Discussiones Mathematicae
Probability and Statistics 30 (2010) 35–51

COMPARISON AT OPTIMAL LEVELS OF CLASSICAL TAIL INDEX ESTIMATORS: A CHALLENGE FOR REDUCED-BIAS ESTIMATION?*

M. Ivette Gomes

DEIO and CEAUL
Faculdade de Ciências, Universidade de Lisboa
1749−016 Lisboa, Portugal

e-mail: ivette.gomes@fc.ul.pt

AND

Lígia Henriques-Rodrigues

Instituto Politécnico de Tomar and CEAUL

To Professor João Tiago Mexia, a token of friendship.

Abstract

In this article, we begin with an asymptotic comparison at optimal levels of the so-called "*maximum likelihood*" (ML) extreme value index estimator, based on the excesses over a high random threshold, denoted PORT-ML, with PORT standing for *peaks over random thresholds*, with a similar ML estimator, denoted PORT-MP, with MP standing for *modified-Pareto*. The PORT-MP estimator is based on the same excesses, but with a trial of accommodation of bias on the Generalized Pareto model underlying those excesses. We next compare the behaviour of these ML implicit estimators with the equivalent behaviour of a few explicit tail index estimators, the *Hill*, the *moment*, the *generalized Hill* and the *mixed moment*. As expected, none of the estimators can always dominate the alternatives, even when we include second-order MVRB tail index estimators, with MVRB standing for *minimum-variance reduced-bias*. However, the asymptotic performance of the MVRB estimators is quite interesting and provides a challenge for a further study of these MVRB estimators at optimal levels.

*Research partially supported by FCT/OE, POCI 2010 and PTDC/FEDER.

Keywords and phrases: statistics of extremes, semi-parametric estimation, bias estimation, heavy tails, optimal levels.

2000 Mathematics Subject Classification: Primary 62G32; Secondary 65C05.

1. Introduction and preliminaries

Heavy-tailed models are quite useful in the most diversified areas of application, like computer science, telecommunication networks, insurance, finance and biostatistics, among others. Power laws, such as the Pareto income distribution (Pareto, 1965) and the Zipf's law for city-size distribution (Zipf, 1941), have been observed a few decades ago in important phenomena in economics and biology and have seriously attracted scientists in recent years.

In *statistics of extremes*, whenever dealing with large values, and with the notation RV_α standing for the class of *regularly varying* functions at infinity with an *index of regular variation* equal to α, i.e., positive measurable functions g such that $\lim_{t\to\infty} g(tx)/g(t) = x^\alpha$, for all $x > 0$, a model F is said to be *heavy-tailed* whenever the right-tail

$$\overline{F} := 1 - F \in RV_{-1/\gamma} \quad \text{for some} \quad \gamma > 0. \tag{1.1}$$

We then have a polynomially decreasing right-tail. Equivalently (Gnedenko, 1943), we are then in the domain of attraction for maxima of a Fréchet-type *extreme value* distribution function (d.f.),

$$EV_\gamma(x) = \exp(-(1+\gamma x)^{-1/\gamma}),\ x \geq -1/\gamma, \quad \text{with} \quad \gamma > 0,$$

and we write $F \in \mathcal{D}_{\mathcal{M}}(EV_{\gamma>0})$. The parameter γ is the *extreme value index* or *tail index*, the primary parameter of extreme events.

For consistent semi-parametric estimation of the tail index γ we need to work with an *intermediate* number k of top order statistics (o.s.'s), i.e., we need to consider a sequence of integers $k = k_n$, $k \in [1, n)$, such that

$$k = k_n \to \infty, \quad \text{and} \quad k_n = o(n), \quad \text{as} \quad n \to \infty. \tag{1.2}$$

1.1. Explicit tail index estimators

Due to its simplicity, the most popular tail index estimator, valid only for $\gamma \geq 0$, is the Hill estimator (Hill, 1975), with the functional form

$$\hat{\gamma}^H_{n,k} := \frac{1}{k}\sum_{i=1}^{k}\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\} =: \frac{1}{k}\sum_{i=1}^{k} V_{ik}, \tag{1.3}$$

where $X_{i:n}$ denotes, as usual, the i-th ascending o.s., $1 \le i \le n$, associated with a random sample $(X_1, X_2, \cdots, X_n)$.

Apart from the Hill estimator, and with the notation

$$M^{(j)}_{n,k} := \frac{1}{k}\sum_{i=1}^{k} V^j_{ik}, \qquad L^{(j)}_{n,k} := 1 - \frac{1}{k}\sum_{i=1}^{k}\left(1 - \frac{X_{n-k:n}}{X_{n-i+1:n}}\right)^j, \quad j \ge 1, \tag{1.4}$$

with V_{ik}, $1 \le i \le k$ defined in (1.3), we shall also consider

- the moment estimator (Dekkers *et al.*, 1989), given by

$$\hat{\gamma}^M_{n,k} := M^{(1)}_{n,k} + \tfrac{1}{2}\left\{1 - \left(M^{(2)}_{n,k}/\big(M^{(1)}_{n,k}\big)^2 - 1\right)^{-1}\right\}, \tag{1.5}$$

- the generalized Hill estimator (Beirlant *et al.*, 1996), based on the Hill estimator in (1.3) and with the functional form

$$\hat{\gamma}^{GH}_{n,k} = \hat{\gamma}^H_{n,k} + \frac{1}{k}\sum_{i=1}^{k}\left\{\ln \hat{\gamma}^H_{n,i} - \ln \hat{\gamma}^H_{n,k}\right\}, \tag{1.6}$$

