Bias reduction of a tail index estimator through an external estimation of the second order parameter*

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

Frederico Caeiro
Universidade Nova de Lisboa, Departamento de Matemática, FCT

and

Fernanda Figueiredo
Universidade do Porto, Faculdade de Economia and CEAUL

Abstract. In this paper we first consider a class of consistent semi-parametric estimators of a positive tail index γ, parameterised in a *tuning* or *control* parameter α. Such a control parameter enables us to have access, for any available sample, to an estimator of the tail index γ with a null dominant component of asymptotic bias, and consequently with a reasonably flat *Mean Squared Error* pattern, as a function of k, the number of top order statistics considered. Such a control parameter depends on a second order parameter, ρ, which will be adequately estimated so that we may achieve a high efficiency relatively to the classical Hill estimator, provided we use a number of top order statistics larger than the one usually required for the estimation through the Hill estimator. An illustration of the behaviour of the estimators is provided, through the analysis of the daily log-returns on the Euro-US Dollar exchange rates.

AMS 1991 subject classification. Primary 62G05, 62E25, 62E20; Secondary 62F35.
Keywords and phrases. *Statistical Theory of Extremes, Semi-parametric estimation.*

1 The class of semi-parametric estimators

In this paper we deal with a semi-parametric estimator of the *tail index* γ, with a null dominant component of asymptotic bias. Such a kind of estimators has revealed nice distributional properties (Peng, 1998; Beirlant *et al.*, 1999; Feuerverger and Hall, 1999; Gomes *et al.*, 2000, 2002; Gomes and Martins, 2001; Caeiro and Gomes, 2002).

*Research partially supported by FCT / POCTI / FEDER.

We shall consider here heavy tails, i.e. $\gamma > 0$ in the *Extreme Value* distribution function (d.f.)

$$EV_\gamma(x) := \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right), \ 1+\gamma x > 0 & \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)), \ x \in \mathbb{R} & \text{if} \quad \gamma = 0 \end{cases},$$

the only non-degenerate d.f. to which $X_{n:n} := \max(X_1, ..., X_n)$ may be attracted, after suitable linear normalization. If this happens we say that the underlying model F is in the max-domain of attraction of EV_γ, and denote this by $F \in \mathcal{D}_M(EV_\gamma)$. As usual, $X_{i:n}$, $1 \leq i \leq n$, denotes the i-th ascending order statistic (o.s.) associated with the random sample $(X_1, \cdots, X_n)$ from the unknown distribution function F.

With $U(t) := F^{\leftarrow}(1 - 1/t)$, $t \geq 1$, $F^{\leftarrow}$ denoting the generalized inverse function of F, we have (Gnedenko, 1943; de Haan, 1970),

$$F \in \mathcal{D}_M(EV_\gamma), \gamma > 0 \quad \text{iff} \quad 1 - F \in RV_{-1/\gamma} \quad \text{iff} \quad U \in RV_\gamma, \tag{1.1}$$

where RV_β stands for the class of *regularly varying* functions at infinity with *index of regular variation* β, i.e., positive measurable functions g such that $\lim_{t\to\infty} g(tx)/g(t) = x^\beta$, for all $x > 0$.

Apart from the first order condition in (1.1) we shall often assume that there exists a function $A(t)$ of constant sign near infinity, measuring the rate of convergence of $\ln U(tx) - \ln U(t)$ towards $\gamma \ln x$ in the first order condition, i.e., a function A such that

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

for every $x > 0$, where ρ (≤ 0) is a *second order parameter.* The limit function in (1.2) must be of the stated form, and $|A(t)| \in RV_\rho$ (Geluk and de Haan, 1987). In this paper we shall restrict ourselves to the case $\rho < 0$.

The first class of semi-parametric estimators of γ, herewith considered under a third order framework to be introduced in section 3, is the same class as in Caeiro and Gomes (2002),

$$\widehat{\gamma}_n^{(\alpha)}(k) := \frac{\Gamma(\alpha)}{M_n^{(\alpha-1)}(k)} \left(\frac{M_n^{(2\alpha)}(k)}{\Gamma(2\alpha+1)}\right)^{1/2}, \quad \alpha \geq 1. \tag{1.3}$$

Such a class is parameterised in the *tuning* parameter α, which may be controlled at our ease. We have $M_n^{(0)} \equiv 1$, and

$$M_n^{(\alpha)}(k) := \frac{1}{k} \sum_{i=1}^{k} \left(\ln X_{n-i+1:n} - \ln X_{n-k:n}\right)^\alpha, \quad \alpha > 0, \tag{1.4}$$

are consistent estimators of $\Gamma(\alpha+1)\gamma^\alpha$, whenever k is *intermediate*, i.e., whenever

$$k = k_n \to \infty, \text{ and } k = o(n), \text{ as } n \to \infty. \tag{1.5}$$

$M_n^{(1)}(k)$, also denoted $\widehat{\gamma}_n^H(k)$, is the classical Hill's estimator (Hill, 1975).

The class of estimators in (1.3), studied in Caeiro and Gomes (2002), generalizes the estimator $\widehat{\gamma}_n^{(1)}(k) := \sqrt{M_n^{(2)}(k)/2}$, studied in Gomes *et al.* (2000). In the class (1.3), and whenever $\rho < 0$, it is possible to find a control parameter α which makes null the dominant component of asymptotic bias of our tail index estimator. Such a control parameter depends on the second order parameter ρ, which, contrarily to what has been done in Caeiro and Gomes (2002), is going to be here properly estimated on the basis of our sample. We may thus get a second class of "asymptotically unbiased" estimators of the tail index γ, now based on an external estimation of the second order parameter ρ, and given by

$$\widehat{\gamma}_n^{(\widehat{\alpha})}(k) := \frac{\Gamma(\widehat{\alpha})}{M_n^{(\widehat{\alpha}-1)}(k)} \left(\frac{M_n^{(2\widehat{\alpha})}(k)}{\Gamma(2\widehat{\alpha}+1)}\right)^{1/2},$$

$$\text{with} \quad \widehat{\alpha} = -\frac{\ln\left[1-\widehat{\rho}-\sqrt{(1-\widehat{\rho})^2-1}\right]}{\ln(1-\widehat{\rho})}, \tag{1.6}$$

where generally $\widehat{\rho}$ may be any consistent estimator of ρ.

Since we get to know from previous papers that these "asymptotically unbiased" estimators behave better than the Hill estimator, we shall also consider, for comparison with the class of estimators in (1.6), an ML-estimator, asymptotically equivalent to the ML-estimator in Gomes and Martins (2002), and given by

$$\widehat{\gamma}_n^{ML(\widehat{\rho})}(k) := \frac{1}{k}\sum_{i=1}^{k} \exp\left(\widehat{D}\ (i/n)^{-\widehat{\rho}}\right)\ U_i \tag{1.7}$$

where

$$\widehat{D} \equiv \widehat{D}(k) := \frac{1}{n^{\widehat{\rho}}}\ \frac{\left(\sum_{i=1}^k i^{-\widehat{\rho}}\right)\left(\sum_{i=1}^k U_i\right) - k\left(\sum_{i=1}^k i^{-\widehat{\rho}}U_i\right)}{\left(\sum_{i=1}^k i^{-\widehat{\rho}}\right)\left(\sum_{i=1}^k i^{-\widehat{\rho}}U_i\right) - k\left(\sum_{i=1}^k i^{-2\widehat{\rho}}U_i\right)},$$

with

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

denoting the scaled log-spacings. Such an estimator is the one with smallest asymptotic variance, among the ones so far considered in the literature. Such an asymptotic variance is $\gamma^2(1-\rho)^2/\rho^2$, the lower limit provided in Drees (1998) for the general class of "asymptotically unbiased" estimators therewith considered.

