A comparative study of two classes of reduced bias' estimators under a third order framework*

Frederico Caeiro

Universidade Nova de Lisboa

M. Ivette Gomes

Universidade de Lisboa, D.E.I.O. and C.E.A.U.L.

and

Lígia Rodrigues

Instituto Politécnico de Tomar and C.E.A.U.L.

June 18, 2005

Abstract. In this paper we are interested in the comparison, under a third order framework, of two classes of reduced bias tail index estimators, recently introduced in the literature. One of those classes is an adequate weighted linear combination of the log-excesses and the other one comes from a direct removal of the estimated dominant component of asymptotic bias of the Hill estimator, the classical tail index estimator for heavy-tailed models. In these classes, the second order parameters in the bias are estimated at a level k_1 of a larger order than that of the level k at which we compute the tail index estimators. Doing this, we are able to keep the asymptotic variance

*Research partially supported by FCT / POCTI / FEDER.

of the new estimators equal to the asymptotic variance of the Hill estimator. The asymptotic distributional properties of the proposed classes of γ-estimators are derived under a third order framework and the estimators are compared with other alternative estimators of γ, not only asymptotically, but also for finite samples through Monte Carlo techniques. An application to the log-exchange rates of the Euro against the USA Dollar is also provided.

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

Keywords and phrases. *Statistics of Extremes*, *Semi-parametric estimation*, *Bias estimation*, *Heavy tails*, *Weighted log-excesses*, *Scaled log-spacings*.

1 The estimators under study and scope of the paper

Let $X_1, X_2, \cdots, X_n$ be independent, identically distributed (i.i.d.) random variables (r.v.'s) with a common distribution function (d.f.) F. Let us denote the corresponding ascending order statistics (o.s.) by $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ and let us assume that there exist sequences of real constants $\{a_n > 0\}$ and $\{b_n \in \mathbb{R}\}$ such that the maximum, linearly normalized, i.e., $(X_{n:n} - b_n)/a_n$, converges in distribution towards a non-degenerate r.v. Then the limit distribution is necessarily an *Extreme Value* d.f., EV_γ, with the functional form

$$EV_\gamma(x) = \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}), \ 1+\gamma x \geq 0 & \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)), & \text{if} \quad \gamma = 0 \end{cases}.$$

The d.f. F is said to belong to the max-domain of attraction of EV_γ, and we write $F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma)$. The parameter γ is the *tail index*, the primary parameter of extreme events. The tail index measures the heaviness of the right tail function $\overline{F} := 1 - F$, and the heavier the tail, the larger the tail index is. In this paper we shall work with Pareto-type distributions, with a strict positive tail index.

1.1 First, second and third order conditions for heavy tails

Heavy-tailed models are quite useful in several areas of application, like computer science, telecommunication networks, insurance and finance. In *Statistics of Extremes*, a model F is said to be *heavy-tailed* whenever the *tail function* $\overline{F}$ is a regularly varying function with a negative index of regular variation equal to $\{-1/\gamma\}$, $\gamma > 0$, or equivalently, the quantile function $U(t) = F^{\leftarrow}(1-1/t),\ t \geq 1$, with $F^{\leftarrow}(x) = \inf\{y : F(y) \geq x\}$, is of regular variation with index γ, i.e.,

$$F \in \mathcal{D}_{\mathcal{M}}\left(EV_{\gamma}\right),\ \gamma > 0 \iff \lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma} \iff \lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\gamma}, \qquad (1.1)$$

for all $x > 0$.

The *second order parameter*, ρ (≤ 0), rules the rate of convergence in the first order condition (1.1), and is the non-positive parameter appearing in the limiting relation

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^{\rho}-1}{\rho}, \qquad (1.2)$$

which we assume to hold for every $x > 0$, and where $|A(t)|$ must then be of regular variation with index ρ (Geluk and de Haan, 1987). We shall assume everywhere in the paper that $\rho < 0$.

In order to obtain information on the asymptotic bias of the "asymptotically unbiased" second order tail index estimators herewith considered, we shall further assume a third order condition, ruling now the rate of convergence in (1.2), and which guarantees that, for all $x > 0$,

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx)-\ln U(t)-\gamma \ln x}{A(t)} - \frac{x^{\rho}-1}{\rho}}{B(t)} = \frac{x^{\rho+\rho'}-1}{\rho+\rho'}, \qquad (1.3)$$

where $|B(t)|$ must then be of regular variation with index ρ'. There appears then this extra third order parameter $\rho' \leq 0$, which we also assume to be negative. Such a condition has already been used in Gomes *et al.* (2002a) and Fraga

Alves *et al.* (2003), for the full derivation of the asymptotic behaviour of the ρ-estimators therewith developed, and in Gomes *el al.* (2004a), for a specific reduced bias' tail index estimator.

Remark 1.1. *For Hall's class of Pareto-type models* (*Hall, 1982; Hall and Welsh, 1985*)*, with a tail function*

$$1 - F(x) = Cx^{-1/\gamma}\left(1 + Dx^{\rho/\gamma} + o\left(x^{\rho/\gamma}\right)\right), \quad as \quad x \to \infty,$$

$C > 0$*,* $D \neq 0$*,* $\rho < 0$*,* (1.2) *holds and we may choose* $A(t) = \alpha\, t^{\rho}$*, for an adequate* α*. If we further specify the term* $\left\{o\left(x^{\rho/\gamma}\right)\right\}$ *and consider a Pareto-type class of models with a tail function*

$$1 - F(x) = Cx^{-1/\gamma}\left(1 + D_1 x^{\rho/\gamma} + D_2 x^{(\rho+\rho')/\gamma} + o\left(x^{(\rho+\rho')/\gamma}\right)\right), \tag{1.4}$$

as $x \to \infty$*, with* $C > 0$*,* $D_1,\ D_2 \neq 0$*,* $\rho,\ \rho' < 0$*,* (1.3) *holds and we may choose* $A(t) = \alpha\, t^{\rho}$*,* $B(t) = \alpha'\, t^{\rho'}$*, for adequate* α *and* α'*.*

Remark 1.2. *Note that for most of the common heavy-tailed models, the third order parameter* ρ' *in* (1.3) *is equal to the second order parameter* ρ *in* (1.2)*. Among those models we mention:*

- *the Fréchet model, with distribution function* (*d.f.*) $F(x) = \exp(-x^{-1/\gamma})$*,* $x \geq 0$*,* $\gamma > 0$*, for which* $\rho' = \rho = -1$*;*
- *the Generalized Pareto* (*GP*) *model, with d.f.* $F(x) = 1 - (1 + \gamma x)^{-1/\gamma}$*,* $x \geq 0$*,* $\gamma > 0$*, for which* $\rho' = \rho = -\gamma$*;*
- *the Burr model, with d.f.* $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$*,* $x \geq 0$*,* $\gamma > 0$*,* $\rho' = \rho < 0$*;*
- *the Student's* t_ν*-model with* ν *degrees of freedom, with a probability density function* (*p.d.f.*)

  $$f_{t_\nu}(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\ \Gamma(\nu/2))}\left[1 + \frac{t^2}{\nu}\right]^{-(\nu+1)/2}, \ t \in \mathbb{R} \quad (\nu > 0),$$

  for which $\gamma = 1/\nu$ *and* $\rho' = \rho = -2/\nu$*.*

In this paper, we shall assume that we are in the class of models in (1.4). Consequently, we may choose

$$A(t) = \alpha\ t^{\rho} =: \gamma\ \beta\ t^{\rho},\ B(t) = \alpha'\ t^{\rho'} =: \beta'\ t^{\rho'}, \quad \beta,\ \beta' \neq 0, \quad \rho,\ \rho' < 0. \quad (1.5)$$

1.2 The estimators under study

For heavy-tailed models, the classical tail index estimator is Hill's estimator (Hill, 1975), the average of the log-excesses or of the scaled log-spacings. We have

$$H_n(k) \equiv H(k) = \frac{1}{k}\sum_{i=1}^{k} V_{ik} = \frac{1}{k}\sum_{i=1}^{k} U_i, \qquad (1.6)$$

where

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \le i \le k < n, \qquad (1.7)$$

are the log-excesses, and

$$U_i := i\left\{\ln X_{n-i+1:n} - \ln X_{n-i:n}\right\}, \quad 1 \le i \le k < n, \qquad (1.8)$$

are the scaled log-spacings.

For intermediate k, i.e., a sequence of integers $k = k_n$, $1 \le k < n$, such that

$$k = k_n \to \infty, \quad k_n = o(n),\ \text{as } n \to \infty, \qquad (1.9)$$

it is well known that the log-excesses $V_{ik},\ 1 \le i \le k$, in (1.7), are approximately the k o.s.'s from an i.i.d. exponential sample of size k, with mean value γ and that the scaled log-spacings U_i, $1 \le i \le k$, in (1.8), are approximately independent and exponential with mean value γ (see, for instance, Caeiro *et al.* (2004), for the sketch of a proof). The Hill estimator in (1.6) is thus consistent for the estimation of γ whenever (1.1) holds and k is intermediate, i.e., (1.9) holds. Under the second order framework in (1.2) the asymptotic distributional representation

$$H_n(k) \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ Z_k^{(1)} + \frac{1}{1-\rho}\ A(n/k)(1 + o_p(1))$$

holds true (de Haan and Peng, 1998), where $Z_k^{(1)} = \sqrt{k}\left(\sum_{i=1}^k E_i/k - 1\right)$, with $\{E_i\}$ a sequence of i.i.d. standard exponential r.v.'s, is an asymptotically standard normal r.v. If we further assume to be under the third order framework in (1.3), we may write,

$$H_n(k) \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\, Z_k^{(1)} + \left(\frac{A(n/k)}{1-\rho} + \frac{A(n/k)\ B(n/k)}{1-\rho-\rho'}\right)(1+o_p(1)).$$

The adequate accommodation of the bias of Hill's estimator has been extensively addressed in recent years by several authors. Beirlant *et al.* (1999) and Feuerverger and Hall (1999) consider exponential regression techniques, based on the exponential approximations $U_i \approx \gamma\left\{1 + b(n/k)(k/i)^\rho\right\} E_i$ and $U_i \approx \gamma \exp\left\{\beta\ (n/i)^\rho\right\} E_i$, respectively, $1 \le i \le k$. They then proceed to the joint estimation of the three unknown parameters or functionals at the same level k. Considering also the scaled log-spacings U_i in (1.8) to be approximately exponential with mean value $\mu_i = \gamma\ \exp\{\beta(n/i)^\rho\}$, $1 \le i \le k$, $\beta \ne 0$, Gomes and Martins (2004) consider the misspecification of ρ at $\rho = -1$ and Gomes and Martins (2002) advance with the external estimation of the second order parameter ρ, together with a first order approximation for the maximum likelihood β-estimator, and then obtain explicit estimators of γ. Related work may be found in Peng (1998), Gomes *et al.* (2000, 2002b), Gomes and Martins (2001), Caeiro and Gomes (2002), Beirlant *et al.* (2002), Gomes *et al.* (2003), Gomes *et al.* (2005), among others. All these methodologies lead to second order reduced bias tail index estimators, with asymptotic variances always larger or equal to $(\gamma\ (1-\rho)/\rho)^2$, the minimal asymptotic variance of an "asymptotically unbiased" estimator in Drees' class of functionals (Drees, 1998).

