

The Jackknife methodology in the estimation of a positive tail index*

M. Ivette Gomes

CEAUL and DEIO (FCUL), Universidade de Lisboa

Cristina Miranda

CEAUL and ISCA, Universidade de Aveiro

and

Clara Viseu

DEIO (FCUL), Universidade de Lisboa

Abstract. In a context of regularly varying tails, we first analyse a generalization of the classical Hill estimator of a positive tail index, with members that are not asymptotically more efficient than the original one. This has led us to propose alternative classical tail index estimators, that may perform asymptotically better than the Hill estimator. Since the improvement is not really significant, we also propose Generalized Jackknife estimators based on any two members of these two classes. These Generalized Jackknife estimators are compared with the Hill estimator, asymptotically, and for finite samples, through the use of Monte Carlo simulation. The finite sample behaviour of the new reduced bias' estimators is also illustrated through a practical example in the field of finance.

AMS 2000 subject classification. Primary 62G32, 62E20; Secondary 65C05.

Keywords and phrases. *Statistical Theory of Extremes, Semi-parametric estimation, Jackknife methodology.*

*Research partially supported by FCT / POCTI / FEDER.

1 Introduction and preliminaries

Let X_1, X_2, \dots, X_n be independent, identically distributed (i.i.d.) random variables (r.v.'s) with common distribution function (d.f.) F , with a Pareto-type tail, i.e., let us assume that there exists a positive constant γ , such that for large x , the tail function

$$\overline{F}(x) := 1 - F(x) = x^{-1/\gamma} L(x),$$

where $L(x)$ is a slowly varying function at infinity, i.e., for every $x > 0$,

$$L(tx)/L(t) \rightarrow 1 \text{ as } t \rightarrow \infty.$$

Consequently, $\overline{F} \in RV_{-1/\gamma}$, where RV_β stands for the class of regularly varying functions at infinity with index of regular variation equal to β , i.e., measurable functions g with infinite right endpoint, and such that $\lim_{t \rightarrow \infty} g(tx)/g(t) = x^\beta$, for all $x > 0$. F is thus in the max-domain of attraction of an *Extreme Value (EV)* d.f.,

$$EV_\gamma(x) := \exp \left\{ -(1 + \gamma x)^{-1/\gamma} \right\}, \quad x \geq -1/\gamma, \quad \gamma > 0.$$

We denote such a fact by $F \in \mathcal{DM}(EV_\gamma)$.

Then the scaled log-spacings

$$U_i := i \{ \ln X_{n-i+1:n} - \ln X_{n-i:n} \}, \quad 1 \leq i \leq k, \quad (1.1)$$

are approximately distributed as k independent exponential r.v.'s, with mean value equal to γ . This leads to the well known Hill estimator for γ (Hill, 1975),

$$\widehat{\gamma}_n^H(k) := \frac{1}{k} \sum_{i=1}^k U_i,$$

the average of the scaled log-spacings in (1.1), and we may think, more generally, on the following generalization of the Hill estimator,

$$\widehat{\gamma}_n^{(\alpha)}(k) := \frac{\alpha}{k} \sum_{i=1}^k \left(\frac{i}{k} \right)^{\alpha-1} U_i, \quad \alpha \geq 1, \quad \left[\widehat{\gamma}_n^{(1)} \equiv \widehat{\gamma}_n^H \right]. \quad (1.2)$$

We shall here also work with the class of estimators,

$$\tilde{\gamma}_n^{(\alpha)}(k) := -\frac{\alpha^2}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} \ln\left(\frac{i}{k}\right) U_i, \quad \alpha \geq 1. \quad (1.3)$$

Apart from the first order condition,

$$F \in \mathcal{D}_{\mathcal{M}}(EV_{\gamma}) \ (\gamma > 0) \quad \text{iff} \quad 1 - F \in RV_{-1/\gamma} \quad \text{iff} \quad U \in RV_{\gamma}, \quad (1.4)$$

where

$$U(t) := F^{\leftarrow}(1 - 1/t), \quad t > 1, \quad F^{\leftarrow}(u) = \inf\{x : F(x) \geq u\},$$

we shall assume, in order to be able to derive the asymptotic normality of the estimators under study, the validity of a second order condition. More specifically, we shall assume that there exists a function $A(\cdot)$ and a parameter $\rho < 0$ such that, for every $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^{\rho} - 1}{\rho}. \quad (1.5)$$

This condition has been widely accepted as the appropriate condition to specify the tail of a Pareto-type distribution. It holds true for most common heavy-tailed models, like the Fréchet, the Generalized Pareto, the Burr and the Student's t .

In section 2 of this paper we deal with the asymptotic behaviour of the classes of estimators in (1.2) and (1.3) and compare them asymptotically at their optimal levels. The value $\alpha = 1$ in (1.2), i.e., the Hill estimator, provides the optimal results within such a class. The estimators in (1.3) may perform better than the Hill estimator, but the improvement is not at all significant. This led us to the introduction, in section 3, of *Generalized Jackknife* estimators of γ associated to any pair of estimators in (1.2) and/or in (1.3). We finally come to a general class of Generalized Jackknife r.v.'s, dependent on a *tuning*

parameter α and on the second order parameter ρ in (1.5). Such a second order parameter is either adequately estimated through $\widehat{\rho}$, to be explicated later on, or misspecified at $\rho = -1$. We advance with different classes of reduced bias tail index estimators, but the one we consider to be more relevant is given by

$$\widehat{\gamma}_{n,\alpha}^{GJ(\widehat{\rho})}(k) = \frac{1}{\widehat{\rho}} \left\{ \alpha \widehat{\gamma}_n^{(\alpha)}(k) - (\alpha - \widehat{\rho}) \widetilde{\gamma}_n^{(\alpha)}(k) \right\}, \quad (1.6)$$

with $\widehat{\gamma}_n^{(\alpha)}(k)$ and $\widetilde{\gamma}_n^{(\alpha)}(k)$ given in (1.2) and (1.3), respectively. Indication on the choice of α in (1.6) is provided and the case $\widehat{\rho} \equiv -1$ in (1.6) is also considered. In section 4, we proceed to a comparison of the proposed estimators, for finite samples, through the use of Monte Carlo techniques. In section 5, we use the developed estimators for the analysis of real data in the field of finance. Proofs of the results in sections 2 and 3 will be postponed to section 6.

2 The asymptotic behaviour of the initial classes of tail index estimators

2.1 Asymptotic comparison at the same threshold k

The main result of this section is given in the following:

Theorem 2.1. *Under the first order condition in (1.4), and for $k = k_n$, an intermediate sequence of integers between 1 and n , i.e., such that*

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

the statistics $\widehat{\gamma}_n^{(\alpha)}(k)$ and $\widetilde{\gamma}_n^{(\alpha)}(k)$ in (1.2) and (1.3), respectively, are weakly consistent for the estimation of the tail index γ . Moreover, under the second order framework in (1.5), we have the validity of the asymptotic distributional representations

$$\widehat{\gamma}_n^{(\alpha)}(k) \stackrel{d}{=} \gamma + \frac{\gamma\alpha}{\sqrt{(2\alpha-1)k}} P_k^{(\alpha)} + \frac{\alpha}{\alpha-\rho} A(n/k)(1+o_p(1)), \quad (2.1)$$

and

$$\tilde{\gamma}_n^{(\alpha)}(k) \stackrel{d}{=} \gamma + \frac{\gamma\sqrt{2}\alpha^2}{(2\alpha-1)\sqrt{(2\alpha-1)k}} Q_k^{(\alpha)} + \frac{\alpha^2}{(\alpha-\rho)^2} A(n/k)(1+o_p(1)), \quad (2.2)$$

where

$$P_k^{(\alpha)} = \sqrt{(2\alpha-1)k} \left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} E_i - \frac{1}{\alpha} \right) \quad (2.3)$$

and

$$Q_k^{(\alpha)} = (2\alpha-1) \sqrt{\frac{(2\alpha-1)k}{2}} \left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} \ln\left(\frac{i}{k}\right) E_i + \frac{1}{\alpha^2} \right) \quad (2.4)$$

are asymptotically standard normal r.v.'s.

Remark 2.1. Note that, asymptotically, we may proceed to the comparison of the two classes of estimators at the same level k and for the same value of α . Indeed, we have the following approximations for bias and variance:

$$\begin{aligned} \text{Bias}_\infty [\hat{\gamma}_n^{(\alpha)}(k)] &= \frac{\alpha A(n/k)}{\alpha-\rho}, & \text{Var}_\infty [\hat{\gamma}_n^{(\alpha)}(k)] &= \frac{\alpha^2\gamma^2}{k(2\alpha-1)}, \\ \text{Bias}_\infty [\tilde{\gamma}_n^{(\alpha)}(k)] &= \frac{\alpha^2 A(n/k)}{(\alpha-\rho)^2}, & \text{Var}_\infty [\tilde{\gamma}_n^{(\alpha)}(k)] &= \frac{2\alpha^4\gamma^2}{k(2\alpha-1)^3}. \end{aligned}$$

Consequently, for every $\alpha \geq 1$ and $\rho < 0$,

$$\left| \text{Bias}_\infty [\tilde{\gamma}_n^{(\alpha)}(k)] \right| < \left| \text{Bias}_\infty [\hat{\gamma}_n^{(\alpha)}(k)] \right|.$$

If $1 \leq \alpha \leq 2$ (for all $\rho \leq 0$) or if $\alpha > 2$ and $\rho \leq -\alpha(\alpha-2)$, we get

$$\left| \text{Bias}_\infty [\tilde{\gamma}_n^{(\alpha)}(k)] \right| \leq \left| \text{Bias}_\infty [\hat{\gamma}_n^H(k)] \right| \leq \left| \text{Bias}_\infty [\hat{\gamma}_n^{(\alpha)}(k)] \right|.$$

If $\alpha > 2$ and $-\alpha(\alpha-2) < \rho \leq 0$,

$$\left| \text{Bias}_\infty [\hat{\gamma}_n^H(k)] \right| \leq \left| \text{Bias}_\infty [\tilde{\gamma}_n^{(\alpha)}(k)] \right| \leq \left| \text{Bias}_\infty [\hat{\gamma}_n^{(\alpha)}(k)] \right|.$$

Concerning the variance, we get, for $1 \leq \alpha < 1 + \sqrt{2}/2$,

$$\text{Var}_\infty [\hat{\gamma}_n^H(k)] \leq \text{Var}_\infty [\tilde{\gamma}_n^{(\alpha)}(k)] \leq \text{Var}_\infty [\hat{\gamma}_n^{(\alpha)}(k)],$$

and

$$\text{Var}_\infty [\widehat{\gamma}_n^H(k)] \leq \text{Var}_\infty [\widetilde{\gamma}_n^{(\alpha)}(k)] \leq \text{Var}_\infty [\widehat{\gamma}_n^{(\alpha)}(k)],$$

for $\alpha \geq 1 + \sqrt{2}/2$.

