$$\overline{\mathrm{RB}}_p(k;\beta,\rho) := \mathrm{H}_p(k)\Big(1 - \frac{\beta(1-\varphi_\rho)}{1-\rho-\varphi_\rho}\left(\frac{n}{k}\right)^\rho\Big) = \xi\Bigg(1 + \frac{(1-p\xi)V_k^{(p)}}{\sqrt{k}\sqrt{1-2p\xi}} + \frac{(1-p\xi)A(n/k)}{\xi(1-p\xi-\rho)}$$
$$+ \frac{(1-p\xi)(\zeta(1-p\xi-\rho)^2+p\xi\rho)A^2(n/k)}{\xi^2(1-p\xi-\rho)^2(1-p\xi-2\rho)} + o_p(A^2(n/k))\Bigg)\Big(1 - \frac{\beta(1-\varphi_\rho)}{1-\rho-\varphi_\rho}\left(\frac{n}{k}\right)^\rho\Big),$$

we furthermore can write

$$\overline{\mathrm{RB}}_p(k;\beta,\rho) = \xi\Bigg(1 + \frac{(1-p\xi)V_k^{(p)}}{\sqrt{k}\sqrt{1-2p\xi}} + \frac{\rho(p\xi-\varphi_\rho)A(n/k)}{\xi(1-p\xi-\rho)(1-\varphi_\rho-\rho)}$$
$$- \frac{A^2(n/k)(1-p\xi)(1-\varphi_\rho)}{\xi^2(1-p\xi-\rho)(1-\varphi_\rho-\rho)} + \frac{(1-p\xi)(\zeta(1-p\xi-\rho)^2+p\xi\rho)A^2(n/k)}{\xi^2(1-p\xi-\rho)^2(1-p\xi-2\rho)} + o_p(A^2(n/k))\Bigg).$$

Consequently (3.4) holds with

$$b^{(1)}_{\overline{\mathrm{RB}}_p} = \frac{\rho(p\xi-\varphi_\rho)}{(1-p\xi-\rho)(1-\varphi_\rho-\rho)}$$

and

$$b^{(2)}_{\overline{\mathrm{RB}}_p} = \frac{1-p\xi}{\xi(1-p\xi-\rho)}\left(\frac{\zeta(1-p\xi-\rho)^2+p\xi\rho}{(1-p\xi-\rho)(1-p\xi-2\rho)} - \frac{1-\varphi_\rho}{1-\varphi_\rho-\rho}\right).$$

If we replace $p\xi$ by φ_ρ, we get

$$b^{(2)}_{\overline{\mathrm{RB}}} = \frac{1-\varphi_\rho}{\xi(1-\varphi_\rho-\rho)}\left(\frac{\zeta(1-\varphi_\rho-\rho)^2+\varphi_\rho\rho}{(1-\varphi_\rho-\rho)(1-\varphi_\rho-2\rho)} - \frac{1-\varphi_\rho}{1-\varphi_\rho-\rho}\right)$$
$$= \frac{1-\varphi_\rho}{\xi(1-\varphi_\rho-\rho)^2(1-\varphi_\rho-2\rho)}\big(\zeta(1-\varphi_\rho-\rho)^2+\rho\varphi_\rho-(1-\varphi_\rho)(1-\varphi_\rho-2\rho)\big),$$

the value associated with ORBMO_1, in (1.9).