In section 2 of this paper we shall briefly refer the asymptotic behaviour of the class of estimators in (1.3), under a second order framework, as derived in Caeiro and Gomes (2002), but using a notation adequate to the introduction of the third order framework. In section 3 we shall work under such a third order framework, and shall make a few comments on a class of estimators of the second order parameter ρ, recently introduced by Fraga Alves et al. (2003). In section 4, we shall incorporate these ρ-estimators in our class of tail index estimators, removing the dominant component of asymptotic bias. In section 5, we shall derive, through simulation techniques, the distributional properties of one of these estimators for finite sample sizes, and for a few heavy-tailed models. Finally, in section 6, we shall consider a case-study related to the daily exchange rates of the Euro against the US Dollar, to illustrate the methodologies advanced in this paper.

2 Asymptotic properties of the first class of estimators

Under the first order condition (1.1) and the restriction (1.5) on the level k, the statistics $\widehat{\gamma}_n^{(\alpha)}(k)$ in (1.3) are consistent for γ, and under some extra mild conditions on the second order behaviour of the model F underlying the data they are asymptotically normal, with an asymptotic bias possibly non-null, and given by $b_\alpha(\rho) \times \lim_{n\to\infty} \sqrt{k}\; A(n/k)$, with A given in (1.2) and where $b_\alpha(\rho)$ will be explicited later.

Let W denote an exponential r.v., with d.f. $F_W(x) = 1 - \exp(-x)$, $x > 0$, and, with the same notation as in Gomes *et al.* (2002) and Fraga Alves *et al.* (2003), let us put

$$\mu_\alpha^{(1)} := \mathbb{E}[W^\alpha] = \Gamma(\alpha+1), \quad \overline{\mu}_\alpha^{(1)} = 1$$

$$\sigma_\alpha^{(1)} := \sqrt{\mathbb{V}ar[W^\alpha]} = \sqrt{\Gamma(2\alpha+1) - \Gamma^2(\alpha+1)}, \quad \overline{\sigma}_\alpha^{(1)} := \sigma_\alpha^{(1)}/\mu_\alpha^{(1)},$$

$$\mu_\alpha^{(2)}(\rho) := \mathbb{E}\left[\frac{W^{\alpha-1}\left(e^{\rho W}-1\right)}{\rho}\right] = \frac{\Gamma(\alpha)(1-(1-\rho)^\alpha)}{\rho\,(1-\rho)^\alpha}, \quad \overline{\mu}_\alpha^{(2)}(\rho) := \mu_\alpha^{(2)}(\rho)/\mu_\alpha^{(1)},$$

where Γ denotes the complete Gamma function, $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx,\ \alpha > 0$.

We first state, without proof, which may be seen in Gomes and Martins (2001), and had already been suggested in Dekkers *et al.* (1989), the main result used to derive the asymptotic properties of $\widehat{\gamma}_n^{(\alpha)}(k)$ in (1.3).

Proposition 1. *Under the validity of the first order condition in (1.1) and for k intermediate, i.e., k such that (1.5) holds, the statistic $M_n^{(\alpha)}(k)$ in (1.4) converges in probability towards $\gamma^\alpha\ \mu_\alpha^{(1)}$. If we further assume the general second order framework in (1.2), the following asymptotic distributional representation,*

$$M_n^{(\alpha)}(k) \stackrel{d}{=} \gamma^\alpha \mu_\alpha^{(1)} + \frac{\gamma^\alpha \overline{\sigma}_\alpha^{(1)}}{\sqrt{k}}\ P_k^{(\alpha)} + \alpha\gamma^{\alpha-1}\overline{\mu}_\alpha^{(2)}(\rho)\ A(n/k)(1+o_p(1)) \tag{2.1}$$

holds, where, with $\{W_i\}_{i\geq 1}$ denoting independent standard exponential r.v.'s, $P_k^{(\alpha)} := \left(\frac{1}{k}\sum_{i=1}^k W_i^\alpha - \mu_\alpha^{(1)}\right)/\sigma_\alpha^{(1)}$ is an asymptotically standard normal r.v. Moreover, the r.v.'s $P_k^{(\alpha)}$ and $P_k^{(\beta)}$ have a covariance structure given by $\sigma_{\alpha,\beta} := \frac{\mu_{\alpha+\beta}^{(1)} - \mu_\alpha^{(1)}\ \mu_\beta^{(1)}}{\sigma_\alpha^{(1)}\sigma_\beta^{(1)}}$.

We next present a result proved in Caeiro and Gomes (2002).

Proposition 2. *Under the conditions and notations of Proposition 1, the asymptotic distributional representation*

$$\widehat{\gamma}_n^{(\alpha)}(k) \stackrel{d}{=} \gamma + \left(\frac{\gamma\sqrt{v_\alpha}}{\sqrt{k}}\ Z_k^{(\alpha)} + b_\alpha(\rho)\ A(n/k)\right)(1+o_p(1))$$

holds true for the statistic $\widehat{\gamma}_n^{(\alpha)}(k)$ in (1.3), where $Z_k^{(\alpha)}$ is asymptotically standard normal, and may be written as

$$Z_k^{(\alpha)} = \frac{1}{\sqrt{v_\alpha}}\left(\frac{\overline{\sigma}_{2\alpha}^{(1)}\ P_k^{(2\alpha)}}{2} - \overline{\sigma}_{\alpha-1}^{(1)}\ P_k^{(\alpha-1)}\right),$$

with $P_k^{(\alpha)}$ given in the distributional representation (2.1) for $M_n^{(\alpha)}(k)$. We have

$$v_\alpha = \frac{1}{4}\left\{\frac{2\Gamma(4\alpha)}{2\alpha\Gamma^2(2\alpha)} + \frac{4\Gamma(2\alpha-1)}{\Gamma^2(\alpha)} - \frac{2\Gamma(3\alpha)}{\alpha\Gamma(2\alpha)\Gamma(\alpha)} - 1\right\} \tag{2.2}$$

and

$$\begin{aligned} b_\alpha(\rho) &= \alpha\ \overline{\mu}_{2\alpha}^{(2)}(\rho) - (\alpha-1)\ \overline{\mu}_{\alpha-1}^{(2)}(\rho) \\ &= \begin{cases} \frac{1}{2\rho}\left\{(1-\rho)^{-2\alpha} - 2(1-\rho)^{-\alpha+1} + 1\right\} & \text{if} \quad \rho < 0 \\ 1 & \text{if} \quad \rho = 0. \end{cases} \end{aligned}$$

Then, since $k = k_n \to \infty$, $\widehat{\gamma}_n^{(\alpha)}(k)$ is consistent for the estimation of γ, and if $\sqrt{k}\ A(n/k) \to \lambda$, finite, $\sqrt{k}\left(\widehat{\gamma}_n^{(\alpha)}(k) - \gamma\right)$ is asymptotically normal, with asymptotic variance $\gamma^2\ v_\alpha$ and asymptotic bias $\lambda\ b_\alpha(\rho)$.

