A Bootstrap-based Method to Achieve Optimality in Estimating the Extreme-value Index

G. Draisma, L. de Haan, L. Peng
Erasmus University Rotterdam
T.T. Pereira
University of Lisbon

Econometric Institute Report EI 2000-18/A

Abstract

Estimators of the extreme-value index are based on a set of upper order statistics. We present an adaptive method to choose the number of order statistics involved in an optimal way, balancing variance and bias components. Recently this has been achieved for the similar but somewhat less involved case of regularly varying tails (Drees and Kaufmann(1997); Danielsson et al.(1996)). The present paper follows the line of proof of the last mentioned paper.

Key words & phrases: Moment estimator, Pickands estimator, bootstrap, mean squared error.

1 Introduction

Suppose we have i.i.d. observations $X_1, X_2, \cdots, X_n$ whose common distribution function F is in the domain of attraction of an extreme-value distribution G_γ (notation: $F \in D(G_\gamma)$). The shape parameter $\gamma \in R$ of this extreme-value distribution (functional form: $\exp(-(1+\gamma x)^{-1/\gamma})$) can be estimated in various ways starting from the sample $X_1, X_2, \cdots, X_n$. Two popular estimators are Pickands' estimator (in its generalized form see e.g. Pereira (1993)):

$$\hat{\gamma}_{n,\theta}(k) := (-\log\theta)^{-1} \log \frac{X_{n,n-[k\theta^2]} - X_{n,n-[k\theta]}}{X_{n,n-[k\theta]} - X_{n,n-k}} \tag{1.1}$$

($\theta \in (0,1)$) where $X_{n,1} \leq \cdots \leq X_{n,n}$ are the order statistics of $X_1, \cdots, X_n$ and $[z]$ denotes the largest integer which is not larger than z, and the moment estimator

$$\hat{\gamma}_{n,2}(k) := M_n^{(1)}(k) + 1 - \frac{1}{2}\left(1 - \frac{(M_n^{(1)}(k))^2}{M_n^{(2)}(k)}\right)^{-1} \tag{1.2}$$

with $M_n^{(j)}(k) := \frac{1}{k}\sum_{i=0}^{k-1}(\log X_{n,n-i} - \log X_{n,n-k})^j$. For this estimator we have to require that the right end point of the distribution is positive.

The estimators from (1.1) and (1.2) are consistent for γ provided $k = k(n) \to \infty$, $k(n) = o(n)(n \to \infty)$. If one increases the speed at which $k(n)$ goes to infinity, the asymptotic variance decreases but the asymptotic bias increases. There is an optimal sequence balancing variance and bias components (see figure 1). This optimal sequence $k_0(n)$ can be determined when the underlying distribution is known, provided the distribution function has a second order expansion involving an extra unknown parameter (Hall (1982); Dekkers and de Haan (1993)). Here we develop a purely sample based way of obtaining the optimal sequence $k_0(n)$ where we assume a second order expansion but do not assume the second order (or first order) characteristic known. The procedure is based on a double bootstrap (see also Hall (1990). Results for the moment estimator and for Pickands' estimator are given in Section 3 and Section 4 respectively. All the proofs are postponed till Section 5. Section 6 reports the result of a small simulation study and section 7 demonstrates the application of the procedure to North Sea wave height data. In an appendix we explain why we use different second order conditions in Section 3 and Section 4.

2 Outline

We want (in the set-up of (1.2)) the value of k minimizing $E_F(\hat{\gamma}_{n,2}(k) - \gamma)^2$ although this is only meant in an asymptotic sense (second moment of the asymptotic distribution). Call this value $k_0(n)$. There are two unknowns in this expression: γ and the distribution function F. The idea is to replace γ by a second estimator $\hat{\gamma}_{n,3}(k)$ and to replace F by the empirical distribution function F_n. This amounts to bootstrapping. It is proved that minimizing the resulting expression, which can be calculated purely on the basis of the sample, still leads to the optimal $k_0(n)$ with the help of a second bootstrap. A similar procedure applies to the estimator from (1.1). Sections 3 and 4 provide the scientific background for the bootstrap procedure. Here we explain step by step how to implement the procedure.

We start with a sample $X_1, \cdots, X_n$.

Step 1: Select randomly and independently n_1 times ($n_1 << n$) a member of the set $\{X_1, X_2, \cdots, X_n\}$, We indicate the result by $X_1^*, \cdots, X_{n_1}^*$. Form the order statistics $X_{n_1,1}^* \leq \cdots \leq X_{n_1,n_1}^*$ and compute $\hat{\gamma}_{n_1,2}^*(k)$ and $\hat{\gamma}_{n_1,3}^*(k)$ (according to the formula after Theorem 3.2 below) for $k = 1, 2, \cdots, n_1$. Form $q_{n_1,k}^* = (\hat{\gamma}_{n_1,4}^*(k))^2$ for $k = 1, 2, \cdots, n_1$.

Step 2: Repeat this procedure r times independently. This results in a sequence $q_{n_1,k,s}^*$, $k = 1, 2, \cdots, n_1$ and $s = 1, 2, \cdots, r$. Calculate $\frac{1}{r}\sum_{s=1}^{r} q_{n_1,k,s}^*$. The number r can be taken as big as necessary.

Step 3: Minimize $\frac{1}{r}\sum_{s=1}^{r} q_{n_1,k,s}^*$ with respect to k. Denote by $\bar{k}_{0,1}^*(n_1)$ the value of k where the minimum is obtained.

Step 4: Repeat Step 1 up to 3 independently with the number n_1 replaced by $n_2 = (n_1)^2/n$. So n_2 is smaller than n_1. This results in $\bar{k}_{0,1}^*(n_2)$.

Step 5: Calculate $\hat{k}_0(n)$ on the basis of $\bar{k}_{0,1}^*(n_1)$ and $\bar{k}_{0,1}^*(n_2)$ according to its definition in Corollary 3.3 below with $\hat{\gamma}_n := \hat{\gamma}_{n,2}([n^{1/2}])$ for example and $\hat{\rho}_n$ according to the formula in the same Corollary.

This $\hat{k}_0(n)$ is the adaptively obtained optimal number of order statistics.

3 Main results for moment estimator

We shall write throughout γ_+ for $\gamma \vee 0$ and γ_- for $\gamma \wedge 0$. Assume $F \in D(G_\gamma)$, i.e. there exists a positive function $a(t)$ such that

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma} \quad \text{for } x > 0,$$

which implies that

$$\lim_{t\to\infty} \frac{\log U(tx) - \log U(t)}{a(t)/U(t)} = \frac{x^{\gamma_-} - 1}{\gamma_-} \quad \text{for } x > 0.$$

Throughout this section we assume $U(\infty) > 0$ and the following second order condition:

$$\lim_{t\to\infty} \frac{\frac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-} - 1}{\gamma_-}}{A(t)} = H(x) \tag{3.1}$$

where $U(t)$ is the inverse function of the function $1/(1-F)$, $a(t)$ is positive and A not changing sign eventually. The function $H(x)$ is assumed not to be a multiple of $(x^{\gamma_-} - 1)/\gamma_-$ and takes the form (supposing the function a

and A are chosen properly)

$$H(x) = \frac{1}{\rho}\left[\frac{x^{\rho+\gamma_-}-1}{\rho+\gamma_-} - \frac{x^{\gamma_-}-1}{\gamma_-}\right]$$
$$= \begin{cases} (\log x)^2/2 & \text{if } \rho = 0, \gamma \geq 0, \\ \frac{1}{\gamma}[x^\gamma \log x - \frac{x^\gamma - 1}{\gamma}] & \text{if } \rho = 0, \gamma < 0, \\ \frac{1}{\rho}[\frac{x^{\rho+\gamma_-}-1}{\rho+\gamma_-} - \frac{x^{\gamma_-}-1}{\gamma_-}] & \text{if } \rho \neq 0, \end{cases} \tag{3.2}$$

depending on a second order parameter $\rho \leq 0$ (see de Haan and Stadtmüller, relation (2.9) page 387).