On the basis of (3.3), we can write

$$1-\rho-p\ \mathrm{H}_p(k) = (1-p\xi-\rho)\Bigg(1 - \frac{\xi p(1-p\xi)V_k^{(p)}}{\sqrt{k}(1-p\xi-\rho)\sqrt{1-2p\xi}} - \frac{p(1-p\xi)A(n/k)}{(1-p\xi-\rho)^2}$$
$$- \frac{p(1-p\xi)(\zeta(1-p\xi-\rho)^2+p\xi\rho)A^2(n/k)}{\xi(1-p\xi-\rho)^3(1-p\xi-2\rho)} + o_p(A^2(n/k)).$$

Moreover,

$$1 - p\ \mathrm{H}_p(k) = (1 - p\xi)\Big(1 - \frac{\xi p V_k^{(p)}}{\sqrt{k}\sqrt{1-2p\xi}} - \frac{pA(n/k)}{1 - p\xi - \rho} - \frac{p\big(\zeta(1-p\xi-\rho)^2 + p\xi\rho\big)A^2(n/k)}{\xi(1-p\xi-\rho)^2(1-p\xi-2\rho)} + o_p(A^2(n/k))\Big).$$

Consequently, if we consider

$$\overline{\overline{\mathrm{RB}}}_p(k;\beta,\rho) := \mathrm{H}_p(k)\Big(1 - \frac{\beta(1-p\ \mathrm{H}_p(k))}{1-\rho-p\ \mathrm{H}_p(k)}\Big(\frac{n}{k}\Big)^{\rho}\Big) = \mathrm{H}_p(k)\Big(1 - \frac{A(n/k)(1-p\ \mathrm{H}_p(k))}{\xi(1-\rho-p\ \mathrm{H}_p(k))}\Big)$$
$$= \mathrm{H}_p(k)\Big(1 - \frac{A(n/k)(1-p\xi)}{\xi(1-p\xi-\rho)} - \frac{p\rho(1-p\xi)A^2(n/k)}{\xi(1-p\xi-\rho)^3} + O_p\Big(\frac{A(n/k)}{\sqrt{k}}\Big) + o_p(A^2(n/k))\Big),$$

we get

$$\overline{\overline{\mathrm{RB}}}_p(k;\beta,\rho) = \xi\Bigg(1 + \frac{(1-p\xi)V_k^{(p)}}{\sqrt{k}\sqrt{1-2p\xi}} + O_p\Big(\frac{A(n/k)}{\sqrt{k}}\Big) + o_p(A^2(n/k))$$
$$- \frac{A^2(n/k)(1-p\xi)^2}{\xi^2(1-p\xi-\rho)^2} - \frac{p\rho(1-p\xi)A^2(n/k)}{\xi(1-p\xi-\rho)^3} + \frac{(1-p\xi)(\zeta(1-p\xi-\rho)^2 + p\xi\rho)A^2(n/k)}{\xi^2(1-p\xi-\rho)^2(1-p\xi-2\rho)}\Bigg).$$

Consequently (3.4) holds with

$$b_{\overline{\overline{\mathrm{RB}}}_p} = -\frac{(1-p\xi)^2}{\xi(1-p\xi-\rho)^2} - \frac{p\rho(1-p\xi)}{(1-p\xi-\rho)^3} + \frac{(1-p\xi)(\zeta(1-p\xi-\rho)^2 + p\xi\rho)}{\xi(1-p\xi-\rho)^2(1-p\xi-2\rho)}$$
$$= \frac{1-p\xi}{(1-p\xi-\rho)^2}\Bigg(\frac{\zeta(1-p\xi-\rho)^3 + p\xi\rho^2 - (1-p\xi)(1-p\xi-\rho)(1-p\xi-2\rho)}{\xi(1-p\xi-\rho)(1-p\xi-2\rho)}\Bigg).$$

■

Proof. (Theorem 4) If we estimate consistently β and ρ through the estimators $\hat\beta$ and $\hat\rho$, we can use Cramer's delta-method, and obtain for any of the RB MOP classes of EVI-estimators, either in (1.10) or in (1.11), generally denoted $\mathrm{RB}_p(k)$,

$$\mathrm{RB}_p\big(k;\hat\beta,\hat\rho\big) - \mathrm{RB}_p\big(k;\beta,\rho\big) \overset{p}{\sim} a_{\mathrm{RB}_p} A(n/k)\left\{\Big(\frac{\hat\beta-\beta}{\beta}\Big) + (\hat\rho-\rho)\Big[\ln(n/k) - b_{\mathrm{RB}_p}\Big]\right\}. \quad (7.1)$$