Moreover, for every $\rho \in \mathbb{R}^-$ and $\alpha > 1$ there is a value $\alpha_0 \equiv \alpha_0(\rho)$, explicitly given by

$$\alpha_0 \equiv \alpha_0(\rho) = -\frac{\ln\left[1-\rho-\sqrt{(1-\rho)^2-1}\right]}{\ln(1-\rho)}, \tag{2.3}$$

such that $b_{\alpha_0}(\rho) = 0$, i.e. $\widehat{\gamma}_n^{(\alpha_0)}(k)$ has a null asymptotic bias, even when $\sqrt{k}\ A(n/k) \to \lambda \neq 0$, as $n \to \infty$.

For values of α close to α_0 in (2.3) we thus get a dominant component of bias close to 0, and consequently we expect to have then, at the optimal level, a high asymptotic efficiency relatively to $\widehat{\gamma}_n^{(1)}(k)$ at its optimal level, and *a fortiori* relatively to the Hill estimator, also at its optimal level, as derived in Caeiro and Gomes (2002). More than this: as illustrated in Caeiro and Gomes (2002), any stability criterion of the sample paths of the estimators in (1.3) may provide a selection of an "optimal" value $\widetilde{\alpha}_0$ (defined in an adequate way, like has been done, for instance, by Gomes and Martins (2001)), which may on its turn provide an estimator of ρ, as a solution of the equation in (2.3). However, such a ρ-estimator does not work well in practice.

Remark 1. *As may be seen in Figure 1 the asymptotic variance of $\widehat{\gamma}_n^{(\alpha_0(\rho))}$, given by $\gamma^2\ v_{\alpha_0}$, with v_α and α_0 given in (2.2) and (2.3), respectively, is larger than the asymptotic variance of $\widehat{\gamma}_n^{ML}(\rho)$ in (1.7), which is given by $\gamma^2(1-\rho)^2/\rho^2$, for every $\rho \leq 0$. However, the reduction in bias may compensate this increase in variance, and, as we shall see in section 5, the "asymptotically unbiased" estimator herewith introduced, compares favourably with the ML-estimator, for a large class of heavy tailed models, when they both are computed at their optimal levels.*

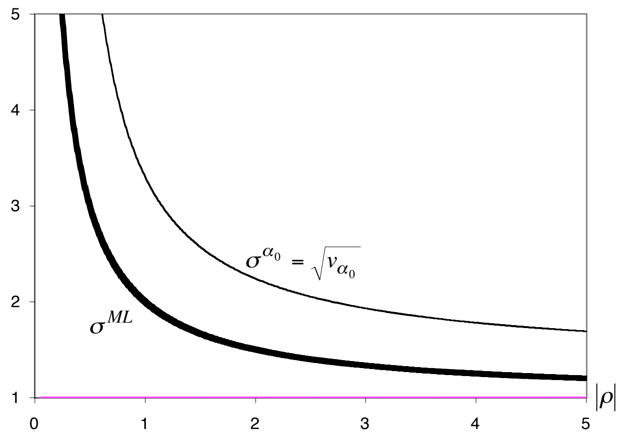

Figure 1: Asymptotic standard deviations of the two "asymptotically unbiased" tail index estimators under consideration, up to a factor γ

3 A third order framework and estimation of the second order parameter

In order to derive asymptotic normality of the "asymptotically unbiased" estimator $\widehat{\gamma}_n^{(\alpha_0)}$ in (1.3), together with information on the asymptotic bias, we need to know the rate of convergence in the second order condition (1.2). We shall thus assume that a third order condition holds, i.e., we assume that there exists $\rho' < 0$ such that

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx)-\ln U(t)-\gamma \ln x}{A(t)} - \frac{x^\rho-1}{\rho}}{B(t)} = \frac{x^{\rho+\rho'}-1}{\rho+\rho'}, \tag{3.1}$$

where $|B|$ must then be of regular variation, with index of regular variation ρ'. For the case $\rho, < 0$, the situation considered throughout the paper, this condition is equivalent to the one assumed in the papers of Gomes and de Haan (1999), Gomes *et al.* (2002) and Fraga Alves *et al.* (2003).

Remark 2. *Note that for most of the well-known heavy-tailed models, like the Fréchet, the Generalized Pareto, the Burr, the Student, and a large variety of models in Hall's class of distributions (Hall, 1982; Hall and Welsh, 1985), the models for which we have, with* $\rho < 0$, *a tail function of the type*

$$1 - F(x) = Cx^{-1/\gamma}\left(1 + D_1 x^{\rho/\gamma} + D_2\, x^{2\rho/\gamma} + o\left(x^{2\rho/\gamma}\right)\right), \quad \textit{as } x \to \infty,$$

the third order condition in (3.1) *holds with* $\rho' = \rho$.

With the same notation as before, i.e., with W denoting a standard exponential r.v., we shall still introduce here some extra notation:

$$\sigma_\alpha^{(2)}(\rho) := \sqrt{\mathbb{V}ar\left[\frac{W^{\alpha-1}\left(e^{\rho W}-1\right)}{\rho}\right]} = \sqrt{\mu_{2\alpha}^{(3)}(\rho) - \left(\mu_\alpha^{(2)}(\rho)\right)^2},$$

$$\overline{\sigma}_\alpha^{(2)}(\rho) := \sigma_\alpha^{(2)}(\rho)/\mu_\alpha^{(1)},$$

with

$$\begin{aligned}
\mu_\alpha^{(3)}(\rho) &:= \mathbb{E}\left[W^{\alpha-2}\left(\frac{e^{\rho W}-1}{\rho}\right)^2\right] = \frac{2\left\{\mu_{\alpha-1}^{(2)}(2\rho) - \mu_{\alpha-1}^{(2)}(\rho)\right\}}{\rho} \\
\overline{\mu}_\alpha^{(3)}(\rho) &= \mu_\alpha^{(3)}(\rho)/\mu_\alpha^{(1)} \\
&= \begin{cases} \frac{1}{\rho^2}\ln\frac{(1-\rho)^2}{1-2\rho} & if \quad \alpha = 1 \\ \frac{1}{\rho^2\ \alpha\ (\alpha-1)}\left\{\frac{1}{(1-2\rho)^{\alpha-1}} - \frac{2}{(1-\rho)^{\alpha-1}} + 1\right\} & if \quad \alpha \neq 1 \end{cases}.
\end{aligned}$$

3.1 The asymptotic distributional behaviour of our first class of tail index estimators under a third order framework

In the lines of Proposition 2 we may further say that:

