Mixed Moment Estimator and Location Invariant Alternatives*

M. ISABEL FRAGA ALVES
CEAUL, DEIO, Faculty of Science
University of Lisbon, Portugal

M. IVETTE GOMES
CEAUL, DEIO, Faculty of Science
University of Lisbon, Portugal

LAURENS DE HAAN
Department of Economics
Erasmus University Rotterdam
The Netherlands

CLÁUDIA NEVES
UIMA, Department of Mathematics
University of Aveiro, Portugal

Abstract. A new class of estimators of the extreme value index is developed. It has a simple form and is very close to the maximum likelihood estimator for a wide class of heavy-tailed models. We also propose an alternative class of estimators, dependent on a tuning parameter $p \in (0, 1)$ and invariant for changes in both scale and/or location.

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

Keywords and phrases. *Extreme value index; semi-parametric estimation; statistics of extremes.*

1 Introduction

The Fisher and Tippett theorem of extreme values (Fisher and Tippett, 1928) states that all possible non-degenerate weak limit distributions of normalized partial maxima of independent, identically distributed (i.i.d.) random variables (r.v.'s) $X_1, X_2, \cdots, X_n \cdots$ are (generalized) extreme value distributions. That is, if there are normalizing constants $a_n > 0$ and b_n such that, for all x,

$$\lim_{n\to\infty} P\left\{a_n^{-1}\left(\max\left(X_1, \cdots, X_n\right) - b_n\right) \le x\right\} = G(x), \tag{1.1}$$

*Research partially supported by FCT / POCTI and POCI / FEDER

where G is some non-degenerate distribution function (d.f.), we can redefine the constants in such a way that the limit G is one of a one-parameter family of distributions,

$$G_\gamma(x) := \begin{cases} \exp\left(-(1+\gamma x)^{-1/\gamma}\right),\ 1+\gamma x > 0 & \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)),\ x \in \mathbb{R} & \text{if} \quad \gamma = 0 \end{cases}, \tag{1.2}$$

the (generalized) *extreme value* distributions, given here in the von Mises-Jenkinson form (von Mises, 1936; Jenkinson, 1955). We say that the d.f. F of the r.v.'s $X_1, X_2, \cdots$ is in the domain of attraction of G_γ if (1.1) holds with $G = G_\gamma$, and use the notation $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$.

For the application of extreme value theory it is necessary to estimate the parameter γ. Several estimators for γ have been proposed in the literature. We mention Hill's estimator (Hill, 1975), valid in the range $\gamma > 0$, Pickand's estimator (Pickands, 1975), the so-called "maximum likelihood" or ML estimator (Smith, 1987), based on the approximate Paretian behaviour of the log-excesses over a high random threshold and valid for $\gamma > -1/2$, the probability weighted moment estimator (Hosking and Wallis, 1987), valid for $\gamma < 1/2$, the moment estimator (Dekkers, Einmahl and de Haan, 1989) and the generalized Hill estimator (Beirlant, Vynckier and Teugels, 1996; Beirlant, Dierckx and Guillou, 2005).

In order to develop a new estimator, consider a combination of Theorems 2.6.1 and 2.6.2 of de Haan, 1970: a d.f. F, with right endpoint

$$x^F := \sup\{t : F(t) < 1\} \in (0, +\infty],$$

is in the domain of attraction of the extreme value d.f. G_γ, i.e., (1.1) holds with $G = G_\gamma$ and $X_1, X_2, \cdots$ have d.f. F, if and only if

$$\lim_{t \to x^F} \frac{(1-F(t)) \int_t^{x^F} \int_y^{x^F} (1-F(x)) x^{-2} dx\, dy}{t^2 \left(\int_t^{x^F} (1-F(x)) x^{-2} dx\right)^2} = \varphi(\gamma) := \begin{cases} 1+\gamma & \text{if} \quad \gamma > 0 \\ \frac{1-\gamma}{1-2\gamma} & \text{if} \quad \gamma \le 0 \end{cases}. \tag{1.3}$$

Let $X_i,\ 1 \le i \le n$, be n i.i.d. r.v.'s with d.f. $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$, and let $X_{i,n},\ 1 \le i \le n$, denote the associated ascending order statistics. We can build a statistic starting from the left hand-side of (1.3), noticing that it can be written as

$$\frac{(1-F(t))^{-1}\left\{\int_t^{x^F} \ln \frac{x}{t} dF(x) - \int_t^{x^F} \left(1 - \frac{t}{x}\right) dF(x)\right\}}{\left\{(1-F(t))^{-1} \int_t^{x^F} \left(1-\frac{t}{x}\right) dF(x)\ \right\}^2} = \frac{E\left(\ln\left(\frac{X}{t}\right) | X > t\right) - E\left(1 - \frac{t}{X} | X > t\right)}{E^2\left(1 - \frac{t}{X} | X > t\right)},$$

and replacing F and t by the empirical d.f. F_n and a high random threshold $X_{n-k,n}$ with $k < n$, respectively. The result is

$$\widehat{\varphi}_n(k) := \frac{M_n^{(1)}(k) - L_n^{(1)}(k)}{\left(L_n^{(1)}(k)\right)^2}, \tag{1.4}$$

where, we define for all $j \geq 1$,

$$L_n^{(j)}(k) := \frac{1}{k}\sum_{i=1}^{k}\left(1 - \frac{X_{n-k,n}}{X_{n-i+1,n}}\right)^j, \; M_n^{(j)}(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\ln \frac{X_{n-i+1,n}}{X_{n-k,n}}\right)^j. \tag{1.5}$$

The statistic in (1.4) is easily transformed into what we call the *Mixed Moment* (*MM*) estimator for the *extreme value index* $\gamma \in \mathbb{R}$:

$$\widehat{\gamma}_n^{MM}(k) \equiv \widehat{\gamma}_n^{MM}(k; X_{n-j+1,n}, 1 \leq j \leq k+1) := \frac{\widehat{\varphi}_n(k) - 1}{1 + 2\min\left(\widehat{\varphi}_n(k) - 1, 0\right)}. \tag{1.6}$$

Since the estimator in (1.6) is not location invariant, we also propose an alternative class of extreme value index estimators, invariant for changes in location, and dependent on a *tuning parameter* p, $0 < p < 1$. Such an estimator has the same functional expression of the estimator in (1.6), but the original sample, $X_i,\ 1 \leq i \leq n$, is replaced everywhere by

$$X_i^* := X_i - X_{[np]+1,n},\ 0 < p < 1,\ 1 \leq i \leq n.$$

We shall thus define for any $p \in (0, 1)$,

$$\widehat{\gamma}_n^{MM}(k; p) := \widehat{\gamma}_n^{MM}(k; X_{n-j+1,n} - X_{[np]+1,n}, 1 \leq j \leq k+1). \tag{1.7}$$

A similar procedure can be applied to estimators like the Hill, the moment and the generalized Hill estimators, and has already been used for quantile estimation in Araújo Santos, Fraga Alves and Gomes, 2006.

The estimator in (1.6) seems an interesting alternative to the most popular *extreme value index* estimators for a general $\gamma \in \mathbb{R}$. The most attractive features of this new estimator are:

- It is valid for any $\gamma \in \mathbb{R}$ and, contrary to the ML estimator (valid only for $\gamma > -1/2$), it has a simple explicit functional form, similar to the ones of Pickands, moment and generalized Hill estimators, also valid for all $\gamma \in \mathbb{R}$.
- It is very close to the ML estimator for a large class of models with $\gamma \geq 0$.
- A shift invariant version with similar properties is available, the one provided in (1.7). The asymptotic variance of the estimators in (1.7) is the same and the dominant component of asymptotic bias is never bigger, provided we keep to adequate k-values and choose an adequate *tuning parameter* p whenever $\gamma \leq 0$.
- There are accompanying shift and scale estimators that make e.g. high quantile estimation straightforward.

The scope of this paper is as follows. In section 2, we state the main results leading to weak consistency and asymptotic normality of the estimator in (1.6) and its location invariant version in (1.7). Section 3 is devoted to a small-scale Monte Carlo simulation, illustrating the performance of the new proposed estimators. In section 4 we state and prove a few auxiliary results needed in section 5, where we provide the proofs of the theorems in section 2.

2 Main results

The following *extended regular variation* property (de Haan, 1984) is a well-known necessary and sufficient condition for $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$:

$$\lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \begin{cases} \frac{x^\gamma-1}{\gamma} & \text{if } \gamma \neq 0 \\ \ln x & \text{if } \gamma = 0 \end{cases}, \tag{2.1}$$

for every $x > 0$ and some positive measurable function a, with U standing for a quantile type function associated to F and defined by

$$U(t) := \left(\frac{1}{1-F}\right)^{\leftarrow}(t) = \inf\left\{x : F(x) \geq 1 - \tfrac{1}{t}\right\}, \quad t \geq 1.$$

2.1 Consistency and asymptotic normality of the MM class of estimators

Theorem 2.1. *Suppose that F has a right endpoint $x^F = U(\infty) > 0$ and* (2.1) *holds for some $\gamma \in \mathbb{R}$. Let $k = k_n$ be an intermediate sequence, i.e., a sequence of positive integers k_n such that $k_n \to \infty$ and $k_n = o(n)$, as $n \to \infty$. Then,*

$$\lim_{n\to\infty} \widehat{\varphi}_n(k) = \varphi(\gamma)$$

in probability, with $\varphi(\gamma)$ defined in (1.3).

Corollary 2.1. *Under the conditions of Theorem 2.1, the mixed moment estimator $\widehat{\gamma}_n^{MM}(k)$ is a consistent estimator of the extreme value index $\gamma \in \mathbb{R}$, i.e., for any intermediate sequence $k = k_n$,*

$$\widehat{\gamma}_n^{MM}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \gamma.$$

Apart from the first order condition in (2.1), we shall now need a second order condition, specifying the rate of convergence in (2.1). We shall assume the existence of a function A, possibly not changing in sign and tending to zero as $t \to \infty$, such that

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma-1}{\gamma}}{A(t)} = H_{\gamma,\rho}(x) := \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right) \tag{2.2}$$

for all $x > 0$, where $\rho \le 0$ is a *second order* parameter controlling the speed of convergence of maximum values, linearly normalized, towards the limit law in (1.2). Under these circumstances, we say that the function U is of *second order extended regular variation*, and use the notation $U \in 2ERV(\gamma, \rho)$. We remark that $\lim_{t\to\infty} A(tx)/A(t) = x^\rho$, for every $x > 0$, i.e., $|A|$ is regularly varying with an index of regular variation equal to ρ. We denote this by $|A| \in RV_\rho$.

From Theorem A in Draisma, de Haan, Peng and Pereira, 1999, together with the modifications in Ferreira, de Haan and Peng, 2003 and de Haan and Ferreira, 2006, we know that if $x^F > 0$ and there exist $a(\cdot)$ and $A(\cdot)$ such that (2.2) holds, with $\rho \le 0$, $\gamma \ne \rho$, then, with

$$\overline{A}(t) := \left(\frac{a(t)}{U(t)} - \gamma_+\right), \qquad \gamma_+ := \max(0,\ \gamma), \tag{2.3}$$

and

$$l := \lim_{t\to\infty}\left(U(t) - \frac{a(t)}{\gamma}\right) \in \mathbb{R} \quad \text{for } \gamma+\rho<0, \tag{2.4}$$

we have

$$\overline{A}(t) \underset{t\to\infty}{\longrightarrow} 0 \text{ (see Lemma 4.1)} \quad \text{and} \quad \frac{\overline{A}(t)}{A(t)} \underset{t\to\infty}{\longrightarrow} c,$$

with

$$c = \begin{cases} 0 & \text{if} \quad \gamma < \rho \le 0 \\ \frac{\gamma}{\gamma+\rho} & \text{if} \quad 0 \le -\rho < \gamma \text{ or } (0<\gamma<-\rho \text{ and } l=0) \\ \pm\infty & \text{if} \quad \gamma+\rho=0 \text{ or } (0<\gamma<-\rho \text{ and } l\ne 0) \text{ or } \rho<\gamma\le 0 \end{cases}. \tag{2.5}$$

Let us define

$$B(t) := \begin{cases} \overline{A}(t) & \text{if} \quad c = \pm\infty \\ A(t) & \quad \text{otherwise} \end{cases}, \tag{2.6}$$

with $\overline{A}(t)$ the function in (2.3).

Theorem 2.2. *Assume that U satisfies* (2.2)*, with the restriction $\gamma \ne \rho$, and that $x^F > 0$. Let $k = k_n$ be an intermediate sequence, let c be the limit in* (2.5) *and let B be the function in* (2.6)*. If we further assume that*

$$\lambda := \lim_{n\to\infty} \sqrt{k}\ B(n/k) \tag{2.7}$$

is finite, we may guarantee that

$$\sqrt{k}\left(\widehat{\varphi}_n(k) - \varphi(\gamma)\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N\left(\lambda\ b_\varphi, \sigma_\varphi^2\right),$$

where

$$b_\varphi = b_\varphi(\gamma,\rho) := \begin{cases} \frac{1-\gamma}{(1-2\gamma)(1-\gamma-\rho)(1-2\gamma-\rho)} & \textit{if} \quad \gamma<\rho\le 0 \\ \frac{-4\gamma(1-\gamma)}{(1-2\gamma)^2(1-3\gamma)} & \textit{if} \quad \rho<\gamma<0 \\ \frac{1+\gamma}{(1-\rho)(1+\gamma-\rho)} & \textit{if} \quad c = \frac{\gamma}{\gamma+\rho} \\ 0 & \quad \textit{otherwise} \end{cases}$$

and

$$\sigma_\varphi^2 = \sigma_\varphi^2(\gamma) := \begin{cases} (1+\gamma)^2 & \text{if} \quad \gamma \geq 0 \\ \frac{(1-\gamma)^2(6\gamma^2-\gamma+1)}{(1-2\gamma)^3(1-3\gamma)(1-4\gamma)} & \text{if} \quad \gamma < 0 \end{cases}.$$

Corollary 2.2. *Under the conditions of Theorem 2.2*

$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k) - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N\left(\lambda\, b, \sigma^2\right),$$

where $b = b_{MM}(\gamma,\rho) := \begin{cases} b_\varphi(\gamma,\ \rho) & \text{if} \quad \gamma \geq 0 \\ (1-2\gamma)^2 b_\varphi(\gamma,\rho) & \text{if} \quad \gamma < 0 \end{cases}$ *and*

$$\sigma^2 = \sigma_{MM}^2(\gamma) := \begin{cases} \sigma_\varphi^2(\gamma) & \text{if} \quad \gamma \geq 0 \\ (1-2\gamma)^4\sigma_\varphi^2(\gamma) & \text{if} \quad \gamma < 0 \end{cases}.$$

We may further specify a non-null asymptotic bias in the region where $b_\varphi = 0$, the region $\mathcal{R} := \{(\gamma,\rho) : \rho < \gamma = 0 \text{ or } (0 < \gamma < -\rho \text{ and } l \neq 0) \text{ or } \gamma = -\rho\}$. It is then necessary to split $\mathcal{R}$ in three regions, $\mathcal{R} = \mathcal{R}_0 \cup \mathcal{R}_1 \cup \mathcal{R}_2$, with $\mathcal{R}_0 := \{(\gamma,\rho) : \gamma = -\rho/2\}$,

$$\begin{aligned} \mathcal{R}_1 &:= \{(\gamma,\rho) : \rho < \gamma = 0 \text{ or } (0 < \gamma < -\rho/2,\ l \neq 0)\}, \\ \mathcal{R}_2 &:= \{(\gamma,\rho) : (-\rho/2 < \gamma < -\rho,\ l \neq 0) \text{ or } \gamma = -\rho\}. \end{aligned}$$

Note that in $\mathcal{R}_1$, $A = o\big(\overline{A}^2\big)$ and in $\mathcal{R}_2$, $\overline{A}^2 = o(A)$.