Recently, Gomes *et al.* (2004*b*), Gomes and Pestana (2004) and Caeiro *et al.* (2004) consider, in different ways and under the second order framework in (1.2), the joint external estimation of both the "scale" and the "shape" parameter in the A function in (1.5), being able to reduce the bias without increasing the asymptotic variance, which is kept at the value γ^2, the asymptotic variance of Hill's estimator. Gomes *et al.* (2004*b*) consider a tail index estimator based on a linear combination of the log-excesses V_{ik} in (1.7), and given by

$$WH_{\widehat{\beta},\,\widehat{\rho}}(k) := \frac{1}{k}\sum_{i=1}^{k} e^{\widehat{\beta}\ (n/k)^{\widehat{\rho}}\ \left((i/k)^{-\widehat{\rho}}-1\right)/(\widehat{\rho}\ln(i/k))}\ V_{ik}, \tag{1.10}$$

WH standing here for *Weighted Hill* estimator. Caeiro *et al.* (2004) consider an estimator of this same type, denoted $\widetilde{\overline{H}}$, but here denoted

$$\overline{H}_{\widehat{\beta},\,\widehat{\rho}}(k) := H(k)\left(1-\frac{\widehat{\beta}}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right), \tag{1.11}$$

for sake of notation's simplicity, where the dominant component of the bias of Hill's estimator, given by $A(n/k)/(1-\rho) = \gamma\ \beta\ (n/k)^{\rho}\ /(1-\rho)$, is thus estimated through $H(k)\ \widehat{\beta}\,(n/k)^{\widehat{\rho}}\ /(1-\widehat{\rho})$, and directly removed from Hill's classical tail index estimator. As before, and both in (1.10) and (1.11), $\widehat{\beta}$ and $\widehat{\rho}$ are adequate consistent estimators of the second order parameters β and ρ, respectively.

1.3 Scope of the paper

In this paper, we intend to derive the asymptotic distributional properties of the two classes of estimators in (1.10) and (1.11), under the third order framework in (1.3), obtaining then information on their asymptotic bias. In section 2, assuming that only γ is unknown, we shall state a theorem that provides an obvious technical motivation for the estimators in (1.10) and (1.11). In section 3, we shall briefly review the estimation of the two second order parameters β and ρ, making explicit the asymptotic bias of the ρ-estimator, and adding some new results related to the estimation of the second order parameter β, under a third order framework. Section 4 is devoted to the derivation of the asymptotic behaviour of the *"Unbiased Hill"* (UH) estimators, $WH_{\widehat{\beta},\,\widehat{\rho}}(k)$ and $\overline{H}_{\widehat{\beta},\,\widehat{\rho}}(k)$, in (1.10) and (1.11), respectively, estimating β and ρ at a larger k value than the one used for the tail index estimation. We also do that only with the estimation of ρ, estimating β at the same level k used for the tail index estimation. In section 5, and through the use of Monte Carlo simulation techniques, we shall exhibit the performance of these *"Unbiased Hill"* estimators, comparatively to the classical Hill estimator and to the *Generalized Jackknife* estimator, studied in Gomes and Martins (2002), with the functional expression

$$GJ_{\widehat{\rho}}(k) := \frac{1}{\widehat{\rho}} \left(\sqrt{2\ M_n^{(2)}(k)} - (2 - \widehat{\rho})\, \frac{M_n^{(2)}(k)}{2\ M_n^{(1)}(k)} \right), \tag{1.12}$$

where

$$M_n^{(j)}(k) := \frac{1}{k} \sum_{i=1}^{k} V_{ik}^{j}, \quad j \geq 1 \quad \left[M_n^{(1)} \equiv H \text{ in } (1.6) \right], \tag{1.13}$$

with V_{ik} the log-excesses in (1.7).

This *Generalized Jackknife* estimator was here considered for comparison, and as a representative of the "asymptotically unbiased" tail index estimators with an asymptotic variance equal to $\gamma^2\,((1-\rho)/\rho)^2$, a value which is, as mentioned before, the minimal asymptotic variance of any "asymptotically unbiased" estimator in Drees' class of functionals. We shall consider only an external estimation of the second order parameter ρ at a level k_1 of a larger order than that of the level k on which we base the tail index estimation. Such a decision is related to the discussion in Gomes and Martins (2002) on the advantages of an external estimation of the second order parameter ρ, or even their misspecification, as in Gomes and Martins (2004), versus an internal estimation at the same level k. In section 6 we provide an illustration of the behaviour of these new reduced bias' estimators through the analysis of the exchange rate of the Euro against the USA Dollar, and some overall conclusions are drawn. Finally, in section 7, we shall provide the proofs of the main results in sections 2, 3 and 4.

2 The asymptotic behaviour of the Unbiased Hill estimators (known β and ρ)

Whenever there is no distinction between the *"Unbiased Hill"* estimators in (1.10) and (1.11), or the corresponding r.v.'s, we shall use UH, generically denoting either WH or $\overline{H}$.

For real values $\alpha \geq 1$, and denoting $\{E_i\}$ a sequence of i.i.d. standard exponential r.v.'s, let us introduce the notations:

$$Z_k^{(\alpha)} := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{\alpha-1} E_i, \qquad \overline{Z}_k^{(\alpha)} := \sqrt{(2\alpha-1)\, k}\,\left(Z_k^{(\alpha)} - \frac{1}{\alpha}\right). \tag{2.1}$$

If we assume that only the tail index parameter γ in unknown:

Theorem 2.1. *Under the second order framework in* (1.2), *further assuming that $A(t)$ may be chosen as in* (1.5), *and for levels k such that* (1.9) *holds, we get, for the "Unbiased Hill" random variables $WH_{\beta,\,\rho}(k)$ and $\overline{H}_{\beta,\,\rho}(k)$, generically denoted $UH_{\beta,\,\rho}$, an asymptotic distributional representation of the type*

$$UH_{\beta,\,\rho}(k) \;\overset{d}{=}\; \gamma + \frac{\gamma}{\sqrt{k}}\,\overline{Z}_k^{(1)} + o_p(A(n/k),$$

where $\overline{Z}_k^{(1)}$ is the asymptotically standard normal r.v. in (2.1) *for $\alpha = 1$. Consequently, $\sqrt{k}\ (UH_{\beta,\,\rho}(k) - \gamma)$ are asymptotically normal with variance equal to γ^2, and with a null mean value not only when $\sqrt{k}\ A(n/k) \longrightarrow 0$, but also when $\sqrt{k}\ A(n/k) \longrightarrow \lambda \neq 0$, finite, as $n \to \infty$. Under the third order framework in* (1.3) *we may further specify the term $o_p(A(n/k))$, writing*

$$\begin{aligned} WH_{\beta,\,\rho}(k) &\overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\,\overline{Z}_k^{(1)} \\ &+ A(n/k)\left(\frac{B(n/k)}{1-\rho-\rho'} - \frac{a_2(\rho)\ A(n/k)}{2\ \gamma} + O_p\left(\frac{1}{\sqrt{k}}\right)\right)(1+o_p(1)), \end{aligned} \tag{2.2}$$

with $a_2(\rho)$ given by

$$a_2(\rho) = -\frac{\ln(1-2\rho) - 2\ln(1-\rho)}{\rho^2}$$

and

$$\begin{aligned}\overline{H}_{\beta,\ \rho}(k) &\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ Z_k^{(1)} \\ &+ A(n/k)\left(\frac{B(n/k)}{1-\rho-\rho'} - \frac{A(n/k)}{\gamma(1-\rho)^2} + O_p\left(\frac{1}{\sqrt{k}}\right)\right)(1+o_p(1)). \quad (2.3)\end{aligned}$$

Consequently, even if $\sqrt{k}\ A(n/k) \rightarrow \infty$, with $\sqrt{k}\ A^2(n/k) \rightarrow \lambda_{_A}$ and $\sqrt{k}\ A(n/k)\ B(n/k) \rightarrow \lambda_{_B}$, $\lambda_{_A}$ and $\lambda_{_B}$ finite, $\sqrt{k}\ (UH_{\beta,\ \rho}(k) - \gamma)$ are asymptotically normal with variance equal to γ^2. The asymptotic bias of $WH_{\beta,\ \rho}(k)$ is

$$b_{_{WH}}(\gamma,\rho,\rho') \equiv ABIAS_{_{WH}} := \frac{\lambda_{_B}}{1-\rho-\rho'} - \frac{\lambda_{_A} a_2(\rho)}{2\ \gamma} \qquad (2.4)$$

and the one of $\overline{H}_{\beta,\ \rho}(k)$ is given by

$$b_{_{\overline{H}}}(\gamma,\rho,\rho') \equiv ABIAS_{_{\overline{H}}} = ABIAS_{_{\overline{\overline{H}}}} := \frac{\lambda_{_B}}{1-\rho-\rho'} - \frac{\lambda_{_A}}{\gamma(1-\rho)^2}. \qquad (2.5)$$