Theorem 1. *If, with the same notation as before, we further assume the general third order framework in (3.1), and if we choose the tuning parameter* α *such that*

$$b_\alpha(\rho) = \alpha\overline{\mu}^{(2)}_{2\alpha}(\rho) - (\alpha-1)\ \overline{\mu}^{(2)}_{\alpha-1}(\rho) = 0,$$

the following asymptotic distributional representation,

$$\begin{aligned}\widehat{\gamma}^{(\alpha)}_n(k) \overset{d}{=} \gamma + \frac{\gamma\sqrt{v_\alpha}}{\sqrt{k}}\ Z^{(\alpha)}_n + \frac{A^2(n/k)}{2\ \gamma}\ &\Big(\alpha(2\alpha-1)\overline{\mu}^{(3)}_{2\alpha}(\rho) - \alpha^2\left(\overline{\mu}^{(2)}_{2\alpha}(\rho)\right)^2 \\ &-(\alpha-1)(\alpha-2)\ \overline{\mu}^{(3)}_{\alpha-1}(\rho) + 2(\alpha-1)^2\left(\overline{\mu}^{(2)}_{\alpha-1}(\rho)\right)^2\Big)\ (1+o_p(1)) \\ &+ A(n/k)\ B(n/k)\left(\alpha\overline{\mu}^{(2)}_{2\alpha}(2\rho) - (\alpha-1)\ \overline{\mu}^{(2)}_{\alpha-1}(2\rho)\right)\ (1+o_p(1)), \quad (3.2)\end{aligned}$$

holds. Consequently $\sqrt{k}\left\{\widehat{\gamma}^{(\alpha)}_n(k) - \gamma\right\}$ *is asymptotically normal with null mean value, not only when* $\sqrt{k}\ A(n/k) \to 0$, *but also whenever* $\sqrt{k}\ A(n/k) \to \lambda$, *finite or infinite, provided that* $\sqrt{k}\ A^2(n/k) \to 0$, *as* $n \to \infty$. *If* $\sqrt{k}\ A^2(n/k) \to \lambda_A$, *finite, and* $\sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_B$, *also finite, there is a non-null asymptotic bias given by*

$$\begin{aligned}\frac{\lambda_A}{2\ \gamma}\ &\Big(\alpha(2\alpha-1)\overline{\mu}^{(3)}_{2\alpha}(\rho) - \alpha^2(\overline{\mu}^{(2)}_{2\alpha}(\rho))^2 - (\alpha-1)(\alpha-2)\ \overline{\mu}^{(3)}_{\alpha-1}(\rho) \\ &+2(\alpha-1)^2\left(\overline{\mu}^{(2)}_{\alpha-1}(\rho)\right)^2\Big) + \lambda_B\left(\alpha\overline{\mu}^{(2)}_{2\alpha}(2\rho) - (\alpha-1)\ \overline{\mu}^{(2)}_{\alpha-1}(2\rho)\right). \quad (3.3)\end{aligned}$$

Proof. Under the third order condition in (3.1), assuming that (1.5) holds, and using the same arguments as in Dekkers *et al.* (1989), in lemma 2 of Draisma *et al.* (1999) and more recently in Gomes *et al.* (2002) and Fraga Alves *et al.* (2003), we may write the distributional representation

$$\begin{aligned}\left(\frac{M^{(2\alpha)}_n(k)}{\mu^{(1)}_{2\alpha}}\right)^{1/2} = \gamma^\alpha\ \Bigg(1 &+ \frac{\overline{\sigma}^{(1)}_{2\alpha}}{2}\ \frac{P^{(2\alpha)}_n}{\sqrt{k}} + \frac{\alpha\ \overline{\mu}^{(2)}_{2\alpha}(\rho)\ A(n/k)}{\gamma} + \frac{\alpha\ \overline{\sigma}^{(2)}_{2\alpha}(\rho)}{\gamma}\ \frac{A(n/k)\ \overline{P}^{(2\alpha)}_n}{\sqrt{k}} \\ &+ \frac{\alpha\ A^2(n/k)}{2\ \gamma^2}\ \left((2\alpha-1)\overline{\mu}^{(3)}_{2\alpha}(\rho) - \alpha\left(\overline{\mu}^{(2)}_{2\alpha}(\rho)\right)^2\right)(1+o_p(1)) \\ &+ \frac{\alpha}{\gamma}\ \overline{\mu}^{(2)}_{2\alpha}(2\rho)\ A(n/k)\ B(n/k)\ (1+o_p(1))\Bigg),\end{aligned}$$

where $P^{(2\alpha)}_n$ and $\overline{P}^{(2\alpha)}_n$ are asymptotically standard Normal r.v.'s.

Also

$$\left(\frac{M_n^{(\alpha-1)}(k)}{\mu_{\alpha-1}^{(1)}}\right)^{-1} = \gamma^{1-\alpha}\left(1-\overline{\sigma}_{\alpha-1}^{(1)}\frac{P_n^{(\alpha-1)}}{\sqrt{k}}\right.$$

$$-\frac{\alpha-1}{\gamma}\ \overline{\mu}_{\alpha-1}^{(2)}(\rho)\ A(n/k) - \frac{\alpha-1}{\gamma}\ \overline{\sigma}_{\alpha-1}^{(2)}(\rho)\ \frac{A(n/k)}{\sqrt{k}}\ \overline{P}_n^{(\alpha-1)}$$

$$-\frac{\alpha-1}{2\gamma^2}\ A^2(n/k)\left((\alpha-2)\ \overline{\mu}_{\alpha-1}^{(3)}(\rho) - 2(\alpha-1)\ (\overline{\mu}_{\alpha-1}^{(2)}(\rho))^2\right)(1+o_p(1))$$

$$\left.-\frac{\alpha-1}{\gamma}\ \overline{\mu}_{\alpha-1}^{(2)}(2\rho)\ A(n/k)\ B(n/k)\ (1+o_p(1))\right).$$

Then

$$\widehat{\gamma}_n^{(\alpha)}(k) \stackrel{d}{=} \gamma\ \left(1+\frac{1}{\sqrt{k}}\left(\frac{\overline{\sigma}_{2\alpha}^{(1)}\ P_n^{(2\alpha)}}{2} - \overline{\sigma}_{\alpha-1}^{(1)}P_n^{(\alpha-1)}\right)\right.$$

$$+\frac{A(n/k)}{\gamma}\ \left(\alpha\overline{\mu}_{2\alpha}^{(2)}(\rho) - (\alpha-1)\ \overline{\mu}_{\alpha-1}^{(2)}(\rho)\right)$$

$$+\frac{A(n/k)}{\gamma\sqrt{k}}\left(\alpha\overline{\sigma}_{2\alpha}^{(2)}(\rho)\ \overline{P}_n^{(2\alpha)} - (\alpha-1)\ \overline{\sigma}_{\alpha-1}^{(2)}(\rho)\ \overline{P}_n^{(\alpha-1)}\right)$$

$$+\frac{A^2(n/k)}{2\ \gamma^2}\ \left(\alpha(2\alpha-1)\overline{\mu}_{2\alpha}^{(3)}(\rho) - \alpha^2\left(\overline{\mu}_{2\alpha}^{(2)}(\rho)\right)^2\right.$$

$$\left.-(\alpha-1)(\alpha-2)\ \overline{\mu}_{\alpha-1}^{(3)}(\rho) + 2(\alpha-1)^2\left(\overline{\mu}_{\alpha-1}^{(2)}(\rho)\right)^2\right)(1+o_p(1))$$

$$\left.+\frac{A(n/k)\ B(n/k)}{\gamma}\left(\alpha\overline{\mu}_{2\alpha}^{(2)}(2\rho) - (\alpha-1)\ \overline{\mu}_{\alpha-1}^{(2)}(2\rho)\right)\ (1+o_p(1))\right).$$