Theorem 2.3. *Under the conditions in Theorem 2.2, if we now further assume that in the region* $\mathcal{R}_1 \cup \mathcal{R}_2$*,*

$$\lambda := \lim_{n\to\infty} \begin{cases} \sqrt{k}\ \overline{A}^2(n/k) & \text{if} \quad (\gamma,\rho) \in \mathcal{R}_1 \\ \sqrt{k}\ A(n/k) & \text{if} \quad (\gamma,\rho) \in \mathcal{R}_2 \end{cases}$$

is finite, we may guarantee that

$$\sqrt{k}\left(\widehat{\varphi}_n(k) - \varphi(\gamma)\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N\left(\lambda\, b_0, \sigma_0^2\right),$$

where

$$b_0 = b_0(\gamma,\rho) := \begin{cases} \frac{2(1+\gamma)}{(1+2\gamma)^2(1+3\gamma)} & \text{if} \quad (\gamma,\rho) \in \mathcal{R}_1 \\ \frac{1+\gamma}{(1-\rho)(1+\gamma-\rho)} \equiv b_\varphi & \text{if} \quad (\gamma,\rho) \in \mathcal{R}_2 \end{cases}$$

and $\sigma_0^2 = \sigma_\varphi^2 = (1+\gamma)^2$.

Remark 2.1. *Corollary 2.2 remains valid under the conditions of Theorem 2.3, with* b_φ *replaced by* b_0.

2.2 Asymptotic relations between estimators

We first compare the estimator in (1.6) with the moment estimator,

$$\widehat{\gamma}_n^M(k) := M_n^{(1)}(k) + \tfrac{1}{2}\left\{1 - (M_n^{(2)}(k)/[M_n^{(1)}(k)]^2 - 1)^{-1}\right\}, \tag{2.8}$$

with $M_n^{(j)}$ given in (1.5). We shall next compare the Mixed Moment estimator in (1.6) with the so-called *Maximum Likelihood* (ML) estimator, $\widehat{\gamma}_n^{ML}(k)$, obtained on the basis of the excesses $X_{n-i+1,n} - X_{n-k,n}$, $1 \le i \le k$. These excesses are approximately distributed as the k top order statistics associated to a sample of size k from a *Generalized Pareto* (GP) d.f., given by

$$GP_\gamma(x;\sigma) = \begin{cases} 1-(1+\gamma x/\sigma)^{-1/\gamma}\,,\ x\ge 0,\ 1+\gamma\, x/\sigma>0 & \text{if} \quad \gamma \neq 0 \\ 1-\exp\left(-x/\sigma\right),\ x\ge 0 & \text{if} \quad \gamma=0 \end{cases},$$

with $\sigma > 0$. The asymptotic behaviour of $\widehat{\gamma}_n^{ML}(k)$ has been worked out by Smith, 1987 and Drees, Ferreira and de Haan, 2004. We shall prove the following result:

Theorem 2.4. *Let $\widehat{\gamma}_n^{ML}(k)$ be a sequence of solutions of the maximum likelihood equations associated to the above mentioned set-up and let us consider the moment estimator $\widehat{\gamma}_n^M(k)$ defined in* (2.8)*. Assume also that $F \in \mathcal{D}_{\mathcal{M}}(G_\gamma)$,* (2.2) *holds with $\gamma \neq \rho$ and $k = k_n$ is an intermediate sequence such that $\sqrt{k}\, B(n/k) \to 0$, as $n \to \infty$, with B the function in* (2.6)*. Then,*

- *if $\gamma \ge 0$,*
$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k) - \widehat{\gamma}_n^{ML}(k)\right) \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0; \tag{2.9}$$
- *if $\gamma \le 0$,*
$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k) - \widehat{\gamma}_n^{M}(k)\right) \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0. \tag{2.10}$$

Furthermore, if $\sqrt{k}\,\overline{A}(n/k) = O(1)$, as $n \to \infty$, and $\gamma \ge 0$ or $\sqrt{k}\,A(n/k) = O(1)$, as $n \to \infty$ and $\big(l = 0$ or $\gamma > -\rho/2, l \neq 0\big)$ then (2.9) *still holds, whereas if $\sqrt{k}\,A(n/k) = O(1)$, as $n \to \infty$ and $\gamma < \rho \le 0$,* (2.10) *is also valid.*

Remark 2.2. *In the region $\{0 < \gamma \le -\rho/2,\ l \neq 0\}$, we can no longer guarantee the validity of* (2.9) *for levels k such that $\sqrt{k}\ A(n/k) = O(1)$, as $n \to \infty$. Indeed, if $\{0 < \gamma < -\rho/2,\ l \neq 0\}$, the bias of the ML estimator is of a smaller order than the bias of the MM estimator. We thus expect, in this region, a better performance of the ML estimator, comparatively to the MM estimator. Note however that in this same region, the asymptotic bias of $MM(k)$ is of the order of $\overline{A}^2(n/k)$, i.e., of a smaller order than $\overline{A}(n/k)$, the asymptotic bias of the Hill, the moment and the generalized Hill estimators.*

Remark 2.3. *For heavy-tailed models, and regarding bias, the new estimator compares favourably with the classical Hill estimator,* $\widehat{\gamma}_n^H(k) \equiv M_n^{(1)}(k)$*, with* $M_n^{(1)}(k)$ *given in* (1.5)*, in the whole* $(\gamma,\ \rho)$*-plane. Indeed, for levels* k *such that* $\sqrt{k}\ \overline{A}(n/k) \to \lambda \neq 0$*, the ones that are usually considered under a heavy-tailed second order framework, the asymptotic bias of* $\sqrt{k}\left(\widehat{\gamma}_n^H(k) - \gamma\right)$ *is equal to* $\lambda/(1-\rho) =: \lambda b_{_H}$*, whereas we are now able to get a null bias in the region* $\gamma + \rho \leq 0$*. Moreover,* $(1+\gamma)(1+\gamma-\rho) < 1$ *for all* $\gamma > 0$*,* $\rho < 0$*, i.e.,* $b_{_{MM}} < b_{_H}$ *in all semi-plane* $\gamma > 0$*. A similar remark applies to the moment's estimator.*

2.3 Asymptotic behavior of the location invariant versions

If we consider the shift invariant version of the extreme value index estimator in (1.6), i.e., the estimator in (1.7), the asymptotic variance of $\widehat{\gamma}_n^{MM}(k;p)$ is kept at the same level and the dominant component of bias changes only in a few cases, as may be seen in the following:

Theorem 2.5. *Assume that all the requirements in Theorem 2.2 hold, except possibly the requirement* $x^F > 0$*. Assume further that* U *has a positive derivative at the point* $1/(1-p)$*. Then, for the same levels* k *as in Theorem 2.2, i.e., intermediate levels* k *such that* (2.7) *holds, for* λ *finite,*

$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k;\ p) - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} N\left(\lambda b^*, \sigma^2\right),$$

with

$$b^* = b^*(\gamma,\rho,p) := \begin{cases} b & \text{if} \quad \gamma > 0 \ \text{or}\ \gamma < \rho \leq 0 \\ \frac{b\ x^F}{x^F - U(1/(1-p))} & \text{if} \quad \rho < \gamma \leq 0 \end{cases},$$

where b *and* σ^2 *are defined in Corollary 2.2 and* $x^F/\left(x^F - U(1/(1-p))\right)$ *is interpreted as 1 when* $x^F = \infty$*.*

Remark 2.4. *Recall that in the region* $0 < \gamma \leq -\rho/2,\ l \neq 0$*, in order to get a possibly non null and finite value for the asymptotic bias, like the one provided in Theorem 2.3, we had to consider* k *values such that* $\sqrt{k}\ \overline{A}^2(n/k) \to \lambda$*, finite, as* $n \to \infty$*. For these values of* k*, the changes in the asymptotic bias are then significant, when we consider the shift invariant estimator. Everything depends on the relative behaviour of the three regularly varying functions,* $\overline{A}^2$*,* $\overline{A}/U$ *and* $1/U^2$*, all with an index of regular variation equal to* -2γ*. This is the main reason why we think sensible to restrict ourselves to levels* k *such that* $\sqrt{k}\ \overline{A}(n/k) \to \lambda$*, getting then a null value for* b*.*

2.4 An estimator for the scale $a(n/k)$

In order to tackle the extreme quantile estimation problem, we shall next introduce an estimator for the scale $a(n/k)$, on the basis of the fact that not only

$$\frac{X_{n-k,n}\ L_n^{(1)}(k)\,(1+|\gamma|)}{a(n/k)} \xrightarrow[n\to\infty]{P} 1$$

(see Proposition 4.2), but also that it is possible to combine $L_n^{(1)}(k)$ and $L_n^{(2)}(k)$, both given in (1.5), in order to obtain a consistent estimator for $1+|\gamma|$ (see Proposition 4.2, again). Specifically,

$$\frac{1}{2}\left(1-\frac{\left(L_n^{(1)}(k)\right)^2}{L_n^{(2)}(k)}\right)^{-1} \xrightarrow[n\to\infty]{P} 1+|\gamma|,$$

and this enables us to introduce the following estimator for $a(n/k)$:

$$\widehat{a}\left(\frac{n}{k}\right) := \frac{X_{n-k,n}}{2} \times \frac{L_n^{(1)}(k)\ L_n^{(2)}(k)}{L_n^{(2)}(k)-\left[L_n^{(1)}(k)\right]^2}. \tag{2.11}$$

Theorem 2.6. *Suppose F satisfies* (2.1) *for some $\gamma \in \mathbb{R}$ and $x^F > 0$. Let $k = k_n$ be an intermediate sequence. Then, with $\widehat{a}(n/k)$ given in* (2.11),

$$\frac{\widehat{a}(n/k)}{a(n/k)} \xrightarrow[n\to\infty]{P} 1. \tag{2.12}$$

Let us further assume that the second order condition (2.2) *holds with $\gamma \neq \rho$, and that, with $\overline{A}$ and c given in* (2.3) *and* (2.5), *respectively,* (2.7) *holds, for λ finite. Then,*

$$\sqrt{k}\left(\frac{\widehat{a}(n/k)}{a(n/k)}-1\right) \xrightarrow[n\to\infty]{d} N\left(\lambda b_*, \sigma_*^2\right),$$

where

$$b_* = b_*(\gamma,\rho) := \begin{cases} -\frac{\rho}{(1-\gamma-\rho)(1-2\gamma-\rho)} & \text{if} \ \ \gamma<\rho\le 0 \\ \frac{2\ \gamma}{(1-2\gamma)(1-3\gamma)} & \text{if} \ \ \rho<\gamma\le 0 \\ \frac{(1+\gamma)(1+2\gamma-2\rho)}{(\gamma+\rho)(1+\gamma-\rho)(1+2\gamma-\rho)} & \text{if} \ \ c=\frac{\gamma}{\gamma+\rho} \\ -\frac{2\gamma}{(1+2\gamma)(1+3\gamma)} & \quad \text{otherwise} \end{cases},$$

and

$$\sigma_*^2 = \sigma_*^2(\gamma) := \left(\gamma_+\right)^2 + \frac{2(1+|\gamma|)^2\left(1+6|\gamma|+12\gamma^2\right)}{(1+2|\gamma|)(1+3|\gamma|)(1+4|\gamma|)}, \quad \gamma\in\mathbb{R}.$$

Theorem 2.7. *Assume that the second order condition* (2.2) *holds, with* $\gamma \neq \rho$, *and that* $x^F > 0$. *Let* $k = k_n$ *be an intermediate sequence such that* (2.7) *holds, for* λ *finite. Then*

$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k) - \gamma,\ \frac{\widehat{a}(n/k)}{a(n/k)} - 1,\ \frac{X_{n-k,n} - U(n/k)}{a(n/k)}\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} (\Gamma,\ \Lambda,\ B),$$

where $(\Gamma,\ \Lambda,\ B)$ *has a joint normal distribution with mean vector* $(\lambda b, \lambda b_*, 0)$, b *and* b_* *provided in Corollary 2.2 and Theorem 2.6, respectively, and covariances given by:* $Cov(\Gamma, B) = 0$,

$$Cov\,(\Gamma, \Lambda) = \begin{cases} -\frac{(1-\gamma)^2(12\gamma^2-4\gamma+1)}{(1-2\gamma)^4(1-3\gamma)(1-4\gamma)} & if\ \ \gamma \le 0 \\ -\frac{(1+\gamma)^2}{1+3\gamma} & if\ \ \gamma > 0 \end{cases} \quad ; \quad Cov\,(\Lambda, B) = \begin{cases} 0 & if\ \ \gamma \le 0 \\ -\gamma & if\ \ \gamma > 0 \end{cases}.$$

3 Finite sample properties

We have run a small-scale Monte Carlo simulation, on the basis of $N = 1000$ runs, for underlying Fréchet(γ) parents, with d.f. $F(x) = \exp(-x^{-1/\gamma}),\ x \ge 0\ (\gamma = 1)$, for a sample size $n = 1000$ ($\rho = -2,\ \tilde{\rho} = -1,\ c = -\infty,\ l = -1/2$). In Figure 1, we picture the mean values (E) and the mean squared errors (MSE) of the mixed moment estimator in (1.6) and its location invariant versions in (1.7) associated to $p = 0.01$ and $p = 0.25$. We use for these three estimators the obvious notation MM, $MM^{(0.01)}$ and $MM^{(0.25)}$, respectively. For comparison, we also picture the same characteristics for the Hill, the Moment, the generalized Hill and the POT-maximum likelihood estimators, denoted H, M, GH and ML, respectively. The generalized Hill estimator can be written as a functional of the Hill estimators, $\widehat{\gamma}_n^H(i) = M_n^{(1)}(i)$, $1 \le i \le k$, with $M_n^{(j)}(k)$ given in (1.5), i.e., we have

$$\widehat{\gamma}_n^{GH}(k) = \widehat{\gamma}_n^H(k) + \frac{1}{k}\sum_{i=1}^{k}\left\{\ln \widehat{\gamma}_n^H(i) - \ln \widehat{\gamma}_n^H(k)\right\}.$$

For a Fréchet model with $\gamma = 1$, we have $\gamma = -\rho/2$ and $l \neq 0$, i.e., we are in the region mentioned in Remark 2.2, where we cannot guarantee asymptotic equivalence between the ML and the MM estimators for levels k such that $\sqrt{k}\ A(n/k) = O(1)$, as $n \to \infty$. However, the mean value of the MM estimator (Figure 2, left) provides some evidence of a better performance of the MM estimator comparatively to all alternative estimators under study in this paper. As expected, for models with a left endpoint at zero like this one, the shift invariant alternatives reveal a poor and poor performance as p increases.