Remark 2.1. *Note that since $\lambda_{_A} \geq 0$ and $2/a_2(\rho) > (1-\rho)^2$ for any ρ, $ABIAS_{_{WH}} \geq ABIAS_{_{\overline{H}}}$. All depends then on the sign of the bias, but we expect the sample paths of WH to be always above the sample paths of $\overline{H}$.*

3 Further results on the second order parameter estimation, under a third order framework

3.1 The estimation of ρ

We have nowadays two general classes of ρ-estimators, which work well in practice, the ones introduced in Gomes *et al.* (2002a) and Fraga Alves *et al.* (2003). We shall consider here particular members of 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 highly stable sample paths as functions of k, the number of top o.s.'s used, for a wide range of large k-values. Such a class of estimators has been first parameterised in a tuning parameter $\tau > 0$, but τ may be more generally considered as a real number (Caeiro and Gomes, 2004), and is defined as,

$$\widehat{\rho}_\tau(k) \equiv \widehat{\rho}_n^{(\tau)}(k) := -\left| \frac{3(T_n^{(\tau)}(k) - 1)}{T_n^{(\tau)}(k) - 3} \right|, \tag{3.1}$$

where

$$T_n^{(\tau)}(k) := \begin{cases} \dfrac{\left(M_n^{(1)}(k)\right)^\tau - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}} & \text{if} \quad \tau \neq 0 \\ \\ \dfrac{\ln\left(M_n^{(1)}(k)\right) - \frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right)}{\frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right) - \frac{1}{3}\ln\left(M_n^{(3)}(k)/6\right)} & \text{if} \quad \tau = 0 \end{cases},$$

with $M_n^{(j)}(k)$ given in (1.13).

We shall here summarize a particular case of the results proved in Fraga Alves *et al.* (2003), making however explicit both the asymptotic bias and variance of this type of ρ-estimators:

Proposition 3.1. *Under the second order framework in* (1.2)*, with* $\rho < 0$*, if* (1.9) *holds, and if* $\sqrt{k}\ A(n/k) \to \infty$*, as* $n \to \infty$*, the statistics* $\widehat{\rho}_n^{(\tau)}(k)$ *in* (3.1) *converge in probability towards* ρ*, as* $n \to \infty$*, for any real* $\tau \neq 0$*. Moreover, under the third order framework in* (1.3)*, if* $\sqrt{k}\ A^2(n/k) \longrightarrow \lambda_{_A}$*, finite, and* $\sqrt{k}\ A(n/k)\ B(n/k) \longrightarrow \lambda_{_B}$*, also finite,* $\sqrt{k}\ A(n/k)\left(\widehat{\rho}_n^{(\tau)}(k) - \rho\right)$ *are asymptotically normal with asymptotic variance*

$$\sigma_\rho^2 \equiv \sigma_\rho^2(\gamma) = \left(\frac{\gamma(1-\rho)^3}{\rho}\right)^2 \left(2\rho^2 - 2\rho + 1\right), \tag{3.2}$$

and a possibly non-null asymptotic bias given by $\{\lambda_{_A} u_\rho + \lambda_{_B} v_\rho\}$*, where*

$$
\begin{aligned}
u_\rho &\equiv u_\rho(\gamma;\tau) \\
&= \frac{\rho\left(\tau(1-2\rho)^2(3-\rho)(3-2\rho) - 6\rho\left(4\rho^3 - 16\rho^2 + 20\rho - 7\right)\right)}{12\,\gamma\left((1-\rho)(1-2\rho)\right)^2}
\end{aligned}
\tag{3.3}
$$

and

$$
v_\rho \equiv v_\rho(\rho') = \rho'\left(1 + \frac{\rho'}{\rho}\right)\left(\frac{1-\rho}{1-\rho-\rho'}\right)^3. \tag{3.4}
$$

Remark 3.1. *The theoretical and simulated results in Fraga Alves et al.* (2003)*, together with the use of these estimators in the Generalized Jackknife statistics of Gomes et al.* (2000)*, as done in Gomes and Martins* (2002)*, as well as their use in the estimators in* (1.10) (*Gomes et al.,* 2004*b*) *and in* (1.11) (*Gomes and Pestana,* 2004*; Caeiro et al.,* 2004)*, lead us again to advise in practice the consideration of the intermediate level*

$$
k_1 = \min\left(n-1, [2n/\ln\ln n]\right) \tag{3.5}
$$

(*not chosen in any optimal way, but under the conditions of Proposition 3.1*)*, and of the tuning parameters* $\tau = 0$ *for the region* $\rho \in [-1, 0)$ *and* $\tau = 1$ *for the region* $\rho \in (-\infty, -1)$*. We however think that practitioners should not choose blindly the value of* τ *in* (3.1)*. It is sensible to draw a few sample paths of* $\widehat{\rho}_\tau(k)$*, as functions of* k*, electing the value of* τ *which provides higher stability for large* k*, by means of any stability criterion, like the ones suggested in Gomes et al.* (2003) *and Gomes and Pestana* (2004)*. For not too small values of* n*, we are most frequently led to the above mentioned choice:* $\widehat{\rho}_0$ *if* $\rho \geq -1$ *and* $\widehat{\rho}_1$ *if* $\rho < -1$*, when we consider only the tuning parameters* $\tau = 0$ *and* $\tau = 1$ *as the possible alternatives.*

Remark 3.2. *When we consider the level* k_1 *in* (3.5)*, together with any of the* ρ*-estimators in this section, computed at the level* k_1*,* $\{\widehat{\rho} - \rho\}$ *is of the order of*

$1/\left(\sqrt{k_1}A(n/k_1)\right) = O\left((\ln_2 n)^{(1-2\rho)/2}/\sqrt{n}\right)$, *with the obvious notation* $\ln_2 n = \ln\ln n$.

Remark 3.3. *Moreover, for any level* k, $(\widehat{\rho}-\rho)\ln(n/k) = o_p(1)$, *and consequently* $\sqrt{k}\ A(n/k)\,(\widehat{\rho}-\rho)\ln(n/k) = o_p(1)$ *whenever* $\sqrt{k}\ A(n/k) \to \lambda$, *finite.*

3.2 Estimation of β based on the scaled log-spacings

We have here considered the β-estimator obtained in Gomes and Martins (2002) and based on the scaled log-spacings U_i in (1.8), $1 \le i \le k$. Let us denote $\widehat{\rho}$ any of the estimators in (3.1) computed at the level k_1 in (3.5). The β-estimator is given by

$$\widehat{\beta}_{\widehat{\rho}}(k) := \left(\frac{k}{n}\right)^{\widehat{\rho}} \frac{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\left(\frac{1}{k}\sum_{i=1}^{k}U_i\right) - \left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\ U_i\right)}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\ U_i\right) - \left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-2\widehat{\rho}}\ U_i\right)}, \tag{3.6}$$

and since the denominator in this expression converges in probability towards

$$\frac{\gamma}{(1-\rho)^2} - \frac{\gamma}{1-2\rho} = -\frac{\gamma\ \rho^2}{(1-\rho)^2(1-2\rho)},$$

$\widehat{\beta}_{\widehat{\rho}}(k)$ is asymptotically equivalent to

$$-\frac{(1-\rho)^2(1-2\rho)}{\gamma\ \rho^2}\left(\frac{k}{n}\right)^{\widehat{\rho}}\left\{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\left(\frac{1}{k}\sum_{i=1}^{k}U_i\right) - \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\ U_i\right\}.$$

Gomes and Martins (2002), keeping up to the second order framework in (1.2), have proved the following result:

Proposition 3.2. *If the second order condition* (1.2) *holds, with* $A(t) = \gamma\ \beta\ t^{\rho}$, $\rho < 0$, *if* $k = k_n$ *is a sequence of intermediate positive integers, i.e.* (1.9) *holds,* *and if* $\sqrt{k}\ A(n/k) \underset{n\to\infty}{\longrightarrow} \infty$, *then* $\widehat{\beta}_{\widehat{\rho}}(k)$ *in* (3.6) *is consistent for the estimation of* β, *whenever* $\widehat{\rho}$ *is consistent for the estimation of* ρ. *Moreover, if* ρ *is known,*

$$\widehat{\beta}_{\rho}(k) \stackrel{d}{=} \beta + \frac{\gamma\ \beta\ (1-\rho)\ \sqrt{1-2\rho}}{\rho\ \sqrt{k}\ A(n/k)}\ W_k^{B}\ + R_k^{B},\ \text{with } R_k^{B} = o_p(1), \tag{3.7}$$

where W_k^{B} *is asymptotically standard normal. More precisely we may write*

$$W_k^{B} = \frac{(1-\rho)\sqrt{1-2\rho}}{|\rho|}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho} - \frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right), \tag{3.8}$$

with $\overline{Z}_k^{(\alpha)}$, $\alpha \geq 1$, *given in* (2.1).

The asymptotic distributional representation (3.7) *holds true as well for* $\widehat{\beta}_{\widehat{\rho}}(k)$, *with* $\widehat{\rho}$ *any of the estimators in* (3.1) *computed at the level* k_1 *in* (3.5).