We present a series of results culminating in Corollary 3.3 that provides a sample based sequence $\hat{k}_0(n)$ such that for any (random or non-random) sequence $k(n)$

$$\limsup_{n\to\infty} \frac{\mathrm{E}\{(\hat{\gamma}_{n,2}(\hat{k}_0(n)) - \gamma)^2\}}{\mathrm{E}\{(\hat{\gamma}_{n,2}(k(n)) - \gamma)^2\}} \leq 1$$

First we restate in slightly greater generality a result from Dekkers and de Haan (1993) providing the optimal number of order statistics for the moment estimator as a function of γ, ρ and the function A.

Theorem 3.1. *Suppose $F \in D(G_\gamma)$ and that (3.1) and (3.2) hold for $\rho < 0$, $\gamma \neq \rho$ and $\gamma \neq 0$. Let $k_0 = k_0(n)$ be a sequence of integers such that the asymptotic second moment of $\hat{\gamma}_{n,2}(k) - \gamma$ is minimal when choosing $k = k_0(n)$. Then*

$$k_0(n)/\{n(\frac{V^2(\gamma)}{b^2(\gamma,\rho)})^{\frac{1}{1-2\rho^*}}(s^-(\frac{1}{n}))^{-1}\} \to 1 \tag{3.3}$$

as $n \to \infty$, where

$$\rho^* = \begin{cases} \rho & \text{if } \gamma > 0, \\ \gamma & \text{if } \rho < \gamma < 0, \\ \rho & \text{if } \gamma < \rho, \end{cases}$$

$$V^2(\gamma) = \begin{cases} \gamma + 1 & \text{if } \gamma > 0, \\ \frac{(1-\gamma)^2(1-2\gamma)(6\gamma^2-\gamma+1)}{(1-3\gamma)(1-4\gamma)} & \text{if } \gamma < 0 \end{cases} \tag{3.4}$$

(the variance component) with $\gamma_+ = 0 \vee \gamma$ and $\gamma_- = 0 \wedge \gamma$ and

$$b(\gamma,\rho) = \begin{cases} (\frac{1}{\rho(1-\rho)})\gamma + \frac{1}{(1-\rho)^2} & \text{for } \gamma > 0, \\ \frac{1}{1-\gamma} & \text{for } \rho < \gamma < 0, \\ \frac{(1-\gamma)(1-2\gamma)}{(1-\rho-\gamma)(1-\rho-2\gamma)} & \text{for } \gamma < \rho \end{cases} \tag{3.5}$$

(the bias component). The function s^- is the inverse function of the decreasing function s satisfying

$$A_0^2(t) = (1+o(1)) \int_t^\infty s(u)\, du, \tag{3.6}$$

where

$$A_0(t) = \begin{cases} A(t) & \text{if } \gamma > 0, \\ \frac{a(t)}{U(t)} & \text{if } \rho < \gamma < 0 \text{ and} \\ A(t) & \text{if } \gamma < \rho. \end{cases}$$

Remark 3.1. Since we only know anything about the asymptotic mean square error for intermediate k, here and in the rest of the paper, when we minimize over k, we only consider values between $\log n$ and $n/(\log n)$, both being intermediate sequences with the optimal value in between.

Remark 3.2. We exclude the two cases $\gamma = 0$ and $\gamma = \rho$ in Theorem 3.1. The reason can be seen from Theorem A in Appendix. This also happens in Dekkers and de Haan (1993).

We are going to turn the asymptotic second moment of $\hat{\gamma}_{n,2}(k) - \gamma$ into something we can handle adaptively, the first step is to replace the unknown γ in the formula by an alternative estimator for γ. The alternative estimator is

$$\hat{\gamma}_{n,3}(k) := \sqrt{M_n^{(2)}(k)/2} + 1 - \frac{2}{3}(1 - \frac{M_n^{(1)}(k) M_n^{(2)}(k)}{M_n^{(3)}(k)})^{-1}.$$

The following theorem is the analogue of theorem 3.1 for the aymptotic second moment of $\hat{\gamma}_{n,2}(k) - \hat{\gamma}_{n,3}(k)$.

Theorem 3.2. *Assume the conditions of Theorem 3.1. Determine $\bar{k}_0 = \bar{k}_0(n)$ such that the asymptotic second moment of $\hat{\gamma}_{n,2}(k) - \hat{\gamma}_{n,3}(k)$ is minimal. Then*

$$\bar{k}_0(n) / \{ n (\frac{\bar{V}^2(\gamma)}{\bar{b}^2(\gamma,\rho)})^{\frac{1}{1-2\rho^*}} (s^-(\frac{1}{n}))^{-1} \} \to 1$$

as $n \to \infty$, where

$$\bar{V}^2(\gamma) = \begin{cases} \frac{1}{4}\,(\gamma^2+1) & \text{if } \gamma > 0, \\ \frac{1}{4}\,\frac{(1-\gamma)^2\,(1-8\gamma+48\gamma^2-154\gamma^3+263\gamma^4-222\gamma^5+72\gamma^6)}{(1-2\gamma)\,(1-3\gamma)\,(1-4\gamma)\,(1-5\gamma)\,(1-6\gamma)} & \text{if } \gamma < 0 \end{cases}$$

and

$$\bar{b}(\gamma,\rho)=\begin{cases}-\frac{\gamma(1-\rho)+\rho}{2\,(1-\rho)^3} & \text{if } \gamma>0,\\ \frac{1-2\gamma-\sqrt{(1-\gamma)\,(1-2\gamma)}}{(1-\gamma)\,(1-2\gamma)} & \text{if } \rho<\gamma<0,\\ \frac{1}{2}\,\frac{-\rho\,(1-\gamma)^2}{(1-\gamma-\rho)\,(1-2\gamma-\rho)\,(1-3\gamma-\rho)} & \text{if } \gamma<\rho.\end{cases}$$