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

- the mixed moment estimator (Fraga Alves *et al.*, 2009), based on the statistics $M^{(1)}_{n,k}$ and $L^{(1)}_{n,k}$ in (1.4), and given by

$$\hat{\gamma}^{MM}_{n,k} := \frac{\hat{\varphi}_{n,k} - 1}{1 + 2\min\big(\hat{\varphi}_{n,k} - 1, 0\big)}, \qquad \text{with} \qquad \hat{\varphi}_{n,k} := \frac{M^{(1)}_{n,k} - L^{(1)}_{n,k}}{\big(L^{(1)}_{n,k}\big)^2}. \tag{1.7}$$

The estimators in (1.5), (1.6) and (1.7) are valid for all $\gamma \in \mathbb{R}$, but will be considered only for $\gamma \ge 0$. None of the estimators in this Section is invariant for changes in location, but they can easily be made location-invariant with the technique used in Araújo Santos *et al.* (2006), Gomes *et al.* (2008a) and Fraga Alves *et al.* (2009).

1.2. PORT-ML and PORT-MP tail index estimators

As mentioned in de Haan and Ferreira (2006), the class of d.f.'s $F\in\mathcal{D}_{\mathcal{M}}(EV_\gamma)$, for some $\gamma > 0$ (or, more generally, for $\gamma \in \mathbb{R}$), cannot be parameterized with a finite number of parameters, and consequently, there does not exist a maximum-likelihood (ML) estimator for γ in such a wide class of models. There exists however an estimator, introduced by Smith (1987), usually denoted as the ML 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 $X_{n-k:n}$,

$$W_{ik} := X_{n-i+1:n} - X_{n-k:n}, \quad 1 \le i \le k < n. \tag{1.8}$$

For models in (1.1), these excesses are approximately distributed as the whole set of the k o.s.'s associated with a sample of size k from a *Generalized Pareto* (GP) model, with d.f.

$$GP(x;\gamma,\alpha) = 1 - (1+\alpha\ x)^{-1/\gamma}\,, \quad x>0 \quad (\alpha,\ \gamma>0),$$

a re-parametrization due to Davison (1984). Indeed, αW_{ik} is well approximated by $Y^{\gamma}_{k-i+1:k} - 1$, with Y a unit Pareto r.v., with d.f. $F_{_Y}(y) = 1 - 1/y,\ y \ge 1$. The solution of the ML equations associated with the above mentioned set-up gives rise to an explicit expression for the ML estimator of γ, a function of the ML-implicit estimator $\hat{\alpha}_{_{ML}}$ of α and the sample of the excesses, given by

$$\hat{\gamma}^{ML}_{n,k} \equiv \hat{\gamma}^{ML}_{n,k,\hat{\alpha}_{ML}} := \frac{1}{k}\sum_{i=1}^{k} \ln(1+\hat{\alpha}_{_{ML}}\ W_{ik}), \tag{1.9}$$

and here called PORT-ML tail index estimator, with PORT standing for *peaks over random threshold*, a terminology introduced in Araújo Santos *et al.* (2006) whenever working with excesses over a central o.s., also adequate to excesses over any intermediate o.s. A comprehensive study of the asymptotic properties of the ML estimator in (1.9) has been undertaken in Drees *et al.* (2004). Weak consistency is attained whenever we work with models in (1.1) and condition (1.2) holds.

Remark 1.1. 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 (1.8), $1 + \hat{\alpha}\ W_{ik} = X_{n-i+1:n}/X_{n-k:n}$, and $\hat{\gamma}^{ML}_{n,k,\hat{\alpha}} = \frac{1}{k}\sum_{i=1}^{k}\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}$ is the average of the log-excesses $V_{ik},\ 1 \le i \le k$, i.e., it is the classical Hill estimator in (1.3).

Dealing with heavy tails only, we are also interested in a similar ML estimator, based on the excesses over a high random threshold, but with a trial of accommodation of bias on the GP model underlying those excesses. Gomes *et al.* (2008b) suggested the use of an adequate weighting of the log-excesses V_{ik} instead of the Hill estimator. These same weights

$$p_{ik} = p_{ik}(\beta, \rho) = e^{-\beta (n/k)^\rho \psi_{ik}} \underset{k \to \infty}{\longrightarrow} 1, \qquad \psi_{ik} = -\frac{(i/k)^{-\rho} - 1}{\rho \ln(i/k)}, \ 1 \le i \le k,$$

dependent on a vector of second-order unknown parameters $(\beta, \rho) \in \mathbb{R} \setminus \{0\} \times \mathbb{R}^-$, made explicit in Section 2 of this paper, are such that, uniformly in i,

$$\alpha W_{ik} - \left(Y_{k-i+1:k}^{\gamma/p_{ik}} - 1\right) = o_p\left(\alpha W_{ik} - \left(Y_{k-i+1:k}^{\gamma} - 1\right)\right).$$