2.2 Asymptotic comparison at optimal levels

We now proceed to an asymptotic comparison of the estimators at their optimal levels in the lines of de Haan and Peng (1998). Suppose $\widehat{\gamma}_n^\bullet(k)$ is a general semi-parametric estimator of the tail index, for which the distributional representation

$$\widehat{\gamma}_n^\bullet(k) = \gamma + \frac{\sigma_\bullet}{\sqrt{k}} Z_n^\bullet + b_\bullet A(n/k) + o_p(A(n/k)) \quad (2.5)$$

holds for any intermediate k , and where Z_n^\bullet is an asymptotically standard normal r.v. Then we have

$$\sqrt{k} [\widehat{\gamma}_n^\bullet(k) - \gamma] \xrightarrow{d} N(\lambda b_\bullet, \sigma_\bullet^2), \text{ as } n \rightarrow \infty,$$

provided k is such that $\sqrt{k} A(n/k) \rightarrow \lambda$, finite, as $n \rightarrow \infty$. In this situation we write $\text{Bias}_\infty [\widehat{\gamma}_n^\bullet(k)] := b_\bullet A(n/k)$ and $\text{Var}_\infty [\widehat{\gamma}_n^\bullet(k)] := \sigma_\bullet^2/k$. The so-called Asymptotic Mean Squared Error (AMSE) is then given by

$$\text{AMSE} [\widehat{\gamma}_n^\bullet(k)] := \frac{\sigma_\bullet^2}{k} + b_\bullet^2 A^2(n/k).$$

Using regular variation theory, it may be proved that, whenever $b_\bullet \neq 0$, there exists a function $\varphi(n)$, dependent only on the underlying model, and not on the estimator, such that

$$\lim_{n \rightarrow \infty} \varphi(n) \text{AMSE} [\widehat{\gamma}_{n_0}^\bullet] = \frac{2\rho - 1}{2\rho} (\sigma_\bullet^2)^{-\frac{2\rho}{1-2\rho}} (b_\bullet^2)^{\frac{1}{1-2\rho}} =: \text{LMSE} [\widehat{\gamma}_{n_0}^\bullet], \quad (2.6)$$

where $\widehat{\gamma}_{n_0}^\bullet := \widehat{\gamma}_n^\bullet(k_0^\bullet(n))$ and $k_0^\bullet(n) := \arg \inf_k \text{AMSE} [\widehat{\gamma}_n^\bullet(k)]$.

It is then sensible to consider the following:

Definition 2.1. Given two biased estimators $\widehat{\gamma}_n^{(1)}(k)$ and $\widehat{\gamma}_n^{(2)}(k)$, for which distributional representations of the type (2.5) hold true, with constants (σ_1, b_1) and (σ_2, b_2) , $b_1, b_2 \neq 0$, respectively, both computed at their optimal levels, the Asymptotic Root Efficiency (AREFF) of $\widehat{\gamma}_{n_0}^{(1)}$ relatively to $\widehat{\gamma}_{n_0}^{(2)}$ is

$$AREFF_{1|2} \equiv AREFF_{\widehat{\gamma}_{n_0}^{(1)}|\widehat{\gamma}_{n_0}^{(2)}} := \sqrt{\frac{LMSE[\widehat{\gamma}_{n_0}^{(2)}]}{LMSE[\widehat{\gamma}_{n_0}^{(1)}]}},$$

with LMSE given in (2.6).

Remark 2.2. Note that this measure was devised so that the higher the AREFF measure, the better the estimator is.

Proposition 2.1. For every $\alpha > 1$, if we compare the estimators within the class $\widehat{\gamma}_n^{(\alpha)}$, in (1.2), with the Hill estimator, which is associated to $\alpha = 1$ in (1.2), we get

$$\widehat{AREFF}_{\alpha|H} = \left(\frac{(\alpha - \rho)(2\alpha - 1)^{-\rho}}{(1 - \rho) \alpha^{1-2\rho}} \right)^{\frac{1}{1-2\rho}} < 1.$$

More than this: both the asymptotic bias and the asymptotic variance of $\widehat{\gamma}_n^{(\alpha)}(k)$, $\alpha > 1$, are bigger than the asymptotic bias and the asymptotic variance, respectively, of the Hill estimator. If we compare $\widetilde{\gamma}_n^{(\alpha)}$ in (1.3) with the Hill estimator, the best estimator within the class $\widehat{\gamma}_n^{(\alpha)}$ in (1.2), we get

$$\widetilde{AREFF}_{\alpha|H} = \left(\frac{(\alpha - \rho)^2(2\alpha - 1)^{-3\rho}}{(1 - \rho) 2^{-\rho} \alpha^{2(1-2\rho)}} \right)^{\frac{1}{1-2\rho}}.$$

This $\widetilde{AREFF}_{\alpha|H}$ measure is presented in Figure 1, being there denoted $AREFF_{\alpha|H}$, for sake of simplicity. As may be seen from Figure 1, the gain in efficiency is smaller than 1.05, when we consider $\widetilde{\gamma}_n^{(\alpha)}$ in (1.3) instead of the Hill estimator. Consequently, in the following section, we are going to consider, for

real tuning parameter $\alpha, \beta \geq 1$, Generalized Jackknife estimators associated to any pair $(\hat{\gamma}_n^{(\alpha)}, \hat{\gamma}_n^{(\beta)})$, $\alpha \neq \beta$, in a way similar to the one considered in Gomes *et al.* (2005), as well as to any of the pairs $(\hat{\gamma}_n^{(\alpha)}, \tilde{\gamma}_n^{(\beta)})$, $\alpha, \beta \geq 1$, and $(\tilde{\gamma}_n^{(\alpha)}, \tilde{\gamma}_n^{(\beta)})$, $\alpha \neq \beta$, with $\hat{\gamma}_n^{(\alpha)}$ and $\tilde{\gamma}_n^{(\alpha)}$ given in (1.2) and (1.3), respectively. For details on the Generalized Jackknife estimation see Gray and Schucany (1972).

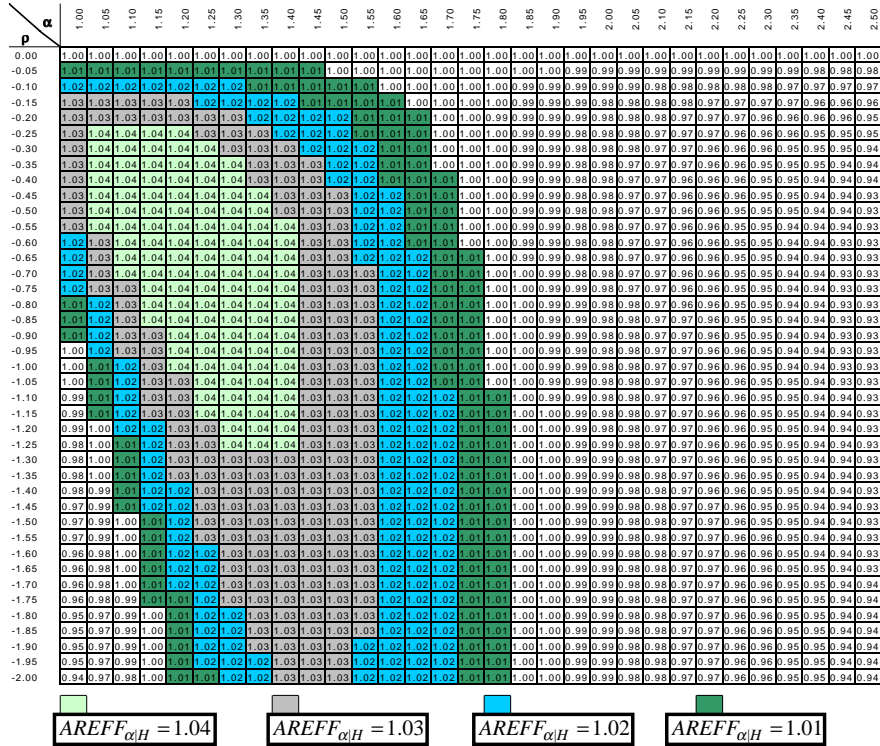


Figure 1: Asymptotic relative efficiency of $\tilde{\gamma}_{n0}^{(\alpha)}$ relatively to $\hat{\gamma}_{n0}^H$.

3 Generalized Jackknife estimators of the tail index

For real values $\alpha, \beta \geq 1$, let us think on the pairs, $(\hat{\gamma}_n^{(\alpha)}(k), \hat{\gamma}_n^{(\beta)}(k))$, $\beta \neq \alpha$, $(\tilde{\gamma}_n^{(\alpha)}(k), \tilde{\gamma}_n^{(\beta)}(k))$, and $(\tilde{\gamma}_n^{(\alpha)}(k), \tilde{\gamma}_n^{(\beta)}(k))$, $\beta \neq \alpha$, with $\hat{\gamma}_n^{(\bullet)}$ and $\tilde{\gamma}_n^{(\bullet)}$ given in (1.2) and (1.3), respectively.

The quotients between the dominant component of bias of the estimators in these pairs are given by

$$q_{\alpha,\beta}^{(1)}(\rho) = \frac{\alpha(\beta - \rho)}{\beta(\alpha - \rho)}, \quad \alpha \neq \beta, \quad q_{\alpha,\beta}^{(2)}(\rho) = \frac{\alpha(\beta - \rho)^2}{\beta^2(\alpha - \rho)}$$

and

$$q_{\alpha,\beta}^{(3)}(\rho) = \frac{\alpha^2(\beta - \rho)^2}{\beta^2(\alpha - \rho)^2}, \quad \alpha \neq \beta,$$

respectively, all dependent on ρ , unknown. We may thus estimate ρ adequately, either internally, at the same level k used for the estimation of γ , as done in Beirlant *et al.* (1999) and Feuerverger and Hall (1999) or externally, at a larger level than the one used for the estimation of γ , as done successfully in Gomes and Martins (2002), Beirlant *et al.* (2002), Caeiro and Gomes (2002b) and Gomes *et al.* (2004), through any of the ρ -estimators available in the literature, like the ones in Gomes *et al.* (2002a) and Fraga Alves *et al.* (2003). We may also misspecify ρ , for instance in -1 , a central and prominent value of this second order parameter, or in $\rho = -\gamma$, like happens for the *Generalized Pareto* models, as done before in several papers, among which we mention Feuerverger and Hall (1999), Gomes *et al.* (2000, 2002b), Caeiro and Gomes (2002a) and Gomes and Martins (2004).