In order to be able to specify the order of R_k^{B}, in (3.7), we need to go into the third order framework, and we then get the following result:

Theorem 3.1. *If we go further into the third order framework in* (1.3), *with the same conditions on* k *as in Proposition 3.2, we get the distributional representation*

$$\begin{aligned}\widehat{\beta}_{\rho}(k)\ &\stackrel{d}{=}\ \beta + \frac{\gamma\ \beta(1-\rho)\sqrt{(1-2\rho)}}{\rho\ \sqrt{k}\ A(n/k)}\ W_k^{B} \\ &+ \left(\frac{\beta(1-\rho)(1-2\rho)(\rho+\rho')B(n/k)}{\rho(1-\rho-\rho')(1-2\rho-\rho')} - \frac{2\beta(1-\rho)A(n/k)}{\gamma(1-3\rho)}\right)(1+o_p(1)),\end{aligned} \tag{3.9}$$

with W_k^{B} *in* (3.8). *Consequently,* R_k^{B} *in* (3.7) *is* $\{O_p(A(n/k)) + O_p(B(n/k))\}$ *and, as stated before in Proposition 3.2,* $\widehat{\beta}_{\rho}(k)$ *converges in probability towards* β, *provided that* $\sqrt{k}\ A(n/k) \to \infty$, *as* $n \to \infty$. *Moreover, if* $\sqrt{k}\ A^2(n/k) \to \lambda_{A}$ *and* $\sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_{B}$, *both finite,*

$$\sqrt{k}\ A(n/k)\left(\frac{\widehat{\beta}_{\rho}(k)-\beta}{\beta}\right)\ \stackrel{a}{\sim}\ Normal\left(\lambda_{A}\overline{u}_{\beta} + \lambda_{B}\overline{v}_{\beta},\ \overline{\sigma}^2_{\beta}\right), \tag{3.10}$$

where

$$\overline{\sigma}^2_{\beta} \equiv \overline{\sigma}^2_{\beta}(\gamma,\ \rho) = \left(\frac{\gamma\ (1-\rho)}{\rho}\right)^2 (1-2\rho), \tag{3.11}$$

$$\overline{u}_\beta \equiv \overline{u}_\beta(\gamma,\ \rho) = -\frac{2(1-\rho)}{\gamma(1-3\rho)}, \quad \overline{v}_\beta \equiv \overline{v}_\beta(\rho,\ \rho') = \frac{(1-\rho)(1-2\rho)(\rho+\rho')}{\rho(1-\rho-\rho')(1-2\rho-\rho')}. \tag{3.12}$$

These same results hold true if we replace ρ *by* $\widehat{\rho}$*, any of the estimators in* (3.1) *computed at the level* k_1 *in* (3.5).

If we replace ρ *by* $\widehat{\rho}(k)$*, the rate of convergence of* $\widehat{\beta}_{\widehat{\rho}(k)}(k)$ *is of the order of* $\left\{\ln(n/k)/(\sqrt{k}\ A(n/k))\right\}$*, which must converge towards zero, so that* $\widehat{\beta}_{\widehat{\rho}(k)}(k)$ *is consistent for the estimation of* β*, and*

$$\frac{\sqrt{k}\ A(n/k)}{\ln(n/k)}\left(\frac{\beta_{\widehat{\rho}(k)}(k)-\beta}{\beta}\right) \overset{p}{\sim} -\sqrt{k}\ A(n/k)\left(\widehat{\rho}(k)-\rho\right). \tag{3.13}$$

If apart from $\sqrt{k}\ A(n/k)/\ln(n/k) \to \infty$*, we further assume that* $\sqrt{k}\ A^2(n/k) \to \lambda_{_A}$ *and* $\sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_{_B}$*, with* $\lambda_{_A}$ *and* $\lambda_{_B}$ *both finite, then*

$$\frac{\sqrt{k}\ A(n/k)}{\ln(n/k)}\left(\frac{\beta-\beta_{\widehat{\rho}(k)}(k)}{\beta}\right) \overset{a}{\sim} Normal\left(\lambda_{_A} u_\rho + \lambda_{_B} v_\rho,\ \sigma_\rho^2\right)$$

with σ_ρ*,* u_ρ *and* v_ρ *given in* (3.2), (3.3) *and* (3.4)*, respectively.*

Remark 3.4. *If, in* (1.3)*,* $\rho = \rho'$*, as happens with most common heavy-tailed models, we have* $\overline{u}_\beta = -\overline{v}_\beta/\gamma$*. Since for a Burr model, we may choose* $B(t) = A(t)/\gamma$*, we then have a null mean value for* $\sqrt{k}\left(\widehat{\beta}_{\widehat{\rho}}(k)-\beta\right)$*, even when* $\lambda_{_A}$ *and* $\lambda_{_B}$ *are non-null, but finite. This result justifies the good performance of this* β*-estimator for Burr models, as detected in Caeiro and Gomes* (2004).

Remark 3.5. *Note that when we consider the level* k_1 *in* (3.5)*, and* $\widehat{\beta} \equiv \widehat{\beta}_{\widehat{\rho}}(k_1)$*, with* $\widehat{\rho}$ *any of the estimators in* (3.1)*, computed also at the same level* k_1*,* $\widehat{\beta}-\beta$ *is thus, from* (3.13)*, of the order of* $\ln(n/k_1)/\left(\sqrt{k_1}A(n/k_1)\right) = O\left(\ln_3 n\ (\ln_2 n)^{(1-2\rho)/2}/\sqrt{n}\right)$.

4 Asymptotic behaviour of the "*Unbiased Hill*" estimators

Let us assume first that we estimate both β and ρ externally at the level k_1 in (3.5). We may state the following:

Theorem 4.1. *Under the conditions of Theorem 2.1, let us consider the tail index estimators in* (1.10) *and* (1.11)*, generically denoted* $UH_{\widehat{\beta},\,\widehat{\rho}}(k)$*, for any of the estimators* $\widehat{\beta}$ *and* $\widehat{\rho}$ *in* (3.1) *and in* (3.6)*, respectively, both computed at the level* k_1 *in* (3.5)*. Then,* $\sqrt{k}\left\{UH_{\widehat{\beta},\,\widehat{\rho}}(k)-\gamma\right\}$ *are asymptotically normal with null mean value, not only when* $\sqrt{k}\ A(n/k) \rightarrow 0$*, but also whenever* $\sqrt{k}\ A(n/k) \rightarrow \lambda$*, finite or infinite, provided that* $\sqrt{k}\ A^2(n/k) \rightarrow 0$ *and* $\sqrt{k}\ A(n/k)\ B(n/k) \rightarrow 0$*, as* $n \rightarrow \infty$*. If* $\sqrt{k}\ A^2(n/k) \rightarrow \lambda_{_A}$ *and* $\sqrt{k}\ A(n/k)\ B(n/k) \rightarrow \lambda_{_B}$*, both finite, perhaps both non-null* (*or at least one of them non-null*)*, the* WH *estimator in* (1.10) *and the* $\overline{H}$ *estimator in* (1.11) *have a non-null asymptotic bias given by* (2.4) *and* (2.5)*, respectively.*

If we consider γ and β estimated at the same level, we are going to have an increase in the variance of our final tail index estimators. Indeed, in Gomes and Martins (2002) for the "quasi-maximum likelihood" tail index estimator therewith introduced, in Gomes *et al.* (2004b) for the tail index estimator $WH_{\widehat{\beta}_{\widehat{\rho}}(k),\,\widehat{\rho}}$ and in Caeiro *et al.* (2004) for the tail index estimator $\overline{H}_{\widehat{\beta}_{\widehat{\rho}}(k),\,\widehat{\rho}}$, an asymptotic normal behaviour has been obtained, being the rate of convergence still of the order of $1/\sqrt{k}$ and the asympotic variance equal to

$$\sigma_2^2 := \left(\gamma(1-\rho))/\rho\right)^2, \tag{4.1}$$

i.e., the asymptotic variance of these reduced bias tail index estimators increases of a factor $\left((1-\rho)/\rho\right)^2 > 1$ for every $\rho \leq 0$.

More generally, we may now obtain information on the asymptotic bias of $WH_{\widehat{\beta}_{\widehat{\rho}}(k),\,\widehat{\rho}}(k)$ and of $\overline{H}_{\widehat{\beta}_{\widehat{\rho}}(k),\,\widehat{\rho}}(k)$, if we go further into the third order framework:

Theorem 4.2. *If the third order condition* (1.3) *holds, if* $k = k_n$ *is a sequence of intermediate integers, i.e.,* (1.9) *holds, and if* $\sqrt{k}\ A(n/k) \underset{n\to\infty}{\longrightarrow} \infty$*, with* $\sqrt{k}\ A^2(n/k)$ *and* $\sqrt{k}\ A(n/k)\ B(n/k)$ *converging both towards zero, as* $n \to \infty$*,* $\sqrt{k}\left(UH_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k) - \gamma\right)$ *is asymptotically normal, with null mean value and variance equal to* σ_2^2*, in* (4.1)*.*

If $\sqrt{k}\ A^2(n/k)\ \to \lambda_A$ *and* $\sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_B$*, both finite, the asymptotic variances of the* $UH_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k)$ *statistics are kept equal to* $(\gamma(1-\rho)/\rho)^2$*. The asymptotic bias of* $WH_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k)$ *and* $\overline{H}_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k)$ *are given by*

$$\frac{\lambda_A(4-(1-3\rho)a_2(\rho))}{2\gamma(1-3\rho)} - \frac{\lambda_B\ \rho'(1-\rho)}{\rho(1-\rho-\rho')(1-2\rho-\rho')} \tag{4.2}$$

and

$$\frac{\lambda_A(2\rho^2-\rho+1)}{\gamma(1-\rho)^2(1-3\rho)} - \frac{\lambda_B\ \rho'(1-\rho)}{\rho(1-\rho-\rho')(1-2\rho-\rho')}, \tag{4.3}$$

respectively.

Remark 4.1. *As already noticed in Caeiro et al.* (2004)*, if we compare Theorems 4.1 and 4.2, we see that the estimation of the two parameters* γ *and* β *at the same level* k *induces an increase in the asymptotic variance of the final* γ*-estimator of a factor given by* $((1-\rho)/\rho)^2$*, greater than* 1*. As noticed before, for instance in Gomes and Martins* (2002)*, the asymptotic variance of the estimator in Feuerverger and Hall* (1999) (*where the three parameters are computed at the same level* k) *is given by* $\sigma_{FH}^2 := \gamma^2\left((1-\rho)/\rho\right)^4$*. The asymptotic variance of the reduced bias' estimators herewith considered is* $\sigma_1^2 = \gamma^2$*, and we have* $\sigma_1 < \sigma_2 < \sigma_{FH}$ *for all* $\rho \le 0$*.*

5 Finite sample behaviour of the estimators — a simulation experiment

We have here implemented simulation experiments, with 5000 runs, based on the estimation of β at the same level $k_1 = \min(n-1, [2n/\ln_2 n])$ we have used for the estimation of ρ, again not chosen in any optimal way. We use the notation $\widehat{\beta}_{j1} = \beta_{\widehat{\rho}_j}(k_1)$, $j = 0, 1$, with $\widehat{\rho}_j$, $j = 0,\ 1$, and $\beta_{\widehat{\rho}}(k)$ given in (3.1) and (3.6), respectively. Similarly to what has been done in Gomes *et al.* (2004b) for the WH-estimator in (1.10) and in Caeiro *et al.* (2004) for the $\overline{H}$-estimator in (1.11), these estimators of ρ and β have been incorporated in the *"Unbiased Hill"* estimators herewith considered, leading to $WH_{\widehat{\beta}_{j1},\ \widehat{\rho}_j}(k)$ or to $\overline{H}_{\widehat{\beta}_{j1},\ \widehat{\rho}_j}(k)$, $j = 0,\ 1$.