This same type of graph appears for other models, and makes clear the importance of the new class of estimators. Note also that the class of estimators in (1.7) is location invariant, and such a nice property has been obtained practically without any kind of payment: the variance of the new class of estimators is the same, and the bias may even be smaller, as may be seen

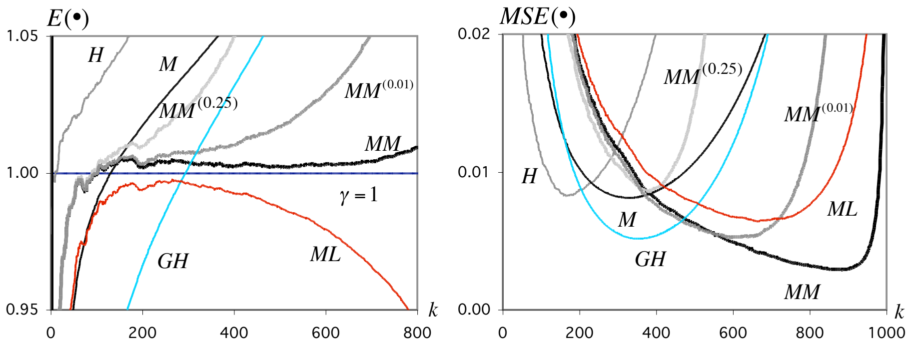

Figure 1: Mean values (*left*) and mean squared errors (*right*) of the MM estimators, together with the alternative estimators H, M, GH and ML, for samples of size $n = 1000$ from a Fréchet(1) model.

in Figure 2, equivalent to Figure 1, but for underlying extreme value G_γ parents, with $\gamma = 0.5$ $(\rho = -1,\ \tilde{\rho} = -\gamma = -0.5,\ l = 1/\gamma = 2)$.

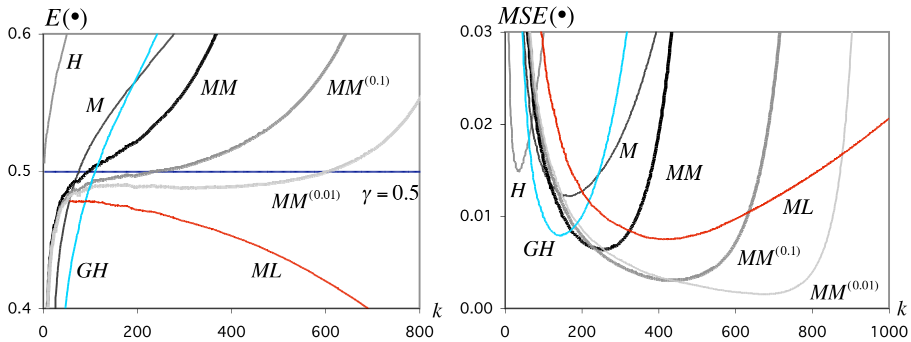

Figure 2: Mean values (*left*) and mean squared errors (*right*) of H, M, ML, GH and MM estimators, for samples of size $n = 1000$ from an extreme value G_γ model with $\gamma = 0.5$.

Note that for this model the left endpoint is $-1/\gamma < 0$, and consequently, the location invariant versions in (1.7) perform much better than the MM estimator for small p.

Indeed, we find particularly interesting the consideration of the tuning parameter p, which may help us on the choice of the "best" estimate, on the basis of any adequate stability criterion. Note however that, as illustrated in Araújo Santos *et al.*, 2006, a shift induced by small values of p when the underlying model, unknown, has an infinite left endpoint, may lead to stable sample paths around a target not close to the true value of γ. We illustrate such a fact with an underlying standard normal model. Figure 3 is thus equivalent to the previous figures, but for a standard normal underlying parent $(\gamma = 0,\ \rho = \tilde{\rho} = 0)$. As p decreases we get stable sample paths, but not so close to the target value γ as the optimal estimate provided via MM. Note that for $p = 0.5$ (a shift through the median), the MSE of $MM(0.5)$ is, as expected due to the symmetry of the normal model, almost overlapping the MSE of the MM estimator and quite interesting: smaller than the MSE of both the ML and the M estimators for all k.

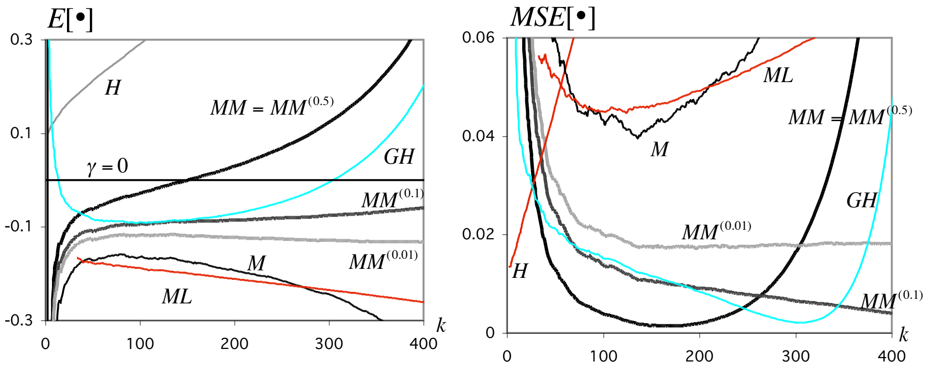

Figure 3: Mean values (*left*) and mean squared errors (*right*) of H, M, IH, ML and MM estimators, for samples of size $n = 1000$ from a standard normal model.

Finally, in Figure 4 we exhibit the behavior of the different estimators under consideration for an extreme value model with $\gamma = -0.1$ ($\rho = \tilde{\rho} = -1$). Note the fact that again as p decreases we get sample paths closer and closer to the target value γ. Note also the fact that despite the much larger bias of the mixed moment estimators, the statistic $MM^{(0.01)}$ clearly overpasses all other estimators considered, regarding mean squared error.

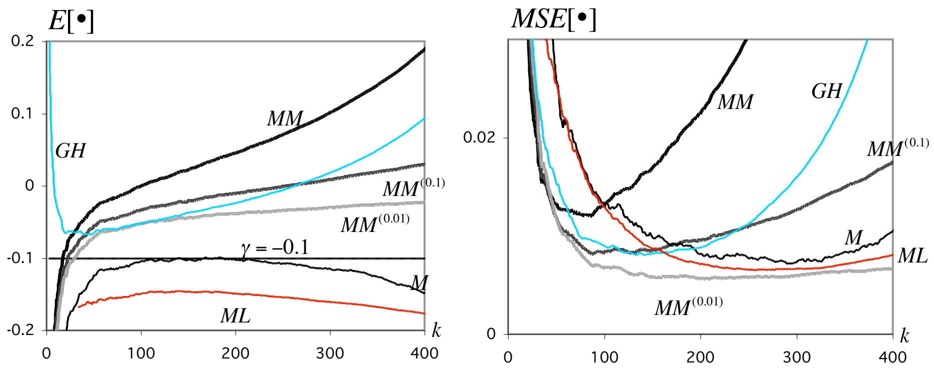

Figure 4: Mean values (*left*) and mean squared errors (*right*) of M, IH, ML and MM estimators, for samples of size $n = 1000$ from an extreme value G_γ model with $\gamma = -0.1$.

For this range of γ values (negative and close to 0), the M estimator has a quite small bias (look at the mean value pattern at Figure 5 (left)). But its variance is very very high and regarding MSE's, the MSE of $M(k)$ is, for all k, higher than the ones of both $ML(k)$ and $MM(p;k)$, provided that p is small.

Finally, in Figures 5 and 6, we try making clear the situation in the region of heavy tails mentioned in Remark 2.2, where the ML estimator should clearly outperform the MM estimator. We have simulated a Burr(γ, ρ, λ) model, with d.f. $F(x) = 1 - (1 + (x - \lambda)^{-\rho})^{-\gamma/\rho}$, $x \geq \lambda$. We then have $\widetilde{\rho} = \rho$, $l = \lambda$, i.e., $l \neq 0$ if $\lambda \neq 0$. In Figures 5 and 6 we have chosen $(\gamma,\ \rho,\ \lambda) = (0.75, -2, 0)$ and $(\gamma,\ \rho,\ \lambda) = (0.75, -2, 3)$, respectively.

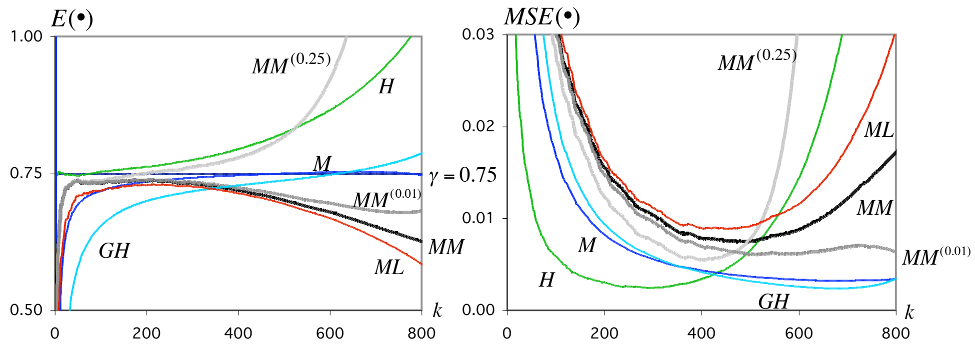

Figure 5: Mean values (*left*) and mean squared errors (*right*) of M, GH, ML and MM estimators, for samples of size $n = 1000$ from a Burr model with $\gamma = 0.75$, $\rho = -2$ and location parameter $\lambda = 0$.

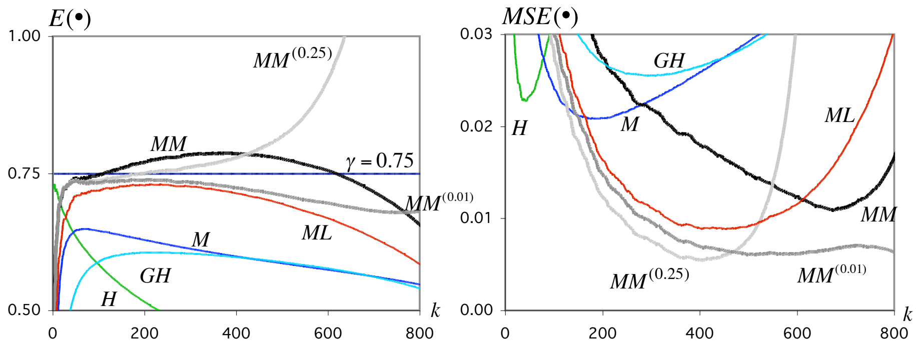

Figure 6: Mean values (*left*) and mean squared errors (*right*) of M, GH, ML and MM estimators, for samples of size $n = 1000$ from a Burr model with $\gamma = 0.75$, $\rho = -2$ and location parameter $\lambda = 3$.

We still make the following remark: even in the situations where, for heavy tails, the ML estimator compares favourably with the MM estimator, a location invariant version $MM(p)$, associated, for instance, to the value $p = 0.01$, outperforms the ML for all models in this region, with a finite left endpoint.

4 Auxiliary results

We state, without proof, the following lemma. For a proof, see de Haan and Ferreira (2006).

Lemma 4.1. *If* (2.1) *holds for some* $\gamma \in \mathbb{R}$, *then the auxiliary function* $a(t)$ *in* (2.1) *is of regular variation at infinity with index* γ, *i.e.,* $a \in RV_\gamma$ *and*

$$\lim_{t\to\infty} \frac{a(t)}{U(t)} = \gamma_+ := \max(0, \gamma).$$

Moreover, if $\gamma > 0$, *both functions* a *and* U *belong to* RV_γ; *if* $\gamma < 0$, *then* $x^F := \lim_{t\to\infty} U(t)$ *exists,* $\lim_{t\to\infty} a(t)/(x^F - U(t)) = -\gamma$ *and* $x^F - U \in RV_\gamma$.

Furthermore, with $\gamma_- := \min(\gamma, 0)$, *and provided that* $x^F = U(\infty) > 0$,

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t)}{a(t)/U(t)} = \frac{x^{\gamma_-} - 1}{\gamma_-}, \quad \textit{for every } x > 0. \tag{4.1}$$

Lemma 4.1 above together with Drees inequality (Drees, 1998) and Proposition 1.7 in Geluk and de Haan, 1987 yield the following uniform bounds:

Lemma 4.2. *If* (2.1) *holds for some* $\gamma \in \mathbb{R}$ *then, for any* $\epsilon > 0$, *there exists* $t_0 = t_0(\epsilon)$ *such that for* $t \geq t_0$ *and* $x \geq 1$,

$$\left| \frac{U(tx) - U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma} \right| \leq \epsilon\ x^{\gamma+\epsilon}, \quad (1-\epsilon)\ x^{-\epsilon} < \frac{x^{\gamma_+} U(t)}{U(tx)} < (1+\epsilon)\ x^\epsilon$$

and

$$\left| \frac{\ln U(tx) - \ln U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-} - 1}{\gamma_-} \right| \leq \epsilon\ x^{\gamma_- + \epsilon}.$$

In Fraga Alves *et al.* (2007), we find a proof of the following:

Proposition 4.1 (Fraga Alves *et al.*, 2007)**.** *Let* $U \in 2ERV(\gamma, \rho)$ *as introduced in* (2.2). *Let* c *be the limit in* (2.5).