The validity of this result led Gomes and Henriques-Rodrigues (2008) to expect to possibly be able to get a "better" estimator of γ, provided that one uses for αW_{ik} the approximation $Y_{k-i+1:k}^{\gamma/p_{ik}} - 1$ instead of the approximation $Y_{k-i+1:k}^{\gamma} - 1$, used to support the PORT-ML estimator. The maximization of the log-likelihood associated with such an approximation, for $1 \le i \le k$, leads us to

$$\hat{\gamma}_{n,k}^{MP} \equiv \hat{\gamma}_{n,k,\hat{\alpha}_{MP},\hat{\beta},\hat{\rho}}^{MP} := \frac{1}{k} \sum_{i=1}^{k} p_{ik}(\hat{\beta}, \hat{\rho}) \ln(1 + \hat{\alpha}_{MP} W_{ik}), \tag{1.10}$$

called the PORT-MP tail index estimator, with MP standing for *modified Pareto*. The estimators $(\hat{\beta}, \hat{\rho})$ need to be adequate consistent estimators of the second-order parameters (β, ρ), essentially such that $\hat{\rho} - \rho = o_p(1/\ln n)$, as $n \to \infty$.

Remark 1.2. If we now replace, in (1.10), $\hat{\alpha}_{MP}$ by the heuristic estimator $\hat{\alpha} = 1/X_{n-k:n}$, we get the *weighted log-excesses* or *weighted-Hill* (WH) estimator,

$$\hat{\gamma}_{n,k}^{WH} \equiv \hat{\gamma}_{n,k,\hat{\alpha},\hat{\beta},\hat{\rho}}^{WH} := \frac{1}{k} \sum_{i=1}^{k} e^{-\hat{\beta}\ (n/k)^{\hat{\rho}}\ \hat{\psi}_{ik}}\ V_{ik}, \tag{1.11}$$

introduced and studied in Gomes *et al.* (2008b).

Remark 1.3. Another bias-corrected Hill (CH) estimator, and the simplest one among the ones so far devised, was introduced in Caeiro *et al.* (2005). It has the functional form

$$\hat{\gamma}^{CH}_{n,k} \equiv \hat{\gamma}^{CH}_{n,k,\hat{\alpha},\hat{\beta},\hat{\rho}} := \hat{\gamma}^{H}_{n,k}\big(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1-\hat{\rho})\big). \tag{1.12}$$

The estimators in (1.11) and (1.12) can both be second-order minimum-variance reduced-bias (MVRB) estimators, for adequate levels k and an adequate external estimation of a vector of second-order parameters, (β,ρ), introduced in Section 2 of this article, i.e., the use of $\hat{\gamma}^{WH}_{n,k}$ or $\hat{\gamma}^{CH}_{n,k}$, and an adequate estimation of (β,ρ), enables us to eliminate the dominant component of bias of the Hill estimator, $\hat{\gamma}^{H}_{n,k}$, keeping its asymptotic variance.

1.2. Scope of the paper

In this article, after reviewing, in Section 2, a few technical details in statistics of extremes related with the topic under consideration, we go on, in Section 3, with the asymptotic comparison at optimal levels of the different classical estimators under consideration, the *Hill*, the *moment*, the *generalized Hill*, the *mixed moment*, the PORT-ML and the PORT-MP, in (1.3), (1.5), (1.6), (1.7), (1.9) and (1.10), respectively. We next see that the consideration of the WH-estimators, in (1.11), or the CH-estimators, in (1.12), enable us to get better estimators at the whole (γ,ρ)-plane, possibly excluding the important region $\gamma+\rho=0$, as well as the region $\gamma=-\rho/(1-\rho)$. This surely provides a challenge for a further comparative study of RB estimators at optimal levels, out of the scope of this paper.

2. Further technical details in statistics of extremes

2.1. First- and second-order framework for heavy-tailed models

In a context of heavy tails, and with the notation

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

the generalized inverse function of the underlying model F, the first order parameter (or *tail index*) γ (>0) appears, for every $x>0$, as the limiting value

$$\gamma = \lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t)}{\ln x} \quad \text{(de Haan, 1970).}$$

Indeed, we can write, equivalently to (1.1),

$$\overline{F} \in RV_{-1/\gamma} \quad \Longleftrightarrow \quad U \in RV_{\gamma}. \tag{2.1}$$

In order to obtain information on the non-degenerate asymptotic behaviour of semi-parametric tail index estimators, we need further assuming a second-order condition, ruling the rate of convergence in the first order condition in (2.1). The *second-order parameter*, ρ (≤ 0), rules such a rate of convergence, and it is the parameter appearing in

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

which we often assume to hold for every $x > 0$, and where $|A|$ must then be in RV_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 reduced-bias estimators, and for technical simplicity, we often even assume that we are working in Hall-Welsh class of models (Hall and Welsh, 1985), with a tail function

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

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

$$A(t) = \gamma\ \beta\ t^\rho,\ \ \rho < 0.$$

Equivalently, we get

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

Models like the log-gamma and the log-Pareto ($\rho = 0$) are thus excluded from this class. The standard Pareto is also excluded. But most heavy-tailed models used in applications, like the Fréchet, the generalized Pareto, the Burr and the Student's t d.f.'s belong to Hall-Welsh class of distributions.

For details on algorithms for the (β, ρ)-estimation, see Gomes and Pestana (2007) 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).