In order to show the convergence of mean squared error, we consider the following quantity

$$\hat{\gamma}_{n,4}(k):=(\hat{\gamma}_{n,2}(k)-\hat{\gamma}_{n,3}(k))\mathbf{1}(|\hat{\gamma}_{n,2}(k)-\hat{\gamma}_{n,3}(k)|\le k^{\delta-1/2}),$$

where $\delta>0$. Then we have

Theorem 3.3. *Assume the conditions of Theorem 3.1. Suppose $\rho<0$. Determine $\bar{k}_{0,1}=\bar{k}_{0,1}(n)$ such that* $\mathrm{E}\{(\hat{\gamma}_{n,4}(k))^2\}$ *is minimal. Then as $n\to\infty$*

$$\bar{k}_{0,1}(n)/\bar{k}_0(n)\to 1.$$

Hence

$$\bar{k}_{0,1}(n)/\{n(\frac{\bar{V}^2(\gamma)}{\bar{b}^2(\gamma,\rho)})^{\frac{1}{1-2\rho^*}}(s^-(\frac{1}{n}))^{-1}\}\to 1.$$

Remark 3.3. Note that Theorem 3.3 holds for any $\delta>0$ in the definition of $\hat{\gamma}_{n,4}(k)$. Thus, in our simulation study, we use $\hat{\gamma}_{n,2}(k)-\hat{\gamma}_{n,3}(k)$ instead of $\hat{\gamma}_{n,4}(k)$.

Next we are going to introduce the bootstrap procedure. One takes n_1 independent drawings from the empirical distribution function of $\mathcal{X}_n:=\{X_1,\cdots,X_n\}$. This results in observations $X_1^*,\cdots,X_{n_1}^*$. We form the order statistics $X_{n_1,1}^*\le\cdots\le X_{n_1,n_1}^*$ and define

$$M_{n_1}^{(j)*}(k):=\frac{1}{k}\sum_{i=1}^{k}(\log X_{n_1,n_1-i+1}^*-\log X_{n_1,n_1-k_1}^*)^j$$

for $k<n_1$ and $j=1,2,3$. Next define

$$\hat{\gamma}_{n_1,2}^*(k):=M_{n_1}^{(1)*}(k)+1-\frac{1}{2}(1-\frac{(M_{n_1}^{(1)*}(k))^2}{M_{n_1}^{(2)*}(k)})^{-1}$$

and

$$\hat{\gamma}_{n_1,3}^*(k):=\sqrt{M_{n_1}^{(2)*}(k)/2}+1-\frac{2}{3}(1-\frac{M_{n_1}^{(1)*}(k)M_{n_1}^{(2)*}(k)}{M_{n_1}^{(3)*}(k)})^{-1}.$$

By bootstrapping we can now estimate

$$Q(n_1, k) := \mathrm{E}\{(\hat{\gamma}^*_{n_1,4}(k))^2 | \mathcal{X}_n\}$$

with

$$\hat{\gamma}^*_{n,4}(k) := (\hat{\gamma}^*_{n,2}(k) - \hat{\gamma}^*_{n,3})\mathbf{1}(|\hat{\gamma}^*_{n,2}(k) - \hat{\gamma}^*_{n,3}(k)| \leq k^{\delta-1/2})$$

as well as we wish.

Now we want to connect the asymptotic behavior of $\arg\inf Q$ with the corresponding quantity for the *asymptotic* expectation as considered in e.g. Theorem 3.1.

Theorem 3.4. *Suppose the conditions of Theorem 3.1 hold and* $n_1 = O(n^{1-\epsilon})$ *for some* $0 < \epsilon < 1/2$. *The random quantity* $\bar{k}^*_{0,1}(n_1)$ *is defined as follows:*

$$\bar{k}^*_{0,1}(n_1) := \arg\inf_k \mathrm{E}\{(\hat{\gamma}^*_{n_1,4}(k))^2 | \mathcal{X}_n\}.$$

Then

$$\bar{k}^*_{0,1}(n_1)/\{n_1(\frac{\bar{V}^2(\gamma)}{\bar{b}^2(\gamma,\rho)})^{\frac{1}{1-2\rho^*}}(s^{-}(\frac{1}{n_1}))^{-1}\} \to 1$$

in probability.

We now use the known quantity $\bar{k}^*_{0,1}$ to estimate $k_0(n)$ and do this via $\bar{k}_0(n)$.

Corollary 3.1. *Suppose the conditions of Theorem 3.4 hold and* $A_0(t) = ct^{\rho^*}$ *with* $c \neq 0$ *and* $\rho^* < 0$. *Then*

$$\bar{k}_0(n)/\{\bar{k}^*_{0,1}(n_1)(\frac{n}{n_1})^{\frac{-2\rho^*}{1-2\rho^*}}\} \to 1$$

in probability.

Remark 3.4. Since A_0 in Theorem 3.1 is a regularly varying function, the extra requirement means that the slowly varying function is in fact a constant.

Next we get rid of the factor $(n/n_1)^{\frac{-2\rho^*}{1-2\rho^*}}$. We do this via an independent second bootstrap procedure with bootstrap sample size n_2.

Theorem 3.5. *Suppose the conditions of Corollary 3.1 hold and* $n_2 = (n_1)^2/n$. *Let*

$$\bar{k}^*_{0,1}(n_2) := \arg\inf_k \mathrm{E}\{(\hat{\gamma}^*_{n_2,4}(k))^2 | \mathcal{X}_n\}.$$

Then

$$\bar{k}_0(n)/\{(\bar{k}^*_{0,1}(n_1))^2/\bar{k}^*_{0,1}(n_2)\} \to 1 \text{ in probability.}$$

Corollary 3.2. *Under the conditions of Theorem 3.5,*

$$k_0(n)/\{\frac{(\bar{k}^*_{0,1}(n_1))^2}{\bar{k}^*_{0,1}(n_2)}(\frac{V^2(\gamma)\bar{b}^2(\gamma,\rho)}{\bar{V}^2(\gamma)b^2(\gamma,\rho)})^{\frac{1}{1-2\rho^*}}\} \to 1 \text{ in probability.}$$

Corollary 3.3. *Suppose the conditions of Theorem 3.5 hold. Define*

$$\hat{k}_0(n) := \frac{(\bar{k}^*_{0,1}(n_1))^2}{\bar{k}^*_{0,1}(n_2)}\Big(\frac{V^2(\hat{\gamma}_n)\,\bar{b}^2(\hat{\gamma}_n,\hat{\rho}_n)}{\bar{V}^2(\hat{\gamma}_n)\,b^2(\hat{\gamma}_n,\hat{\rho}_n)}\Big)^{\frac{1}{1-2\hat{\rho}_n}}$$

with $\bar{k}^*_{0,1}(n_1)$ and $\bar{k}^*_{0,1}(n_2)$ as defined in Theorem 3.4 and Theorem 3.5 respectively and with $\hat{\gamma}_n$ any consistent estimator of γ (for instance $\hat{\gamma}_{n,2}(k)$ with $k = k(n)$ any sequence with $k \to \infty, k/n \to 0$), $\hat{\rho}_n$ any consistent estimator for ρ^*, for instance

$$\hat{\rho}_n := \frac{\log \bar{k}^*_{0,1}(n_1)}{-2\log n_1 + 2\log \bar{k}^*_{0,1}(n_1)}.$$

Then

$$\hat{k}_0(n)/k_0(n) \to 1 \text{ in probability,}$$

hence the asymptotic second moment of $\hat{\gamma}_{n,2}(\hat{k}_0(n)) - \gamma$ is asymptotically equal to the asymptotic second moment of $\hat{\gamma}_{n,2}(k_0(n)) - \gamma$.