3.1 Estimation of ρ

Let us assume first that we estimate ρ through an estimator $\hat{\rho}$, adequately chosen so that $\hat{\rho} - \rho = o_p(1)$ at any of the levels on which we are going to base the estimation of the tail index γ . We thus get the Generalized Jackknife classes of estimators, dependent on the tuning parameters α and β ($\alpha, \beta \geq 1$):

$$\begin{aligned} \hat{\gamma}_{n,\alpha,\beta}^{GJ_1(\hat{\rho})}(k) &:= \frac{\hat{\gamma}_n^{(\alpha)}(k) - q_{\alpha,\beta}^{(1)}(\hat{\rho}) \hat{\gamma}_n^{(\beta)}(k)}{1 - q_{\alpha,\beta}^{(1)}(\hat{\rho})} \\ &= \frac{\beta(\alpha - \hat{\rho}) \hat{\gamma}_n^{(\alpha)}(k) - \alpha(\beta - \hat{\rho}) \hat{\gamma}_n^{(\beta)}(k)}{\hat{\rho}(\alpha - \beta)}, \quad \alpha \neq \beta, \end{aligned} \quad (3.1)$$

$$\begin{aligned}
\widehat{\gamma}_{n,\alpha,\beta}^{GJ_2(\widehat{\rho})}(k) &:= \frac{\widehat{\gamma}_n^{(\alpha)}(k) - q_{\alpha,\beta}^{(2)}(\widehat{\rho}) \widetilde{\gamma}_n^{(\beta)}(k)}{1 - q_{\alpha,\beta}^{(2)}(\widehat{\rho})} \\
&= \frac{\beta^2(\alpha - \widehat{\rho}) \widehat{\gamma}_n^{(\alpha)}(k) - \alpha(\beta - \widehat{\rho})^2 \widetilde{\gamma}_n^{(\beta)}(k)}{\rho(2\alpha\beta - \beta^2 - \alpha\widehat{\rho})} \quad (3.2)
\end{aligned}$$

and

$$\begin{aligned}
\widehat{\gamma}_{n,\alpha,\beta}^{GJ_3(\widehat{\rho})}(k) &:= \frac{\widetilde{\gamma}_n^{(\alpha)}(k) - q_{\alpha,\beta}^{(3)}(\widehat{\rho}) \widetilde{\gamma}_n^{(\beta)}(k)}{1 - q_{\alpha,\beta}^{(3)}(\widehat{\rho})} \\
&= \frac{\beta^2(\alpha - \widehat{\rho})^2 \widetilde{\gamma}_n^{(\alpha)}(k) - \alpha^2(\beta - \widehat{\rho})^2 \widetilde{\gamma}_n^{(\beta)}(k)}{\widehat{\rho}(\alpha - \beta)(2\alpha\beta - \widehat{\rho}(\alpha + \beta))}, \quad \alpha \neq \beta. \quad (3.3)
\end{aligned}$$

Remark 3.1. For $\widehat{\rho}$ we may choose the estimator first used in Gomes and Martins (2002), which is based on the class of estimators in Fraga Alves et al. (2003). More specifically, we shall here consider, for any $\tau > 0$, the estimator

$$\widehat{\rho}_\tau := - \left| \frac{3(T_n^{(\tau)}(k_1) - 1)}{(T_n^{(\tau)}(k_1) - 3)} \right|, \quad k_1 = \min \left(n - 1, \left\lceil \frac{2n}{\ln \ln n} \right\rceil \right), \quad (3.4)$$

where

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

with

$$M_n^{(j)}(k) := \frac{1}{k} \sum_{i=1}^k \left(\ln \frac{X_{n-i+1:n}}{X_{n-k:n}} \right)^j, \quad j = 1, 2, 3.$$

We have $\widehat{\rho}_\tau - \rho = O_p(1/(\sqrt{k_1} A(n/k_1))) = o_p(\sqrt{k} A(n/k))$ for any τ and for any level k such that $\sqrt{k} A(n/k) \rightarrow \lambda$, finite. Indeed, we may even go further on, and consider adequate levels k such that $\sqrt{k} A(n/k) \rightarrow \infty$, as $n \rightarrow \infty$. The complete study of the estimators under this context would then imply the need of

a third order framework, like the one considered in Gomes and de Haan (1999), Caeiro and Gomes (2002b), Gomes et al. (2002a), Fraga Alves et al. (2003) and Gomes et al. (2004).

Theorem 3.1. *Let us assume that we are under the conditions of Theorem 2.1, and consider $\hat{\rho}$ replaced by ρ in (3.1), (3.2) and (3.3). We have, for $i = 1, 2, 3$,*

$$\widehat{\gamma}_{n,\alpha,\beta}^{GJ_i(\rho)}(k) \stackrel{d}{=} \gamma + \frac{\gamma \sigma_{\alpha,\beta}^{GJ_i(\rho)}}{\sqrt{k}} Z_{k,\alpha,\beta}^{GJ_i(\rho)} + o_p(A(n/k)), \quad (3.5)$$

with $Z_{k,\alpha,\beta}^{GJ_i(\rho)}$ asymptotically standard normal. The asymptotic standard deviations are given by

$$\sigma_{\alpha,\beta}^{GJ_1(\rho)} = \frac{\alpha \beta}{|\rho|(\beta - \alpha)} \sqrt{\frac{d_{\alpha,\beta}^{(1)}(\rho)}{(2\alpha - 1)(2\beta - 1)(\alpha + \beta - 1)}}, \quad (3.6)$$

$$\sigma_{\alpha,\beta}^{GJ_2(\rho)} = \frac{\alpha \beta^2}{|\rho|(\beta^2 - 2\alpha\beta + \alpha\rho)|(2\beta - 1)(\alpha + \beta - 1)} \sqrt{\frac{d_{\alpha,\beta}^{(2)}(\rho)}{(2\alpha - 1)(2\beta - 1)}}, \quad (3.7)$$

and

$$\sigma_{\alpha,\beta}^{GJ_3(\rho)} = \frac{\sqrt{2}\alpha^2 \beta^2}{|\rho|(\beta - \alpha)(2\alpha\beta - \rho(\alpha + \beta))} \sqrt{\frac{d_{\alpha,\beta}^{(3)}(\rho)}{(2\alpha - 1)^3(2\beta - 1)^3(\alpha + \beta - 1)^3}}, \quad (3.8)$$

where

$$\begin{aligned} d_{\alpha,\beta}^{(1)}(\rho) &= (\alpha - \rho)^2(2\beta - 1)(\alpha + \beta - 1) + (\beta - \rho)^2(2\alpha - 1)(\alpha + \beta - 1) \\ &\quad - 2(\alpha - \rho)(\beta - \rho)(2\alpha - 1)(2\beta - 1), \end{aligned}$$

$$\begin{aligned} d_{\alpha,\beta}^{(2)}(\rho) &= (\alpha - \rho)^2(2\beta - 1)^3(\alpha + \beta - 1)^2 + 2(\beta - \rho)^4(2\alpha - 1)(\alpha + \beta - 1)^2 \\ &\quad - 2(\alpha - \rho)(\beta - \rho)^2(2\alpha - 1)(2\beta - 1)^3 \end{aligned}$$

and

$$d_{\alpha,\beta}^{(3)}(\rho) = (\alpha - \rho)^4(2\beta - 1)^3(\alpha + \beta - 1)^3 + (\beta - \rho)^4(2\alpha - 1)^3(\alpha + \beta - 1)^3 \\ - 2(\alpha - \rho)^2(\beta - \rho)^2(2\alpha - 1)^3(2\beta - 1)^3.$$

The same distributional results in (3.5) hold true for the tail index estimators in (3.1), (3.2) and (3.3), if we use an estimator $\hat{\rho}$ like the ones in (3.4), i.e., such that $\hat{\rho} - \rho = o_p(1)$ for levels k such that $\sqrt{k} A(n/k) \rightarrow \lambda$, finite.

In Figures 2 and 3, we present the asymptotic standard deviations in (3.6) and (3.7), respectively, for $\rho = -1$. Similar patterns have been obtained for other values of ρ .

If we look at the Figures 2 and 3, we notice that the use of the GJ_1 and the GJ_2 statistics enables us to reach, for $\sqrt{k} \left(\hat{\gamma}_{n,\alpha,\beta}^{GJ_i}(k) - \gamma \right)$, $i = 1, 2$, an asymptotic standard deviation equal to 2.00, for $\gamma = 1$ and $\rho = -1$. The use of the GJ_3 statistic does not enable us to reach such a value: this is the reason why we have decided not to present a figure with the standard deviation in (3.8) and why this estimator has been discarded from this study. In the following, due to the previous considerations and in order to work only with one *tuning* parameter, instead of two, we are going to consider the following particular cases of the classes of estimators in (3.1) and (3.2):

$$\hat{\gamma}_{n,\alpha}^{GJ_1(\hat{\rho})} \equiv \hat{\gamma}_{n,\alpha,1}^{GJ_1(\hat{\rho})} := \frac{(\alpha - \hat{\rho}) \hat{\gamma}_n^{(\alpha)}(k) - \alpha(1 - \hat{\rho}) \hat{\gamma}_n^{(1)}(k)}{\hat{\rho}(\alpha - 1)}, \quad \alpha > 1, \quad (3.9)$$

and

$$\hat{\gamma}_{n,\alpha}^{GJ_2(\hat{\rho})} \equiv \hat{\gamma}_{n,\alpha,\alpha}^{GJ_2(\hat{\rho})} := \frac{1}{\hat{\rho}} \left(\alpha \hat{\gamma}_n^{(\alpha)}(k) - (\alpha - \hat{\rho}) \hat{\gamma}_n^{(\alpha)}(k) \right), \quad \alpha \geq 1. \quad (3.10)$$