2.2. Motivation for the PORT-MP estimators – only γ is unknown

Let us assume that everything is known, apart from γ. Then,

Theorem 2.1 (Gomes and Henriques-Rodrigues, 2008)**.** *For models in Hall-Welsh class, in* (2.3)*, and for intermediate levels k, i.e. if* (1.2) *holds, we get for $\hat{\gamma}^{MP}_{n,k,\alpha,\beta,\rho}$, with $\hat{\gamma}^{MP}_{n,k,\hat{\alpha}_{MP},\hat{\beta},\hat{\rho}}$ provided in* (1.10)*, an asymptotic distributional representation of the type*

$$\hat{\gamma}^{MP}_{n,k,\alpha,\beta,\rho} \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\, N_k + o_p(A(n/k)),$$

where N_k is asymptotically standard normal. Consequently, and as $n \to \infty$, $\sqrt{k}\big(\hat{\gamma}^{MP}_{n,k,\alpha,\beta,\rho} - \gamma\big)$ is asymptotically normal, with a null mean value, not only when $\sqrt{k}\ A(n/k) \to 0$, but also when $\sqrt{k}\ A(n/k) \to \lambda \neq 0$, finite.

The main problems to be dealt with are related with the study of how the estimation of (α, β, ρ) affects the asymptotic distributional behaviour of $\hat{\gamma}^{MP}_{n,k,\alpha,\beta,\rho}$. Theorem 2.1 still holds for $\hat{\gamma}^{MP}_{n,\alpha,\hat{\beta},\hat{\rho}}$, i.e. we still have MVRB tail index estimators, if we assume α known and we estimate β and ρ externally, in an adequate way, i.e., so that $\hat{\rho} - \rho = o_p(1/\ln n)$ and $\hat{\beta} - \beta = o_p(1)$. If we estimate α and γ jointly through the maximum likelihood procedure, we no longer have a MVRB estimator, as can be seen in the following Section, Theorem 2.2.

2.3. Asymptotic behaviour of the tail index estimators

Under the validity of the second-order condition in (2.2), trivial adaptations of the results in Beirlant *et al.* (2005), Caeiro *et al.* (2005), de Haan and Ferreira (2006), Gomes and Henriques-Rodrigues (2008), Gomes *et al.* (2008b) and Fraga Alves *et al.* (2009) enable us to state, without proof, the following theorem, again for models with $\gamma > 0$. Let the notation $\mathcal{N}\left(\mu, \sigma^2\right)$ stand for a normal r.v. with mean value μ and variance σ^2.

Theorem 2.2. *Assume that condition* (2.2) *holds. Let $k = k_n$ be an intermediate sequence, i.e.* (1.2) *holds, and let us 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

$$\sqrt{k}\left(\hat{\gamma}^{\bullet}_{n,k} - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} \mathcal{N}\left(\lambda b_{\bullet}, \sigma^2_{\bullet}\right),$$

where

$$b_{H} = \frac{1}{1-\rho}, \qquad b_{M} = b_{GH} = \frac{\gamma - \gamma\rho + \rho}{\gamma(1-\rho)^2},$$

$$b_{MM} = b_{ML} = \frac{(1+\gamma)(\gamma+\rho)}{\gamma(1-\rho)(1+\gamma-\rho)},$$

$$\sigma^2_{H} = \gamma^2, \qquad \sigma^2_{M} = \sigma^2_{GH} = 1+\gamma^2, \qquad \textit{and} \qquad \sigma^2_{MM} = \sigma^2_{ML} = (1+\gamma)^2.$$

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 $\hat{\rho}$, in such a way that $\hat{\rho} - \rho = o_p(1/\ln n)$, we get*

$$b_{WH} = b_{CH} = 0,$$

$$b_{MP} = -\frac{(1+\gamma)(1+2\gamma)}{\gamma^3}\Big(\frac{1}{\rho}\ln\frac{(1+\gamma)(1-\rho)}{1+\gamma-\rho} + \frac{\gamma}{1+\gamma-\rho}\Big),$$

$$\sigma^2_{WH} = \sigma^2_{CH} = \sigma^2_{H} = \gamma^2 \qquad \textit{and} \qquad \sigma^2_{MP} = \sigma^2_{MM} = \sigma^2_{ML} = (1+\gamma)^2.$$

Remark 2.1. As can be seen from Theorem 2.2, the PORT-MP tail index estimator is no longer a MVRB estimator or even a second-order reduced-bias tail index estimator, i.e., the estimation of α through maximum-likelihood gives rise to a dominant component of bias of the order of $A(n/k)$. Note however that the estimation of α through the simple heuristic estimator $\hat{\alpha} = 1/X_{n-k:n}$ leads us to the $MVRB$ estimator (1.11), already mentioned in Remark 1.2. We then have

$$\hat{\gamma}^{WH}_{n,k} \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\, N_k + o_p(A(n/k)),$$

provided that we estimate β and ρ externally, in an adequate way, i.e., so that $\hat{\rho} - \rho = o_p(1/\ln n)$ and $\hat{\beta} - \beta = o_p(1)$ for all k on which we base $\hat{\gamma}^{WH}_{n,k} \equiv \hat{\gamma}^{WH}_{n,k,\hat{\beta},\hat{\rho}}$. A similar asymptotic behaviour holds for the estimator $\hat{\gamma}^{CH}_{n,k}$ in (1.12).