4 Main results for Pickands' estimator

Throughout this section we assume that F is in the differentiable domain of attraction of G_γ (notation: $F \in D_{\text{dif}}(G_\gamma)$), i.e., F is differentiable in a left neighborhood of $x_\infty := \sup\{x : F(x) < 1\}$ and there exist $a_n > 0$ and $b_n \in R$ such that

$$\lim_{n\to\infty} \frac{\partial}{\partial x}[F^n(a_n x + b_n)] = G'_\gamma(x) \tag{4.1}$$

locally uniformly for all $x \in R$. This is mainly done for convenience. In fact not much is lost of the general case and the computations are more simple. The differentiable domains of attraction were introduced by Pickands (1986). Clearly $F \in D_{\text{dif}}(G_\gamma)$ implies $F \in D(G_\gamma)$ for the same normalizing constants a_n and b_n. Define $U(t) := (1/(1-F))^{-}(t)$. The following proposition characterizes the differentiable domain of attraction of G_γ.

Proposition 1. $F \in D_{\text{dif}}(G_\gamma)$ for some $\gamma \in R$ if and only if $U(t)$ is differentiable for all sufficiently large t and $U'(t) \in RV_{\gamma-1}$.

Proof. See Pickands (1986). □

In order to get the limit distribution function of estimator $\hat{\gamma}_{n,\theta}(k)$ we have to require some kind of second order condition. Because of Proposition 1 it is quite natural to assume that there is a positive function $A^*(t)(\to 0$ as $t \to \infty)$ such that

$$\lim_{t\to\infty} \frac{\frac{U'(tx)}{U'(t)} - x^{\gamma-1}}{A^*(t)}$$

exists for every $x > 0$. In order to avoid trivialities we also assume that the limit function is not a multiple of $x^{\gamma-1}$. Then the limit function must be of the form $c'x^{\gamma-1}\frac{x^\rho-1}{\rho}$ for constants $\rho \leq 0$ and $c' \neq 0$ (see Theorem 1.9 of Geluk and de Haan (1987) or Lemma 3.2.1 of Bingham et al. (1987); $(x^0-1)/0$ is defined as $\log x$). We can and will subsume the constant c' in the function A^*. So suppose there is a function A with $\lim_{t\to\infty} A(t) = 0$ and not changing sigh near infinity, such that

$$\lim_{t\to\infty} \frac{\frac{U'(tx)}{U'(t)} - x^{\gamma-1}}{A(t)} = x^{\gamma-1}\frac{x^\rho-1}{\rho} \tag{4.2}$$

for all $x > 0$. The function $|A|$ is then regularly varying with index $\rho(notation:$ $|A| \in RV_\rho)$. It can be proved (see Pereira(1993) or de Haan and Stadtmüller(1996)) that (3.2) is equivalent to

$$\lim_{t\to\infty} \frac{U(tx) - U(t) - tU'(t)\frac{x^\gamma-1}{\gamma}}{tU'(t)A(t)} = h_{\gamma,\rho}(x) := \frac{1}{\rho}[\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}]. \tag{4.3}$$

First we determine the theoretically optimal value $k_0(n)$ asymptotically.

Theorem 4.1. *Assume $F \in D_{dif}(G_\gamma)$ and (4.3) holds for $A(t) = ct^\rho$ with $c \neq 0$ and $\rho < 0$. Determine $k_0 = k_0(n)$ such that the asymptotic second moment of $\hat{\gamma}_{n,\theta}(k) - \gamma$ is minimal. Then*

$$k_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}})^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}\} \to 1$$

as $n \to \infty$.

Next we compute the optimum with γ replaced by $\hat{\gamma}_{n,\theta}(k\theta^2)$.

Theorem 4.2. *Assume $F \in D_{dif}(G_\gamma)$ and (4.3) holds for $A(t) = ct^\rho$ with $c \neq 0$ and $\rho < 0$. Determine $\bar{k}_0 = \bar{k}_0(n)$ such that the asymptotic second moment of $\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta}(k\theta^2)$ is minimal. Then*

$$\bar{k}_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})(1+\theta^{-2})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}\} \to 1$$

as $n \to \infty$.

Corollary 4.1. *Assume $F \in D_{dif}(G_\gamma)$ and (4.3) holds for $A(t) = ct^\rho$ with $c \neq 0$ and $\rho < 0$. Determine $k_0(n)$ such that the asymptotic second moment of $\hat{\gamma}_{n,\theta}(k) - \gamma$ is minimal and $\bar{k}_0(n)$ such that the asymptotic second moment of $\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta}(k\theta^2)$ is minimal. Then*

$$\frac{\bar{k}_0(n)}{k_0(n)} \to \left(\frac{1+\theta^{-2}}{(1-\theta^{-2\rho})^2}\right)^{\frac{1}{1-2\rho}}$$

as $n \to \infty$.

In order to show the convergence of mean squared error, we consider the following quantity

$$\bar{\gamma}_{n,\theta}(k) := (\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta^2}(k\theta^2))\mathbf{1}(|\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta^2}(k\theta^2)| \leq k^{\delta-1/2}),$$

where $\delta > 0$. Then we have

Theorem 4.3. *Assume $F \in D_{dif}(G_\gamma)$ and (4.3) holds for $A(t) = ct^\rho$ with $c \neq 0$ and $\rho < 0$. Determine $\bar{k}_{0,1} = \bar{k}_{0,1}(n)$ such that $\mathrm{E}\{(\bar{\gamma}_{n,\theta}(k))^2\}$ is minimal. Then as $n \to \infty$*

$$\bar{k}_{0,1}(n)/\bar{k}_0(n) \to 1.$$

As in Section 3, we draw resamples $\mathcal{X}_{n_1}^* = \{X_1^*, \cdots, X_{n_1}^*\}$ from $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ with replacement. Let $n_1 < n$ and $X_{n_1,1}^* \leq \cdots \leq X_{n_1,n_1}^*$ denote the order statistics of $\mathcal{X}_{n_1}^*$ and define

$$\hat{\gamma}_{n_1,\theta}^*(k_1) := (-\log\theta)^{-1} \log \frac{X_{n_1,n_1-[k_1\theta^2]}^* - X_{n_1,n_1-[k_1\theta]}^*}{X_{n_1,n_1-[k_1\theta]}^* - X_{n_1,n_1-k_1}^*}.$$

Then we propose to use the following bootstrap estimate of the mean square error

$$\mathrm{E}\{(\bar{\gamma}_{n_1,\theta}^*(k_1))^2|\mathcal{X}_n\}.$$

We can prove

Theorem 4.4. *Assume $F \in D_{dif}(G_\gamma)$ and (4.3) holds for $A(t) = ct^\rho$ with $c \neq 0$ and $\rho < 0$. Let $n_1 = O(n^{1-\epsilon})$ for some $\epsilon \in (0,1)$. Determine $k_{1,0}^*(n_1)$ such that $\mathrm{E}\{(\bar{\gamma}_{n_1,\theta}^*(k))^2|\mathcal{X}_n\}$ is minimal. Then*

$$k_{1,0}^*(n_1)/\{\left(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})(1+\theta^{-2})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}(1-\theta^{-2\rho})^2}\right)^{\frac{1}{1-2\rho}} n_1^{\frac{-2\rho}{1-2\rho}}\} \xrightarrow{p} 1$$

as $n \to \infty$.