Remark 3.2. Note that under the conditions in Theorem 2.1, and whenever $\sqrt{k} A(n/k) \rightarrow \lambda$, finite, not necessarily null, we get a limiting normal behaviour,

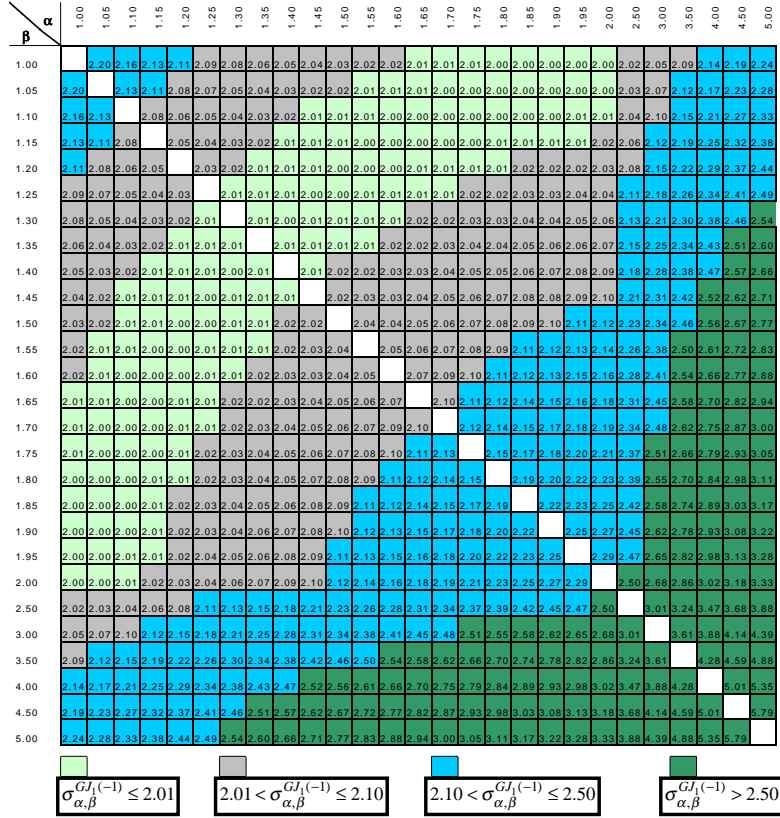


Figure 2: Asymptotic standard deviation of $\widehat{\gamma}_{n,\alpha,\beta}^{GJ_1^{(-)}}$.

with null mean value, for $\sqrt{k} \left(\widehat{\gamma}_{n,\alpha}^{GJ_i(\widehat{\rho})}(k) - \gamma \right)$, $i = 1, 2$. We may then slightly simplify the formulas in (3.6) and (3.7), getting the asymptotic standard deviations,

$$\sigma_{\alpha}^{GJ_1(\rho)} \equiv \sigma_{\alpha,1}^{GJ_1(\rho)} = \gamma \sqrt{\frac{\alpha(\alpha - 2\rho(1 - \rho))}{\rho^2(2\alpha - 1)}}, \quad (3.11)$$

and

$$\sigma_{\alpha}^{GJ_2(\rho)} \equiv \sigma_{\alpha,\alpha}^{GJ_2(\rho)} = \frac{\gamma \alpha^2}{(2\alpha - 1)} \sqrt{\frac{(2\alpha^2 - 2\alpha + 2\rho^2 - 2\rho + 1)}{\rho^2(2\alpha - 1)}}, \quad (3.12)$$

respectively.

Remark 3.3. If we consider the minimization of the variance term in (3.11), we get $\alpha_0^{(1)} := \arg \min_{\alpha} \sigma_{\alpha,1}^{GJ_1(\rho)} = 1 - \rho$. This suggests the consideration of

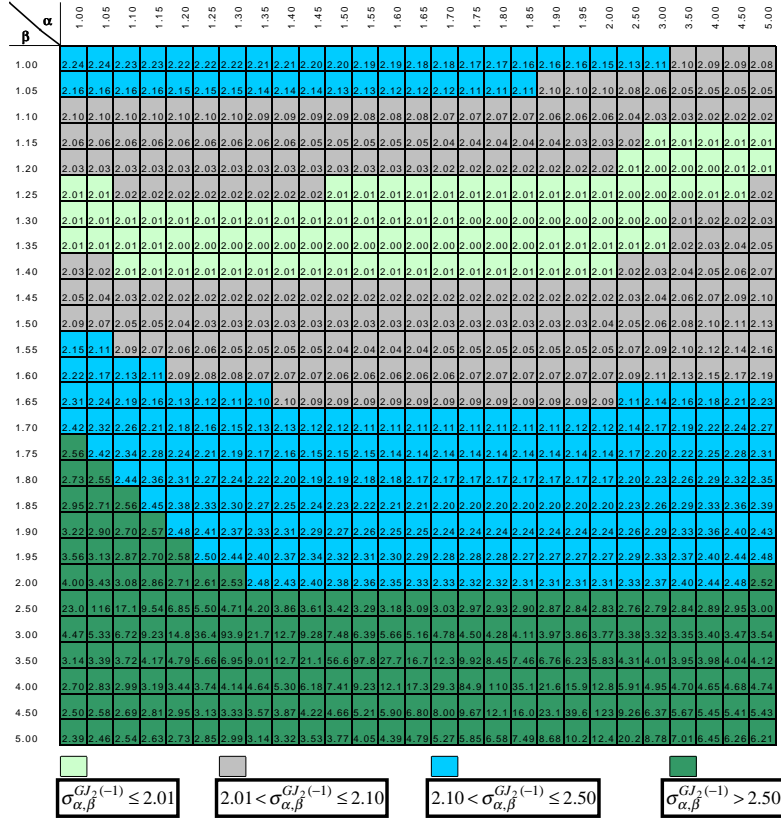


Figure 3: Asymptotic standard deviation of $\widehat{\gamma}_{n,\alpha,\beta}^{GJ_2^{(-)}}$.

an estimator derived from (3.9), where we plug in this value $\alpha_0^{(1)}$, i.e., of the estimator

$$\widehat{\gamma}_{n,\widehat{\alpha}_0^{(1)}}^{GJ_1(\widehat{\rho})} := \frac{(1 - \widehat{\rho})^2 \widehat{\gamma}_n^{(1)}(k) - (1 - 2\widehat{\rho}) \widehat{\gamma}_n^{(1-\widehat{\rho})}(k)}{\widehat{\rho}^2}, \quad \text{where } \widehat{\alpha}_0^{(1)} = 1 - \widehat{\rho}, \quad (3.13)$$

for which the distributional representation in Theorem 3.1 holds true. We then have $\sigma_{1-\widehat{\rho}}^{GJ_1(\widehat{\rho})} = \gamma (1 - \widehat{\rho})/\widehat{\rho}$.

Similarly, let us consider the minimization of the variance term in (3.12), i.e., the value

$$\alpha_0 \equiv \alpha_0^{(2)} := \arg \min_{\alpha} \frac{\alpha^4 (2\alpha^2 - 2\alpha + 2\rho^2 - 2\rho + 1)}{(2\alpha - 1)^3}$$

$$\iff 3\alpha_0^3 - 5\alpha_0^2 + \alpha_0(\rho^2 - \rho + 3) - (2\rho^2 - 2\rho + 1) = 0.$$

With the obvious notation $\hat{\alpha} \equiv \hat{\alpha}_0 = \alpha_0^{(2)}$, it is thus sensible to consider, now on the basis of the estimator in (3.10), the estimator

$$\widehat{\gamma}_{n,\hat{\alpha}}^{GJ_2(\hat{\rho})}(k) = \frac{1}{\hat{\rho}} \left(\hat{\alpha} \widehat{\gamma}_n^{(\hat{\alpha})}(k) - (\hat{\alpha} - \hat{\rho}) \widehat{\gamma}_n^{(\hat{\alpha})}(k) \right),$$

$$\hat{\alpha} : 3\hat{\alpha}^3 - 5\hat{\alpha}^2 + \hat{\alpha}(\hat{\rho}^2 - \hat{\rho} + 3) - (2\hat{\rho}^2 - 2\hat{\rho} + 1) = 0, \quad (3.14)$$

for which the distributional representation in Theorem 3.1 also holds.

In Table 1, we present the value $\alpha_0 = \alpha_0^{(2)}$, associated to the minimum possible asymptotic variances of the estimators in (3.10), for a few values of ρ . For $\gamma = 1$, we also present the values of $\sigma_{\alpha_0}^{GJ_2(\rho)}$ and $\sigma_{1-\rho}^{GJ_1(\rho)} = \gamma(1-\rho)/\rho$, which are very close to each other, showing that there is practically no difference between the asymptotic variances of the two estimators in (3.9) and (3.10), when we choose the α -values that minimize the corresponding asymptotic variances.

Table 1:

ρ	-0.1	-0.2	-0.3	-0.4	-0.5	-1.0	-1.5	-2.0
$\alpha_0 = \alpha_0^{(2)}$	1.0477	1.0913	1.1314	1.1687	1.2034	1.3476	1.4571	1.5428
$\sigma_{\alpha_0}^{GJ_2(\rho)}$	11.0000	6.0002	4.3337	3.5008	3.0013	2.0048	1.6758	1.5137
$(1-\rho)/\rho$	11.0000	6.0000	4.3333	3.5000	3.0000	2.0000	1.6667	1.5000

Remark 3.4. As mentioned before, for the class of estimators in (3.9), we are indeed able to reach an asymptotic variance equal to $(\gamma(1-\rho)/\rho)^2$, the minimal asymptotic variance of an “asymptotically unbiased” estimator in Drees’ class of functionals (Drees, 1998). For the class in (3.10) we are able to reach an asymptotic variance quite close to $(\gamma(1-\rho)/\rho)^2$, as may be seen from Figure 4, where we present the asymptotic standard deviations in (3.11) and (3.12), for $\rho = -1$ and $\gamma = 1$.