If we choose α such that $\alpha\overline{\mu}_{2\alpha}^{(2)}(\rho) - (\alpha-1)\ \overline{\mu}_{\alpha-1}^{(2)}(\rho) = 0$, the result in (3.2) follows immediately: the rate of convergence is still of the order of $\frac{1}{\sqrt{k}}$, and the asymptotic bias is null not only when $\sqrt{k}\ A(n/k) \to 0$ but also when $\sqrt{k}\ A(n/k) \to \lambda \neq 0$. The bias in (3.3) comes also straigthforwardly, and is equal to $\lambda_A\ u_A$ if $|\rho| < |\rho'|$, since then $B(t) = o(A(t))$, being equal to $\lambda_B\ u_B$ whenever $|\rho| > |\rho'|$, due to the fact that then $A(t) = o(B(t))$, as $t \to \infty$.

□

Remark 3. *The experience we have from the second order framework leads us to state that the minimum asymptotic mean squared error is now going to be attained whenever* $\sqrt{k}\ A(n/k) \to \infty$ *and*

$$\begin{cases} \sqrt{k}\ A^2(n/k) \to \lambda_A \neq 0,\ \textit{finite} & \textit{if}\ \ |\rho| \leq |\rho'| \\ \sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_B \neq 0,\ \textit{finite} & \textit{if}\ \ |\rho| > |\rho'| \end{cases}.$$

Note that if $\rho = \rho'$ *and* $\sqrt{k}\ A^2(n/k) \to \lambda_A \neq 0$, *finite, then,* $\sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_B \neq 0$, *also finite.*

3.2 The estimation of the second order parameter

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, which show highly stable sample paths as functions of k, the number of top o.s. used, for a wide range of large k-values. Such a class of estimators is parameterised in a tuning parameter τ and depends on the statistics

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

which converge towards $3(1-\rho)/(3-\rho)$, independently of τ, whenever the second order condition (1.2) holds and k is such that $k = o(n)$ and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$. The estimators are thus given by

$$\widehat{\rho}_n^{(\tau)}(k) := \min\left(0,\ \frac{3(T_n^{(\tau)}(k) - 1)}{T_n^{(\tau)}(k) - 3}\right). \tag{3.4}$$

We shall formalize, without proofs, the main distributional results of the estimators in (3.4). Proofs may be found in Fraga Alves *et al.* (2003).

Proposition 3. *If the first order condition (1.1) holds, k is a sequence of intermediate integers, i.e., (1.5) holds, and if $\lim_{n\to\infty} \sqrt{k}\ A(n/k) = \infty$, then $\widehat{\rho}_n^{(\tau)}(k)$ in (3.4) is consistent for the estimation of ρ, i.e., converges in probability towards ρ, as $n \to \infty$.*

Proposition 4. *If the third order condition (3.1) holds, k is a sequence of intermediate integers, i.e., (1.5) holds, if $\lim_{n\to\infty} \sqrt{k}\ A(n/k) = \infty$, but $\lim_{n\to\infty} \sqrt{k}\ A^2(n/k) = \lambda_1$, finite and $\lim_{n\to\infty} \sqrt{k}\ A(n/k)\ B(n/k) = \lambda_2$, finite, we may guaranteee asymptotic normality of the estimators $\widehat{\rho}_n^{(\tau)}(k)$ in (3.4), and $\widehat{\rho}_n^{(\tau)}(k) - \rho = O_p\left(1/\left(\sqrt{k}A(n/k)\right)\right)$.*

Remark 4. *The theoretical and simulated results in Fraga Alves et al. (2003), together with the use in these estimators in the Generalized Jackknife statistics of Gomes et al. (2000) (Gomes and Martins, 2002) led us to advise in practice the consideration of the level*

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

(not chosen in any optimal way), and of the tuning parameters $\tau = 0$ for the region $\rho \in [-1, 0)$ and $\tau = 1$ for the region $\rho \in (-\infty, -1)$. Anyway, we advise practitioners not to choose blindly the value of τ in (3.4). It is sensible to draw a few sample paths of $\widehat{\rho}_n^{(\tau)}(k)$ in (3.4), as functions of k, and for a few values of τ, like $\tau = 0, 0.5, 1$, electing the value of τ which provides higher stability for large k, by means of any stability criterion. The value of k_1 in (3.5)*, although not chosen in any optimal way, may be considered for any chosen value of the tuning parameter τ.*

4 Distributional properties of the "asymptotically unbiased" estimators of the tail index, based on the external estimation of ρ

We shall now consider the class of estimators in (1.6), where $\widehat{\rho} \equiv \widehat{\rho}_\tau = \widehat{\rho}_n^{(\tau)}(k_1)$ is any of the estimators in (3.4), with k_1 in (3.5).

Theorem 2. *If the second order condition (1.2) holds, if $k = k_n$ is a sequence of intermediate positive integers, i.e., (1.5) holds, and if $\sqrt{k}\ A(n/k) \underset{n\to\infty}{\longrightarrow} \lambda$, finite, non necessarily null, then $\sqrt{k}\left(\widehat{\gamma}_n^{(\widehat{\alpha})}(k) - \gamma\right)$ are asymptotically normal with null mean value. Indeed, with Z_k denoting an asymptotically standard normal r.v., we have, for the tail index estimators under study,*

$$\widehat{\gamma}_n^{(\widehat{\alpha})}(k) \stackrel{d}{=} \gamma + \frac{\gamma\ \varphi(\rho)}{\sqrt{k}}\ Z_k + o_p(A(n/k)), \tag{4.1}$$

with $\varphi(\rho) = v_{\alpha_0}$, v_α given in (2.2) and $\alpha_0 = \alpha_0(\rho)$ given in (2.3).
More than this: the results in Theorem 1 hold true if we replace $\widehat{\gamma}_n^{(\alpha)}(k)$ by $\widehat{\gamma}_n^{(\widehat{\alpha})}(k)$ in (1.6)*, where $\widehat{\rho} \equiv \widehat{\rho}_\tau = \widehat{\rho}_n^{(\tau)}(k_1)$ is any of the estimators in (3.4), with k_1 in (3.5).*