Finally we connect $k_0(n)$ with $k_{1,0}^*$ and $k_{2,0}^*$ asymptotically.

Theorem 4.5. *Assume $F \in D_{dif}(G_\gamma)$ and (3.3) holds for $A(t) = ct^\rho$ ($\rho < 0$). Let $n_1 = O(n^{1-\epsilon})$ for some $\epsilon \in (0, 1/2)$ and $n_2 = (n_1)^2/n$. Determine $k_{i,0}^*(n_i)$ such that $\mathrm{E}\{(\bar{\gamma}_{n_i,\theta}^*(k_i))^2|\mathcal{X}_n\}$ is minimal ($i = 1, 2$). Define $f_\theta(\rho) = (\frac{1+\theta^{-2}}{(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}}$. Then*

$$\frac{(k_{1,0}^*)^2}{k_{2,0}^* f_\theta(\frac{\log k_{1,0}^*}{2(\log k_{1,0}^* - \log n_1)})} / k_0(n) \xrightarrow{p} 1$$

as $n \to \infty$.

So as before we get an estimator for $k_0(n)$ which leads to an estimator for θ which has asymptotically the lowest mean squared error.

5 Proofs

We shall give some lemmas first.

Lemma 5.1. *Let $Y_1, \cdots, Y_n$ be i.i.d. random variables with common distribution function $1 - x^{-1}$ ($x > 1$) and $Y_{n,1} \le \cdots \le Y_{n,n}$ be the order statistics. Assume $k \to \infty$, $k/n \to 0$. Then*

(i) $Y_{n,n-k}/\frac{n}{k} \to 1$ in probability

(ii) Define

$$\begin{cases} P_n := \frac{1}{k}\sum_{i=1}^k \frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-}-1}{\gamma_-} - \frac{1}{1-\gamma_-} \\ Q_n := \frac{1}{k}\sum_{i=1}^k (\frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-}-1}{\gamma_-})^2 - \frac{2}{(1-\gamma_-)(1-2\gamma_-)} \\ R_n := \frac{1}{k}\sum_{i=1}^k (\frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-}-1}{\gamma_-})^3 - \frac{6}{(1-\gamma_-)(1-2\gamma_-)(1-3\gamma_-)}. \end{cases}$$

We have $\sqrt{k}(P_n, Q_n, R_n)$ converges in distribution to (P, Q, R), say, which is normally distributed with mean vector zero and covariance matrix

$$\begin{cases} \mathrm{E}\,P^2 = \frac{1}{(1-\gamma_-)^2(1-2\gamma_-)} \\ \mathrm{E}\,Q^2 = \frac{4(5-11\gamma_-)}{(1-\gamma_-)^2(1-2\gamma_-)^2(1-3\gamma_-)(1-4\gamma_-)} \\ \mathrm{E}\,R^2 = \frac{36(19-105\gamma_-+146\gamma_-^2)}{(1-\gamma_-)^2(1-2\gamma_-)^2(1-3\gamma_-)^2(1-4\gamma_-)(1-5\gamma_-)(1-6\gamma_-)} \\ \mathrm{E}(PQ) = \frac{4}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)} \\ \mathrm{E}(PR) = \frac{18}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)(1-4\gamma_-)} \\ \mathrm{E}(QR) = \frac{12(9-21\gamma_-)}{(1-\gamma_-)^2(1-2\gamma_-)^2(1-3\gamma_-)(1-4\gamma_-)(1-5\gamma_-)} \end{cases}$$

Moreover,

$$\begin{cases} k\,\mathrm{E}\,P_n^2 \to \mathrm{E}\,P^2 \\ k\,\mathrm{E}\,Q_n^2 \to \mathrm{E}\,Q^2 \\ k\,\mathrm{E}\,R_n^2 \to \mathrm{E}\,R^2. \end{cases}$$

(iii) Define for $j = 1, 2, 3$

$$d_n^{(j)} := \frac{1}{k}\sum_{i=1}^{k} j\,H(Y_{n,n-i+1}/Y_{n,n-k})\,(\frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-}-1}{\gamma_-})^{j-1}$$

Then by the law of large numbers

$$d_n^{(j)} \xrightarrow{p} d_j = \int_1^\infty j\,H(y)\,(\frac{y^{\gamma_-}-1}{\gamma_-})^{j-1}\,\frac{dy}{y^2}, \quad j = 1, 2, 3$$

or explicitly

$$d_1 = \frac{1}{(1-\gamma_-)(1-\rho-\gamma_-)},$$

$$d_2 = \frac{2(3-2\rho-4\gamma_-)}{(1-\gamma_-)(1-2\gamma_-)(1-\rho-\gamma_-)(1-\rho-2\gamma_-)} \text{ and}$$

$$d_3 = \frac{18\gamma_-^2 - 22\gamma_- + 15\rho\,\gamma_- + 3\rho^2 - 8\rho + 6}{(1-\gamma_-)\,(1-2\gamma_-)\,(1-3\gamma_-)\,(1-\rho-\gamma_-)\,(1-\rho-2\gamma_-)\,(1-\rho-3\gamma_-)}.$$

Proof. Similar to the proof of Theorem 3.4 of Dekkers et al. (1989) by writing (P_n, Q_n, R_n) as a sum of i.i.d. random vectors. □

The following is an extension of a result by Drees (1998).

Lemma 5.2. *Let f be a measurable function. Suppose there exist a real parameter α and functions $a_1(t) > 0$ and $A_1(t) \to 0$ such that for all $x > 0$*

$$\lim_{t\to\infty} \frac{\frac{f(tx)-f(t)}{a_1(t)} - \frac{x^\alpha - 1}{\alpha}}{A_1(t)} = H_1(x)$$

where

$$H_1(x) = \frac{1}{\beta}[\frac{x^{\alpha+\beta}-1}{\alpha+\beta} - \frac{x^\alpha-1}{\alpha}] \quad (\beta \le 0).$$

Then for any $\epsilon > 0$ there exists $t_0 > 0$ such that for all $t \ge t_0$, $tx \ge t_0$,

$$|\frac{\frac{f(tx)-f(t)}{a_1(t)} - \frac{x^\alpha-1}{\alpha}}{A_1(t)} - H_1(x)| \le \epsilon[1 + x^\alpha + 2x^{\alpha+\beta}e^{\epsilon|\log x|}].$$

Proof. Suppose $\alpha \neq 0$. Then from relation (2.2) of Theorem 1 of de Haan and Stadmüller (1996), we have

$$\frac{(tx)^{-\alpha}a_1(tx) - t^{-\alpha}a_1(t)}{t^{-\alpha}a_1(t)A_1(t)} \to \frac{x^\beta - 1}{\beta}.$$