Remark 2.2. Relatively to Smith's result, rephrased in this context in Theorem 2.2 (i.e. with the replacement of a fixed threshold u by a random threshold $X_{n-k:n}$), we have for the PORT-MP the same asymptotic variance we had for the PORT-ML tail index estimator, the value $(1+\gamma)^2$, but a change in bias, although both bias are of the same order if $\gamma + \rho \neq 0$. If $\gamma + \rho = 0$ the PORT-ML estimator, being a second-order reduced-bias estimator of γ, is expected to outperform the PORT-MP estimator.

3. Asymptotic comparison at optimal levels

We now proceed to an asymptotic comparison of the estimators at their optimal levels in the lines of de Haan and Peng (1998), Gomes and Martins (2001), Gomes *et al.* (2005; 2007) and Gomes and Neves (2008). Suppose that $\hat{\gamma}^{\bullet}_{n,k}$ is any general semi-parametric estimator of the tail index, with distributional representation,

$$\hat{\gamma}^{\bullet}_{n,k} = \gamma + \frac{\sigma_{\bullet}}{\sqrt{k}} Z^{\bullet}_n + b_{\bullet} A(n/k) + o_p(A(n/k)),$$

which holds for any intermediate k, and where $Z^{\bullet}_n$ is an asymptotically standard normal r.v. Then we have,

$$\sqrt{k}\big(\hat{\gamma}^{\bullet}_{n,k} - \gamma\big) \xrightarrow{d} \mathcal{N}(\lambda b_{\bullet}, \sigma^2_{\bullet}), \quad \text{as } n \to \infty,$$

provided k is such that $\sqrt{k}A(n/k) \to \lambda$, finite, as $n \to \infty$.

The Asymptotic Mean Square Error (AMSE) is defined as

$$AMSE\big(\hat{\gamma}^{\bullet}_{n,k}\big) := \frac{\sigma^2_{\bullet}}{k} + b^2_{\bullet} A^2(n/k),$$

where $Bias_{\infty}\big(\hat{\gamma}^{\bullet}_{n,k}\big) := b_{\bullet}A(n/k)$ and $Var_{\infty}\big(\hat{\gamma}^{\bullet}_{n,k}\big) := \sigma^2_{\bullet}/k$.

Let $k_0^\bullet = k_0^\bullet(n) := \arg\min_k AMSE\big(\hat\gamma^\bullet_{n,k}\big)$ be the optimal level for the estimation of γ through $\hat\gamma^\bullet_{n,k}$, i.e., the level associated with a minimum AMSE, and let us denote $\hat\gamma^\bullet_{n0} := \hat\gamma^\bullet_{n,k_0^\bullet}$, the estimator computed at its optimal level. The use of regular variation theory (Bingham *et al.* 1987) enabled Dekkers and de Haan (1993) to prove that, whenever $b_\bullet \neq 0$, $\exists\ \varphi(n) = \varphi(n;\rho,\gamma)$, dependent only on the underlying model, and not on the estimator, such that

$$\lim_{n\to\infty} \varphi(n) AMSE\big(\hat\gamma^\bullet_{n0}\big) = \frac{2\rho-1}{\rho}\left(\sigma_\bullet^2\right)^{-\frac{2\rho}{1-2\rho}}\left(b_\bullet^2\right)^{\frac{1}{1-2\rho}} =: LMSE\big(\hat\gamma^\bullet_{n0}\big).$$

It is then sensible to consider the following:

Definition 3.1. Given $\hat\gamma^{(1)}_{n0} = \hat\gamma^{(1)}_{n,k_0^{(1)}}$ and $\hat\gamma^{(2)}_{n0} = \hat\gamma^{(2)}_{n,k_0^{(2)}}$, based on two biased estimators $\hat\gamma^{(1)}_{n,k}$ and $\hat\gamma^{(2)}_{n,k}$ for which distributional representations of the above-mentioned type hold 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 $\hat\gamma^{(1)}_{n0}$ relatively to $\hat\gamma^{(2)}_{n0}$ is

$$AREFF_{1|2} \equiv AREFF_{\hat\gamma^{(1)}_{n0}|\hat\gamma^{(2)}_{n0}} := \sqrt{\frac{LMSE\big(\hat\gamma^{(2)}_{n0}\big)}{LMSE\big(\hat\gamma^{(1)}_{n0}\big)}},$$

with LMSE given before.

Remark 3.1. Note that this measure was devised so that the higher the AREFF measure, the better the first estimator is.

Remark 3.2. If $b_\bullet = 0$, the squared bias summand is $o(A^2(n/k))$, and consequently the associated $LMSE\big(\hat\gamma^\bullet_{n0}\big) = 0$.

In the (γ,ρ)-plane, the $AREFF$ of $\hat\gamma^{MP}_{n0}$ relatively to $\hat\gamma^{ML}_{n0}$ is presented in Figure 1. As can be seen from Figure 1, the gain in efficiency for the PORT-MP estimator happens for two regions of values of $(\gamma > 0, \rho < 0)$, away from $\gamma + \rho = 0$ and close to either $\gamma = 0$ or to $\rho = 0$. In the region $\gamma + \rho = 0$, the PORT-ML estimator is a second-order reduced-bias tail index estimator, and consequently, it is expected to outperform the PORT-MP estimator at optimal levels. These results claim for a semi-parametric test of the hypothesis $H_0 : \eta = \gamma + \rho = 0$. The non-rejection of such an hypothesis would lead us to the consideration of the PORT-ML estimator, things working in favor of the PORT-MP estimator, in case of rejection of H_0. This is however an open subject, out of the scope of this paper.