The simulations show that the tail index estimators $WH_{\widehat{\beta}_{j1},\ \widehat{\rho}_j}(k)$ and $\overline{H}_{\widehat{\beta}_{j1},\ \widehat{\rho}_j}(k)$, j equal to either 0 or 1, according as $|\rho| \le 1$ or $|\rho| > 1$, seem to work reasonably well, as illustrated in Figures from 1 till 4. In these figures we picture for Fréchet and Generalized Pareto underlying models, and a sample of size $n = 1000$, the mean values ($E[\bullet]$) and the mean squared errors ($MSE[\bullet]$) of the Hill estimator H and the Generalized Jacknife estimator $GJ_j = GJ_{\widehat{\rho}_j}$ in (1.12), together with $WH_j = WH_{\widehat{\beta}_{j1},\ \widehat{\rho}_j}$ and $\overline{H}_j = \overline{H}_{\widehat{\beta}_{j1},\ \widehat{\rho}_j}$ (*left*), $j = 0$ or $j = 1$, according as $|\rho| \le 1$ or $|\rho| > 1$. The r.v.'s $WH = WH_{\beta,\ \rho}$ are also pictured, so that we may see that, for some models, there exists still a significant difference between the behaviour of the statistics under study and that of the r.v.'s. Such a discrepancy suggests that some improvement in the estimation of second order parameters β and ρ is still welcome.

Remark 5.1. *Note that the comment made in Remark 2.1 is coherent with the pictures of the estimators' mean values.*

Remark 5.2. *For the Fréchet model* (*Figure 1*), *and among the two UH-*

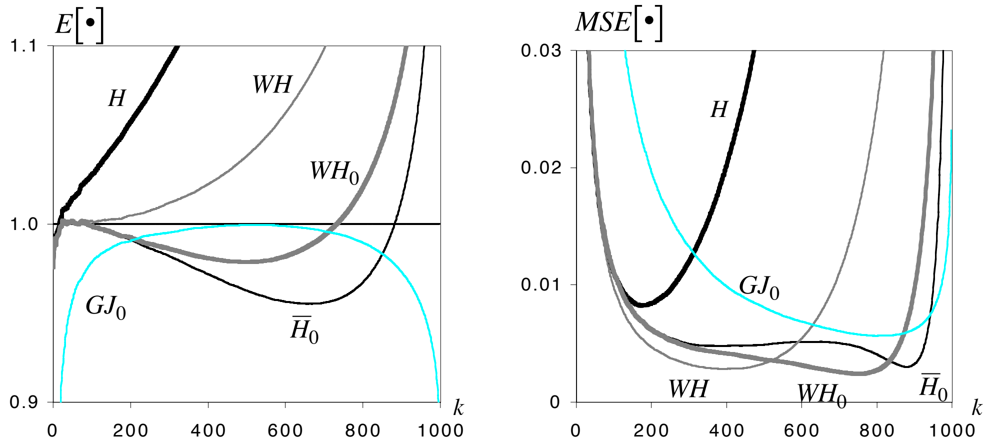

Figure 1: Underlying *Fréchet* parent with $\gamma = 1$ $(\rho = -1)$.

estimators herewith considered, the WH-estimator exhibits the best performance.

Remark 5.3. *For a Generalized Pareto (GP) model, we may further draw the following comments:*

- *For $\rho = -1$ (Figure 2), the $\overline{H}$ statistic is the best one regarding MSE at the optimal level, but merely because the associated sample paths cross the target at a lower threshold. The WH-statistic is however the one with the smallest bias for not too large values of k.*
- *For values of $|\rho| < 1$ (Figure 3), the $\overline{H}$-statistic behaves slightly better than the WH-statistic, but the behaviour of the GJ-statistic is more interesting for this region of ρ-values.*
- *For $|\rho| > 1$ (Figure 4), we need to use $\widehat{\rho}_1$ (instead of $\widehat{\rho}_0$) or any of the hybrid estimators suggested in Gomes and Pestana (2004). In all the simulated cases, the $\overline{H}$ and the WH-statistics are almost equivalent. The GJ statistic, although exhibiting a better performance than the Hill estimator, when both statistics are compared at their optimal levels, is not able to overpass the UH-statistics.*

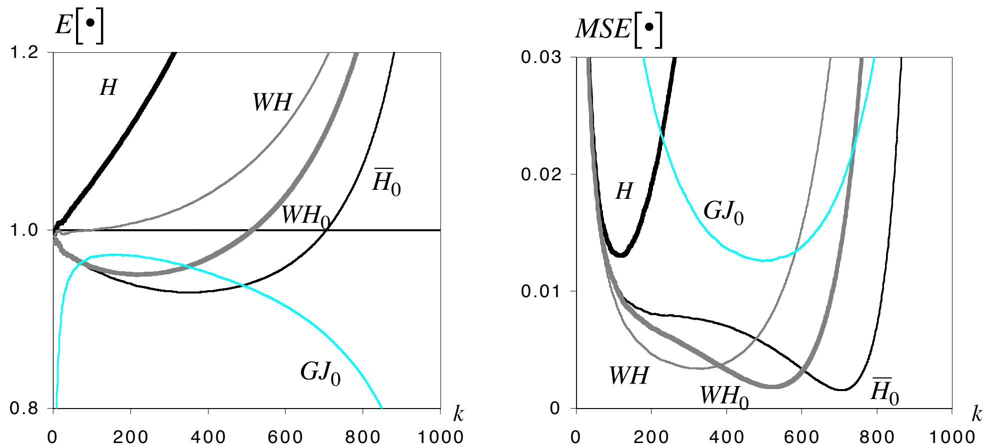

Figure 2: Underlying GP parent with $\gamma = 1$ $(\rho = -1)$.

6 A case-study and overall conclusions

6.1 Log-exchange rates of Euro against USA Dollar

When analysing heavy-tailed data, quite common in financial time series, one never knows how much the underlying model differs from a strict Pareto model. And this is the unique situation where the Hill estimator is "perfect". Otherwise, all depends on the specificity of the underlying heavy-tailed model and on the practitioner's objectives. If we want to use (or have only access to) a very small number of top o.s., the Hill estimator has been considered the most adequate one, among the tail index estimators so far available in the literature. This is no longer true now: these new estimators are similar to Hill's estimator from small to moderate values of k, being much better than the Hill estimator, when we consider larger values of k, although intermediate.

After taking a decision on the estimate of ρ, and assuming that $|\rho| \leq 1$, a situation which seems to appear often in practice, we should simultaneously picture the sample path of a few tail index estimators, with different specificities. On the basis of those sample paths we may then get, in a more appropriate

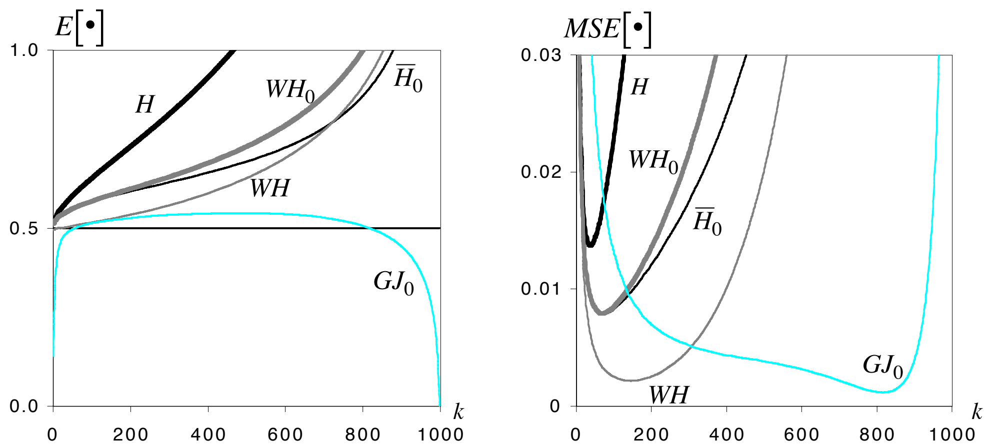

Figure 3: Underlying GP parent with $\gamma = 0.5$ ($\rho = -0.5$).

way, a sensible estimate of the tail index γ, like we shall see later on, in the application provided. And the new estimators herewith presented should be taken as the adequate substitutes of Hill's estimator.

We shall now consider an illustration of the performance of the above mentioned estimators, through the analysis of the Euro-USA Dollar daily exchange rates from January 4, 1999 till December 15, 2004. In Figure 10, working with the $n^+ = 748$ positive log-returns, we present the sample path of the $\widehat{\rho}_\tau$ estimates in (3.1) (*left*), as function of k, for $\tau = 0$ and $\tau = 1$. Note that the sample paths of the ρ-estimates associated to $\tau = 0$ and $\tau = 1$ lead us to choose, on the basis of any stability criterion for large k, the estimate associated to $\tau = 0$. From previous experience with this type of estimates, we conclude that the underlying ρ-value is larger or equal to -1, and the consideration of $\tau = 0$ is then advisable. The estimate of ρ is in this case $\widehat{\rho}_0 = -0.7$. In this same figure (*right*), we also present the sample paths of the classical Hill estimator, H, of $\overline{H}_0 = \overline{H}_{\widehat{\beta}_{01},\,\widehat{\rho}_0}$ and of $WH_0 = WH_{\widehat{\beta}_{01},\,\widehat{\rho}_0}$.