Hence

$$\begin{aligned} & \frac{f(tx)-a_1(tx)/\alpha-(f(t)-a_1(t)/\alpha)}{a_1(t)A_1(t)/\alpha} \\ = & \frac{f(tx)-f(t)-a_1(t)\frac{x^\alpha-1}{\alpha}}{a_1(t)A_1(t)/\alpha} \\ & -x^\alpha\frac{(tx)^{-\alpha}a_1(tx)-t^{-\alpha}a_1(t)}{t^{-\alpha}a_1(t)A_1(t)} \\ \to & \alpha H_1(x) - x^\alpha\frac{x^\beta-1}{\beta} = -\frac{x^{\alpha+\beta}-1}{\alpha+\beta}. \end{aligned}$$

Similar to the proof of Lemma 2.2 of de Haan and Peng (1997), we get

$$\begin{aligned} & |x^\alpha\frac{(tx)^{-\alpha}a_1(tx)-t^{-\alpha}a_1(t)}{t^{-\alpha}a_1(t)A_1(t)/\alpha} - x^\alpha\frac{x^\beta-1}{\beta}| \\ \leq & x^\alpha\epsilon[1+x^\beta e^{\epsilon|\log x|}] \end{aligned}$$

and

$$\begin{aligned} & |\frac{f(tx)-a_1(tx)/\alpha-(f(t)-a_1(t)/\alpha)}{a_1(t)A_1(t)/\alpha} + \frac{x^{\alpha+\beta}-1}{\alpha+\beta}| \\ \leq & \epsilon[1+x^{\alpha+\rho}e^{\epsilon|\log x|}]. \end{aligned}$$

Hence

$$\begin{aligned} & |\frac{\frac{f(tx)-f(t)}{a_1(t)}-\frac{x^\alpha-1}{\alpha}}{A_1(t)} - H_1(x)| \\ = & |\frac{f(tx)-a_1(tx)/\alpha-(f(t)-a_1(t)/\alpha)}{a_1(t)A_1(t)} + \frac{x^{\alpha+\beta}-1}{\alpha(\alpha+\beta)} \\ & +x^\alpha\frac{(tx)^{-\alpha}a_1(tx)/\alpha-t^{-\alpha}a_1(t)/\alpha}{t^{-\alpha}a_1(t)A_1(t)} - x^\alpha\frac{x^\beta-1}{\alpha\beta}| \\ \leq & \frac{\epsilon}{|\alpha|}[1+x^\alpha+2x^{\alpha+\beta}e^{\epsilon|\log x|}]. \end{aligned}$$

Suppose $\alpha = 0$ and $\beta < 0$. Then from the proof of Theorem 2 (iii) of de Haan and Stadtmüller (1996) we have $a_1(t) \to c_0 \in (0,\infty)$ and $\frac{c_0-a_1(t)}{a_1(t)A_1(t)} \to -1/\beta$. Hence

$$\frac{f(tx) - c_0\log(tx) - (f(t) - c_0\log t)}{a_1(t)A(t)} \to \frac{1}{\beta}\frac{x^\beta-1}{\beta}.$$

The rest of proof for $\alpha = 0$ and $\beta < 0$ is similar to the case $\alpha \neq 0$.

Suppose $\alpha = \beta = 0$. Write

$$g(t) := f(t) - \frac{1}{t}\int_0^t f(s)\,ds$$

which implies

$$f(t) = g(t) + \int_0^t \frac{g(s)}{s}\, ds$$

(see Corollary 1.2.1 of de Haan (1970)). From Omey and Willekens (1988) we have

$$\frac{g(tx) - g(t)}{a_1(t)A_1(t)} \to \log x.$$

Note that

$$\begin{aligned} & \frac{f(tx)-f(t)-a_1(t)\log x}{a_1(t)A_1(t)} \\ = \; & \frac{g(tx)-g(t)}{a_1(t)A_1(t)} + \int_1^x \frac{g(ts)-a_1(t)}{sa_1(t)A_1(t)}\, ds. \end{aligned}$$

Hence

$$\frac{g(tx) - a_1(t)}{a_1(t)A_1(t)} \to \log x - 1.$$

Furthermore

$$\frac{g(t) - a_1(t)}{a_1(t)A_1(t)} \to -1.$$

Using Proposition 1.19.4 of Geluk and de Haan (1987), we can easily see the lemma holds. Thus we complete the proof. $\square$

Let F_n denote the empirical distribution function of $\mathcal{X}_n$ and $U_n = (\frac{1}{1-F_n})^-$.

Lemma 5.3. *If (3.1) and (3.2) hold and* $n_1 = O(n^{1-\epsilon_0})$ *for some* $\epsilon_0 \in (0,1)$. *Then for any* $0 < \epsilon < 1$ *there exists* $t_0 > 0$ *such that for all* $t_0 \le t \le n_1(\log n_1)^2$ *and* $t_0 \le tx \le n_1(\log n_1)^2$

$$\begin{aligned} & \left|\frac{\frac{\log U_n(tx)-\log U_n(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)} - H(x)\right| \\ \le \; & [\frac{\sqrt{tx}\log n}{n} + \epsilon] d(\gamma_-,\rho) x^\rho e^{\epsilon|\log x|} \\ & + [\frac{\sqrt{t}\log n}{n} + \epsilon] d(\gamma_-,\rho) \\ & + \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_-+\rho}e^{\epsilon|\log x|}] \\ & + \frac{d(\gamma_-,\rho)}{|A(t)|}\frac{\log n}{n}[\sqrt{tx} + \sqrt{t}] \qquad a.s. \end{aligned} \tag{5.1}$$

where $d(\gamma_-,\rho) > 0$ *is a constant which only depends on* γ_- *and* ρ.

Proof. Let G_n denote the empirical distribution function of n independent, uniformly distributed random variables. As n is large enough and $n_1 = O(n^{1-\epsilon_0})$, we have

$$1/2 \leq \sup_{t \leq n_1 (\log n_1)^2} |tG_n^-(\frac{1}{t})| \leq 2 \quad \text{a.s.} \tag{5.2}$$

and

$$\sup_{t \geq 2} |\sqrt{t}(G_n(\frac{1}{t}) - \frac{1}{t})| \leq \frac{\log n}{\sqrt{n}} \quad \text{a.s.}$$

(see equations (10) and (17) of Chapter 10.5 of Shorack and Wellner (1986)). Hence

$$\sup_{4 \leq t \leq n_1 (\log n_1)^2} \sqrt{\frac{1}{G_n^-(\frac{1}{t})}} |G_n(G_n^-(\frac{1}{t})) - G_n^-(\frac{1}{t})| \leq \frac{\log n}{\sqrt{n}} \quad \text{a.s.}$$

Therefore for all $4 \leq t \leq n_1 (\log n_1)^2$

$$|tG_n^-(\frac{1}{t}) - 1| \leq \frac{2\sqrt{t}\log n}{\sqrt{n}} \quad \text{a.s.} \tag{5.3}$$