Proof. The proof for the r.v. $\widehat{\gamma}_n^{(\alpha_0)}(k)$, $\alpha_0 = \alpha_0(\rho)$ follows the same lines of the proof of Theorem 1, and we get straightforwardly, as a particular case of (3.2),

$$\widehat{\gamma}_n^{(\alpha_0)}(k) \stackrel{d}{=} \gamma + \frac{\gamma\ \varphi(\rho)}{\sqrt{k}}\ Z_k + o_p(A(n/k)).$$

The result in (4.1) is due to the fact that we may write

$$\widehat{\gamma}_n^{(\widehat{\alpha})}(k) = \gamma_n^{(\alpha_0)}(k) + (\widehat{\rho}_\tau - \rho)\ \xi(k)\ (1 + o_p(1)),$$

with $\widehat{\rho}_\tau - \rho = o_p(1)$ for all k in the conditions of the theorem, and $\xi(k) = O_p\left(1/\sqrt{k}\right)$.
The last part of the theorem comes from the fact that, due to the choice of the

level k_1 used in the estimation of ρ, $\widehat{\rho} - \rho = O\left((\ln\ln n)^{1/(1-2\rho)}/\sqrt{n}\right)$. Consequently, if $|\rho| \leq |\rho'|$ and $\sqrt{k}\ A^2(n/k) \to \lambda_{_A}$, finite, $n/k = O\left(n^{1/(1-4\rho)}\right)$ and $\widehat{\rho} - \rho = o_p\left(A^2(n/k)\right)$. Analogously, if $|\rho| > |\rho'|$ and $\sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_{_B}$, finite, $n/k = O\left(n^{1/(1-2(\rho+\rho'))}\right)$ and $\widehat{\rho} - \rho = o_p\left(A(n/k)\ B(n/k)\right)$. □

Remark 5. *Note that should we have worked with $\widehat{\rho}(k)$, computed at the same level k, the fact that $\widehat{\rho}(k) - \rho = O_p\left(1/\left(\sqrt{k}\ A(n/k)\right)\right) = O_p(1)$ would lead to an increase in the asymptotic variance of our tail index estimator.*

5 Exact distributional properties of our second class of estimators — a simulation study

We are going to consider here the *Fréchet* model, $F(x) = e^{-x^{-1/\gamma}}$, $x \geq 0$, with $\gamma = 1$ (for which $\rho = -1$) and *Burr* models, $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, $\rho < 0$, with $\gamma = 1$. The comparison will be performed at the optimal levels, i.e., at levels k_0 minimizing the mean squared errors, considering the r.v.'s $\widehat{\gamma}_{n0}^{(\alpha_0)} \equiv \widehat{\gamma}_{n0}^{(\alpha_0(\rho))}$ and $\widehat{\gamma}_{n0}^{ML(\rho)}$ and the tail index estimators $\widehat{\gamma}_{n0}^{(\widehat{\alpha}_0)} \equiv \widehat{\gamma}_{n0}^{(\alpha_0(\widehat{\rho}_0))}$ and $\widehat{\gamma}_{n0}^{ML(\widehat{\rho}_0)}$.

The measure of efficiency is

$$REFF_{\widehat{\gamma}_n^{\bullet}|\widehat{\gamma}_n^H} := \sqrt{MSE\left(\widehat{\gamma}_{n0}^{H}\right)/MSE\left(\widehat{\gamma}_{n0}^{\bullet}\right)}.$$

In Table 1 we present, for each model, the relative efficiencies of $\widehat{\gamma}_{n0}^{(\alpha_0(\rho))}$ and of $\widehat{\gamma}_{n0}^{(\alpha_0(\widehat{\rho}))}$, in the first and second row, respectively. The relative efficiencies of $\widehat{\gamma}_{n0}^{ML(\rho)}$ and of $\widehat{\gamma}_{n0}^{ML(\widehat{\rho})}$ are presented in the third and fourth rows, respectively. We have chosen $\tau = 0$ for all models, although we get to know, from practice, that $\tau = 1$ provides better results (in the sense of more stable sample paths as function of k, whenever k is large) for values of $|\rho| > 1$. The simulation is based on 10 replicates of 1000 runs each.

In Figures 2, 3 and 4 we illustrate for a sample size $n = 1000$, and for the Fréchet($\gamma = 1$), the Burr($\gamma = 1$, $\rho = -0.5$) and the Burr($\gamma = 1$, $\rho = -2$) models, respectively, the mean value and the mean squared errors functions of $\widehat{\gamma}_n^{(\alpha_0(\rho))}$, $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}))}$, $\widehat{\gamma}_n^{ML(\rho)}$ and $\widehat{\gamma}_n^{ML(\widehat{\rho})}$, denoted $\alpha_0(\rho)$, $\alpha_0(\widehat{\rho})$, $ML(\rho)$ and $ML(\widehat{\rho})$, respectively. For comparison, we present also the Hill estimator, denoted H.

To illustrate the type of the sample paths (the usual situation in practice), we have drawn Figure 5. From this figure it is possible to see that even for small values of ρ, like $\rho = -2$, the choice of $\widehat{\rho}_0$ instead of $\widehat{\rho}_1$ is also possible, and leads to similar

Table 1: Relative efficiencies for Fréchet and Burr models.

n	100	200	500	1000	2000
Fréchet parent: $\gamma = 1 \quad (\rho = -1)$					
	1.524 ± 0.034	1.668 ± 0.049	1.866 ± 0.030	1.989 ± 0.043	2.170 ± 0.050
	1.104 ± 0.009	1.125 ± 0.024	1.186 ± 0.016	1.242 ± 0.020	1.151 ± 0.019
	1.415 ± 0.015	1.468 ± 0.016	1.544 ± 0.014	1.601 ± 0.014	1.657 ± 0.014
	1.221 ± 0.012	1.215 ± 0.015	1.249 ± 0.014	1.278 ± 0.010	1.225 ± 0.013
Burr parent: $\gamma = 1, \quad \rho = -0.5$					
	2.447 ± 0.070	2.310 ± 0.090	2.217 ± 0.057	2.212 ± 0.036	2.184 ± 0.039
	1.762 ± 0.048	1.634 ± 0.057	1.521 ± 0.039	1.476 ± 0.022	1.409 ± 0.029
	1.310 ± 0.023	1.388 ± 0.016	1.490 ± 0.020	1.555 ± 0.018	1.603 ± 0.018
	1.10 ± 0.019	1.133 ± 0.013	1.170 ± 0.015	1.208 ± 0.014	1.213 ± 0.016
Burr parent: $\gamma = 1, \quad \rho = -1$					
	1.590 ± 0.051	1.507 ± 0.053	1.463 ± 0.028	1.444 ± 0.024	1.437 ± 0.021
	1.639 ± 0.047	1.646 ± 0.056	1.722 ± 0.032	1.814 ± 0.034	1.822 ± 0.036
	1.255 ± 0.015	1.309 ± 0.017	1.358 ± 0.011	1.395 ± 0.014	1.421 ± 0.014
	1.352 ± 0.015	1.437 ± 0.020	1.554 ± 0.015	1.654 ± 0.020	1.598 ± 0.012
Burr parent: $\gamma = 1, \quad \rho = -2$					
	1.173 ± 0.019	1.116 ± 0.029	1.092 ± 0.022	1.085 ± 0.019	1.082 ± 0.018
	1.008 ± 0.013	1.045 ± 0.025	1.140 ± 0.024	1.210 ± 0.020	1.224 ± 0.020
	1.155 ± 0.007	1.182 ± 0.015	1.203 ± 0.012	1.223 ± 0.013	1.244 ± 0.015
	1.266 ± 0.013	1.342 ± 0.018	1.485 ± 0.016	1.607 ± 0.020	1.445 ± 0.017

conclusions.