(*i*) *If* $\gamma > 0$,

$$\lim_{t\to\infty} \frac{\frac{U(t)}{U(tx)} - x^{-\gamma}}{\widetilde{A}(t)} = K_{\gamma,\rho}(x) := \begin{cases} -x^{-\gamma}\ \frac{x^\rho - 1}{\rho} & \textit{if} \quad c = \frac{\gamma}{\gamma+\rho} \\ -x^{-\gamma}\ \frac{x^{-\gamma}-1}{-\gamma} & \textit{if} \quad c = \pm\infty \end{cases}, \tag{4.2}$$

and

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{\widetilde{A}(t)} = \widetilde{K}_{\gamma,\rho}(x) := \begin{cases} \frac{x^\rho - 1}{\rho} & \textit{if} \quad c = \frac{\gamma}{\gamma+\rho} \\ \frac{x^{-\gamma}-1}{-\gamma} & \textit{if} \quad c = \pm\infty \end{cases}, \tag{4.3}$$

for all $x > 0$, *where, with* $\overline{A}(t)$ *given in* (2.3),

$$\widetilde{A}(t) := \begin{cases} \frac{\gamma\ A(t)}{\gamma+\rho} & \textit{if} \quad c = \frac{\gamma}{\gamma+\rho} \\ \overline{A}(t) & \textit{if} \quad c = \pm\infty \end{cases}, \tag{4.4}$$

Necessarily, $\left|\widetilde{A}\right| \in RV_{\widetilde{\rho}}$*, with*

$$\widetilde{\rho} = \begin{cases} \rho & \text{if} \quad c = \frac{\gamma}{\gamma+\rho} \\ -\gamma & \text{if} \quad c = \pm\infty \end{cases}. \tag{4.5}$$

(ii) If $\gamma \le 0$*, we need further to assume that* $\gamma \neq \rho$*. Then,*

$$\lim_{t\to\infty} \frac{\frac{U(t)}{a^*(t)}\left(1-\frac{U(t)}{U(tx)}\right) - \frac{x^\gamma-1}{\gamma}}{A^*(t)} = K^*_{\gamma,\,\rho}(x) = \begin{cases} x^\gamma \ln x & \text{if} \quad \gamma<\rho=0 \\ \frac{x^{\gamma+\rho}-1}{\gamma+\rho} & \text{if} \quad \gamma<\rho<0 \\ \frac{x^{2\gamma}-1}{2\gamma} & \text{if} \quad \rho<\gamma<0 \\ \ln^2 x & \text{if} \quad \rho<\gamma=0 \end{cases}, \tag{4.6}$$

where

$$A^*(t) = \begin{cases} \frac{A(t)}{\gamma} & \text{if} \quad \gamma<\rho=0 \\ \frac{A(t)}{\rho} & \text{if} \quad \gamma<\rho<0 \\ -\frac{2\,\overline{A}(t)}{\gamma} & \text{if} \quad \rho<\gamma<0 \\ -\overline{A}(t) & \text{if} \quad \rho<\gamma=0 \end{cases}, \tag{4.7}$$

and

$$a^*(t) = \begin{cases} a(t)\left(1-\frac{A(t)}{\gamma}\right) & \text{if} \quad \gamma<\rho=0 \\ a(t)\left(1-\frac{A(t)}{\rho}\right) & \text{if} \quad \gamma<\rho<0 \\ a(t)\left(1+\frac{2\,\overline{A}(t)}{\gamma}\right) & \text{if} \quad \rho<\gamma<0 \\ a(t) & \text{if} \quad \rho<\gamma=0 \end{cases}. \tag{4.8}$$

Necessarily, $|A^*| \in RV_{\rho^*}$*, with*

$$\rho^* = \begin{cases} \rho & \text{if} \quad \gamma<\rho\le 0 \\ \gamma & \text{if} \quad \rho<\gamma\le 0 \end{cases}. \tag{4.9}$$

We shall now provide in more detail the general behaviour of $\ln U(tx) - \ln U(t)$ and $U(t)/U(tx)$ in the region where the limiting value c in (2.5) is $\pm\infty$, i.e., the region where $A(t) = o\left(\overline{A}(t)\right)$. Recall that such a region is $\{(\gamma,\rho) : \rho<\gamma\le 0$ or $\gamma=-\rho$ or $(0<\gamma<-\rho,\ l\neq 0)\}$, with l given in (2.4). The results in the following lemma come straightforwardly from Taylor's expansion.

Lemma 4.3. *Assume that* (2.2) *holds. If* $\rho<\gamma\le 0$*, then*

$$\frac{U(t)}{U(tx)} = 1 - \overline{A}(t)\left(\tfrac{x^\gamma-1}{\gamma}\right) + \overline{A}^2(t)\left(\tfrac{x^\gamma-1}{\gamma}\right)^2 - \overline{A}^3(t)\left(\tfrac{x^\gamma-1}{\gamma}\right)^3(1+o(1)) - \frac{A(t)\,\overline{A}(t)}{\rho}\left(\tfrac{x^{\gamma+\rho}-1}{\gamma+\rho} - \tfrac{x^\gamma-1}{\gamma}\right)(1+o(1)). \tag{4.10}$$

If $0<\gamma\le-\rho$, *since* A *may dominate* $\overline{A}^2$ *as well as the other way round, we write*

$$\ln U(tx)-\ln U(t)=\gamma\ln x+\overline{A}(t)\left(\tfrac{x^{-\gamma}-1}{-\gamma}\right)+\frac{\gamma\ A(t)}{\gamma+\rho}\left(\tfrac{x^{\rho}-1}{\rho}-\tfrac{x^{-\gamma}-1}{-\gamma}\right)(1+o(1))$$
$$-\frac{1}{2}\ \overline{A}^2(t)\left(\tfrac{x^{-\gamma}-1}{-\gamma}\right)^2+o\left(\overline{A}^2(t)\right)\quad(4.11)$$

and

$$\frac{U(t)}{U(tx)}=x^{-\gamma}\left\{1-\overline{A}(t)\left(\tfrac{x^{-\gamma}-1}{-\gamma}\right)-\frac{\gamma\ A(t)}{\gamma+\rho}\left(\tfrac{x^{\rho}-1}{\rho}-\tfrac{x^{-\gamma}-1}{-\gamma}\right)(1+o(1))\right.$$
$$\left.+\overline{A}^2(t)\left(\tfrac{x^{-\gamma}-1}{-\gamma}\right)^2+o\left(\overline{A}^2(t)\right)\right\}.\quad(4.12)$$

After the statements in Proposition 4.1, and again on the basis of Drees' inequality (Drees, 1998), we may now state the following:

Lemma 4.4. *If* (2.2) *holds for some* $\gamma>0$ *then, for any* $\epsilon>0$, *there exists* $t_0=t_0(\epsilon)$ *such that for* $t\ge t_0$ *and* $x\ge1$,

$$\left|\frac{\frac{U(t)}{U(tx)}-x^{-\gamma}}{\widetilde{A}(t)}-K_{\gamma,\rho}(x)\right|\le\epsilon x^{-\gamma+\widetilde{\rho}+\epsilon},$$

$$\left|\frac{\left(1-\frac{U(t)}{U(tx)}\right)^2-\left(1-x^{-\gamma}\right)^2}{\widetilde{A}(t)}-2(x^{-\gamma}-1)K_{\gamma,\rho}(x)\right|\le\epsilon x^{-2\gamma+\widetilde{\rho}+\epsilon},$$

$$\left|\frac{\ln U(tx)-\ln U(t)-\gamma\ln x}{\widetilde{A}(t)}-\widetilde{K}_{\gamma,\rho}(x)\right|\le\epsilon x^{\widetilde{\rho}+\epsilon},$$

with $A(t)$, $K_{\gamma,\rho}(x)$, $\widetilde{A}(t)$, $\widetilde{\rho}$ *and* $\widetilde{K}_{\gamma,\rho}(x)$ *given in* (2.2), (4.2), (4.4), (4.5) *and* (4.3), *respectively.*

If (2.2) *holds for some* $\gamma\le0$ *then also, for any* $\epsilon>0$, *there exists* $t_0=t_0(\epsilon)$ *such that for* $t\ge t_0$ *and* $x\ge1$,

$$\left|\frac{\frac{U(t)}{a^*(t)}\left(1-\frac{U(t)}{U(tx)}\right)-\frac{x^{\gamma}-1}{\gamma}}{A^*(t)}-K^*_{\gamma,\rho}(x)\right|\le\epsilon\ x^{\gamma+\rho^*+\epsilon}$$

and

$$\left|\frac{\left(\frac{U(t)}{a^*(t)}\left(1-\frac{U(t)}{U(tx)}\right)\right)^2-\left(\frac{x^{\gamma}-1}{\gamma}\right)^2}{A^*(t)}-2\left(\frac{x^{\gamma}-1}{\gamma}\right)K^*_{\gamma,\rho}(x)\right|\le\epsilon x^{2\gamma+\rho^*+\epsilon},$$

now with $K^*_{\gamma,\rho^*}(x)$, $A^*(t)$, $a^*(t)$ *and* ρ^* *given in* (4.6), (4.7), (4.8) *and* (4.9), *respectively.*

Here and throughout the paper, let $\{Y_{i,n}\}_{i=1}^{n}$ be the ascending order statistics associated to the independent r.v.'s $\{Y_i\}_{i=1}^{n}$ with common standard Pareto d.f., $1-y^{-1}$, for all $y \geq 1$. For the proofs, bear in mind the equality in distribution

$$\{X_{i,n}\}_{i=1}^{n} \stackrel{d}{=} \{U(Y_{i,n})\}_{i=1}^{n}.$$

Since $k = k_n$ is an intermediate sequence, $Y_{n-k,n} \underset{n\to\infty}{\overset{\text{a.s.}}{\longrightarrow}} \infty$, Rényi's representation enables us to write

$$Q_{i,n} := \frac{Y_{n-i+1,n}}{Y_{n-k,n}} \stackrel{d}{=} Y_{k-i+1:k}, \; 1 \leq i \leq k, \tag{4.13}$$

and for any measurable function g,

$$\frac{1}{k}\sum_{i=1}^{k} g(Q_{i,n}) \stackrel{d}{=} \frac{1}{k}\sum_{i=1}^{k} g(Y_i). \tag{4.14}$$

Proposition 4.2. *Under the first order condition* (2.1)*, for intermediate sequences* $k = k_n$*, with* $M_n^{(j)}(k)$ *and* $L_n^{(j)}(k)$*,* $j \geq 1$*, given in* (1.5)*, the following limits in probability hold:*

$$\begin{aligned}
&\left(\frac{X_{n-k,n}}{a(n/k)}\right) M_n^{(1)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{1}{1-\gamma_-},\\
&\left(\frac{X_{n-k,n}}{a(n/k)}\right) L_n^{(1)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{1}{1+|\gamma|}, \quad \left(\frac{X_{n-k,n}}{a(n/k)}\right)^2 L_n^{(1)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{\gamma_+^{-1}}{1+|\gamma|},\\
&\left(\frac{X_{n-k,n}}{a(n/k)}\right)^2 L_n^{(2)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{2}{(1+|\gamma|)(1+2|\gamma|)},\\
&\left(\frac{X_{n-k,n}}{a(n/k)}\right)^3 L_n^{(3)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{6}{(1+|\gamma|)(1+2|\gamma|)(1+3|\gamma|)},\\
&\left(\frac{X_{n-k,n}}{a(n/k)}\right)^4 L_n^{(4)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{24}{(1+|\gamma|)(1+2|\gamma|)(1+3|\gamma|)(1+4|\gamma|)}.
\end{aligned}$$

Proof. First define

$$L_{k,n} := \left(\frac{k}{n}\right) Y_{n-k,n}.$$

Since $k = k_n$ is an intermediate sequence,

$$L_{k,n} \underset{n\to\infty}{\overset{P}{\longrightarrow}} 1 \tag{4.15}$$

(cf. e.g. Smirnov, 1952). If (2.1) holds, then, as stated in Lemma 4.1, $a \in RV_\gamma$. It thus follows from the uniform convergence theorem for regularly varying functions (see Geluk and de Haan, 1987, Theorem 1.3) that

$$\frac{a(Y_{n-k,n})}{a\left(\frac{n}{k}\right)} = \frac{a\left(\frac{n}{k}\, L_{k,n}\right)}{a\left(\frac{n}{k}\right)} \underset{n\to\infty}{\overset{P}{\longrightarrow}} 1. \tag{4.16}$$

Concerning $M_n^{(1)}(k)$, we may write

$$\left(\frac{X_{n-k,n}}{a(n/k)}\right) M_n^{(1)}(k) \overset{d}{=} \frac{a\left(Y_{n-k,n}\right)}{a(n/k)} \frac{1}{k} \sum_{i=1}^{k} \frac{\ln U\left(Y_{n-i+1,n}\right) - \ln U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)/U\left(Y_{n-k,n}\right)}$$

and the core of this part of the proof lies at relation (4.1). Hence, using (4.13) and applying the result in Lemma 4.2, related to $\ln U$, we get

$$\frac{1}{k} \sum_{i=1}^{k} \frac{\ln U\left(Y_{n-i+1,n}\right) - \ln U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)/U\left(Y_{n-k,n}\right)} < \frac{1}{k} \sum_{i=1}^{k} \frac{Q_{i,n}^{\gamma_-} - 1}{\gamma_-} + \frac{\epsilon}{k} \sum_{i=1}^{k} Q_{i,n}^{\gamma_- + \epsilon}. \tag{4.17}$$

Let us look at the right hand-side of (4.17): for an intermediate sequence $k = k_n$, the law of large numbers and (4.14) ensures that the first term converges in probability towards $\int_1^\infty y^{-2}\left(y^{\gamma_-} - 1\right)\ dy/\gamma_- = (1-\gamma_-)^{-1}$, as $n \to \infty$, while for any $\epsilon > 0$, the second term is such that

$$\frac{1}{k} \sum_{i=1}^{k} Q_{i,n}^{\gamma_- + \epsilon} \xrightarrow[n\to\infty]{P} \int_1^\infty y^{-2}\ y^{\gamma_- + \epsilon}\ dy < \infty.$$

A similar reasoning can be applied to the obvious lower bound that comes from Lemma 4.2, and we get

$$\frac{1}{k} \sum_{i=1}^{k} \frac{\ln U\left(Y_{n-i+1,n}\right) - \ln U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)/U\left(Y_{n-k,n}\right)} \xrightarrow[n\to\infty]{P} \frac{1}{1-\gamma_-}.$$

Hence, (4.16) and Slutsky's theorem yield the result.

The moment statistics $L_n^{(j)}(k)$, $j = 1, 2, 3$ are based on the random terms

$$\tfrac{X_{n-k,n}}{a(n/k)}\left(1 - \tfrac{X_{n-k,n}}{X_{n-i+1,n}}\right) \overset{d}{=} \tfrac{a\left(Y_{n-k,n}\right)}{a(n/k)}\ \tfrac{U(Y_{n-i+1,n}) - U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)}\ \tfrac{U\left(Y_{n-k,n}\right)}{U(Y_{n-i+1,n})},$$

for $i = 1, 2, \cdots, k$. Under condition (2.1), Lemma 4.2 guarantees that, for any $\epsilon > 0$, and sufficiently large n,

$$\frac{1}{k} \sum_{i=1}^{k} \frac{U\left(Y_{n-i+1,n}\right) - U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)} \frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)} < \frac{1+\epsilon}{k} \sum_{i=1}^{k} \left(\frac{Q_{i.n}^{\gamma} - 1}{\gamma} + \epsilon\ Q_{i,n}^{\gamma+\epsilon}\right) Q_{i,n}^{-\gamma_+ + \epsilon}.$$

The law of large numbers and (4.14) ensures that this upper bound is equal in distribution to

$$\frac{1+\epsilon}{k} \sum_{i=1}^{k} \frac{Y_i^{\gamma} - 1}{\gamma}\ Y_i^{-\gamma_+ + \epsilon} + \frac{\epsilon}{k} \sum_{i=1}^{k} Y_i^{\gamma_- + \epsilon}$$

and converges in probability towards

$$\frac{1+\epsilon}{\gamma} \int_1^\infty y^{-2}\left(y^{\gamma - \gamma_+ + \epsilon} - y^{\gamma_+ - \epsilon}\right)\ dy + \epsilon \int_1^\infty y^{-2} y^{\gamma_- + \epsilon}\ dy.$$

Since $\epsilon > 0$ is arbitrary, Lebesgue' s dominated convergence theorem may be applied to the first integral with dominating function given by $g(y) = ((1+\epsilon_0)/\gamma)\left(y^{\gamma_-+\epsilon_0} - y^{-(\gamma_+-\epsilon_0)}\right)$, for all $y \geq 1$ and $\epsilon < \epsilon_0 < 1+|\gamma|$, integrable against dy/y^2 and such that

$$\lim_{\epsilon\downarrow 0} \frac{1+\epsilon}{\gamma} \int_1^\infty \left(y^{\gamma_-+\epsilon_0} - y^{-(\gamma_+-\epsilon_0)}\right) \frac{dy}{y^2}$$
$$= \frac{1}{\gamma} \int_1^\infty \left(y^{\gamma_-} - y^{-\gamma_+}\right) \frac{dy}{y^2} = \frac{1}{(1-\gamma_-)(1+\gamma_+)} = \frac{1}{1+|\gamma|}.$$