Remark 3.5. Note that the choice $\alpha = 1 - \rho$ and the consideration of the tail index estimator $\widehat{\gamma}_{n,\hat{\alpha}_0^{(1)}}^{GJ_1(\hat{\rho})}$ in (3.13), leads us to the asymptotic distributional

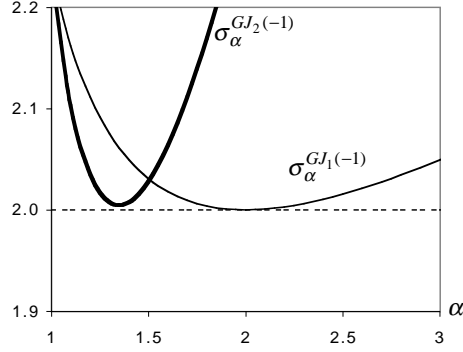


Figure 4: Asymptotic standard deviation indicators of $\widehat{\gamma}_{n,\alpha}^{GJ_1(-1)}$ and $\widehat{\gamma}_{n,\alpha}^{GJ_2(-1)}$, for $\gamma = 1$.

representation,

$$\widehat{\gamma}_{n,\widehat{\alpha}_0^{(1)}}^{GJ_1(\widehat{\rho})} \stackrel{d}{=} \gamma + \frac{\gamma(1-\rho)}{|\rho|} \frac{Z_{k,\alpha,1}^{GJ_1(\rho)}}{\sqrt{k}} + o_p(A(n/k)),$$

for levels k such that $\sqrt{k} A(n/k) \rightarrow \lambda$, finite, as $n \rightarrow \infty$, and whenever $\widehat{\rho}$ is under the conditions of Theorem 3.1.

Consequently, the estimator with smallest asymptotic variance in the class of estimators in (3.9), i.e., the estimator in (3.13), has the same asymptotic variance and second order bias of the “Maximum Likelihood” estimator studied in Gomes and Martins (2002), and given by

$$\widehat{\gamma}_n^{ML(\widehat{\rho})}(k) := \frac{1}{k} \sum_{i=1}^k U_i - \left(\frac{1}{k} \sum_{i=1}^k i^{-\widehat{\rho}} U_i \right) \frac{\left(\sum_{i=1}^k i^{-\widehat{\rho}} \right) \left(\sum_{i=1}^k U_i \right) - k \left(\sum_{i=1}^k i^{-\widehat{\rho}} U_i \right)}{\left(\sum_{i=1}^k i^{-\widehat{\rho}} \right) \left(\sum_{i=1}^k i^{-\widehat{\rho}} U_i \right) - k \left(\sum_{i=1}^k i^{-2\widehat{\rho}} \right)},$$

where U_i , $1 \leq i \leq k$, are the scaled log-spacings in (1.1).

3.2 Misspecification of ρ

If we misspecify $\rho = -1$, we have the bias’ quotients

$$q_{\alpha,\beta}^{(1)} \equiv q_{\alpha,\beta}^{(1)}(-1) = \frac{\alpha(\beta+1)}{\beta(\alpha+1)}, \quad q_{\alpha,\beta}^{(2)} \equiv q_{\alpha,\beta}^{(2)}(-1) = \frac{\alpha(\beta+1)^2}{\beta^2(\alpha+1)},$$

$$q_{\alpha,\beta}^{(3)} \equiv q_{\alpha,\beta}^{(3)}(-1) = \frac{\alpha^2(\beta+1)^2}{\beta^2(\alpha+1)^2},$$

and we thus get the Generalized Jackknife classes of estimators $\widehat{\gamma}_{n,\alpha,\beta}^{GJ_i(-1)}(k)$, $i = 1, 2, 3$, given in (3.1), (3.2) and (3.3), respectively, with ρ replaced by -1 .

We shall also consider the particular cases in (3.9) and (3.10), but now with ρ replaced by -1 , i.e., we shall consider,

$$\widehat{\gamma}_{n,\alpha}^{GJ_1(-1)} \equiv \widehat{\gamma}_{n,\alpha,1}^{GJ_1(-1)} := \frac{2\alpha \widehat{\gamma}_n^{(1)}(k) - (\alpha + 1) \widehat{\gamma}_n^{(\alpha)}(k)}{(\alpha - 1)}, \quad \alpha > 1, \quad (3.15)$$

and

$$\widehat{\gamma}_{n,\alpha}^{GJ_2(-1)}(k) \equiv \widehat{\gamma}_{n,\alpha,\alpha}^{GJ_2(-1)} := -\alpha \widehat{\gamma}_n^{(\alpha)}(k) + (\alpha + 1) \widetilde{\gamma}_n^{(\alpha)}(k), \quad \alpha \geq 1. \quad (3.16)$$

On the basis of Theorem 3.1, we get straightforwardly the general distributional behaviour of the new estimators $\widehat{\gamma}_{n,\alpha,\beta}^{GJ_\bullet(-1)}(k)$. Indeed, with the misspecification of ρ at -1 , we are going to have no change in the asymptotic variances, which are given by the expressions in (3.6), (3.7) and (3.8), with $\rho = -1$, but whenever $\rho \neq -1$, we get a dominant component of bias of the order of $A(n/k)$, like in the classical tail index estimators. We may state the following result:

Theorem 3.2. *Under the conditions of Theorem 2.1 and with the same notation of Theorem 3.1, we get the following asymptotic distributional representations:*

$$\widehat{\gamma}_{n,\alpha,\beta}^{GJ_1(-1)}(k) \stackrel{d}{=} \gamma + \frac{\gamma \sigma_{\alpha,\beta}^{GJ_1(-1)}}{\sqrt{k}} Z_{k,\alpha,\beta}^{GJ_1(-1)} + \frac{\alpha\beta(1+\rho)}{(\alpha-\rho)(\beta-\rho)} A(n/k)(1+o_p(1)),$$

$$\begin{aligned} \widehat{\gamma}_{n,\alpha,\beta}^{GJ_2(-1)}(k) &\stackrel{d}{=} \gamma + \frac{\gamma \sigma_{\alpha,\beta}^{GJ_2(-1)}}{\sqrt{k}} Z_{k,\alpha,\beta}^{GJ_2(-1)} \\ &\quad + \frac{\alpha\beta^2(1+\rho)(\alpha(\rho-1) - 2\alpha\beta + \beta^2 + \rho)}{(\beta^2 - 2\alpha\beta - \alpha)(\alpha-\rho)(\beta-\rho)^2} A(n/k)(1+o_p(1)) \end{aligned}$$

and

$$\begin{aligned} \widehat{\gamma}_{n,\alpha,\beta}^{GJ_3(-1)}(k) &\stackrel{d}{=} \gamma + \frac{\gamma \sigma_{\alpha,\beta}^{GJ_3(-1)}}{\sqrt{k}} Z_{k,\alpha,\beta}^{GJ_3(-1)} \\ &\quad + \frac{\alpha^2 \beta^2 (1+\rho)(2\alpha\beta + (1-\rho)(\beta+\alpha) - 2\rho)}{2\alpha\beta + \alpha + \beta} A(n/k)(1 + o_p(1)). \end{aligned}$$

For the particular case $\beta = 1$ in $\widehat{\gamma}_{n,\alpha,\beta}^{GJ_1(-1)}$, we are dealing with the class of estimators in (3.15), and we get,

$$\begin{aligned} \widehat{\gamma}_{n,\alpha}^{GJ_1(-1)}(k) &\stackrel{d}{=} \gamma + \gamma \sqrt{\frac{\alpha(\alpha+4)}{(2\alpha-1)k}} Z_{k,\alpha,1}^{GJ_1(-1)} \\ &\quad + \frac{\alpha(1+\rho)}{(1-\rho)(\alpha-\rho)} A(n/k)(1 + o_p(1)), \end{aligned}$$

For the particular case $\alpha = \beta$ in $\widehat{\gamma}_{n,\alpha,\beta}^{GJ_2(-1)}$, we are dealing with the class of estimators in (3.16), and we get,

$$\begin{aligned} \widehat{\gamma}_{n,\alpha}^{GJ_2(-1)} &\stackrel{d}{=} \gamma + \frac{\gamma \alpha^2}{(2\alpha-1)} \sqrt{\frac{2\alpha^2 - 2\alpha + 5}{2\alpha-1}} \frac{Z_{k,\alpha,\alpha}^{GJ_2(-1)}}{\sqrt{k}} \\ &\quad + \frac{\alpha^2(1+\rho)}{(\alpha-\rho)^2} \frac{A(n/k)}{(1 + o_p(1))}. \end{aligned}$$

Remark 3.6. The estimator with smallest asymptotic variance in the class of estimators herewith considered, i.e.,

$$\widehat{\gamma}_{n,2}^{GJ_1(-1)}(k) \equiv \widehat{\gamma}_{n,2,1}^{GJ_1(-1)}(k) := 4\widehat{\gamma}_n^{(1)}(k) - 3\widehat{\gamma}_n^{(2)}(k)$$

is, as expected from the result in Remark 3.5, asymptotically equivalent to the “Maximum Likelihood” estimator studied in Gomes and Martins (2004), and given by

$$\widehat{\gamma}_n^{ML}(k) := \frac{1}{k} \sum_{i=1}^k U_i - \left(\frac{1}{k} \sum_{i=1}^k i U_i \right) \frac{\sum_{i=1}^k (2i - k - 1) U_i}{\sum_{i=1}^k i(2i - k - 1) U_i},$$

where U_i , $1 \leq i \leq k$, are the scaled log-spacings in (1.1), i.e., both estimators have the same asymptotic variance and second order bias.

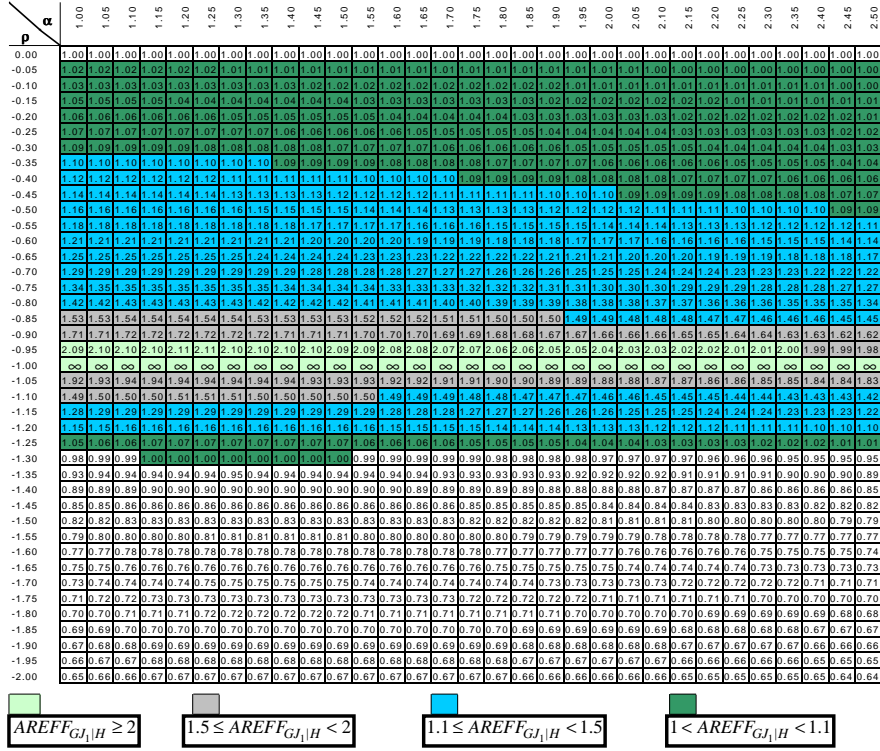


Figure 5: Asymptotic efficiency of $\hat{\gamma}_{n,\alpha}^{GJ_1(-1)}$ relatively to the Hill estimator, at their optimal levels.