Now we use Lemma 5.2, (5.2), (5.3),

$$|y^\gamma - 1| \leq |\gamma|(2^{\gamma-1} \vee 2^{-\gamma+1})|y-1| \quad \text{for} \quad 1/2 \leq y \leq 2$$

and $U_n \stackrel{d}{=} U(\frac{t}{tG_n^-(\frac{1}{t})})$. It follows that for any $\epsilon \in (0,1)$ there exists $t_0 > 4$ such that for all $t_0 \le t \le n_1(\log n_1)^2$ and $t_0 \le tx \le n_1(\log n_1)^2$

$$
\begin{aligned}
& |\frac{\frac{\log U_n(tx)-\log U_n(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)} - H(x)| \\
\stackrel{d}{=} \quad & |\frac{\log U(\frac{tx}{txG_n^-(\frac{1}{tx})}) - \log U(tx) - \frac{a(tx)}{U(tx)}\frac{(txG_n^-(\frac{1}{tx}))^{-\gamma_-}-1}{\gamma_-}}{A(tx)a(tx)/U(tx)} \frac{A(tx)a(tx)/U(tx)}{A(t)a(t)/U(t)} \\
& - \frac{\log U(\frac{t}{tG_n^-(\frac{1}{t})}) - \log U(t) - \frac{a(t)}{U(t)}\frac{(tG_n^-(\frac{1}{t}))^{-\gamma_-}-1}{\gamma_-}}{A(t)a(t)/U(t)} \\
& + \frac{\log U(tx) - \log U(t) - \frac{a(t)}{U(t)}\frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)a(t)/U(t)} - H(x) \\
& + \frac{\frac{a(tx)}{U(tx)}\frac{(txG_n^-(\frac{1}{tx}))^{-\gamma_-}-1}{\gamma_-}}{A(t)a(t)/U(t)} - \frac{(tG_n^-(\frac{1}{t}))^{-\gamma_-}-1}{\gamma_- A(t)}| \\
\le \quad & \{|H(\frac{1}{txG_n^-(\frac{1}{tx})})| + \epsilon[1 + (txG_n^-(\frac{1}{tx}))^{-\gamma_-} \\
& + 2(txG_n^-(\frac{1}{tx}))^{-\gamma_-+\rho} e^{\epsilon|\log(txG_n^-(\frac{1}{tx}))|}]\}(1+\epsilon)x^\rho e^{\epsilon|\log x|} \\
& + |H(\frac{1}{tG_n^-(\frac{1}{t})})| + \epsilon[1 + (tG_n^-(\frac{1}{t}))^{-\gamma_-} \\
& + 2(tG_n^-(\frac{1}{t}))^{-\gamma_-+\rho} e^{\epsilon|\log(tG_n^-(\frac{1}{t}))|}] \\
& + \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_-+\rho}e^{\epsilon|\log x|}] \\
& + (1+\epsilon)|\frac{(txG_n^-(\frac{1}{tx}))^{-\gamma_-}-1}{\gamma_- A(t)}| + |\frac{(tG_n^-(\frac{1}{t}))^{-\gamma_-}-1}{\gamma_- A(t)}| \quad \text{a.s.} \\
\le \quad & [d(\gamma_-,\rho)\frac{\sqrt{tx}\log n}{\sqrt{n}} + \epsilon d(\gamma_-,\rho)]x^\rho e^{\epsilon \log x|} \\
& + d(\gamma_-,\rho)\frac{\sqrt{t}\log n}{\sqrt{n}} + \epsilon d(\gamma_-,\rho) \\
& + \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_-+\rho}e^{\epsilon|\log x|}] \\
& + \frac{d(\gamma_-,\rho)}{|A(t)|}\frac{\sqrt{tx}\log n}{\sqrt{n}} + \frac{d(\gamma_-,\rho)}{|A(t)|}\frac{\sqrt{t}\log n}{\sqrt{n}} \quad \text{a.s.}
\end{aligned}
$$

where $d(\gamma_-,\rho) > 0$ is a constant only depending on γ_- and ρ. The lemma follows. □

Proof of Theorem 3.1. A full proof of a somewhat restricted case has been given in Dekkers and de Haan (1993). We shall give a sketch of the proof.

By Lemma 5.2, for any $\epsilon > 0$ there exists $t_0 > 0$ such that for all $t \ge t_0$, $tx \ge t_0$

$$|\frac{\frac{\log U(tx)-\log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)} - H(x)| \le \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_-+\rho}e^{\epsilon|\log x|}].$$

Applying this relation with t replaced by $Y_{n,n-k}$ and x by $Y_{n,n-i}/Y_{n,n-k}$, adding the inequalities for $i = 0, 1, \cdots, k-1$ and dividing by k we get

$$
\begin{aligned}
& \frac{M_n^{(1)}(k)}{a(Y_{n,n-k})/U(Y_{n,n-k})} \\
\le \quad & \frac{1}{1-\gamma_-} + P_n + A(Y_{n,n-k})\frac{1}{k}\sum_{i=1}^k H(Y_{n,n-i+1}/Y_{n,n-k}) \\
& + \epsilon A(Y_{n,n-k})\frac{1}{k}\sum_{i=1}^k \{1 + (Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-} \\
& + 2(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-+\rho} e^{\epsilon|\log(Y_{n,n-i+1}/Y_{n,n-k})|}\}
\end{aligned}
$$

Note that $\{Y_{n,n-i+1}/Y_{n,n-k}\}_{i=1}^k \stackrel{d}{=} \{Y_i'\}_{i=1}^k$ with $Y_1', \cdots, Y_k'$ i.i.d. with common distribution function $1 - 1/x \quad (x > 1)$. We apply the law of large numbers to the third and fourth terms. Also note that $\frac{k}{n} Y_{n,n-k} \to 1$ in probability, so that since $|A|$ is regularly varying, we have $(A(n/k))^{-1} A(Y_{n,n-k}) \to 1$ in probability. As a result

$$\frac{M_n^{(1)}(k)}{a(Y_{n,n-k})/U(Y_{n,n-k})} = \frac{1}{1-\gamma_-} + P_n + A(n/k)\, d_1 + o_p(A(n/k)).$$

Hence

$$\begin{aligned} & \frac{(M_n^{(1)}(k))^2}{a^2(Y_{n,n-k})/U^2(Y_{n,n-k})} \\ = \; & \frac{1}{(1-\gamma_-)^2} + \frac{2P_n}{1-\gamma_-} + \frac{2A(n/k)\, d_1}{1-\gamma_-} + o_p(A(n/k)). \end{aligned}$$

Similarly

$$\begin{aligned} & \frac{M_n^{(2)}(k)}{a^2(Y_{n,n-k})/U^2(Y_{n,n-k})} \\ = \; & \frac{2}{(1-\gamma_-)(1-2\gamma_-)} + Q_n + A(n/k)\, d_2 + o_p(A(n/k)). \end{aligned}$$