Regarding the tail index estimation, note that the Hill estimator exhibits

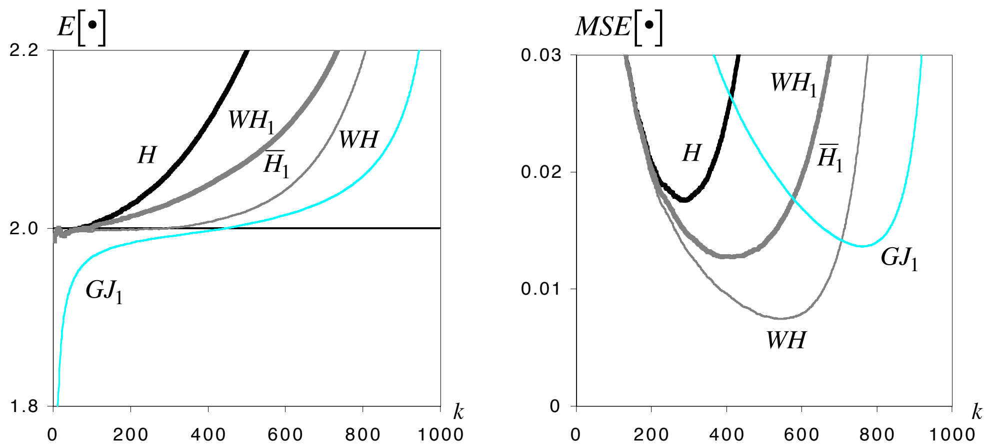

Figure 4: Underlying GP parent with $\gamma = 2$ $(\rho = -2)$.

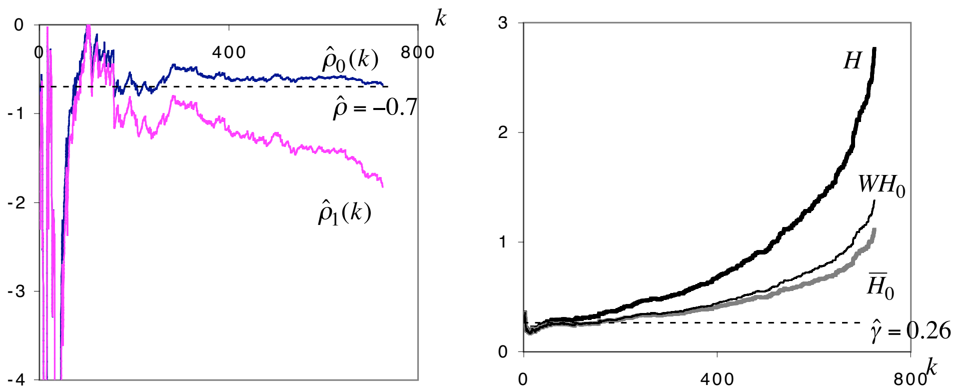

Figure 5: Estimates of the second order parameter ρ (*left*) and of the tail index γ (*right*) for the Daily Log-Returns of the Euro-UK Pound.

here a relevant bias. We are thus a long way from a strict Pareto underlying model. The other estimators, which are "asymptotically unbiased" up to the second order, reveal a smaller bias, and enable us to take a decision upon the estimate of γ to be used, with the help of any stability criterion or any heuristic procedure, like a largest run method as the one described in the sequel, and already suggested in Gomes *et al.* (2003): Let us consider a set of reduced bias' tail index estimates $\widehat{\gamma}_i(k)$, $1 \le k < n$, $i \in \mathcal{I}$, with a small number r of decimal figures. Let us denote them $\widehat{\gamma}_{i|r}(k)$. Then, for any value $i \in \mathcal{I}$ and for any possible value a in the domain of $\widehat{\gamma}_{i|r}(k)$, consider the largest run associated with a, i.e., $R_i(a)$, the maximum number of consecutive k values such that $\widehat{\gamma}_{i|r}(k) = a$. Next, compute $a_i^M := \arg\max_a R_i(a)$, and consider as a data-driven estimate of the tail index γ, $\widehat{\gamma} = a_{i_0}^M$ with $i_0 := \arg\max_i a_i^M$.

Here, if we consider the tail index estimates with one decimal figure, the largest run is achieved by the sample path of the $\overline{H}$-estimator in (1.11). Such a largest run has a size equal to 266, for k between 68 and 333, and is associated to $\overline{H} = 0.3$. For the WH-estimate, and also with one decimal figure, we would also get a tail index estimate equal to 0.3, but with a run of size 235 ($60 \leq k \leq 294$). With this same criterion, the Hill estimator would provide an estimate also equal to 0.3, with a run of size 126 ($43 \leq k \leq 168$). According to the previous heuristic procedure we would thus be led to the choice of the $\overline{H}$ estimator, computed at any of the levels from $k = 68$ till $k = 333$, all providing the same estimate $\widehat{\gamma} = 0.3$. Should we consider this same criterion, but the estimates with two decimal figures, would we be led to an estimate equal to 0.26, provided by any of the reduced bias' estimators. This value, $\widehat{\gamma} = 0.26$, is the one pictured in Figure 5.

6.2 Some overall conclusions

- Generally, we may say that there is not any significant difference between the estimators, WH, and $\overline{H}$ in (1.10) and (1.11), respectively. Anyway, whenever confronted with real data, the drawing of sample paths of a few alternative estimates may help us in the choice of the most adequate estimate of the tail index.

- The Generalized Jacknife statistic in (1.12) exhibits, for some of the models, sample paths more stable around the target value γ for a wider region of k-values than the statistics herewith studied, but at the expenses of mean squared errors much higher than those of the Hill statistic for small up to moderate values of k.

- Indeed, the main advantage of the so-called *Unbiased Hill* estimators WH and $\overline{H}$ in (1.10) and (1.11), respectively, lies on the fact that we may estimate β and ρ adequately through $\widehat{\beta}$ and $\widehat{\rho}$ so that the MSE of the new estimator is smaller than the MSE of Hill's estimator for all k, even when $|\rho| > 1$, a region where has been difficult to find alternatives for the Hill estimator. And this happens together with a higher stability of the sample paths around the target value γ. These new estimators work indeed better than the Hill estimator for all values of k, contrarily to the previous alternatives available in the literature, like the Generalized Jackknife estimator in (1.12).

7 Proofs

Let us further introduce the notation,

$$\psi_{ik}(\rho) \equiv \psi_\rho(i/k) := -\frac{(i/k)^{-\rho} - 1}{\rho \ln(i/k)} \quad [\psi_{kk} \equiv 1], \tag{7.1}$$

and for any $\alpha \geq 1$,

$$P_k^{(\alpha-1)}(\rho) := \frac{1}{k} \sum_{i=1}^{k} \psi_{ik}^{\alpha-1}(\rho)\ E_{k-i+1:k}, \tag{7.2}$$

$$Q_k^{(\alpha)}(\rho) := \frac{1}{k} \sum_{i=1}^{k} \psi_{ik}^{\alpha-1}(\rho)\ \frac{\exp\left(\rho E_{k-i+1:k}\right) - 1}{\rho}. \tag{7.3}$$

Gomes et al. (2004b) proved that, as $n \to \infty$, for k such that (1.9) holds and for any $\alpha \geq 1$, the limit in probability of both $P_k^{(\alpha)}(\rho)$ and $Q_k^{(\alpha)}(\rho)$ is given by

$$a_\alpha(\rho) = -\int_0^1 \psi_\rho^\alpha(v)\ lnv\ dv < \infty. \tag{7.4}$$

For $0 \leq \alpha < 1$, $P_k^{(\alpha)}(\rho)$ also converges in probability towards $a_\alpha(\rho)$ in (7.4), with $a_0 \equiv 1$.

We may easily derive the results in the following lemma:

Lemma 7.1. *Under the third order framework in* (1.3), *for levels k such that* (1.9) *holds, and with U_i given in* (1.8), *the distributional representations*

$$\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{\alpha-1} U_i \;\stackrel{d}{=}\; \frac{\gamma}{\alpha} + \frac{\gamma}{\sqrt{(2\alpha-1)\,k}}\;\overline{Z}_k^{(\alpha)} + \frac{A(n/k)}{\alpha-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) + \frac{A(n/k)\;B(n/k)}{\alpha-\rho-\rho'}\;(1+o_p(1)) \quad (7.5)$$

hold true for any $\alpha \geq 1$, where $\overline{Z}_k^{(\alpha)}$ in (2.1) *are asymptotically standard normal r.v.'s.*

Similarly, now with V_{ik} and ψ_{ik} given in (1.7) *and* (7.1), *respectively, the distributional representations*

$$\frac{1}{k}\sum_{i=1}^{k}\psi_{ik}^{\alpha}(\rho)\;V_{ik} \;\stackrel{d}{=}\; \gamma\;a_\alpha(\rho) + \frac{\gamma\;\sigma_\alpha(\rho)\;\overline{P}_k^{(\alpha)}(\rho)}{\sqrt{k}} + a_{\alpha+1}(\rho)\;A(n/k) + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) + b_{\alpha+1}(\rho,\rho')\;A(n/k)B(n/k)(1+o_p(1)) \quad (7.6)$$

hold true for any $\alpha \geq 0$, with

$$\begin{aligned} b_\alpha(\rho,\rho') &= \lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\psi_{ik}^{\alpha-1}(\rho)\;\mathbb{E}\left[\frac{Y_{k-i+1:k}^{\rho+\rho'}-1}{\rho+\rho'}\right] \\ &= -\int_0^1 \psi_\rho^{\alpha-1}(v)\;\psi_{\rho+\rho'}(v)\ln v\;dv < \infty, \end{aligned} \quad (7.7)$$

$$\begin{aligned} \sigma_\alpha^2 &= \lim_{n\to\infty} k\;\mathbb{V}ar\left(P_k^{(\alpha)}(\rho)\right) \\ &= \frac{2}{\rho^{2\alpha}}\iint_{0\leq u<v\leq 1}(\psi_\rho(u)\;\psi_\rho(v))^\alpha\left(\frac{1-v}{v}\right)du\;dv < \infty \quad [\sigma_0 = 1], \end{aligned} \quad (7.8)$$

being

$$\overline{P}_k^{(\alpha)}(\rho) := \sqrt{k}\left(P_k^{(\alpha)}(\rho) - a_\alpha(\rho)\right)/\sigma_\alpha(\rho) \quad (7.9)$$

an asymptotically standard normal r.v.