The second integral is finite for any $\epsilon < 1-\gamma_-$. Consequently, the second sum in the upper bound is stochastically bounded. We can establish a similar asymptotic lower bound, and apply a similar reasoning to that lower bound. From (4.16) we thus get

$$\left(\frac{X_{n-k,n}}{a(n/k)}\right) L_n^{(1)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{1}{1+|\gamma|}. \tag{4.18}$$

Now, Lemma 4.1, jointly with (4.16), implies the convergence

$$\frac{X_{n-k,n}}{a(n/k)} \overset{d}{=} \frac{a\left(Y_{n-k,n}\right)}{a(n/k)} \frac{U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)} \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{1}{\gamma_+} \tag{4.19}$$

and (4.18) together with (4.19) imply that

$$\left(\frac{X_{n-k,n}}{a(n/k)}\right)^2 L_n^{(1)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{\gamma_+^{-1}}{1+|\gamma|}.$$

The proofs of the other limiting results are similar and are therefore omitted. ■

Lemma 4.5. *Let $k = k_n$ be an intermediate sequence. If the underlying quantile function U satisfies* (2.1) *with $\gamma \leq 0$, then, for any $j \geq m \geq 1$,*

$$\left(\frac{U\left(Y_{n-k,n}\right)}{a(n/k)}\right)^m \frac{1}{k} \sum_{i=1}^k \frac{U\left(Y_{n-i+1,n}\right)}{U\left(Y_{n-k,n}\right)} \left(1 - \frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)}\right)^j = O_p\left(\left(\frac{a(n/k)}{U\left(Y_{n-k,n}\right)}\right)^{j-m}\right).$$

Proof. Since $U\left(Y_{n-k,n}\right) \leq U\left(Y_{n-i+1,n}\right)$, for $i = 1, 2, \cdots, k$, we may write

$$\left(\frac{U\left(Y_{n-k,n}\right)}{a(n/k)}\right)^{m-j} \frac{1}{k} \sum_{i=1}^k \left(\frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)}\right)^{j-1} \left(\frac{U\left(Y_{n-i+1,n}\right) - U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)}\right)^j$$
$$\leq \left(\frac{U\left(Y_{n-k,n}\right)}{a(n/k)}\right)^{m-j} \left\{\frac{1}{k} \sum_{i=1}^k \left(\frac{U\left(Y_{n-i+1,n}\right) - U\left(Y_{n-k,n}\right)}{a\left(Y_{n-k,n}\right)}\right)^j\right\}.$$

It follows from the first inequality of Lemma 4.2 that the second factor is $O_p(1)$. ■

Remark 4.1. *If we consider $a^*(t)$ in (4.8), since $a^*(t) \sim a(t)$, as $t \to \infty$, the results in Proposition 4.2 and Lemma 4.5 obviously hold if we replace $a(n/k)$ by $a^*(n/k)$.*

The two following lemmas follow immediately from the central limit theorem (cf. Billingsley, 1979, Theorem 29.5):

Lemma 4.6. *For $\gamma > 0$,*

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\ln Y_i - 1,\ \frac{1}{k}\sum_{i=1}^{k}\left(1 - Y_i^{-\gamma}\right) - \frac{\gamma}{1+\gamma},\right.$$
$$\left.\frac{1}{k}\sum_{i=1}^{k}\frac{\left(1-Y_i^{-\gamma}\right)^2}{2} - \frac{\gamma^2}{(1+\gamma)(1+2\gamma)}\right) \xrightarrow[k\to\infty]{d} (P_0, P_1, P_2),$$

where (P_0, P_1, P_2) has a joint normal distribution with mean zero and covariance matrix given by

$$\begin{array}{lll} E\left(P_0^2\right) = 1, & E\left(P_0P_1\right) = \frac{\gamma}{(1+\gamma)^2}, & E\left(P_0P_2\right) = \frac{\gamma^2(2+3\gamma)}{(1+\gamma)^2(1+2\gamma)^2}, \\ & E\left(P_1^2\right) = \frac{\gamma^2}{(1+\gamma)^2(1+2\gamma)}, & E\left(P_1P_2\right) = \frac{2\gamma^3}{(1+\gamma)^2(1+2\gamma)(1+3\gamma)} \\ & & E\left(P_2^2\right) = \frac{\gamma^4(5+11\gamma)}{(1+\gamma)^2(1+2\gamma)^2(1+3\gamma)(1+4\gamma)} \end{array}.$$

Lemma 4.7. *For $\gamma \leq 0$,*

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{Y_i^{\gamma}-1}{\gamma}\right) - \frac{1}{1-\gamma}, \frac{1}{k}\sum_{i=1}^{k}\frac{1}{2}\left(\frac{Y_i^{\gamma}-1}{\gamma}\right)^2 - \frac{1}{(1-\gamma)(1-2\gamma)}\right) \xrightarrow[k\to\infty]{d} (P_1^*, P_2^*),$$

where (P_1^*, P_2^*) has a joint normal distribution with mean zero and covariance matrix given by

$$\begin{array}{ll} E\left((P_1^*)^2\right) = \frac{1}{(1-\gamma)^2(1-2\gamma)}, & E\left((P_2^*)^2\right) = \frac{5-11\gamma}{(1-\gamma)^2(1-2\gamma)^2(1-3\gamma)(1-4\gamma)}, \\ E\left(P_1^*P_2^*\right) = \frac{2}{(1-\gamma)^2(1-2\gamma)(1-3\gamma)} & . \end{array}$$

Again with Y a unit Pareto r.v, let us denote, for any $\alpha,\ \beta \leq 0$,

$$d_{\alpha,\beta} := E\left(Y^{\alpha}\left(\frac{Y^{\beta}-1}{\beta}\right)\right) = \frac{1}{(1-\alpha)(1-\alpha-\beta)}. \tag{4.20}$$

Proposition 4.3. *Assume that the second order condition (2.2) holds with $\gamma > 0$. If $k = k_n$ is an intermediate sequence, then, as $n \to \infty$,*

$$
\begin{aligned}
\frac{X_{n-k,n}}{a(n/k)} &= \frac{1}{\gamma} + \frac{B}{\sqrt{k}} - \frac{\overline{A}(n/k)}{\gamma^2}(1+o_p(1)) + o_p\left(\widetilde{A}(n/k)\right) + o_p\left(1/\sqrt{k}\right), \\
M_n^{(1)}(k) &= \gamma + \frac{\gamma\ P_0}{\sqrt{k}} + d_{M_1}\widetilde{A}(n/k) + o_p\left(\widetilde{A}(n/k)\right) + o_p\left(1/\sqrt{k}\right), \\
L_n^{(1)}(k) &= \frac{\gamma}{1+\gamma} + \frac{P_1}{\sqrt{k}} + d_{L_1}\widetilde{A}(n/k) + o_p\left(\widetilde{A}(n/k)\right) + o_p\left(1/\sqrt{k}\right), \\
\left\{L_n^{(1)}(k)\right\}^2 &= \frac{\gamma^2}{(1+\gamma)^2} + \frac{2\gamma}{1+\gamma}\ \frac{P_1(1+o_p(1))}{\sqrt{k}} + \frac{2\gamma\ d_{L_1}\widetilde{A}(n/k)(1+o_p(1))}{1+\gamma}, \\
L_n^{(2)}(k) &= \frac{2\ \gamma^2}{(1+\gamma)(1+2\gamma)} + \frac{2\ P_2(1+o_p(1))}{\sqrt{k}} + d_{L_2}\widetilde{A}(n/k)(1+o_p(1)),
\end{aligned}
$$

where $\widetilde{A}(t)$ is given in (4.4) *and B is a standard normal r.v. independent of* $(P_0,\ P_1,\ P_2)$, *the normally distributed vector in Lemma 4.6. Moreover, with $\widetilde{\rho}$ and $d_{\alpha,\beta}$ given in* (4.5) *and* (4.20)*, respectively, we get*

$$
d_{M_1} = d_{0,\widetilde{\rho}} = \frac{1}{1-\widetilde{\rho}}, \qquad d_{L_1} = d_{-\gamma,\widetilde{\rho}} = \frac{1}{(1+\gamma)(1+\gamma-\widetilde{\rho})},
$$

$$
d_{L_2} = 2\left(d_{-\gamma,\widetilde{\rho}} - d_{-2\gamma,\widetilde{\rho}}\right) = \frac{2\gamma(2+3\gamma-\widetilde{\rho})}{(1+\gamma)(1+2\gamma)(1+\gamma-\widetilde{\rho})(1+2\gamma-\widetilde{\rho})}.
$$

Proof. Since (4.2) holds, $(k\ Y_{n-k,n}/n)^\gamma \ = \ 1+\gamma\ B/\sqrt{k} + o_p\left(\widetilde{A}(n/k)\right) + o_p\left(1/\sqrt{k}\right)$, with B a standard normal r.v. (see de Haan and Ferreira, 2006, Theorem 2.4.2) and $X_{n-k,n} \stackrel{d}{=} U\left(Y_{n-k,n}\right)$, we may write, as $n\to\infty$,

$$
\frac{X_{n-k,n}}{a(n/k)} \ = \ \frac{U(n/k)}{a(n/k)}\left(1+\frac{\gamma\ B}{\sqrt{k}} + o_p\left(\widetilde{A}(n/k)\right) + o_p\left(1/\sqrt{k}\right)\right).
$$

Notice now that from (2.3), $a(n/k)/U(n/k) = \gamma + \overline{A}(n/k)$, and consequently,

$$
\frac{U(n/k)}{a(n/k)} = \frac{1}{\gamma}\left(1-\frac{\overline{A}(n/k)}{\gamma}(1+o(1))\right).
$$

Consequently,

$$
\frac{X_{n-k,n}}{a(n/k)} \ = \ \frac{1}{\gamma} + \frac{B}{\sqrt{k}} - \frac{\overline{A}(n/k)}{\gamma^2}(1+o_p(1)) + o_p\left(\widetilde{A}(n/k)\right) + o_p\left(1/\sqrt{k}\right).
$$

The use of (4.14) and (4.15), together with the inequalities in Lemma 4.4 for $\gamma>0$, and the notation $R_k := \frac{1}{k}\sum_{i=1}^k Y_i^{-\gamma+\tilde{\rho}+\epsilon}$, leads us to

$$
\begin{aligned}
M_n^{(1)}(k) &= \tfrac{\gamma}{k}\sum_{i=1}^{k}\ln Y_i + \widetilde{A}\left(\tfrac{n}{k}\right)\left(\tfrac{1}{k}\sum_{i=1}^{k}\widetilde{K}_{\gamma,\rho}(Y_i) + \tfrac{o_p(1)}{k}\sum_{i=1}^{k}Y_i^{\widetilde{\rho}+\epsilon}\right),\\
L_n^{(1)}(k) &= \tfrac{1}{k}\sum_{i=1}^{k}\left(1-Y_i^{-\gamma}\right) - \widetilde{A}\left(\tfrac{n}{k}\right)\left(\tfrac{1}{k}\sum_{i=1}^{k}K_{\gamma,\rho}(Y_i) + o_p(R_k)\right)\\
L_n^{(2)}(k) &= \tfrac{1}{k}\sum_{i=1}^{k}\left(1-Y_i^{-\gamma}\right)^2 - \widetilde{A}\left(\tfrac{n}{k}\right)\left(\tfrac{2}{k}\sum_{i=1}^{k}\left(1-Y_i^{-\gamma}\right)K_{\gamma,\rho}(Y_i) + o_p(R_k)\right),
\end{aligned}
$$

with $\left(L_n^{(j)}(k), M_n^{(j)}(k)\right)$, $K_{\gamma,\rho}(x)$, $\widetilde{A}(t)$ and $\widetilde{K}_{\gamma,\rho}(x)$ given in (1.5), (4.2), (4.4) and (4.3), respectively. The law of large numbers implies that the partial sums associated with the $o_p(1)$ terms are negligible. Hence, all the results in the proposition follow straightforwardly. ■

Let us now define, with $a^*(t)$ given in (4.8) for $\gamma \le 0$, $a^*(t) = a(t)$ if $\gamma > 0$, and for $j \ge m \ge 1$,

$$
M_1^* := \left(\tfrac{X_{n-k,n}}{a^*(n/k)}\right) M_n^{(1)}(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{1}{1-\gamma_-}, \tag{4.21}
$$

$$
L_j^* := \left(\tfrac{X_{n-k,n}}{a^*(n/k)}\right)^j L_n^{(j)}(k), \tag{4.22}
$$

$$
R_{k,n}^{(m,j)} := \left(\tfrac{U\left(Y_{n-k,n}\right)}{a^*(n/k)}\right)^m \frac{1}{k}\sum_{i=1}^{k}\tfrac{U(Y_{n-i+1,n})}{U\left(Y_{n-k,n}\right)}\left(1-\tfrac{U\left(Y_{n-k,n}\right)}{U(Y_{n-i+1,n})}\right)^j. \tag{4.23}
$$

Proposition 4.4. *Under the conditions in Proposition 4.1, for $\gamma \le 0$, with L_j^*, $j = 1, 2$ given in (4.22), and if $k = k_n$ is an intermediate sequence,*

$$
\begin{aligned}
L_1^* &\overset{d}{=} \frac{1}{1-\gamma} + \left(\frac{P_1^*}{\sqrt{k}} + d^*_{L_1} A^*\left(\frac{n}{k}\right)\right)(1+o_p(1)),\\
(L_1^*)^2 &\overset{d}{=} \frac{1}{(1-\gamma)^2} + \frac{2}{1-\gamma}\left(\frac{P_1^*}{\sqrt{k}} + d^*_{L_1} A^*\left(\frac{n}{k}\right)\right)(1+o_p(1)),\\
L_2^* &\overset{d}{=} \frac{2}{(1-\gamma)(1-2\gamma)} + 2\left(\frac{P_2^*}{\sqrt{k}} + d^*_{L_2} A^*\left(\frac{n}{k}\right)\right)(1+o_p(1)),
\end{aligned}
$$

as $n \to \infty$, where $(P_1^*,\ P_2^*)$ is normally distributed with mean vector zero and covariance matrix given in Lemma 4.7, and with ρ^* and $d_{\alpha,\beta}$ given in (4.9) and (4.20), respectively,

$$
d^*_{L_1} = \begin{cases} d_{\gamma,0} = \frac{1}{(1-\gamma)^2} & \textit{if} \quad \gamma < \rho = 0 \\ 2 & \textit{if} \quad \rho < \gamma = 0 \\ d_{0,\gamma+\rho^*} = \frac{1}{1-\gamma-\rho^*} & \textit{otherwise} \end{cases}, \tag{4.24}
$$

$$
d^*_{L_2} = \begin{cases} \frac{d_{2\gamma,0}-d_{\gamma,0}}{\gamma} = \frac{2-3\gamma}{(1-\gamma)^2(1-2\gamma)^2} & \textit{if} \quad \gamma < \rho = 0 \\ 6 & \textit{if} \quad \rho < \gamma = 0 \\ \frac{d_{\gamma,\gamma+\rho^*}-d_{0,\gamma+\rho^*}}{\gamma} = \frac{2-2\gamma-\rho^*}{(1-\gamma)(1-\gamma-\rho^*)(1-2\gamma-\rho^*)} & \textit{otherwise} \end{cases}. \tag{4.25}
$$