Table 1 and the figures suggest us the following comments related with the estimator in (1.6), with $\widehat{\rho} = \widehat{\rho}_0$:

1. For the *Fréchet* model (Figure 2) we need to use practically all o.s.'s, both whenever we assume ρ known or we estimate ρ adequately — for this model the bias component of the estimator seems to be almost negligible. At the optimal levels we are thus able to improve greatly the performance of the Hill estimator. The ML-estimator has a higher bias, but its smaller variance makes it slightly better than the estimator herewith proposed, when both are considered at their optimal levels.

2. For models with ρ close to 0 (Figure 3) — the most problematic zone for the Hill estimator —, the pattern of the mean squared error of our estimator is nicely U-shaped, and even with the second order parameter estimated in a non-optimal way, the estimator $\widehat{\gamma}_n^{(\widehat{\alpha})}$ is highly efficient relatively to the Hill estimator, when they are both considered at their optimal levels. More than that: the estimator herewith considered is more efficient than the Hill estimator at its optimal level for a wide range of k-values. It also compares favourably with the ML-estimator. As mentioned in Remark 1, the increase in variance is here highly compensated by a decrease in squared bias.

3. For small values of ρ, like the value $\rho = -2$, used here for illustration in Figure 4, the behaviour of $\widehat{\gamma}_n^{(\widehat{\alpha})}$ is not much better than that of $\widehat{\gamma}_n^H$, when we think on minimum mean squared errors. But just to be slightly better, in a region

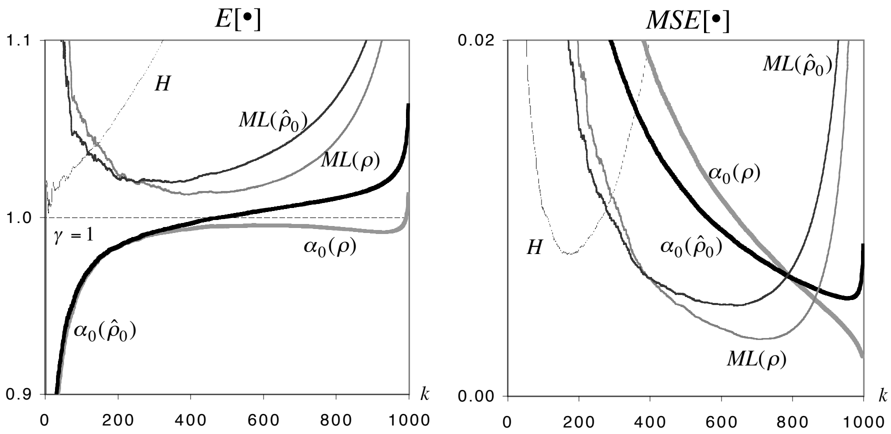

Figure 2: Mean values (*left*) and mean squared errors (*right*) of $\widehat{\gamma}_n^H(k)$, of the estimators $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))}$, $\widehat{\gamma}_n^{ML(\widehat{\rho}_0)}$ and of the r.v.'s $\widehat{\gamma}_n^{(\alpha_0(\rho))}$, $\widehat{\gamma}_n^{ML(\rho)}$, for sample size $n = 1000$ from a *Fréchet*$(\gamma = 1)$ model.

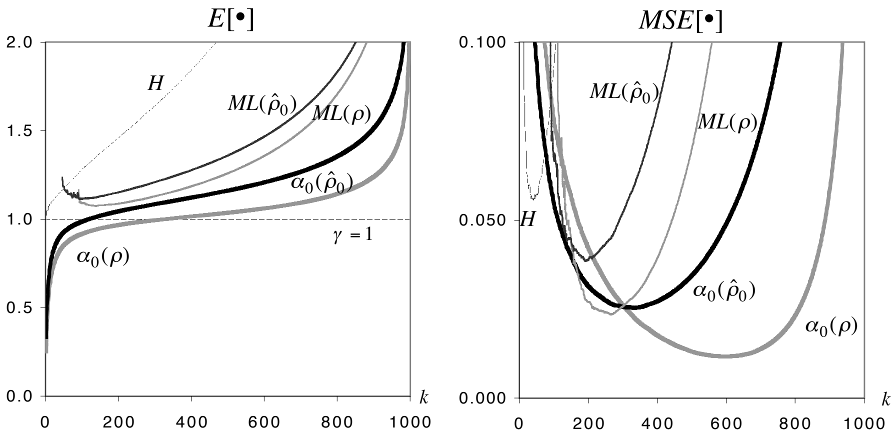

Figure 3: Mean values (*left*) and mean squared errors (*right*) of $\widehat{\gamma}_n^H(k)$, of the estimators $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))}$, $\widehat{\gamma}_n^{ML(\widehat{\rho}_0)}$ and of the r.v.'s $\widehat{\gamma}_n^{(\alpha_0(\rho))}$, $\widehat{\gamma}_n^{ML(\rho)}$, for sample size $n = 1000$ from a *Burr*$(\gamma = 1,\ \rho = -0.5)$ model.

of ρ-values where has been difficult to find alternatives to the Hill estimator which works pretty well, is already relevant. In this region the ML estimator performs better than the estimator in (1.6), but due to a very sharp mean squared error structure.

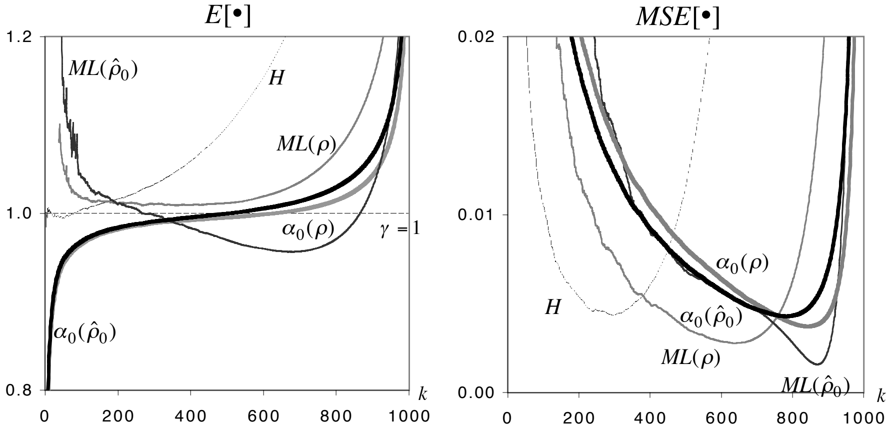

Figure 4: Mean values (*left*) and mean squared errors (*right*) of $\widehat{\gamma}_n^H(k)$, of the estimators $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))}$, $\widehat{\gamma}_n^{ML(\widehat{\rho}_0)}$ and of the r.v.'s $\widehat{\gamma}_n^{(\alpha_0(\rho))}$, $\widehat{\gamma}_n^{ML(\rho)}$, for sample size $n = 1000$ from a *Burr* model with $\gamma = 1$ and $\rho = -2$.