Proof. The summand $O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)$ is due to the replacement of $A(Y_{n-k:n})$ by $A(n/k)$, together with the fact that $Y_{n-k:n} - n/k = O_p\left(n/\left(k\sqrt{k}\right)\right)$ and $A'(n/k) = O\left(k\ A(n/k)/n\right)$. The remaining part of the lemma comes from the third order set-up in (1.3) together with the same kind of reasoning as in Gomes and Martins (2004), Gomes *et al.* (2003, 2004b) and Caeiro *et al.* (2004). □

7.1 Proof of the theorem in section 2

Proof. (Theorem 2.1). For the estimator in (1.10), with $P_k^{(\alpha)}(\rho)$ and $Q_k^{(\alpha)}(\rho)$ given in (7.2) and (7.3), respectively, and noticing that $\overline{P}_k^{(0)}(\rho) = \overline{Z}_k^{(1)}$, $\overline{Z}_k^{(\alpha)}$ and $\overline{P}_k^{(\alpha)}(\rho)$ given in (2.1) and (7.9), respectively, we get,

$$\begin{aligned}
WH_{\beta,\rho}(k) &:= \frac{1}{k}\sum_{i=1}^{k} e^{-A(n/k)\psi_{ik}(\rho)/\gamma}\ V_{ik} \\
&\stackrel{d}{=} \frac{1}{k}\sum_{i=1}^{k} V_{ik} - \frac{A(n/k)}{\gamma}\left(\frac{1}{k}\sum_{i=1}^{k}\psi_{ik}(\rho)\ V_{ik}\right) \\
&\qquad\qquad + \frac{A^2(n/k)}{2\gamma^2}\left(\frac{1}{k}\sum_{i=1}^{k}\psi_{ik}^2(\rho)\ V_{ik}\right)(1+o_p(1)).
\end{aligned}$$

The use of (7.6) enables us to write,

$$\begin{aligned}
WH_{\beta,\rho}(k) \stackrel{d}{=} \gamma &+ \frac{\gamma\ \overline{P}_k^{(0)}}{\sqrt{k}} + A(n/k)\left(a_1(\rho) + b_1(\rho,\rho')\ B(n/k) + O_p\left(\frac{1}{\sqrt{k}}\right)\right) \\
&- \frac{A(n/k)}{\gamma}\left(\gamma\ a_1(\rho) + \frac{\gamma\sigma_1(\rho)\overline{P}_k^{(1)}}{\sqrt{k}} + a_2(\rho)A(n/k)(1+o_p(1))\right) \\
&+ \frac{a_2(\rho)A^2(n/k)}{2\gamma}(1+o_p(1)).
\end{aligned}$$

Since $b_1(\rho,\rho') = a_1(\rho+\rho') = 1/(1-\rho-\rho')$, we get

$$\begin{aligned}
WH_{\beta,\rho}(k) &\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ \overline{Z}_k^{(1)} - A(n/k)\left(\frac{a_2(\rho)A(n/k)}{2\gamma} - \frac{B(n/k)}{1-\rho-\rho'}\right)(1+o_p(1)) \\
&\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ \overline{Z}_k^{(1)} + o_p(A(n/k),
\end{aligned}$$

and (2.2) holds.

Regarding the estimator in (1.11), since (7.5) holds true, with $\alpha = 1$,

$$
\begin{aligned}
\overline{H}_{\beta,\,\rho}(k) &\stackrel{d}{=} \left(\gamma + \frac{\gamma}{\sqrt{k}}\,\overline{Z}_k^{(1)} + \frac{A(n/k)}{1-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\right. \\
&\qquad \left. + \frac{A(n/k)B(n/k)}{1-\rho-\rho'}(1+o_p(1))\right) \times \left(1 - \frac{A(n/k)}{\gamma(1-\rho)}\right) \\
&\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\,\overline{Z}_k^{(1)} + \left(\frac{A(n/k)B(n/k)}{1-\rho-\rho'} - \frac{A^2(n/k)}{\gamma(1-\rho)^2}\right)(1+o_p(1)) \\
&\qquad + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)(1+o(1)) \\
&\stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\,\overline{Z}_k^{(1)} + o_p(A(n/k),
\end{aligned}
$$

i.e., (2.3) holds.

Note that since $\sqrt{k}\ O_p\left(A(n/k)/\sqrt{k}\right) = O_p(A(n/k)) \to 0$, the summands $O_p\left(A(n/k)/\sqrt{k}\right)$ are totally irrelevant for the asymptotic bias in (2.4) and (2.5), that follow straightforwardly from the above obtained distributional representations. □

7.2 Proof of the theorem in section 3

Proof. (Theorem 3.1). The distributional representation in (3.9) comes directly from (7.5). Indeed, if we write

$$
\widehat{\beta}_\rho(k) =: \left(\frac{k}{n}\right)^\rho \times \frac{\varphi(\rho)}{\psi(\rho)} = \frac{\gamma\ \beta}{A(n/k)} \times \frac{\varphi(\rho)}{\psi(\rho)},
$$

we get

$$
\begin{aligned}
\varphi(\rho) \stackrel{d}{=} \frac{\gamma}{\sqrt{k}}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho} - \frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right) - \frac{\rho^2 A(n/k)}{(1-\rho)^2(1-2\rho)} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) \\
- \frac{\rho(\rho+\rho')A(n/k)\ B(n/k)}{(1-\rho)(1-\rho-\rho')(1-2\rho-\rho')}(1+o_p(1)) + O_p(1/k),
\end{aligned}
$$

and

$$\psi(\rho) \stackrel{d}{=} -\frac{\gamma\ \rho^2}{(1-\rho)^2(1-2\rho)}+\frac{\gamma}{\sqrt{k}}\left(\frac{\overline{Z}_k^{(1-\rho)}}{(1-\rho)\sqrt{1-2\rho}}-\frac{\overline{Z}_k^{(1-2\rho)}}{\sqrt{1-4\rho}}\right)+O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)$$
$$-\frac{2\rho^2 A(n/k)}{(1-\rho)(1-2\rho)(1-3\rho)}-\frac{\rho(2\rho+\rho')A(n/k)B(n/k)}{(1-\rho)(1-2\rho-\rho')(1-3\rho-\rho')}(1+o_p(1))+O_p(1/k).$$

Consequently, if $\sqrt{k}\ A(n/k)\to\infty$, i.e., $1/\sqrt{k}=o(A(n/k))$,

$$\frac{1}{\psi(\rho)} \stackrel{d}{=} -\frac{(1-\rho)^2(1-2\rho)}{\gamma\ \rho^2}\left(1-\frac{2(1-\rho)A(n/k)}{\gamma(1-3\rho)}+o_p(A(n/k))\right),$$

$$\frac{\varphi(\rho)}{\psi(\rho)} \stackrel{p}{\sim} \frac{A(n/k)}{\gamma}-\frac{(1-\rho)^2(1-2\rho)}{\rho^2\sqrt{k}}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho}-\frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right)$$
$$-\frac{2(1-\rho)A^2(n/k)}{\gamma^2(1-3\rho)}+\frac{(1-\rho)(1-2\rho)(\rho+\rho')A(n/k)B(n/k)}{\gamma\rho(1-\rho-\rho')(1-2\rho-\rho')},$$

$$\widehat{\beta}_\rho(k) \stackrel{p}{\sim} \beta\left(1-\frac{\gamma\ (1-\rho)^2(1-2\rho)}{\rho^2\sqrt{k}\ A(n/k)}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho}-\frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right)-\frac{2(1-\rho)\ A(n/k)}{\gamma(1-3\rho)}\right.$$
$$\left.+\frac{(1-\rho)(1-2\rho)(\rho+\rho')\ B(n/k)}{\rho(1-\rho-\rho')(1-2\rho-\rho')}\right),$$

and since the asymptotic covariance between $\overline{Z}_k^{(1)}$ and $\overline{Z}_k^{(1-\rho)}$ is given by $\sqrt{1-2\rho}/(1-\rho)$, (3.9) follows, as well as (3.10), (3.11) and (3.12).

The remaining of the proposition comes from the fact that, since

$$\frac{d}{d\rho}\widehat{\beta}_\rho(k)/\widehat{\beta}_\rho(k)=-\ln(n/k)(1+o_p(1)),$$

we may write

$$\widehat{\beta}_{\widehat{\rho}}(k)=\widehat{\beta}_\rho(k)-\widehat{\beta}_\rho(k)\,(\widehat{\rho}-\rho)\ln(n/k)(1+o_p(1))$$
$$=\beta-\frac{\gamma\ \beta(1-\rho)^2(1-2\rho)}{\rho^2\sqrt{k}\ A(n/k)}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho}-\frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right)$$
$$+\left(\frac{\beta(1-\rho)(1-2\rho)(\rho+\rho')B(n/k)}{\rho(1-\rho-\rho')(1-2\rho-\rho')}-\frac{2\beta(1-\rho)A(n/k)}{\gamma(1-3\rho)}\right)(1+o_p(1))$$
$$-\beta\ (\widehat{\rho}-\rho)\ln(n/k)(1+o_p(1)). \quad (7.10)$$

Now, if we estimate ρ through any of the estimators in (3.1) computed at the level k_1 in (3.5), and under the conditions of the proposition, i.e., if $\sqrt{k}\ A(n/k) \to \infty$, we have $\sqrt{k}\ A(n/k)\,(\widehat{\rho}-\rho)\ln(n/k) \to 0$, as $n \to \infty$, provided that $\sqrt{k}\ A^2(n/k) \to \lambda_{_A}$, finite. Indeed, if $\sqrt{k}\ A(n/k) \to \infty$, k is of a larger order than $n^{-2\rho/(1-2\rho)}$, and then

$$\frac{n}{k} < O\left(n^{1/(1-2\rho)}\right), \qquad \frac{\ln n/k}{A(n/k)} < O\left(\frac{\ln n}{n^{\rho/(1-2\rho)}}\right).$$

From the result in Remark 3.2, we thus get

$$0 \le \left|\frac{(\widehat{\rho}-\rho)\ln n/k}{A(n/k)}\right| < O\left((\ln_2 n)^{(1-2\rho)/2}\,\frac{\ln n}{n^{1/(2(1-2\rho))}}\right) \underset{n\to\infty}{\longrightarrow} 0.$$

Consequently, the results in (3.9) and (3.10) hold true when we replace ρ by $\widehat{\rho}$ in (3.1) computed at the level k_1 in (3.5).