Proof. The results in Lemma 4.4 for $\gamma \le 0$, with the notation $S_k := \frac{1}{k}\sum_{i=1}^k Q_{i,n}^{\gamma+\rho^*+\epsilon}$, lead us to the following distributional representations,

$$L_1^* \overset{d}{=} \frac{1}{k}\sum_{i=1}^k \frac{Q_{i,n}^\gamma - 1}{\gamma} + A^*\left(Y_{n-k,n}\right)\left(\frac{1}{k}\sum_{i=1}^k K_{\gamma,\rho}^*\left(Q_{i,n}\right) + o_p(S_k)\right)$$

and

$$L_2^* \overset{d}{=} \frac{1}{k}\sum_{i=1}^k \left(\frac{Q_{i,n}^\gamma - 1}{\gamma}\right)^2 + A^*\left(Y_{n-k,n}\right)\left(\frac{2}{k}\sum_{i=1}^k \left(\frac{Q_{i,n}^\gamma - 1}{\gamma}\right) K_{\gamma,\rho}^*\left(Q_{i,n}\right) + o_p(S_k)\right),$$

with $K_{\gamma,\rho}^*(x)$ and $A^*(t)$ given in (4.6) and (4.7), respectively. Since (4.15) holds for any intermediate sequence $k = k_n$, the use of (4.13) and (4.14), and now with the notation $S_k^* := \frac{1}{k}\sum_{i=1}^k Y_i^{\gamma+\rho^*+\epsilon}$ enables us to write:

$$\begin{aligned} L_1^* &\overset{d}{=} \frac{1}{k}\sum_{i=1}^k \frac{Y_i^\gamma - 1}{\gamma} + A^*\left(\frac{n}{k}\right)\left(\frac{1}{k}\sum_{i=1}^k K_{\gamma,\rho}^*\left(Y_i\right) + o_p(S_k^*)\right) \\ L_2^* &\overset{d}{=} \frac{1}{k}\sum_{i=1}^k \left(\frac{Y_i^\gamma - 1}{\gamma}\right)^2 + A^*\left(\tfrac{n}{k}\right)\left(\frac{1}{k}\sum_{i=1}^k 2\left(\frac{Y_i^\gamma - 1}{\gamma}\right) K_{\gamma,\rho}^*\left(Y_i\right) + o_p(S_k^*)\right). \end{aligned}$$

The random term associated to $o_p(1)$ is stochastically bounded. Indeed the integral appearing in the limit of that term is finite for any $\epsilon < 1 - \gamma - \rho^*$. The law of large numbers ensures then that $d_{L_1}^* = E\left(K_{\gamma,\rho}^*(Y)\right)$ and $d_{L_2}^* = E\left(\left(\frac{Y^\gamma - 1}{\gamma}\right) K_{\gamma,\rho}^*(Y)\right)$. Taking P_1^* and P_2^* as the normal limiting r.v.'s of Lemma 4.7 and noticing that

$$(L_1^*)^2 - \frac{1}{(1-\gamma)^2} = \left(L_1^* - \frac{1}{1-\gamma}\right)\left(L_1^* + \frac{1}{1-\gamma}\right) = \frac{2}{1-\gamma}\left(L_1^* - \frac{1}{1-\gamma}\right)(1 + o_p(1)),$$

we bring the proof to an end. ■

5 Proof of the main results

Proof. [**Theorem 2.1**]. The proof of Theorem 2.1 relies essentially on the results in Proposition 4.2. Again with $a^*(t)$ given in (4.8) for $\gamma \le 0$, $a^*(t) = a(t)$ if $\gamma > 0$, note first that the estimator in (1.4) can be written as

$$\widehat{\varphi}_n(k) = \frac{\left(\frac{X_{n-k,n}}{a^*(n/k)}\right) M_n^{(1)}(k) - \left(\frac{X_{n-k,n}}{a^*(n/k)}\right) L_n^{(1)}(k)}{\left(\frac{X_{n-k,n}}{a^*(n/k)}\right)\left(L_n^{(1)}(k)\right)^2} = \frac{M_1^* - L_1^*}{\left(a^*(n/k)/X_{n-k,n}\right)(L_1^*)^2}.$$

We shall now consider the cases $\gamma > 0$ and $\gamma \le 0$ separately. If $\gamma > 0$, the use of Proposition 4.2, together with Slutsky's argument, leads us to:

$$\widehat{\varphi}_n(k) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{1 - (1+\gamma)^{-1}}{\gamma(1+\gamma)^{-2}} = 1 + \gamma.$$

If $\gamma \leq 0$,

$$
\begin{aligned}
M_n^{(1)}(k) - L_n^{(1)}(k) &\stackrel{d}{=} \frac{1}{k}\sum_{i=1}^{k}\left\{\ln\frac{U\left(Y_{n-i+1,n}\right)}{U\left(Y_{n-k,n}\right)} - \left(1-\frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)}\right)\right\} \\
&= \frac{1}{k}\sum_{i=1}^{k}\left\{-\ln\left(1-\left(1-\frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)}\right)\right) - \left(1-\frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)}\right)\right\} \qquad (5.1)
\end{aligned}
$$

Note now that for $0 < x < 1$,

$$
0 < -\ln(1-x) - x - \frac{x^2}{2} < \frac{x^3}{3(1-x)}
$$

and use it to deal with (5.1), after assigning $x := 1 - \frac{U\left(Y_{n-k,n}\right)}{U\left(Y_{n-i+1,n}\right)}$. Then, with M_1^* and L_j^*, $j = 1,2$ given in (4.21) and (4.22), respectively, we get, with $R_{k.n}^{(m,j)}$ defined in (4.23),

$$
\frac{L_2^*}{2} < \frac{U\left(Y_{n-k,n}\right)}{a^*(n/k)}\left(M_1^* - L_1^*\right) < \frac{L_2^*}{2} + \frac{R_{k,n}^{(2,3)}}{3}.
$$

Then, since $a^*\left(n/k\right)/U\left(Y_{n-k,n}\right) \xrightarrow[n\to\infty]{P} 0$, $R_{k,n}^{(2,3)} = O_p\left(a^*\left(n/k\right)/U\left(Y_{n-k,n}\right)\right) = o_p(1)$. From (4.22) and Proposition 4.2 we know that $(L_1^*)^2 \xrightarrow[n\to\infty]{P} 1/(1-\gamma)^2$ and $L_2^*/2 \xrightarrow[n\to\infty]{P} 1/\left((1-\gamma)(1-2\gamma)\right)$. Therefore, Slutsky's theorem leads us to the desired result, i.e.,

$$
\widehat{\varphi}_n(k) = \frac{X_{n-k,n}}{a^*(n/k)}\,\frac{M_1^* - L_1^*}{\left(L_1^*\right)^2} \stackrel{d}{=} \frac{L_2^*}{2\,\left(L_1^*\right)^2} \xrightarrow[n\to\infty]{P} \frac{1-\gamma}{1-2\gamma}, \quad \gamma \leq 0. \qquad (5.2)
$$

■

Proof. [**Corollary 2.1**]. Theorem 2.1 ensures the consistency of $\widehat{\varphi}_n(k)$ as an estimator of the strictly increasing function $\varphi(\gamma)$, in (1.3). Hence, by simple inversion and a Slutsky's argument, we get the consistency of $\widehat{\gamma}_n^{MM}(k)$, in (1.6), for any real γ. ■

Proof. [**Theorem 2.2**]. Let us consider first the case $\gamma > 0$. Due to the definition of $\widehat{\varphi}_n(k)$ in (1.4) and given the limits in probability provided in Proposition 4.2, we write

$$
\begin{aligned}
\sqrt{k}\left\{\widehat{\varphi}_n(k) - (1+\gamma)\right\} &= \sqrt{k}\,\frac{M_n^{(1)}(k) - L_n^{(1)}(k) - (1+\gamma)\left(L_n^{(1)}(k)\right)^2}{\left(L_n^{(1)}(k)\right)^2} \\
&= \left(\frac{1+\gamma}{\gamma}\right)^2\left[\sqrt{k}\,\left(M_n^{(1)}(k) - \gamma\right) - \sqrt{k}\left(L_n^{(1)}(k) - \frac{\gamma}{1+\gamma}\right)\right. \\
&\qquad \left. -(1+\gamma)\sqrt{k}\left(\left(L_n^{(1)}(k)\right)^2 - \left(\frac{\gamma}{1+\gamma}\right)^2\right)\right](1+o_p(1)).
\end{aligned}
$$

Now, the results in Proposition 4.3 enable us to write

$$\sqrt{k}\,\{\widehat{\varphi}_n(k)-(1+\gamma)\} = \left(\frac{1+\gamma}{\gamma}\right)^2 \Big[(\gamma\ P_0-(1+2\gamma)P_1) \\ +\sqrt{k}\ \widetilde{A}\left(\frac{n}{k}\right)\left(d_{_{M_1}}-(1+2\gamma)d_{_{L_1}}\right)\Big] + o_p\left(\sqrt{k}\ \widetilde{A}\left(\frac{n}{k}\right)\right)+o_p(1),$$

and, on the basis of Lemma 4.6, we easily get an asymptotic variance equal to $(1+\gamma)^2$.

If $c=\gamma/(\gamma+\rho)$, $\widetilde{\rho}=\rho$, $\widetilde{A}(n/k)=\gamma\ A(n/k)/(\gamma+\rho)$. We then get

$$\sqrt{k}\,\{\widehat{\varphi}_n(k)-(1+\gamma)\} = \left(\frac{1+\gamma}{\gamma}\right)^2 \Big[(\gamma\ P_0-(1+2\gamma)P_1) \\ +\frac{\sqrt{k}\ A\left(\frac{n}{k}\right)\gamma^2}{(1-\rho)(1+\gamma)(1+\gamma-\rho)}\Big] + o_p\left(\sqrt{k}\ A\left(\frac{n}{k}\right)\right)+o_p(1),$$

and the asymptotic bias provided for the region where $c=\gamma/(\gamma+\rho)$ follows.

If $c=\pm\infty$, $\widetilde{\rho}=-\gamma$, $\widetilde{A}(t)=\overline{A}(t)$ and $d_{_{M_1}}-(1+2\gamma)d_{_{L_1}}\equiv 0$. Consequently, the asymptotic bias of $\sqrt{k}\,\{\widehat{\varphi}_n(k)-(1+\gamma)\}$ is null provided we consider levels such that $\sqrt{k}\ \overline{A}(n/k)\to\lambda$, finite, as $n\to\infty$.

If $\gamma\le 0$, we use the distributional representation in (5.2). Similarly to the proof of Theorem 2.1, if we now use the inequalities, $0<-\ln(1-x)-x-x^2/2-x^3/3<x^4/(4(1-x))$, $0<x<1$, for $x=1-U\left(Y_{n-k,n}\right)/U\left(Y_{n-i+1,n}\right)$, we get

$$\frac{L_2^*}{2}+\frac{a^*\left(n/k\right)}{U\left(Y_{n-k,n}\right)}\ \frac{L_3^*}{3} < \frac{U\left(Y_{n-k,n}\right)}{a^*(n/k)}\left(M_1^*-L_1^*\right) < \frac{L_2^*}{2}+\frac{a^*\left(n/k\right)}{U\left(Y_{n-k,n}\right)}\ \frac{L_3^*}{3}+\frac{R_{k,n}^{(2,4)}}{4},$$

with $R_{k,n}^{(m,j)}$ in (4.23). Due to the limits in probability provided in Proposition 4.2, if we consider levels k such that $\sqrt{k}\ \overline{A}(n/k)\to\lambda$, finite, and with $\overline{A}$ given in (2.3),

$$\sqrt{k}\ \overline{A}(n/k)\ \frac{L_3^*}{3}\ \underset{n\to\infty}{\overset{P}{\longrightarrow}}\ \frac{2\lambda}{(1-\gamma)(1-2\gamma)(1-3\gamma)}.$$

The same condition and Lemma 4.5 imply $\sqrt{k}\ R_{k,n}^{(2,4)}\ \underset{n\to\infty}{\overset{P}{\longrightarrow}}\ 0$. From Proposition 4.4 we get the asymptotic normality of L_2^* and $(L_1^*)^2$ and we easily derive

$$\sqrt{k}\ \left(\widehat{\varphi}_n(k)-\varphi(\gamma)\right)\ \overset{d}{=}\ (1-\gamma)^2(1+o_p(1))\left(\sqrt{k}\left(\frac{L_2^*}{2}-\frac{1}{(1-\gamma)(1-2\gamma)}\right)\right. \\ \left.-\frac{1-\gamma}{1-2\gamma}\sqrt{k}\left((L_1^*)^2-\frac{1}{(1-\gamma)^2}\right)+\sqrt{k}\ \overline{A}(n/k)\frac{L_3^*}{3}\right)+o_p(1),$$

which finally leads to

$$\sqrt{k}\ \left(\widehat{\varphi}_n(k)-\varphi(\gamma)\right) \overset{d}{=} (1-\gamma)^2\left(P_2^*-\frac{2\ P_1^*}{1-2\gamma}\right)$$
$$+\sqrt{k}\ A^*(n/k)(1-\gamma)^2\ \left(d^*_{L_2}-\frac{2\ d^*_{L_1}}{1-2\gamma}\right)(1+o_p(1))$$
$$+\sqrt{k}\ \overline{A}(n/k)\ \frac{2(1-\gamma)}{(1-2\gamma)(1-3\gamma)}(1+o_p(1))+o_p(1).$$

Lemma 4.7 leads then to the asymptotic variance in the theorem. From (4.24) and (4.25), we get

$$(1-\gamma)^2\left(d^*_{L_2}-\frac{2\ d^*_{L_1}}{1-2\gamma}\right)=\begin{cases}\frac{\rho(1-\gamma)}{(1-2\gamma)(1-\gamma-\rho)(1-2\gamma-\rho)} & \text{if}\quad \gamma<\rho<0\\ \frac{\gamma(1-\gamma)}{(1-2\gamma)^2(1-3\gamma)} & \text{if}\quad \rho<\gamma<0\\ \frac{\gamma}{(1-2\gamma)^2} & \text{if}\quad \gamma<\rho=0\\ 2 & \text{if}\quad \rho<\gamma=0\end{cases}.$$

From Proposition 4.1, and from the fact that $c=0$ in the region $\gamma<\rho\le 0$ (see (2.5)), we get

$$\lim_{t\to\infty}\frac{\overline{A}(t)}{A^*(t)}=\begin{cases}0 & \text{if}\quad \gamma<\rho\le 0\\ -\frac{\gamma}{2} & \text{if}\quad \rho<\gamma<0\\ -1 & \text{if}\quad \rho<\gamma=0\end{cases}.$$

This finally enables us to get, for the asymptotic bias, the value λb_φ in the theorem, equal to 0 whenever $\rho<\gamma=0$. ■