for suitable functions $a_{\mathrm{RB}_p} \equiv a_{\mathrm{RB}_p}(\xi,\rho)$ and $b_{\mathrm{RB}_p} \equiv b_{\mathrm{RB}_p}(\xi,\rho)$, The first part of the theorem, related with levels k such that $\sqrt{k}\ A(n/k) \to \lambda$, finite, thus follows in a straight forward way from (7.1).

Next, since (2.14) holds, i.e., $(\hat{\beta}-\beta)/\beta \overset{p}{\sim} -\ln(n/k_1)\,(\hat{\rho}-\rho)$, we have

$$\mathrm{RB}_p(k;\hat{\beta},\hat{\rho}) - \mathrm{RB}_p(k;\beta,\rho) \overset{p}{\sim} -a_{\mathrm{RB}_p}\,(\hat{\rho}-\rho)\,A(n/k)\,\Big(\ln(k/k_1)+b_{\mathrm{RB}_p}\Big) =: W_{k,k_1}. \tag{7.2}$$

Under the conditions in the theorem, i.e., with k_1 optimal for the ρ-estimation, $\hat{\rho}-\rho = O_p\big(1/\big(\sqrt{k_1}\;A(n/k_1)\big)\big)$,

$$\sqrt{k}\;W_{k,k_1} = O_p\Big(\frac{\sqrt{k}\;A(n/k)}{\sqrt{k_1}\;A(n/k_1)}\ln\Big(\frac{k}{k_1}\Big)\Big) = O_p\Big(\Big(\frac{k}{k_1}\Big)^{\frac{1}{2}-\rho}\ln\Big(\frac{k}{k_1}\Big)\Big) = o_p(1) \quad \text{if} \quad k/k_1\to 0,$$

and the second part of the theorem follows.

If we further assume that $\hat{\rho}-\rho = o_p(\ln(n/k)/(\sqrt{k}A(n/k))$, we are able to prove the final results in the theorem. ■

Acknowledgements. Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2014 (CEAUL), PEst-OE/MAT/UI0297/2014 (CMA/UNL) and National Research Foundation of South Africa.

References

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

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

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

[4] 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.

[5] Beran, J., Schell, D. and Stehlík, M. (2014). The harmonic moment tail index estimator: asymptotic distribution and robustness. *Ann. Inst. Statist. Math.* **66**, 193–220.

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

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

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

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

[10] Caeiro, F. and Gomes, M.I. (2008). Minimum-variance reduced-bias tail index and high quantile estimation. *Revstat* **6**:1, 1–20.

[11] Caeiro, F. and Gomes, M.I. (2011). Asymptotic comparison at optimal levels of reduced-bias extreme value index estimators. *Statistica Neerlandica* **65**:4, 462–488.

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

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

[14] Dell'Aquila, R. & Embrechts, P. (2006). Extremes and robustness: a contradiction? *Fin. Mkts. Portfolio Mgmt.* **20**, 103–118.

[15] 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.

[16] 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.

[17] 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.

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

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

[20] Gomes, M.I. and Guillou, A. (2014). Extreme Value Theory and Statistics of Univariate Extremes: A Review. *International Statistical Review*, doi:10.1111/insr.12058.

[21] 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.

[22] 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.

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

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

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

[26] Gomes, M.I., Brilhante, F. and Pestana, D. (2013). New reduced-bias estimators of a positive extreme value index. *Communications in Statistics—Simulation and Computation*, DOI:10.1080/03610918.2013.875567

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

[28] 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.

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

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

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

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

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

[34] Stehlík M., Potocký R., Waldl H. and Fabián Z. (2010). On the favourable estimation of fitting heavy tailed data. *Computational Statistics* **25**, 485–503.