Under the conditions of the theorem, if we estimate ρ through $\widehat{\rho}(k)$, $(\widehat{\rho}(k)-\rho)\ln(n/k)$ is the dominant term among the terms in (7.10), dependent on k. Then the behaviour of $\widehat{\beta}_{\widehat{\rho}(k)}$ is directly related to the behaviour of $\{\widehat{\rho}(k)-\rho\}$, as stated in (3.13). □

7.3 Proof of theorems in section 4

Proof. (Theorem 4.1). If we estimate consistently β and ρ through the estimators $\widehat{\beta}$ and $\widehat{\rho}$ in the conditions of the theorem, we may use Taylor's expansion series, and obtain for any of the estimators in (1.10) and (1.11), generically denoted UH

$$UH_{\widehat{\beta},\ \widehat{\rho}}(k) - UH_{\beta,\ \rho}(k) \overset{p}{\sim} A(n/k)\left\{\left(\widehat{\beta}-\beta\right)a_{_{UH}}/\beta \right.$$
$$\left. + (\widehat{\rho}-\rho)\left(a_{_{UH}}\ \ln(n/k) + b_{_{UH}}\right)\right\}, \quad (7.11)$$

where $a_{_{UH}} = -1/(1-\rho)$ and $b_{_{UH}} = -1/(1-\rho)^2$. Indeed,

$$
\begin{aligned}
\frac{\partial \overline{H}_{\beta,\,\rho}}{\partial \beta} &\overset{p}{\sim} -\frac{A(n/k)}{\beta(1-\rho)},\\
\frac{\partial \overline{H}_{\beta,\,\rho}}{\partial \rho} &\overset{p}{\sim} -A(n/k)\left(\frac{1}{1-\rho}\ln\left(\frac{n}{k}\right)+\frac{1}{(1-\rho)^2}\right),\\
\frac{\partial WH_{\beta,\,\rho}}{\partial \beta} &\overset{p}{\sim} -\frac{A(n/k)}{\gamma\,\beta}\left(\frac{1}{k}\sum_{i=1}^{k}\psi_{_{ik}}(\rho)\;V_{_{ik}}\right) \overset{p}{\sim} -\frac{A(n/k)}{\beta(1-\rho)},\\
\frac{\partial WH_{\beta,\,\rho}}{\partial \rho} &\overset{p}{\sim} -\frac{A(n/k)}{\gamma}\left(\frac{1}{k}\sum_{i=1}^{k}\left\{\frac{1}{\rho}+\psi_{_{ik}}(\rho)\left[\ln\left(\frac{n}{k}\right)-\frac{1}{\rho}-\ln\left(\frac{i}{k}\right)\right]\right\}V_{_{ik}}\right)\\
&\overset{p}{\sim} -A(n/k)\left(\frac{1}{1-\rho}\ln\left(\frac{n}{k}\right)+\frac{1}{(1-\rho)^2}\right),
\end{aligned}
$$

since

$$\frac{1}{k}\sum_{i=1}^{k}\psi_{_{ik}}(\rho)\;V_{_{ik}} \underset{k\to\infty}{\overset{p}{\longrightarrow}} -\gamma\int_0^1 \psi_\rho(v)\ln v\;dv = \frac{\gamma}{1-\rho}$$

and

$$\frac{1}{k}\sum_{i=1}^{k}\psi_{_{ik}}(\rho)\ln\left(\frac{i}{k}\right)\;V_{_{ik}} \underset{k\to\infty}{\overset{p}{\longrightarrow}} -\gamma\int_0^1 \psi_\rho(v)\ln^2 v\;dv = -\frac{\gamma(2-\rho)}{(1-\rho)^2}.$$

Consequently, since (3.13) holds true, i.e., $\left(\widehat{\beta}-\beta\right)/\beta \overset{p}{\sim} -\ln(n/k_1)\,(\widehat{\rho}-\rho)$, we have

$$UH_{\widehat{\beta},\,\widehat{\rho}}(k) - UH_{\beta,\,\rho}(k) = O_p\left((\widehat{\rho}-\rho)\,A(n/k)\ln(k/k_1)\right). \tag{7.12}$$

If $\sqrt{k}\;A^2(n/k) \to \lambda_{_A}$, finite, k is at most of the order of $n^{-4\rho/(1-4\rho)}$. Then, k/k_1 is at most of the order of $n^{-1/(1-4\rho)}\ln_2 n$. Consequently,

$$\sqrt{k}\;(\widehat{\rho}-\rho)\ln(k/k_1) \le O_p\left(\frac{n^{\frac{-2\rho}{1-4\rho}}(\ln_2 n)^{\frac{1}{2}-\rho}\ln n}{n^{\frac{1}{2}}}\right) = O_p\left(\frac{(\ln_2 n)^{\frac{1}{2}-\rho}\ln n}{n^{\frac{1}{2(1-4\rho)}}}\right).$$

Hence, $\sqrt{k}\;(\widehat{\rho}-\rho)\;A(n/k)\;\ln(k/k_1)$ converges towards zero, as $n\to\infty$, and the results in the theorem, related to either the WH or the $\overline{H}$ estimators follow. □

Proof. (Theorem 4.2). If we consider $WH_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k)$, we get, under the third order framework in (1.3),

$$\begin{aligned} WH_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k) &= \gamma + \frac{\gamma}{\sqrt{k}}\left(\overline{Z}_k^{(1)} + \frac{(1-\rho)(1-2\rho)}{\rho^2}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho} - \frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right)\right) \\ &+ \left(\frac{A(n/k)\ B(n/k)}{1-\rho-\rho'} - \frac{(1-2\rho)(\rho+\rho')A(n/k)B(n/k)}{\rho(1-\rho-\rho')(1-2\rho-\rho')}\right)(1+o_p(1)) \\ &- \left(\frac{a_2(\rho)A^2(n/k)}{2\ \gamma} - \frac{2\ A^2(n/k)}{\gamma(1-3\rho)}\right)(1+o_p(1)). \end{aligned}$$

We thus get for the WH-estimator an asymptotic variance equal to $(\gamma(1-\rho)/\rho)^2$ and the asymptotic bias in (4.2).

If we next consider

$$\overline{H}_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k) := H(k)\left(1 - \frac{\widehat{\beta}_{\widehat{\rho}}(k)}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right),$$

we get, under the third order framework in (1.3),

$$\begin{aligned} \overline{H}_{\widehat{\beta}_{\widehat{\rho}}(k),\ \widehat{\rho}}(k) &= \gamma + \frac{\gamma}{\sqrt{k}}\left(\overline{Z}_k^{(1)} + \frac{(1-\rho)(1-2\rho)}{\rho^2}\left(\frac{\overline{Z}_k^{(1)}}{1-\rho} - \frac{\overline{Z}_k^{(1-\rho)}}{\sqrt{1-2\rho}}\right)\right) \\ &+ \left(\frac{A(n/k)\ B(n/k)}{1-\rho-\rho'} - \frac{(1-2\rho)(\rho+\rho')A(n/k)B(n/k)}{\rho(1-\rho-\rho')(1-2\rho-\rho')}\right)(1+o_p(1)) \\ &- \left(\frac{A^2(n/k)}{\gamma(1-\rho)^2} - \frac{2\ A^2(n/k)}{\gamma(1-3\rho)}\right)(1+o_p(1)) + o_p\left(\ln(n/k)A(n/k)\left(\widehat{\rho}-\rho\right)\right). \end{aligned}$$

Consequently, we get for the $\overline{H}$-estimator the same asymptotic variance as before, but a different asymptotic bias, the one in (4.3). □

References

[1] Beirlant, J., Dierckx, G., Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177-200.

[2] Beirlant, J., Dierckx, G., Guillou, A. and Starica, C. (2002). On exponential representations of log-spacings of extreme order statistics. *Extremes* **5**: 2, 157-180.

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

[4] Caeiro, F. and Gomes, M. I. (2004). *A new class of estimators of the "scale" second order parameter*. Notas e Comunicações CEAUL 20/04. Submitted.

[5] Caeiro, F. Gomes, M. I. and Pestana, D. D. (2004). *Direct reduction of bias of the classical Hill estimator.* Notas e Comunicações CEAUL 18/04. Submitted.

[6] Drees, H. (1998). A general class of estimators of the extreme value index. *J. Statist. Planning and Inference* **98**, 95-112.

[7] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.

[8] Fiavlová, A., Jurecková, J. and Picek, J. (2004). Estimating Pareto tail index based on sample means. *RevStat* **2**(1), 75-100.

[9] 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.

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

[11] Gomes, M. I., Caeiro, F. and Figueiredo, F. (2004a). Bias reduction of a tail index estimator trough an external estimation of the second order parameter. *Statistics* **38**(6), 497-510.

[12] Gomes, M. I. Figueiredo, F. and Mendonça, S. (2003). Asymptotically best linear unbiased tail estimators under a second order regular variation condition. Accepted at *J. Statist. Planning and Inference.*

[13] Gomes, M. I. Haan, L. de and Henriques, L. (2004b). *Tail index estimation through accommodation of bias in the weighted log-excesses.* Notas e Comunicações C.E.A.U.L. 14/2004. *Submitted.*

[14] Gomes, M. I., de Haan, L. and Peng, L. (2002a). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour. *Extremes* **5**:4, 387-414.

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

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

[17] Gomes, M. I. and Martins, M. J. (2004). Bias reduction and explicit estimation of the tail index. *J. Statist. Planning and Inference* **124**, 361-378.

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

[19] Gomes, M. I., Martins, M. J. and Neves, M. (2002b). Generalized Jackknife semi-parametric estimators of the tail index. *Portugaliae Mathematica* **59**:4, 393-408.

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

[21] Gomes, M. I. and Pestana, D. (2004). *A simple second order reduced bias' tail index estimator.* Submitted.

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

[23] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *J. Royal Statist. Soc.* **B44**, 37-42.

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

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

[26] Peng, L. (1998). Asymptotically unbiased estimator for the extreme-value index. *Statistics and Probability Letters* **38**(2), 107-115.