An asymptotic comparison of $\hat{\gamma}_{n,\alpha}^{GJ_\bullet(-1)}$ and the Hill estimator, $\hat{\gamma}_n^H$, computed at their optimal levels, enable us to state the following result:

Theorem 3.3. For $\rho \neq -1$, we get for $\hat{\gamma}_{n,\alpha}^{GJ_i(-1)}$, $i = 1, 2$, the following asymptotic efficiencies relatively to the Hill estimator:

$$AREFF_{GJ_1|H}(\alpha) = \left(\frac{(\alpha - \rho)(2\alpha - 1)^{-\rho}}{|1 + \rho|\alpha^{1-\rho}(\alpha + 4)^{-\rho}} \right)^{\frac{1}{1-2\rho}},$$

$$AREFF_{GJ_2|H}(\alpha) = \left(\frac{1}{\alpha^{2-4\rho}} \left(\frac{2\alpha^2 - 2\alpha + 5}{(2\alpha - 1)^3} \right)^\rho \frac{(\alpha - \rho)^2}{(1 - \rho)|1 + \rho|} \right)^{\frac{1}{1-2\rho}}.$$

These two $AREFF$ -measures are presented in Figures 5 and 6.

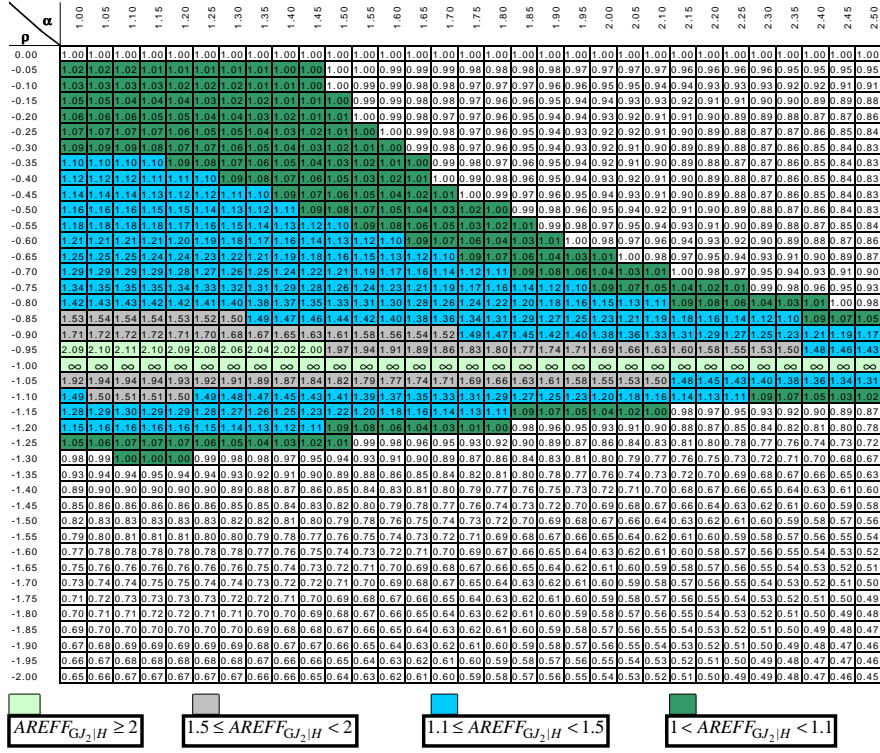


Figure 6: Asymptotic efficiency of $\hat{\gamma}_{n,\alpha}^{GJ_2(-1)}$ relatively to the Hill estimator, at their optimal levels.

It is possible to see that we cannot practically distinguish asymptotically the two estimators in (3.15) and (3.16), when they are computed at the level α_0 that minimizes their asymptotic variance. Anyway, we get a region of (α, ρ) values where the GJ_1 estimator behaves better than the GJ_2 estimator, when they are computed at the same tuning parameter α . However, for every ρ there exists always unique values

$$\tilde{\alpha}_{01} := \arg \min_{\alpha} LMSE \left(\hat{\gamma}_{n0,\alpha}^{GJ_1(-1)} \right) \text{ and } \tilde{\alpha}_{02} := \arg \min_{\alpha} LMSE \left(\hat{\gamma}_{n0,\alpha}^{GJ_2(-1)} \right).$$

If we then compute $AREFF \left(\hat{\gamma}_{n0,\alpha}^{GJ_2(-1)} \mid \hat{\gamma}_{n0,\alpha}^{GJ_1(-1)} \right)$, we get, for every ρ , a value larger than 1. We are thus led to advise the use of the estimator in (3.16), among the estimators with ρ misspecified at $\rho = -1$.

4 Simulated distributional behaviour

In Figures 7, 8 and 9 we present the mean value and mean squared error patterns, as functions of k , the number of top o.s. used, of the estimator in (3.14), denoted in the figures $GJ(\hat{\rho}_0)$ and $GJ(\hat{\rho}_1)$, when we estimate ρ through $\hat{\rho}_0$ and $\hat{\rho}_1$, respectively, with $\hat{\rho}_\tau$ given in (3.4). We also represent the same features of the estimator in (3.16), computed at the value $\tilde{\alpha}_{02}$ that minimizes its *LMSE*, and denoted $GJ(-1)$ in the figures. We have considered the Burr model, with d.f. $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, $\gamma = 1$ and $\rho = -0.5, -1$ and -2 . The behaviour of the Hill estimator is pictured for reference, and the simulation is based on 5000 runs.

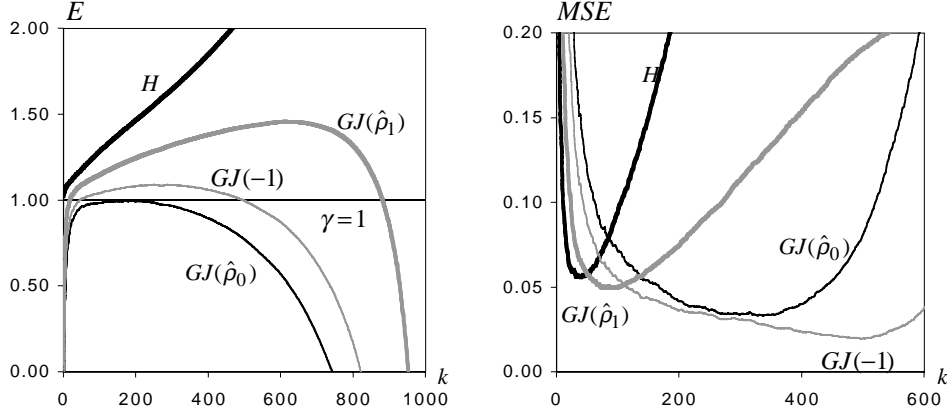


Figure 7: Burr parent with $(\gamma, \rho) = (1, -0.5)$

In Table 1 we present for different simulated models, and $\tau = 0$ or 1 according as $|\rho| < 1$ or $|\rho| \geq 1$, simulated measures of efficiency and bias reduction of the estimator $\hat{\gamma}_{n, \hat{\alpha}_0}^{GJ_2(\hat{\rho}_\tau)}$ relatively to the Hill estimator, both computed at their optimal levels, i.e.,

$$REFF_{GJ_2(\hat{\rho})|H} = \sqrt{\frac{MSE[\hat{\gamma}_{n0}^H]}{MSE[\hat{\gamma}_{n0}^{GJ_2(\hat{\rho})}]}}}, \quad BRI_{GJ_2(\hat{\rho})|H} = \left| \frac{E[\hat{\gamma}_{n0}^H - \gamma]}{E[\hat{\gamma}_{n0}^{GJ_2(\hat{\rho})} - \gamma]} \right|. \quad (4.1)$$

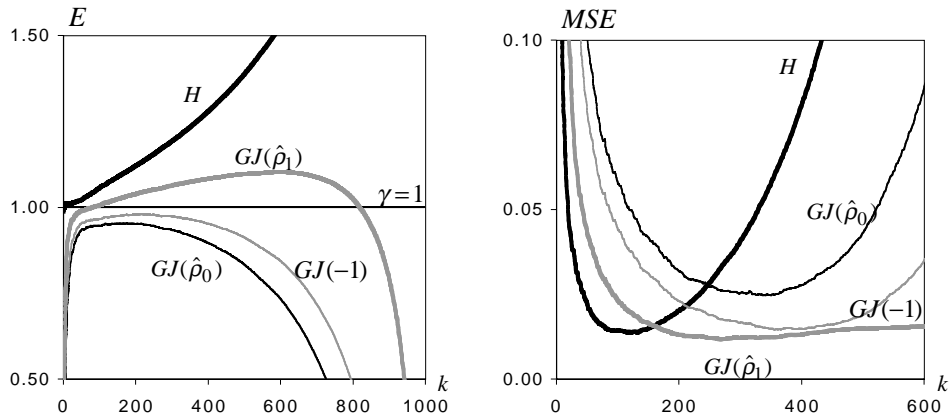


Figure 8: Burr parent with $(\gamma, \rho) = (1, -1)$

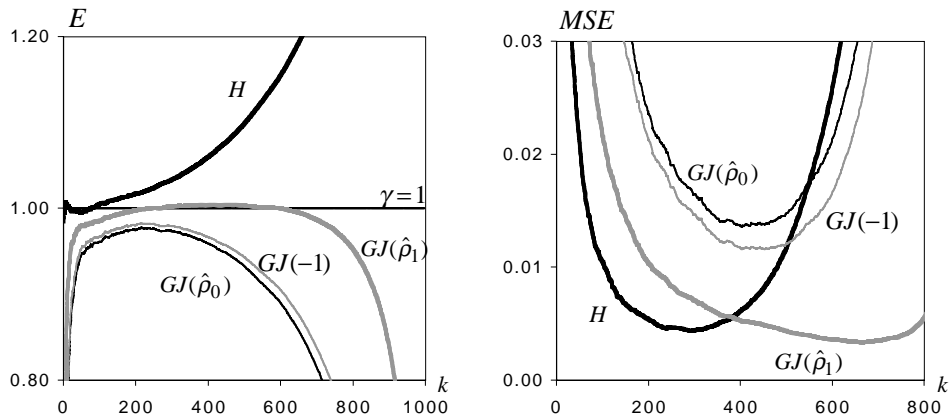


Figure 9: Burr parent with $(\gamma, \rho) = (1, -2)$

This measure was computed on the basis of a multi-sample simulation of size 5000×10 . For details on multi-sample simulation see Gomes and Oliveira (2001).