Proof. [**Corollary 2.2**]. Under the conditions of Theorem 2.2, $\widehat{\varphi}_n(k)$ is a consistent estimator of $\varphi(\gamma)$. The extreme value index is given by the inverse function φ^{-1}, independent of the sample size n, with positive and continuous derivative everywhere. In fact $\frac{d}{dx}\ \varphi^{-1}(x)=1$ if $x>0$ and $\frac{d}{dx}\ \varphi^{-1}(x)=(1-2x)^{-2}$, otherwise. Hence, Cramér's delta-method yields that the Mixed Moment estimator in (1.6) satisfies

$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k)-\gamma\right)\ \underset{n\to\infty}{\overset{d}{\longrightarrow}}\begin{cases}P & \text{if}\quad \gamma>0\\ (1-2\gamma)^2P & \text{if}\quad \gamma\le 0\end{cases},$$

where P stands for the normal limiting r.v. of $\sqrt{k}\left(\widehat{\varphi}_n(k)-\varphi(\gamma)\right)$. ■

Proof. [**Theorem 2.3**]. In order to get a possibly non-null asymptotic bias for the case $\gamma>0$, $c=\pm\infty$, it is imperative to consider the terms provided in the expansions (4.11) and (4.12). From (4.11), and again with $d_{\alpha,\beta}$ given in (4.20), we get

$$\sqrt{k}\left(M_n^{(1)}(k)-\gamma\right)\overset{d}{=}\gamma\ P_0+\sqrt{k}\ \overline{A}\left(\frac{n}{k}\right)\ d_{0,-\gamma}$$
$$+\sqrt{k}\left(\frac{\gamma\left(d_{0,\rho}-d_{0,-\gamma}\right)A(n/k)}{\gamma+\rho}+\overline{A}^2\left(\frac{n}{k}\right)\frac{d_{0,-2\gamma}-d_{0,-\gamma}}{\gamma}\right)(1+o_p(1)). \quad (5.3)$$

From (4.12), we get

$$\sqrt{k}\left(L_n^{(1)}(k)-\frac{\gamma}{1+\gamma}\right) \stackrel{d}{=} P_1+\sqrt{k}\,\overline{A}\left(\frac{n}{k}\right)\,d_{-\gamma,-\gamma}$$
$$+\sqrt{k}\left(\frac{\gamma\,(d_{-\gamma,\rho}-d_{-\gamma,-\gamma})\,A(n/k)}{\gamma+\rho}+\overline{A}^2\left(\frac{n}{k}\right)\frac{2\,(d_{-\gamma,-2\gamma}-d_{-\gamma,-\gamma})}{\gamma}\right)(1+o_p(1)), \quad (5.4)$$

and consequently,

$$\sqrt{k}\left(\left(L_n^{(1)}(k)\right)^2-\frac{\gamma^2}{(1+\gamma)^2}\right) \stackrel{d}{=} \frac{2\gamma\,P_1}{1+\gamma}+\frac{2\gamma\,d_{-\gamma,-\gamma}}{1+\gamma}\,\sqrt{k}\,\overline{A}\left(\frac{n}{k}\right)$$
$$+\frac{2\gamma^2\,\sqrt{k}\,A(n/k)}{(1+\gamma)(\gamma+\rho)}\,(d_{-\gamma,\rho}-d_{-\gamma,-\gamma})\,(1+o_p(1))$$
$$+\sqrt{k}\,\overline{A}^2\left(\frac{n}{k}\right)\left(\frac{4\,(d_{-\gamma,-2\gamma}-d_{-\gamma,-\gamma})}{1+\gamma}+d^2_{-\gamma,-\gamma}\right)(1+o_p(1)). \quad (5.5)$$

Since $|A|\in RV_\rho$ and $\overline{A}^2\in RV_{-2\gamma}$, A dominates $\overline{A}^2$ if $\gamma>-\rho/2$, but $\overline{A}^2$ dominates A if $\gamma\le-\rho/2$. In the region $-\rho/2<\gamma\le-\rho$, $l\neq 0$, with l in (2.4), we may still report the asymptotic bias of $\sqrt{k}\,\{\widehat{\varphi}_n(k)-(1+\gamma)\}$ to the function A. Whenever $\sqrt{k}\,A(n/k)\to\lambda$, the $\overline{A}$ term vanishes, i.e., $\sqrt{k}\,\overline{A}(n/k)\to 0$, and we get the bias λb_0, with

$$\begin{aligned} b_0 &= \left(\tfrac{1+\gamma}{\gamma}\right)^2\left(\tfrac{\gamma}{\gamma+\rho}\right)[(d_{0,\rho}-(1+2\gamma)d_{-\gamma,\rho})-(d_{0,-\gamma}-(1+2\gamma)d_{-\gamma,-\gamma})]\\ &= \left(\tfrac{1+\gamma}{\gamma}\right)^2\left(\tfrac{\gamma}{\gamma+\rho}\right)(d_{0,\rho}-(1+2\gamma)d_{-\gamma,\rho})=\frac{1+\gamma}{(1-\rho)(1+\gamma-\rho)}\equiv b_\varphi,\end{aligned}$$

the same value we got for the region where $c=\gamma/(\gamma+\rho)$.

If we are working with a model such that $0<\gamma<-\rho/2$, $l\neq 0$, $\overline{A}^2$ dominates A and a non-null bias is related to levels k such that $\sqrt{k}\,\overline{A}^2(n/k)\to\lambda$. The asymptotic bias is then related to the scale factor associated to $\sqrt{k}\,\overline{A}^2\,(n/k)$ in the distributional representation of $\sqrt{k}\,(\widehat{\varphi}_n(k)-\varphi(\gamma))$. Such a factor comes directly from the distributional representations in (5.3), (5.4) and (5.5), and, up to the scale factor $\left(\frac{1+\gamma}{\gamma}\right)^2$, is given by

$$\frac{d_{0,-2\gamma}-d_{0,-\gamma}}{\gamma}-\frac{2\,(d_{-\gamma,-2\gamma}-d_{-\gamma,-\gamma})}{\gamma}-4\,(d_{-\gamma,-2\gamma}-d_{-\gamma,-\gamma})-(1+\gamma)d^2_{-\gamma,-\gamma}$$
$$=\frac{1}{\gamma}\left(d_{0,-2\gamma}+(1+2\gamma)d_{-\gamma,-\gamma}-2(1+2\gamma)d_{-\gamma,-2\gamma}-\gamma(1+\gamma)d^2_{-\gamma,-\gamma}\right)$$
$$=\frac{2\gamma^2(1+\gamma)}{(1+\gamma)^2(1+2\gamma)^2(1+3\gamma)}.$$

Hence the result.

If $\rho<\gamma=0$, $A(t)=o\left(\overline{A}^2(t)\right)$ and we get, from (4.10),

$$\frac{U(t)}{a(t)}\left(1-\frac{U(t)}{U(tx)}\right)=\ln x-\overline{A}(t)\,\ln^2 x+\overline{A}^2(t)\,\ln^3 x(1+o(1)).$$

Again in the lines of the proofs of Theorems 2.1 and 2.2, we need now to use the inequalities, $0 < -\ln(1-x) - x - x^2/2 - x^3/3 - x^4/4 < x^5/(5(1-x))$, $0 < x < 1$, for $x = 1 - U(Y_{n-k,n})/U(Y_{n-i+1,n})$, and we get

$$\frac{L_2^*}{2} + \frac{a^*(n/k)}{U(Y_{n-k,n})}\frac{L_3^*}{3} + \left(\frac{a^*(n/k)}{U(Y_{n-k,n})}\right)^2\frac{L_4^*}{4} < \frac{U(Y_{n-k,n})}{a^*(n/k)}(M_1^* - L_1^*)$$
$$< \frac{L_2^*}{2} + \frac{a^*(n/k)}{U(Y_{n-k,n})}\frac{L_3^*}{3} + \left(\frac{a^*(n/k)}{U(Y_{n-k,n})}\right)^2\frac{L_4^*}{4} + \frac{R_{k,n}^{(2,5)}}{5},$$

again with $R_{k,n}^{(m,j)}$ in (4.23). Due to the limit in probability provided in Proposition 4.2 and related to L_4^*, if we consider levels k such that $\sqrt{k}\,\overline{A}^2(n/k) \to \lambda$, finite, and with $\overline{A}$ given in (2.3),

$$\sqrt{k}\,\overline{A}^2(n/k)\,\frac{L_4^*}{4} \underset{n\to\infty}{\overset{P}{\longrightarrow}} \frac{6\lambda}{(1-\gamma)(1-2\gamma)(1-3\gamma)} = 6\,\lambda.$$

The same condition and Lemma 4.5 imply $\sqrt{k}\,R_{k,n}^{(2,5)} \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0$. The contributions to the asymptotic bias provided by $L_2^*/2$, $(L_1^*)^2$, $L_3^*/3$ and $L_4^*/4$ are then given by $36\,\overline{A}^2$, $-16\,\overline{A}^2$, $-24\,\overline{A}^2$ and $6\,\overline{A}^2$, respectively. Hence, the limiting value 2λ, whenever $\sqrt{k}\,\overline{A}^2(n/k) \to \lambda$, finite, as $n \to \infty$. ∎

Remark 5.1. *The case $\gamma = -\rho/2$ has been excluded: then everything depends on the relative behaviour of A and $\overline{A}^2$, both regularly varying with the same index of regular variation ρ.*

Proof. [**Theorem 2.4**]. We are going to show that the normalized mixed moment estimator can be approximated by functionals of a sequence of Brownian motions. Then we show that for $\gamma \geq 0$ these functionals are the same as for the maximum likelihood estimator and for $\gamma \leq 0$ they are the same as for the moment estimator. Note first that, with $\{W_n(s)\}_{s\geq 0}$ denoting a sequence of Brownian motions, Lemma 2.4.10 of de Haan and Ferreira, 2006 gives for $\gamma \neq 0$, and as $n \to \infty$,

$$\left(\frac{k}{n}\,Y_{n-k,n}\right)^\gamma = 1 + \frac{\gamma}{\sqrt{k}}\left(W_n(1) + o_p(1)\right).$$

The same lemma *ibidem* implies for $0 < s \leq 1$,

$$\left(\frac{k}{n}\,Y_{n-[ks],n}\right)^\gamma = s^{-\gamma}\left(1 + \frac{\gamma}{\sqrt{k}}\,s^{-1}W_n(s)\right) + \frac{o_p(1)}{\sqrt{k}}\max\left(1, s^{-\gamma-\frac{1}{2}-\epsilon}\right),$$

where the o_p-term is uniform in s. Defining $Q_n(s) := Y_{n-[ks],n}/Y_{n-k,n}$, it follows that, for any $\epsilon > 0$,

$$Q_n^\gamma(s) = \left(\frac{Y_{n-[ks],n}}{Y_{n-k,n}}\right)^\gamma = s^{-\gamma}\left(1 + \frac{\gamma}{\sqrt{k}}\left(s^{-1}W_n(s) - W_n(1)\right)\right)$$
$$+ \frac{o_p(1)}{\sqrt{k}}\max\left(1, s^{-\gamma-\frac{1}{2}-\epsilon}\right). \quad (5.6)$$

Similarly, we get

$$\ln Q_n(s) = -\ln s + \frac{1}{\sqrt{k}}\left(s^{-1}W_n(s) - W_n(1)\right) + \frac{o_p(1)}{\sqrt{k}}\max\left(1, s^{-\frac{1}{2}-\epsilon}\right). \tag{5.7}$$

For $\gamma > 0$, we get from Proposition 4.1 (i) (in connection with Lemma 4.4) and the substitution explained just before Proposition 4.2,

$$\sqrt{k}\left(M_n^{(1)}(k) - \gamma\right) = \gamma\sqrt{k}\left(\int_0^1 \ln Q_n(s)ds - 1\right) + o_p(1),$$

$$\sqrt{k}\left(L_n^{(1)}(k) - \frac{\gamma}{1+\gamma}\right) = \gamma\sqrt{k}\left(\int_0^1 \left(1 - Q_n^{-\gamma}(s)\right)ds - \frac{1}{1+\gamma}\right) + o_p(1).$$

Then, by (5.6) and (5.7),

$$\sqrt{k}\left(M_n^{(1)}(k) - \gamma\right) = \gamma\int_0^1 \left(s^{-1}W_n(s) - W_n(1)\right)ds + o_p(1),$$

$$\sqrt{k}\left(L_n^{(1)}(k) - \frac{\gamma}{1+\gamma}\right) = \gamma\int_0^1 s^{\gamma}\left(s^{-1}W_n(s) - W_n(1)\right)ds + o_p(1).$$

Now, using Cramer's delta method as in the previous proof, we obtain under the conditions of the theorem

$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k) - \gamma\right) = \frac{(1+\gamma)^2}{\gamma}\int_0^1 \left(s^{-1} - (1+2\gamma)s^{\gamma-1}\right)W_n(s)ds + (1+\gamma)W_n(1) + o_p(1), \tag{5.8}$$

and both the functional of the Brownian motion and the bias are the same as the ones got for $\widehat{\gamma}_n^{ML}$ in Theorem 2.1 of Drees *et al.*, 2004, page 1183. Hence the result for $\gamma > 0$.

If $\gamma = 0$, (5.8) becomes

$$\sqrt{k}\,\widehat{\gamma}_n^{MM}(k) = \int_0^1 (2 + \ln s)\ s^{-1}W_n(s)ds - W_n(1) + o_p(1),$$

again the same approximation as for $\widehat{\gamma}_n^{ML}$. The latter, combined with Theorem 2.2 of Drees *et al.*, 2004 ascertains the full result with respect to $\gamma = 0$.

Finally, if $\gamma < 0$, we get for each $j = 1, 2$,

$$\sqrt{k}\left(L_j^* - \frac{j}{(1-j\gamma)(1-(j-1)\gamma)}\right) = \sqrt{k}\left(\int_0^1 \left(\frac{Q_n^{\gamma}(s) - 1}{\gamma}\right)^j ds - \int_0^1 \left(\frac{s^{-\gamma} - 1}{\gamma}\right)^j ds\right) + o_p(1).$$

Defining,

$$\widetilde{P}_j := j \int_0^1 \Big(\frac{s^{-\gamma}-1}{\gamma}\Big)^{j-1} \big(s^{-\gamma-1} W_n(s) - W_n(1)\big)\, ds, \quad j = 1, 2 \tag{5.9}$$

and using (5.6), we obtain that

$$\sqrt{k}\left(L_j^* - \frac{j}{\big(1-j\gamma\big)\big(1-(j-1)\gamma\big)}\right) - \widetilde{P}_j \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0.$$

Note that $(\widetilde{P}_1, \widetilde{P}_2)$ is equal in distribution to the limit random vector (P, Q) from Lemma 3.5.5 of de Haan and Ferreira, 2006. Now, as shown in the proof of Theorem 2.2 and Corollary 2.2,

$$\sqrt{k}\left(\widehat{\gamma}_n^{MM}(k) - \frac{\frac{L_2^*}{2} - (L_1^*)^2}{L_2^* - (L_1^*)^2}\right) \underset{n\to\infty}{\overset{p}{\longrightarrow}} 0.$$

Note also that

$$\frac{\frac{L_2^*}{2} - (L_1^*)^2}{L_2^* - (L_1^*)^2} = 1 - \frac{1}{2}\left(1 - \frac{(L_1^*)^2}{L_2^*}\right)^{-1},$$

which is the same formula as in Corollary 3.5.6 of de Haan and Ferreira, 2006. Hence, the (asymptotic) correlation coefficient between $\widehat{\gamma}_n^{MM}(k)$ and $\widehat{\gamma}_n^{M}(k)$ is equal to 1.