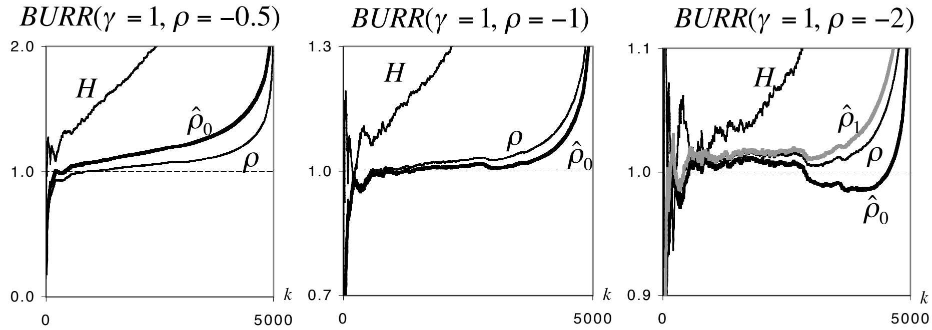

Figure 5: Sample paths of $\widehat{\gamma}_n^H(k)$, of the estimator $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))}(k)$ and of the r.v. $\widehat{\gamma}_n^{(\alpha_0(\rho))}(k)$, for samples of size $n = 5000$, from *Burr* models with $\gamma = 1$ and $\rho = -0.5,\ -1,\ -2$.

6 A case-study

We shall finally consider an illustration of the performance of the above mentioned estimators, reporting results associated to the Euro-US$ daily exchange rates from January 4, 1999 till December 15, 2003. This data has been collected by the European System of Central Banks, and was obtained from $http://www.bportugal.pt/rates/cambtx/$.

In Figure 6 we present the Daily Exchange Rates x_t over the above mentioned period and the Log-Returns, $r_t = 100 \times (\ln(x_t) - \ln(x_{t-1}))$, the data to be analyzed. The sample of log-returns has a size $n = 1266$, and these log-returns clearly exhibit graphically an indication of an underlying heavy-tailed model. In Figure 7, and working merely with the $n^+ = 611$ positive log-returns, we now present the sample path of the ρ_τ estimates (*left*), as function of k, for $\tau = 0$ and $\tau = 1$, together with the sample paths of the classical tail index estimate (H) and of the two "asymptotically unbiased" tail index estimates discussed in this paper, $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))}$ and $\widehat{\gamma}_n^{ML(\widehat{\rho}_0)}$ (*right*), associated to the value $\widehat{\rho} = \widehat{\rho}_0(611) = -0.69$.

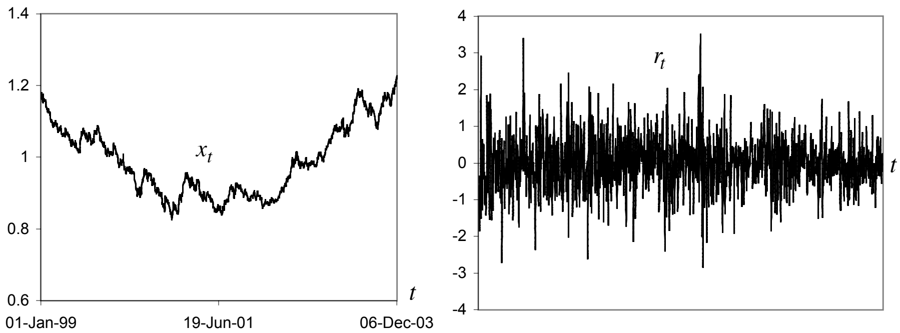

Figure 6: Daily Exchange Rates (*left*) and Daily Log-Returns (*right*) on Euro-US$ Exchange Rate.

Remark 6. *Note that the sample path of the ρ-estimates associated to $\tau = 0$ and $\tau = 1$ lead us to chosse, on the basis of any stability criterion for large k, the estimate associated to $\tau = 0$. From the experience we have with this type of estimates, this means that the underlying ρ-value is larger or equal to -1, and the consideration of $\tau = 0$ is then advisable.*

Remark 7. *Regarding the tail index estimation, note that whereas the Hill estimator is unbiased for the estimation of the tail index γ, whenever the underlying model is a Pareto model, with distribution $F(x) = 1 - x^{-1/\gamma}$, $x \geq 0$ ($\gamma > 0$), it exhibits a relevant bias when we have only Pareto-like tails, as happens here, and may be seen from Figure 7. The other estimators, which are "asymptotically unbiased" 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 the "largest run":*

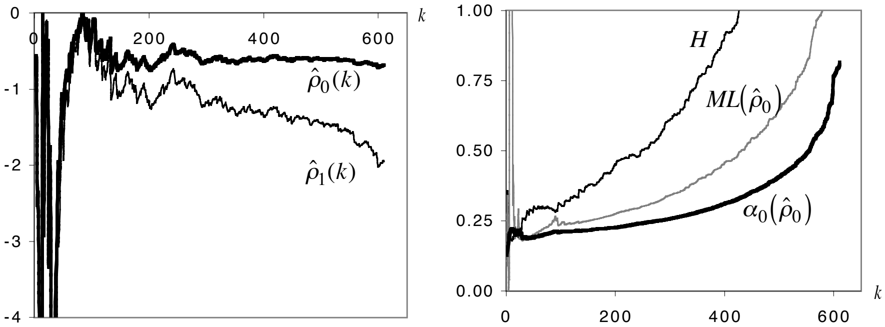

Figure 7: Estimates of the second order parameter ρ (*left*) and of the tail index γ (*right*) for the Daily Log-Returns on the Euro-US$.

here, if we consider the tail index estimates with 2 *decimal figures, the largest run is achieved by the sample path of the estimator in* (1.6)*. Such a run has a size equal to* 43 $(86 \leq k \leq 128)$*, and is associated to* $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))} = 0.21$*. Should we be happy with one decimal figure, would we also get the largest run for the same estimator: a run of size* 278 $(4 \leq k \leq 281)$ *associated to* $\widehat{\gamma}_n^{(\alpha_0(\widehat{\rho}_0))} = 0.2$*. For the ML-estimate, and with one decimal figure, we would also get a tail index estimate equal to* 0.2*, but with a run of size* 130*. With this same criterion, the Hill estimator would provide an estimate equal to* 0.3*, with a run of size* 102*.*

References

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

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

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

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

[5] Drees, H. (1998). A general class of estimators of the extreme value index. *J. Statist. Planning and Inference* **66**, 95-112.

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

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

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

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

[10] Gomes, M. I. and L. de Haan (1999). Approximation by penultimate extreme value distributions. *Extremes* **2**:1, 71-85.

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

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

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

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

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

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

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

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

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

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