Some overall remarks:

1. As ρ approaches 0, the new estimators, computed at their optimal levels, compare favourably with the Hill estimator. This is indeed a kind of behaviour shared by all reduced bias' estimators considered before, and it is essentially due to the high bias of the Hill estimator in this region of

Table 2: Relative efficiencies and bias reduction indicators, $REF_{GJ_2(\hat{\rho})|H}/BRJ_{GJ_2(\hat{\rho})|H}$.

n	1000	2000	5000	10000	20000
$STU(\rho = -.5)$	1.31/3.19	1.51/5.70	1.71/5.70	1.93/36.43	2.19/72.74
$BU(\rho = -.5)$	1.31/3.19	1.51/8.80	1.74/7.00	1.93/32.72	2.21/47.53
$FRE(\rho = -1)$	1.35/7.13	1.37/10.07	1.28/6.27	1.32/5.39	1.39/7.47
$STU(\rho = -1)$	1.47/30.06	1.58/30.63	1.66/32.65	1.77/32.51	1.90/31.91
$BU\rho = (-1)$	1.47/30.06	1.57/31.18	1.67/25.01	1.79/28.36	1.91/27.16
$STU(\rho = -2)$	0.90/1.61	0.99/1.97	1.05/1.99	1.09/2.04	1.12/1.95
$BU(\rho = -2)$	1.10/3.54	1.13/4.23	1.12/13.17	1.14/9.38	1.18/45.28

ρ -values. Regarding minimum mean squared error, the best performance is achieved when we misspecify ρ at $\rho = -1$. Regarding bias, the best performance is achieved whenever we use $\hat{\rho}_0$ as the ρ -estimate — then the sample path is quite close to the target value γ , for a wide region of k -values (see Figure 7, left).

- For models with $\rho = -1$, as illustrated in Figure 8, the new estimator in (3.14), at its optimal level, overpasses the Hill estimator, also at its optimal level, if we estimate ρ through $\hat{\rho}_1$. The best results, regarding sample paths' stability, are achieved by the estimator $GJ(-1)$. If we consider $\hat{\rho}_0$, the reduction of bias is too big, and the new estimator, in (3.14), has a negative bias for all k . Similar conclusions may be drawn for all simulated models with $\rho = -1$, like the Fréchet model, $F(x) = \exp(-x^{-1/\gamma})$, $x \geq 0$, a typical heavy-tailed model.
- For small values of ρ it has been difficult to find competitors to the Hill estimator. However, with this class we are indeed able to overpass the Hill estimator for large n , again through the use $\hat{\rho}_1$. An illustration of this is presented in Figure 9.

5 An application in the field of finance

We shall herewith consider an illustration of the performance of the above mentioned estimators, through the analysis of the Euro-UK Pound daily exchange rates from January 4, 1999 till December 15, 2003. In Figure 10, working with the $n_+ = 593$ positive log-returns, we present the sample path of the $\hat{\rho}_\tau$ estimates in (3.4), as function of k , for $\tau = 0$ and $\tau = 1$ (*left*), together with the sample paths of the classical Hill estimator in (1), denoted H and the estimator $\hat{\gamma}_{n,\hat{\alpha}}^{GJ(\hat{\rho})}$ in (3.14) for $\hat{\rho} = \hat{\rho}_0$, denoted $GJ(\hat{\rho}_0)$ (*right*).

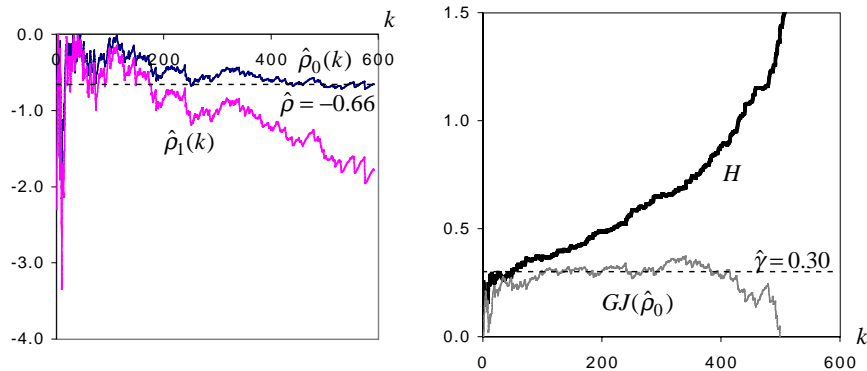


Figure 10: Estimates of the second order parameter ρ (*left*) and of the tail index γ (*right*) for the Daily Log-Returns of the Euro-UK Pound.

We may then draw the following conclusions:

1. 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 may 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 $\hat{\rho}_0 = -0.66$, and we should consider the estimator $\hat{\gamma}_{n,\hat{\alpha}}^{GJ_2(\hat{\rho}_0)}$ in (3.14).
2. Regarding the tail index estimation, note that whereas the Hill estimator

is unbiased for the estimation of the tail index γ when the underlying model is a strict Pareto model, it exhibits a relevant bias when we have only Pareto-like tails, as happens here, and may be seen from Figure 10 (right). The *Generalized Jackknife* estimator, which is “asymptotically unbiased” reveals 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”, considered, for instance, in Gomes *et al.* (2003).

Here we have merely considered, as a possible γ -estimate, the median of the *GJ*-estimates in the region (k_a, k_b) , with $k_a = \left[n_+^{-2\hat{\rho}_0/(1-2\hat{\rho}_0)} \right] = 37$, $k_b = \left[5 n_+^{-2\hat{\rho}_0/(1-2\hat{\rho}_0)} \right] = 189$. We have then got the value $\hat{\gamma} = 0.30$, provided in Figure 10, the same value we had obtained before in the above mentioned paper, where this same set of data has been analysed.

6 Proofs

We first state the following lemmas, where E_i , $i \geq 1$, denotes a sequence of i.i.d. unit exponential r.v.'s, with d.f. $F_E(x) = 1 - \exp(-x)$, $x \geq 0$. The notation $a_k \sim b_k$ means that $a_k/b_k \rightarrow 1$, as $k \rightarrow \infty$.

Lemma 6.1. (Chernoff *et al.*, 1967). Let $Z_k = \frac{1}{k} \sum_{i=1}^k \alpha_{ik} E_i$ and let us write,

$$v_k = \sqrt{\frac{1}{k} \sum_{i=1}^k \alpha_{ik}^2} = \sqrt{k \text{Var}(Z_k)}.$$

Then

$$\frac{\sqrt{k} \left(Z_k - \frac{1}{k} \sum_{i=1}^k \alpha_{ik} \right)}{v_k} \xrightarrow[k \rightarrow \infty]{d} \text{Normal}(0, 1) \iff \max_{1 \leq i \leq k} |\alpha_{ik}| = o\left(\sqrt{k} v_k\right),$$

as $k \rightarrow \infty$. If we further have $\frac{1}{k} \sum_{i=1}^k \alpha_{ik} = \mu + o_p(1/\sqrt{k})$, $v_k \rightarrow \sigma > 0$, μ and σ finite, and $\max_{1 \leq i \leq k} |\alpha_{ik}| = o(\sqrt{k})$, as $k \rightarrow \infty$, then

$$\frac{\sqrt{k}(Z_k - \mu)}{\sigma} \xrightarrow[k \rightarrow \infty]{d} \text{Normal}(0, 1).$$

Lemma 6.2. *Let us denote*

$$V_k^{(\alpha)} := \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} E_i, \quad \alpha \geq 1, \quad (6.1)$$

and

$$W_k^{(\alpha)} := \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} \ln\left(\frac{i}{k}\right) E_i, \quad \alpha \geq 1, \quad (6.2)$$

where E_i , $1 \leq i \leq k$, is a standard exponential random sample of size k . As $k \rightarrow \infty$,

$$\mathbb{E} \left[V_k^{(\alpha)} \right] = \frac{1}{\alpha} + o(1/\sqrt{k}), \quad \mathbb{E} \left[W_k^{(\alpha)} \right] = -\frac{1}{\alpha^2} + o(1/\sqrt{k}), \quad (6.3)$$

and

$$\text{Var} \left[V_k^{(\alpha)} \right] \sim \frac{1}{(2\alpha - 1)k}, \quad \text{Var} \left[W_k^{(\alpha)} \right] \sim \frac{2}{(2\alpha - 1)^3 k}. \quad (6.4)$$

Moreover, both r.v.'s, $P_k^{(\alpha)} = \sqrt{(2\alpha - 1)k} (V_k^{(\alpha)} - 1/\alpha)$ in (2.3) and $Q_k^{(\alpha)} = (2\alpha - 1)\sqrt{(2\alpha - 1)k/2} (W_k^{(\alpha)} + 1/\alpha^2)$ in (2.4), are asymptotically standard normal, and the second order structure between these r.v.'s is given by

$$\text{Cov} \left[P_k^{(\alpha)}, P_k^{(\beta)} \right] \sim \frac{\sqrt{(2\alpha - 1)(2\beta - 1)}}{\alpha + \beta - 1}, \quad (6.5)$$

$$\text{Cov} \left(Q_k^{(\alpha)}, Q_k^{(\beta)} \right) \sim \frac{((2\alpha - 1)(2\beta - 1))^{3/2}}{(\alpha + \beta - 1)^3}, \quad (6.6)$$

and

$$\text{Cov} \left(P_k^{(\alpha)}, Q_k^{(\beta)} \right) \sim -\frac{(2\beta - 1)\sqrt{(2\alpha - 1)(2\beta - 1)}}{\sqrt{2}(\alpha + \beta - 1)^2}. \quad (6.7)$$