ρ \ γ	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
-0.10	0.0	1.6	2.7	3.6	4.4	5.2	5.9	6.6	7.2	7.8	8.4	9.0	9.6	10.1	10.7	11.2	11.8	12.3	12.8	13.3
-0.20	0.9	0.0	0.9	1.4	1.8	2.1	2.4	2.7	3.0	3.2	3.5	3.7	3.9	4.1	4.4	4.6	4.8	4.9	5.1	5.3
-0.30	1.1	0.7	0.0	0.7	1.0	1.2	1.5	1.6	1.8	2.0	2.1	2.3	2.4	2.5	2.7	2.8	2.9	3.0	3.1	3.2
-0.40	1.2	0.9	0.6	0.0	0.6	0.8	1.0	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.2	2.3
-0.50	1.2	1.0	0.8	0.5	0.0	0.5	0.7	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.5	1.6	1.7	1.7	1.8	1.8
-0.60	1.2	1.1	0.9	0.7	0.5	0.0	0.5	0.7	0.8	0.9	1.0	1.1	1.2	1.2	1.3	1.3	1.4	1.4	1.5	1.5
-0.70	1.2	1.1	1.0	0.8	0.7	0.5	0.0	0.5	0.7	0.8	0.9	0.9	1.0	1.1	1.1	1.2	1.2	1.3	1.3	1.3
-0.80	1.2	1.1	1.0	0.9	0.8	0.7	0.5	0.0	0.5	0.6	0.7	0.8	0.9	0.9	1.0	1.0	1.1	1.1	1.2	1.2
-0.90	1.2	1.1	1.0	0.9	0.8	0.8	0.6	0.5	0.0	0.5	0.6	0.7	0.8	0.8	0.9	0.9	1.0	1.0	1.1	1.1
-1.00	1.2	1.1	1.0	1.0	0.9	0.8	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.8	0.8	0.9	0.9	1.0	1.0	1.0
-1.10	1.2	1.1	1.0	1.0	0.9	0.9	0.8	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.8	0.8	0.8	0.9	0.9	1.0
-1.20	1.2	1.1	1.0	1.0	0.9	0.9	0.8	0.8	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.7	0.8	0.8	0.9	0.9
-1.30	1.2	1.1	1.1	1.0	1.0	0.9	0.9	0.8	0.8	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.7	0.8	0.8	0.9
-1.40	1.1	1.1	1.1	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.7	0.8	0.8
-1.50	1.1	1.1	1.1	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.7	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.7	0.8
-1.60	1.1	1.1	1.1	1.0	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.7	0.7	0.6	0.5	0.0	0.5	0.6	0.7	0.7
-1.70	1.1	1.1	1.1	1.0	1.0	1.0	0.9	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.6	0.5	0.0	0.5	0.6	0.7
-1.80	1.1	1.1	1.1	1.0	1.0	1.0	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.6	0.5	0.0	0.5	0.6
-1.90	1.1	1.1	1.1	1.0	1.0	1.0	1.0	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.6	0.5	0.0	0.5
-2.00	1.1	1.1	1.1	1.0	1.0	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.6	0.6	0.0

$AREFF_{MP|ML} < 1$ $\quad 1 \le AREFF_{MP|ML} < 2$ $\quad AREFF_{MP|ML} \ge 2$

Figure 1. AREFF indicator of $\hat{\gamma}_{n0}^{MP}$ relatively to $\hat{\gamma}_{n0}^{ML}$.

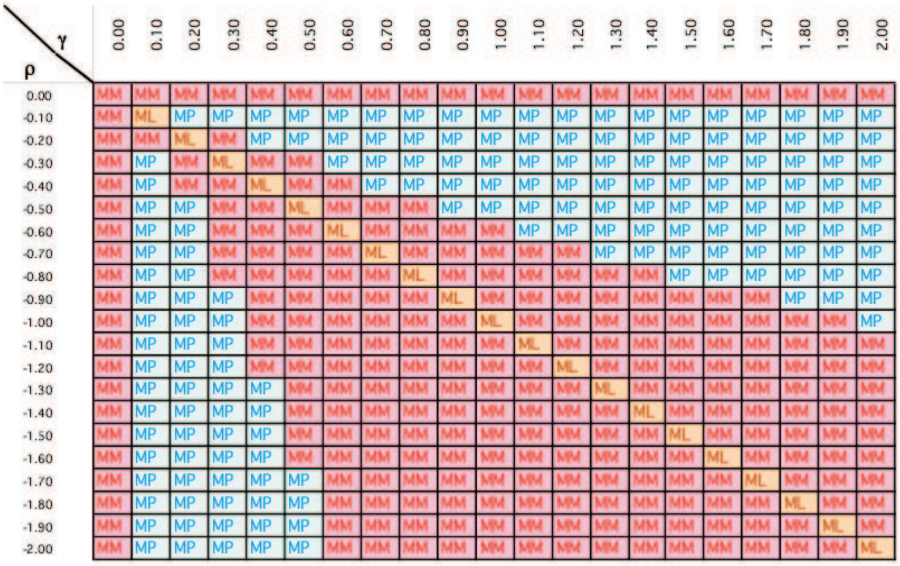
Next, in Figure 2 we present for the (γ, ρ)-plane, with $\gamma \ge 0$, $\rho \le 0$, the region where the moment estimator in (1.5) (denoted Mo), or equivalently the generalized Hill estimator in (1.6), beats the Hill estimator in (1.3). As can be seen in Figure 3, the MM-estimator in (1.7) (asymptotically equivalent to the ML-estimator in (1.9), unless $\gamma + \rho = 0$ and $(\gamma, \rho) \ne (0, 0)$), and the ML-estimator can outperform the Hill and/or the moment estimators at optimal levels. In Figure 4 we show the comparative behaviour at optimal levels of the PORT-ML and PORT-MP tail index estimators, in (1.9) and (1.10), respectively, already done in Gomes and Henriques-Rodrigues (2008), now also comparatively with the equivalent behaviour of the MM- and ML-estimators. In Figure 5 we exhibit the comparative behaviour of all "classical" tail index estimators under study.