The second part of the Theorem, namely checking asymptotic bias, is left to the reader since it follows straightforwardly by taking into account the relation among the the second order auxiliary functions A, $\widetilde{A}$ and $\widetilde{A}^2$ and the terms depending on γ and ρ that trail along with. ■

Proof. [**Theorem 2.5**]. We first prove that the estimator $\widehat{\gamma}_n^{MM}(k;p)$ has exactly the same asymptotic behavior as $\widehat{\widehat{\gamma}}_n^{MM}(k;p)$, defined as $\widehat{\gamma}_n^{MM}(k)$ in (1.6), but with $X_{n-i+1,n}$ replaced everywhere by $X_{n-i+1,n} - Q(p)$, $1 \le i \le n$, with Q the inverse of the d.f. F. In fact this statement is true for all the statistics in (1.5). We shall prove it for $M_n^{(1)}(k)$. For $i = 1, 2, \cdots, k$,

$$\ln\big(X_{n-i+1,n} - X_{[np]+1:n}\big) = \ln\big(X_{n-i+1,n} - Q(p)\big) + \ln\left(1 - \tfrac{X_{[np]+1,n} - Q(p)}{X_{n-i+1,n} - Q(p)}\right).$$

The last term is at most

$$\ln\left(1 - \frac{X_{[np]+1,n} - Q(p)}{X_{n,n} - Q(p)}\right).$$

Since $X_{[np]+1,n} - Q(p) = O_p\big(n^{-1/2}\big)$ and $X_{n,n} - Q(p) \underset{n\to\infty}{\overset{P}{\longrightarrow}} Q(1) - Q(p) \in (0, \infty]$, we have

$$\begin{aligned}
\sqrt{k}\ \ln\left(1 - \frac{X_{[np]+1,n} - Q(p)}{X_{n,n} - Q(p)}\right) &= (1 + o_p(1))\sqrt{\frac{k}{n}}\ \frac{\sqrt{n}\,\big(X_{[np]+1,n} - Q(p)\big)}{Q(1) - Q(p)} \\
&= \sqrt{\frac{k}{n}}\ O_p(1) \underset{n\to\infty}{\overset{P}{\longrightarrow}} 0.
\end{aligned}$$

Similarly for $L_n^{(1)}(k)$, $L_n^{(2)}(k)$ and thus for $\widehat{\gamma}_n^{MM}(k)$. Now, $\widehat{\widehat{\gamma}}_n^{MM}(k;p)$ differs from $\widehat{\gamma}_n^{MM}(k)$ only by a shift: we replace $U(t)$ by $\widetilde{U}(t) := U(t) - U\left((1-p)^{-1}\right)$. Hence condition (2.2) is valid for $\widetilde{U}$ and Theorems 2.1 and 2.2 are valid for the adjusted estimators albeit that, for $\rho < \gamma < 0$, the bias term has to be multiplied by $x^F/\left(x^F - U\left(1/(1-p)\right)\right)$. ■

Proof. [**Theorem 2.6**]. The proof of (2.12) follows immediatelly from Proposition 4.2 and Slutsky's theorem. The asymptotic normality result is based on the distributional representations provided in Propositions 4.3 and 4.4. We shall here consider again separately the cases $\gamma > 0$ and $\gamma \le 0$.

If $\gamma > 0$, we may write

$$\frac{\widehat{a}(n/k)}{a(n/k)} - 1 = \frac{\frac{X_{n-k,n}}{a(n/k)} \frac{L_n^{(1)}(k)\ L_n^{(2)}(k)}{2} - \left(L_n^{(2)}(k) - \left(L_n^{(1)}(k)\right)^2\right)}{L_n^{(2)}(k) - \left(L_n^{(1)}(k)\right)^2}.$$

Proposition 4.3, enables us to write,

$$\begin{aligned}\frac{X_{n-k,n}}{a(n/k)}\left(\frac{L_n^{(1)}(k)\ L_n^{(2)}(k)}{2}\right) &= \frac{\gamma^2}{(1+\gamma)^2(1+2\gamma)} + \frac{\gamma^3\ B}{(1+\gamma)^2(1+2\gamma)\sqrt{k}} + \frac{\gamma\ P_1}{(1+\gamma)(1+2\gamma)\sqrt{k}}\\ &+ \frac{P_2}{(1+\gamma)\sqrt{k}} + o_p\left(\frac{1}{\sqrt{k}}\right) + \frac{\widetilde{A}(n/k)}{1+\gamma}\left(d_{L_2} + \frac{\gamma\ d_{L_1}}{1+2\gamma}\right)(1+o_p(1)) - \frac{\gamma\ \overline{A}(n/k)}{(1+\gamma)^2(1+2\gamma)}(1+o_p(1)),\end{aligned}$$

and

$$\begin{aligned}L_n^{(2)}(k) - \left(L_n^{(1)}(k)\right)^2 &= \frac{\gamma^2}{(1+\gamma)^2(1+2\gamma)} - \frac{2\gamma\ P_1}{(1+\gamma)\sqrt{k}} + \frac{2\ P_2}{\sqrt{k}} + o_p\left(\frac{1}{\sqrt{k}}\right)\\ &\quad + 2\widetilde{A}(n/k)\left(d_{L_2} - \frac{\gamma\ d_{L_1}}{1+\gamma}\right)(1+o_p(1)).\end{aligned}$$

Hence,

$$\begin{aligned}\sqrt{k}\left(\frac{\widehat{a}(n/k)}{a(n/k)} - 1\right) &= \tfrac{(1+\gamma)(3+4\gamma)P_1}{\gamma} - \tfrac{(1+\gamma)(1+2\gamma)^2 P_2}{\gamma^2} + \gamma\ B + o_p(1))\\ &\quad + \sqrt{k}\left(\widetilde{A}(n/k)\left(\tfrac{(1+\gamma)(3+4\gamma)\ d_{L_1}}{\gamma} - \tfrac{(1+\gamma)(1+2\gamma)^2\ d_{L_2}}{\gamma^2}\right) - \frac{\overline{A}(n/k)}{\gamma}\right)(1+o_p(1)).\end{aligned}$$

The asymptotic bias follows then from the results in Proposition 4.3 and the asymptotic variance can be calculated on the basis of Lemma 4.6.

If $\gamma \le 0$, and on the basis of the results in Proposition 4.2, we may write

$$\widehat{a}(n/k) = a^*(n/k)\ \frac{L_1^*\ L_2^*/2}{L_2^* - (L_1^*)^2} = a(n/k)\ \frac{a^*(n/k)}{a(n/k)}\ \frac{L_1^*\ L_2^*/2}{L_2^* - (L_1^*)^2},$$

and consequently,

$$\sqrt{k}\left(\frac{\widehat{a}(n/k)}{a(n/k)}-1\right) \stackrel{d}{=} \sqrt{k}\ \left(\frac{\left(\frac{a^*(n/k)}{a(n/k)}\right)\ \frac{L_1^*\ L_2^*}{2} - L_2^* + (L_1^*)^2}{L_2^* - (L_1^*)^2}\right)$$

with $a^*(t)$ given in (4.8), i.e., with

$$\frac{a^*(t)}{a(t)} = \begin{cases} 1 - A^*(t) & \text{if} \quad \gamma < 0,\ \rho \le 0 \\ 1 & \text{if} \quad \rho < \gamma = 0 \end{cases}$$

Proposition 4.4 ensures that

$$\begin{aligned} \frac{L_1^*\ L_2^*}{2} \ &= \ \tfrac{1}{(1-\gamma)^2(1-2\gamma)} + \tfrac{P_1^*}{(1-\gamma)(1-2\gamma)\sqrt{k}} + \tfrac{P_2^*}{(1-\gamma)\sqrt{k}} \\ &\qquad + o_p\left(\frac{1}{\sqrt{k}}\right) + A^*(n/k)\left(\tfrac{d^*_{L_1}}{(1-\gamma)(1-2\gamma)} + \tfrac{d^*_{L_2}}{1-\gamma}\right)(1+o_p(1)). \end{aligned} \tag{5.10}$$

Consequently, in the region $\gamma < 0,\ \rho \le 0,$

$$\begin{aligned} \frac{a^*(n/k)}{a(n/k)}\ \frac{L_1^*\ L_2^*}{2} \ &= \ \tfrac{1}{(1-\gamma)^2(1-2\gamma)} + \tfrac{P_1^*}{(1-\gamma)(1-2\gamma)\sqrt{k}} + \tfrac{P_2^*}{(1-\gamma)\sqrt{k}} + o_p\left(\tfrac{1}{\sqrt{k}}\right) \\ &\qquad + A^*(n/k)\left(\tfrac{d^*_{L_1}}{(1-\gamma)(1-2\gamma)} + \tfrac{d^*_{L_2}}{1-\gamma} - \tfrac{1}{(1-\gamma)^2(1-2\gamma)}\right)(1+o_p(1)). \end{aligned}$$

Since

$$\begin{aligned} L_2^* - (L_1^*)^2 \ &= \ \frac{1}{(1-\gamma)^2(1-2\gamma)} - \frac{2\ P_1^*}{(1-\gamma)\sqrt{k}} + \frac{2\ P_2^*}{\sqrt{k}} + o_p\left(\frac{1}{\sqrt{k}}\right) \\ &\qquad + 2A^*(n/k)\left(d^*_{L_2} - \frac{d^*_{L_1}}{1-\gamma}\right)(1+o_p(1)), \end{aligned} \tag{5.11}$$

$$\begin{aligned} \frac{a^*(n/k)}{a(n/k)}\ \frac{L_1^*\ L_2^*}{2} - L_2^* + (L_1^*)^2 &= \frac{(3-4\gamma)P_1^*}{(1-\gamma)(1-2\gamma)\sqrt{k}} - \frac{(1-2\gamma)P_2^*}{(1-\gamma)\sqrt{k}} + o_p\left(\frac{1}{\sqrt{k}}\right) \\ + A^*(n/k)&\left(\frac{(3-4\gamma)\ d^*_{L_1}}{(1-\gamma)(1-2\gamma)} - \frac{(1-2\gamma)\ d^*_{L_2}}{1-\gamma} - \frac{1}{(1-\gamma)^2(1-2\gamma)}\right)(1+o_p(1)). \end{aligned}$$

We thus get

$$\begin{aligned} \sqrt{k}\left(\frac{\widehat{a}(n/k)}{a(n/k)}-1\right) &= (1-\gamma)(3-4\gamma)\ P_1^* - (1-\gamma)(1-2\gamma)^2\ P_2^* + o_p(1) \\ &\quad + \sqrt{k}\ A^*(n/k)\left((1-\gamma)(3-4\gamma)d^*_{L_1} - (1-\gamma)(1-2\gamma)^2\ d^*_{L_2} - 1\right)(1+o_p(1)), \end{aligned}$$

where the normal r.v.'s (P_1^*, P_2^*) and the constants $d^*_{L_1}$ and d_{L_2} are defined in Lemma 4.7, (4.24) and (4.25), respectively. The asymptotic bias and variance are then obtained by straightforward calculations. In the region $\rho < \gamma = 0$, directly from (5.10) and (5.11), we get

$$\frac{L_1^* \ L_2^*}{2} - L_2^* + (L_1^*)^2 = \left(\frac{3\ P_1^*}{\sqrt{k}} - \frac{P_2^*}{\sqrt{k}} + A^*(n/k)\left(3\ d^*_{L_1} - d^*_{L_2}\right)\right)(1 + o_p(1))$$

and, $3\ d^*_{L_1} - d^*_{L_2} \equiv 0$. Hence the null asymptotic bias associated to levels k such that $\sqrt{k}\ \overline{A}(n/k) \to \lambda$, finite, as $n \to \infty$. ■

Proof. [**Theorem 2.7**]. We only need to take into account the following expansion, obtained under the second order condition (2.2):

$$\begin{aligned}
\frac{X_{n-k,n} - U(n/k)}{a(n/k)} &\stackrel{d}{=} \frac{U\left(Y_{n-k,n}\right) - U(n/k)}{a(n/k)} \\
&= \frac{\left(\frac{k\ Y_{n-k,n}}{n}\right)^{\gamma} - 1}{\gamma} + A(n/k)\left(H_{\gamma,\rho}\left(\frac{k\ Y_{n-k,n}}{n}\right) + o_p(1)\right) \\
&= \frac{B}{\sqrt{k}} + o_p\left(\frac{1}{\sqrt{k}}\right) + o_p\left(A(n/k)\right),
\end{aligned}$$

as $n \to \infty$, where B is the standard normal r.v. in Proposition 4.3. The result comes from Theorems 2.2 and 2.6, after straightforward calculations. ■

References

[1] Araújo Santos, P., Fraga Alves, M. I. and Gomes, M. I. (2006). Peaks Over Random Threshold Methodology for Tail Index and Quantile Estimation. *Revstat* **4** (3), 227-247.

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

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

[4] Billingsley, P. (1979). *Probability and Measure.* John Wiley and Sons, New York, Toronto, London.

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

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

[7] Drees, H. (1998). On smooth statistical tail functions. *Scand. J. Statist.* **25**: 187-210.

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

[9] Ferreira, A., Haan, L. de and Peng, L. (2003). On optimizing the estimation of high quantiles of a probability distribution. *Statistics* **37** (5): 401-434.

[10] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency of the largest or smallest member of a sample. *Proc. Cambridge Phil. Soc.* **24**: 180-190.

[11] Fraga Alves, M. I., Gomes M. I., de Haan, L. and Neves (2007). A Note on Second Order Conditions in Extreme Value Theory: Linking General and Heavy Tails Conditions. Notas e Comunicações CEAUL ??/2007.

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

[13] Haan, L. de (1970). *On Regular Variation and its Application to Weak Convergence of Sample Extremes.* Mathematisch Centrum Amsterdam.

[14] Haan, L. de (1984). Slow variation and characterization of domains of attraction. In Tiago de Oliveira, ed., *Statistical Extremes and Applications*, D. Reidel, Dordrecht, 31-48.

[15] Haan, L. de and Ferreira, A. (2006). *Extreme Value Theory: an Introduction.* Springer Series in Operations Research and Financial Engineering.

[16] Hosking, J. R. M. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics* **29**, 339–349

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

[18] Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quart. J. Meteorol. Soc.* **81**: 158-171.

[19] Mises, R. von (1936). La distribution de la plus grande de n valeurs. *Revue Math. Union Interbalcanique* **1**: 141-160. Reprinted in *Selected Papers of Richard von Mises*, Amer. Soc. 2 (1964): 271-294.

[20] Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119–131.

[21] Smirnov, N. (1952). Limit distributions for terms of a variational series. *Amer. Math. Soc. Transl. Ser. I* **11**: 82-143.

[22] Smith, R. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15**: 1174-1207.