Proof. Since $\mathbb{E}[E_i] = \text{Var}[E_i] = 1$,

$$\begin{aligned}\mathbb{E}\left(V_k^{(\alpha)}\right) &= \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} \xrightarrow{k \rightarrow \infty} \int_0^1 x^{\alpha-1} dx = \frac{1}{\alpha} \\ \mathbb{E}\left(W_k^{(\alpha)}\right) &= \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} \ln\left(\frac{i}{k}\right) \xrightarrow{k \rightarrow \infty} \int_0^1 x^{\alpha-1} \ln x dx = -\frac{1}{\alpha^2}\end{aligned}$$

More than this: the differences $\mathbb{E}\left(V_k^{(\alpha)}\right) - 1/\alpha$ and $\mathbb{E}\left(W_k^{(\alpha)}\right) + 1/\alpha^2$ are $o_p\left(1/\sqrt{k}\right)$, and (6.3) follows. Since the weights $\alpha_{ik} = (i/k)^{\alpha-1}$ and $\alpha_{ik} = (i/k)^{\alpha-1} \ln(i/k)$ are under the conditions of Lemma 6.1, the asymptotic normality of the r.v.'s in (2.3) and in (2.4) comes thus from Lemma 6.1, together with the fact that,

$$\begin{aligned}k \text{Var}\left(V_k^{(\alpha)}\right) &= \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{2(\alpha-1)} \xrightarrow{k \rightarrow \infty} \int_0^1 x^{2\alpha-2} dx = \frac{1}{2\alpha-1} \\ k \text{Var}\left(W_k^{(\alpha)}\right) &= \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{2(\alpha-1)} \ln^2\left(\frac{i}{k}\right) \xrightarrow{k \rightarrow \infty} \int_0^1 x^{2\alpha-2} \ln^2 x dx \\ &= \frac{2}{(2\alpha-1)^3},\end{aligned}$$

i.e., (6.4) holds true.

Similar computations lead us to (6.5), (6.6) and (6.7). Indeed,

$$\begin{aligned}\text{Cov}\left(P_k^{(\alpha)}, P_k^{(\beta)}\right) &\xrightarrow{k \rightarrow \infty} \frac{\sqrt{(2\alpha-1)(2\beta-1)}}{\alpha+\beta-1} \int_0^1 x^{\alpha+\beta-2} dx \\ &= \frac{\sqrt{(2\alpha-1)(2\beta-1)}}{\alpha+\beta-1}, \\ \text{Cov}\left(Q_k^{(\alpha)}, Q_k^{(\beta)}\right) &\xrightarrow{k \rightarrow \infty} \frac{((2\alpha-1)(2\beta-1))^{3/2}}{2} \int_0^1 x^{\alpha+\beta-2} \ln^2 x dx \\ &= \frac{((2\alpha-1)(2\beta-1))^{3/2}}{(\alpha+\beta-1)^3}, \\ \text{Cov}\left(P_k^{(\alpha)}, Q_k^{(\beta)}\right) &\xrightarrow{k \rightarrow \infty} \frac{((2\beta-1)\sqrt{(2\alpha-1)(2\beta-1)})}{\sqrt{2}} \int_0^1 x^{\alpha+\beta-2} \ln x dx \\ &= -\frac{(2\beta-1)\sqrt{(2\alpha-1)(2\beta-1)}}{\sqrt{2}(\alpha+\beta-1)^2}.\end{aligned}$$

□

Proof. (Theorem 2.1). We may write

$$\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} U_i \stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} i \left\{ \ln \frac{U(Y_{n-i+1:n})}{U(Y_{n-i:n})} \right\},$$

where $Y_{i:n}$, $1 \leq i \leq n$, are the ascending o.s. associated to an i.i.d. standard Pareto sample of size n , from a parent $F_Y(y) = 1 - y^{-1}$, $y \geq 1$.

Since $Y_{n-i+1:n}/Y_{n-i:n} \stackrel{d}{=} Y_{1:i}$, $1 \leq i \leq k$, $\ln Y_{1:i} \stackrel{d}{=} E_{1:i} = E_i/i$, and from (1.5), which enables us to write

$$\ln \frac{U(tx)}{U(t)} = \gamma \ln x + \frac{x^\rho - 1}{\rho} A(t)(1 + o(1)),$$

for every $x > 0$, and as $t \rightarrow \infty$, we get

$$\begin{aligned} \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} U_i &\stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} i \left\{ \ln \frac{U(Y_{n-i:n} Y_{1:i})}{U(Y_{n-i:n})} \right\} \\ &\stackrel{d}{=} \frac{\gamma}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} E_i + A(n/k) \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-\rho} \left(\frac{e^{\rho E_i/i} - 1}{\rho} \right) (1 + o_p(1)) \\ &\stackrel{d}{=} \gamma V_k^{(\alpha)} + A(n/k) \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-\rho} \left(\frac{e^{\rho E_i/i} - 1}{\rho} \right) (1 + o_p(1)), \end{aligned}$$

with $V_k^{(\alpha)}$ given in (6.1). The use of the inequality,

$$\frac{\rho E_i^2}{i^2} < \frac{e^{\rho E_i/i} - 1}{\rho} - \frac{E_i}{i} < 0,$$

enables us to guarantee that

$$\sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-\rho} \left(\frac{e^{\rho E_i/i} - 1}{\rho} \right) \xrightarrow[k \rightarrow \infty]{p} \frac{1}{\alpha - \rho}.$$

Next, the use of this relation and of (2.3) enables us to write

$$\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} U_i \stackrel{d}{=} \frac{\gamma}{\alpha} + \frac{\gamma P_k^{(\alpha)}}{\sqrt{(2\alpha - 1)k}} + \frac{1}{(\alpha - \rho)} A(n/k)(1 + o_p(1)),$$

where $P_k^{(\alpha)}$, given in (2.3), is asymptotically standard normal, and (2.1) follows.

The distributional representation (2.2) follows in a similar way. \square

Proof. (Theorem 3.1). The proof of (3.5), $i = 1, 2, 3$, follows the lines of the proof of Theorem 2.1. The second part of the theorem follows from the fact that

$$\varphi_\rho^{(1)}(k) := \frac{\partial}{\partial \rho} \widehat{\gamma}_{n,\alpha}^{GJ_1(\rho)}(k) = O_p \left(\widehat{\gamma}_n^{(\alpha)}(k) - \widehat{\gamma}_n^{(\beta)}(k) \right) = O_p \left(1/\sqrt{k} \right),$$

$$\varphi_\rho^{(2)}(k) := \frac{\partial}{\partial \rho} \widehat{\gamma}_{n,\alpha}^{GJ_2(\rho)}(k) = O_p \left(\widehat{\gamma}_n^{(\alpha)}(k) - \widehat{\gamma}_n^{(\beta)}(k) \right) = O_p \left(1/\sqrt{k} \right)$$

and

$$\varphi_\rho^{(3)}(k) := \frac{\partial}{\partial \rho} \widehat{\gamma}_{n,\alpha}^{GJ_3(\rho)}(k) = O_p \left(\widehat{\gamma}_n^{(\alpha)}(k) - \widehat{\gamma}_n^{(\beta)}(k) \right) = O_p \left(1/\sqrt{k} \right).$$

Hence,

$$\widehat{\gamma}_{n,\alpha}^{GJ_\bullet(\widehat{\rho})}(k) = \widehat{\gamma}_{n,\alpha}^{GJ_\bullet(\rho)}(k) + (\widehat{\rho} - \rho) \varphi_\rho^\bullet(k)(1 + o_p(1)),$$

and then, whenever $\sqrt{k} A(n/k) \rightarrow \lambda$, finite, i.e., $A(n/k) = O_p(1/\sqrt{k})$, the conditions in the theorem enable us to guarantee that $(\widehat{\rho} - \rho) \varphi_\rho^\bullet(k) = o_p(A(n/k))$. \square

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] 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 M. I. Gomes (2002a). A class of asymptotically unbiased semi-parametric estimators of the tail index. *Test* **11**:2, 345-364.
- [4] Caeiro, F. and M. I. Gomes (2002b). Bias reduction in the estimation of parameters of rare events. *Theory of Stochastic Processes* **8** (24), 1-2, 67-76.
- [5] Chernoff, H. Gastwirth, J. L. and Johns, M. V. Jr. (1967). Asymptotic distribution of linear combinations of functions of order statistics with application to estimation. *Ann. Math. Statist.* **38**, 52-72.

- [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 P. Hall (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.
- [8] 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**:1, 193-213.
- [9] Gomes, M. I., Caeiro, F. and Figueiredo, F. (2004). Bias reduction of a tail index estimator through an external estimation of the second order parameter. *Statistics* **38**(6), 497-510.
- [10] 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*.
- [11] Gomes, M. I. and de Haan, L. (1999). Approximation by penultimate extreme value distributions. *Extremes* **2**:1, 71-85.
- [12] 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.
- [13] Gomes, M. I. and M. J. Martins (2001). Generalizations of the Hill estimator — asymptotic versus finite sample behaviour. *J. Statistical Planning and Inference* **93**, 161-180.
- [14] 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.
- [15] Gomes, M.I. and Martins, M.J. (2004). Bias reduction and explicit semi-parametric estimation of the tail index. *J. Statistical Planning and Inference* **124**, 361-378.
- [16] 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, 2001.

- [17] Gomes, M.I., Martins, M.J. and M. Neves (2002b). Generalized Jackknife semi-parametric estimators of the tail index. *Portugaliae Mathematica* **59**:4, 393-408.
- [18] Gomes, M. I., Miranda, M. C. and Pereira, H. (2005). Revisiting the role of the Jackknife methodology in the estimation of a positive tail index. *Commun. Statist. — Theory and Methods* **34**, 1-20.
- [19] Gomes, M. I. and Oliveira, O. (2001). The bootstrap methodology in Statistical Extremes — choice of the optimal sample fraction. *Extremes* **4**:4, 331-358, 2002.
- [20] Gray, H.L. and W.R. Schucany (1972). *The Generalized Jackknife Statistic*. Marcel Dekker.
- [21] Haan, L. de and L. Peng (1998). Comparison of tail index estimators. *Statistica Neerlandica* **52**, 60-70.
- [22] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.