ρ \ γ	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	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
-0.10	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.20	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.30	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.40	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.50	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.60	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.70	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.80	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.90	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-1.00	H	H	H	H	H	Mo	Mo	Mo	Mo	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-1.10	H	H	H	H	H	Mo	Mo	Mo	Mo	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo
-1.20	H	H	H	H	H	Mo	Mo	Mo	H	H	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo
-1.30	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	H	H	H	Mo	Mo	Mo
-1.40	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	H	H	H	H	Mo	Mo
-1.50	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	H	H	H	H	H	Mo
-1.60	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	H	H	H	H	H	H
-1.70	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	H	H	H	H	H	H
-1.80	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	H	H	H	H	H	H
-1.90	H	H	H	H	H	H	H	Mo	H	H	H	H	H	H	H	H	H	H	H	H	H
-2.00	H	H	H	H	H	H	H	Mo	H	H	H	H	H	H	H	H	H	H	H	H	H

Figure 2. The moment can outperform the Hill estimator at optimal levels.

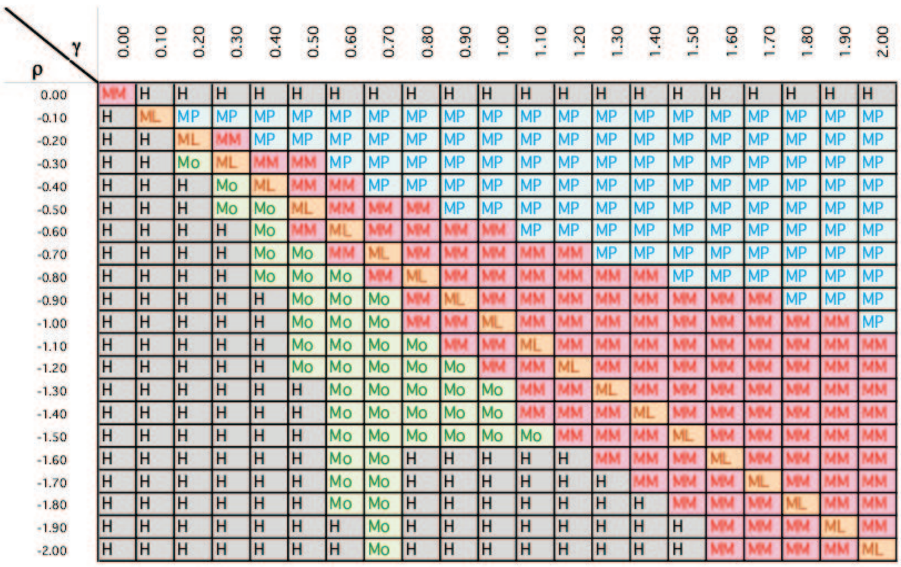
ρ \ γ	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	1.90	2.00
0.00	MM	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
-0.10	H	ML	MM	MM	MM	MM	MM	MM	MM	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.20	H	H	ML	MM	MM	MM	MM	MM	MM	MM	MM	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.30	H	H	Mo	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	Mo	Mo	Mo	Mo	Mo	Mo	Mo	Mo
-0.40	H	H	H	Mo	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	Mo	Mo	Mo	Mo	Mo	Mo
-0.50	H	H	H	Mo	Mo	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	Mo	Mo	Mo	Mo
-0.60	H	H	H	H	Mo	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	Mo	Mo
-0.70	H	H	H	H	Mo	Mo	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-0.80	H	H	H	H	Mo	Mo	Mo	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-0.90	H	H	H	H	H	Mo	Mo	Mo	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.00	H	H	H	H	H	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.10	H	H	H	H	H	Mo	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.20	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM
-1.30	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM
-1.40	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM
-1.50	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	MM	MM	MM	ML	MM	MM	MM	MM	MM
-1.60	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	MM	MM	MM	ML	MM	MM	MM	MM
-1.70	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	MM	MM	MM	ML	MM	MM	MM
-1.80	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	MM	MM	MM	ML	MM	MM
-1.90	H	H	H	H	H	H	H	Mo	H	H	H	H	H	H	H	H	MM	MM	MM	ML	MM
-2.00	H	H	H	H	H	H	H	Mo	H	H	H	H	H	H	H	H	MM	MM	MM	MM	ML

Figure 3. The MM ($\equiv$ML, unless $\gamma + \rho = 0$, $(\gamma, \rho) \neq (0, 0)$) and the ML can outperform the Hill and/or the moment estimators at optimal levels.



ρ \ γ	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	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM
-0.10	MM	ML	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.20	MM	MM	ML	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.30	MM	MP	MM	ML	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.40	MM	MP	MM	MM	ML	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.50	MM	MP	MP	MM	MM	ML	MM	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.60	MM	MP	MP	MM	MM	MM	ML	MM	MM	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.70	MM	MP	MP	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP
-0.80	MM	MP	MP	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM	MP	MP	MP	MP	MP	MP
-0.90	MM	MP	MP	MP	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MP	MP	MP
-1.00	MM	MP	MP	MP	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MP
-1.10	MM	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.20	MM	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM
-1.30	MM	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM
-1.40	MM	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM
-1.50	MM	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM	MM
-1.60	MM	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM	MM
-1.70	MM	MP	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM	MM
-1.80	MM	MP	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM	MM
-1.90	MM	MP	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML	MM
-2.00	MM	MP	MP	MP	MP	MP	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	MM	ML

Figure 4. The PORT-MP can outperform the MM estimator at optimal levels.



ρ \ γ	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	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.20	H	H	ML	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.30	H	H	Mo	ML	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.40	H	H	H	Mo	ML	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.50	H	H	H	Mo	Mo	ML	MM	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.60	H	H	H	H	Mo	MM	ML	MM	MM	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP	MP	MP
-0.70	H	H	H	H	Mo	Mo	MM	ML	MM	MM	MM	MM	MM	MP	MP	MP	MP	MP	MP	MP	MP
-0.80	H	H	H	H	Mo	Mo	Mo	MM	ML	MM	MM	MM	MM	MM	MM	MP	MP	MP	MP	MP	MP
-0.90	H	H	H	H	H	Mo	Mo	Mo	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MP	MP	MP
-1.00	H	H	H	H	H	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM	MP
-1.10	H	H	H	H	H	Mo	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM	MM
-1.20	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM	MM
-1.30	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	MM	MM	ML	MM	MM	MM	MM	MM	MM	MM
-1.40	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	MM	MM	MM	ML	MM	MM	MM	MM	MM	MM
-1.50	H	H	H	H	H	H	Mo	Mo	Mo	Mo	Mo	Mo	MM	MM	MM	ML	MM	MM	MM	MM	MM
-1.60	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	MM	MM	MM	ML	MM	MM	MM	MM
-1.70	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	MM	MM	MM	ML	MM	MM	MM
-1.80	H	H	H	H	H	H	Mo	Mo	H	H	H	H	H	H	H	MM	MM	MM	ML	MM	MM
-1.90	H	H	H	H	H	H	H	Mo	H	H	H	H	H	H	H	H	MM	MM	MM	ML	MM
-2.00	H	H	H	H	H	H	H	Mo	H	H	H	H	H	H	H	H	MM	MM	MM	MM	ML

Figure 5. Comparative overall behaviour of the classical tail index estimators under comparison.

Finally, we enhance the fact that in the region $\gamma+\rho \neq 0$ and $\gamma \neq -\rho/(1-\rho)$ the WH-estimators in (1.11), as well as the CH-estimators in (1.12), overpass all other classical estimators under consideration. The region $\gamma+\rho=0$ (where $b_{ML}=b_{MM}=b_{WH}=b_{CH}=0$) as well as the region $\gamma=-\rho/(1-\rho)$, (where $b_M=b_{GH}=b_{WH}=b_{CH}=0$) are "technically difficult" to handle and deserve further attention. The MM, the ML, the WH and the CH estimators, in (1.7), (1.9), (1.11) and (1.12), respectively, are all second-order reduced-bias estimators in the region $\gamma+\rho=0$. The MM and the ML estimators have an asymptotic variance equal to $\gamma^2+1>\gamma^2$, the asymptotic variance of WH and CH. However, 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. A similar comment applies to the behaviour of the M, the GH, the WH and the CH-estimators in the region $\gamma=-\rho/(1-\rho)$. Again, despite of the fact that the M and the GH estimators have an asymptotic variance equal to $(1+\gamma)^2>\gamma^2$, the asymptotic variance of WH and CH, all depends on the comparative behaviour of the mean squared errors.

References

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

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

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

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

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

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

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

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

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

[10] M.I. Fraga Alves, M.I. Gomes and L. de Haan, *A new class of semi-parametric estimators of the second order parameter*, Portugaliae Mathematica **60** (2) (2003), 193–214.

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

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

[13] B.V. Gnedenko, *Sur la distribution limite du terme maximum d'une série aléatoire*, Ann. Math. **44** (1943), 423–453.

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

[15] M.I. Gomes, L. de Haan and L. Henriques-Rodrigues, *Tail index estimation through accommodation of bias in the weighted log-excesses*, J. Royal Statistical Society **B70** (1) (2008b), 31–52.

[16] M.I. Gomes and L. Henriques-Rodrigues, *Tail index estimation for heavy tails: accommodation of bias in the excesses over a high threshold*, Extremes, **11** (3) (2008), 303–328.

[17] M.I. Gomes and M.J. Martins, *Alternatives to Hill's estimator — asymptotic versus finite sample behaviour*, J. Statist. Planning and Inference **93** (2001), 161–180.

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

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

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

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

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

[23] L. de Haan, *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Amsterdam 1970.

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

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

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

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

[28] V. Pareto, Oeuvres Complétes, Geneva, Droz. 1965.

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

[30] G.K. Zipf, National Unity and Disunity, Blomington, in: Principia Press 1941.

Received 15 January 2010