Combining these expansions we get

$$\begin{aligned} & \hat{\gamma}_{n,2}(k) - \gamma \\ = \; & M_n^{(1)}(k) - \gamma_+ + \frac{M_n^{(2)}(k) - 2(M_n^{(1)}(k))^2}{2M_n^{(2)}(k) - 2(M_n^{(1)}(k))^2} - \gamma_- \\ = \; & (\gamma_+ + \frac{a(Y_{n,n-k})}{U(Y_{n,n-k})} - \gamma_+)[\frac{1}{1-\gamma_-} + P_n + d_1 A(n/k)] - \gamma_+ \\ & + \frac{(1-\gamma_-)^2(1-2\gamma_-)}{2}\{(1-2\gamma_-)Q_n - 4P_n + (d_2 - 2\gamma_- d_2 - 4d_1)A(n/k)\} \\ & + o_p(A(n/k)) \\ = \; & (\frac{a(Y_{n,n-k})}{U(Y_{n,n-k})} - \gamma_+)\frac{1}{1-\gamma_-} + (\frac{a(Y_{n,n-k})}{U(Y_{n,n-k})} - \gamma_+)[P_n + d_1 A(n/k)] \\ & + \frac{(1-\gamma_-)^2(1-2\gamma_-)}{2}\{(\frac{2\gamma_+}{(1-\gamma_-)^2(1-2\gamma_-)} - 4)P_n + (1-2\gamma_-)Q_n \\ & + (d_2 - 2\gamma_- d_2 - 4d_1 + \frac{2\gamma_+}{(1-\gamma_-)^2(1-2\gamma_-)} d_1)A(n/k)\} + o_p(A(n/k)). \end{aligned}$$

From the proof of Lemma 5.2 and Theorem 2, part (iii) of de Haan and Stadtmüller (1996) we can prove

$$\frac{\frac{a(t)}{U(t)} - \gamma_+}{A(t)} = \begin{cases} \frac{\frac{a(t)}{U(t)} - \gamma}{A(t)} \to \frac{\gamma}{\rho} & \text{if } \gamma > 0 \\ \frac{a(t)}{U(t)}/|A(t)| \to \infty & \text{if } \rho < \gamma \le 0 \\ \frac{a(t)}{U(t)}/A(t) \to 0 & \text{if } \gamma < \rho. \end{cases}$$

Consequently, by Lemma 5.1, we have that the asymptotic second moment of $\hat{\gamma}_{n,2}(k) - \gamma$ equals

$$\begin{aligned} & (V^2(\gamma)/k + b^2(\gamma,\rho)A_0^2(n/k)) \\ = \; & (V^2(\gamma)r/n + b^2(\gamma,\rho)A_0^2(r)) \end{aligned}$$

with $r := n/k$. One obtains the minimum with respect to r by using (3.6) and equating the derivative to zero (for details see Dekkers and de Haan (1993)). Since we have assumed in the derivation that $k(n)$ is an intermediate sequence, we still have to show that the resulting $k_0(n)$ is really the optimum. But it is easy to see that for any $k(n)$ with $k(n)/k_0(n) \to 0$ or ∞ the asymptotic second moment of $\hat{\gamma}_{n,2}(k) - \gamma$ is of large order as long as $k(n) \to \infty$ and $k(n)/n \to 0$ $(n \to \infty)$. In order to stay within these bounds, one can add the extra restriction $\log n \le k(n) \le n/\log n$ in the optimization procedure. The theorem follows. $\square$

Proof of Theorem 3.2. First we develop an asymptotic expansion for the alternative estimator $\hat{\gamma}_{n,3}(k)$. By the same arguments as in the proof of Theorem 3.1 we may show

$$\begin{aligned} & \frac{M_n^{(3)}(k)}{a^3(Y_{n,n-k})/U^3(Y_{n,n-k})} \\ = \;& \frac{6}{(1-\gamma_-)(1-2\gamma_-)(1-3\gamma_-)} + R_n + d_3 A(n/k) + o_p(A(n/k)) \end{aligned}$$

and

$$\begin{aligned} & \frac{M_n^{(1)}(k)M_n^{(2)}(k)}{a^3(Y_{n,n-k})/U^3(Y_{n,n-k})} \\ = \;& \frac{2}{(1-\gamma_-)^2(1-2\gamma_-)} + \frac{2P_n}{(1-\gamma_-)(1-2\gamma_-)} + \frac{Q_n}{1-\gamma_-} \\ & + \left(\frac{2d_1}{(1-\gamma_-)(1-2\gamma_-)} + \frac{d_2}{1-\gamma_-}\right)A(n/k) + o_p(A(n/k)). \end{aligned}$$

Hence

$$
\begin{aligned}
&\hat{\gamma}_{n,3}(k)-\gamma \\
= \; & \sqrt{M_n^{(2)}(k)/2}-\gamma_+ + \tfrac{M_n^{(3)}(k)-3M_n^{(1)}(k)M_n^{(2)}(k)}{3M_n^{(3)}(k)-3M_n^{(1)}(k)M_n^{(2)}(k)}-\gamma_- \\
= \; & \tfrac{a(Y_{n,n-k})}{U(Y_{n,n-k})}\sqrt{\tfrac{1}{(1-\gamma_-)(1-2\gamma_-)}+\tfrac{Q_n}{2}+\tfrac{d_2}{2}A(n/k)}-\gamma_+ \\
& +\tfrac{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}{12}\{-\tfrac{6}{1-2\gamma_-}P_n-3Q_n+(1-3\gamma_-)R_n \\
& +((1-3\gamma_-)d_3-\tfrac{6}{1-2\gamma_-}d_1-3d_2)A(n/k)\}+o_p(A(n/k)) \\
= \; & (\gamma_+ + \tfrac{a(Y_{n,n-k})}{U(Y_{n,n-k})}-\gamma_+)\{\tfrac{1}{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}+\tfrac{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{4}Q_n \\
& +\tfrac{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{4}d_2A(n/k)\}-\gamma_+ \\
& +\tfrac{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}{12}\{-\tfrac{6}{1-2\gamma_-}P_n-3Q_n+(1-3\gamma_-)R_n \\
& +((1-3\gamma_-)d_3-\tfrac{6}{1-2\gamma_-}d_1-3d_2)A(n/k)\}+o_p(A(n/k)) \\
= \; & (\tfrac{a(Y_{n,n-k})}{U(Y_{n,n-k})}-\gamma_+)\tfrac{1}{\sqrt{(1-\gamma_-)(1-2\gamma_-)}} \\
& +(\tfrac{a(Y_{n,n-k})}{U(Y_{n,n-k})}-\gamma_+)[\tfrac{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{4}Q_n+\tfrac{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{4}d_2A(n/k)] \\
& +\tfrac{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}{12}\{-\tfrac{6}{1-2\gamma_-}P_n+(\tfrac{3\gamma_+\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}-3)Q_n \\
& +(1-3\gamma_-)R_n+[(1-3\gamma_-)d_3-\tfrac{6}{1-2\gamma_-}d_1-3d_2 \\
& +\tfrac{3\gamma_+\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}d_2]A(n/k)\}+o_p(A(n/k)).
\end{aligned}
$$

Combining the above expansion with the expansion of $\hat{\gamma}_{n,2}(k)-\gamma$ in the proof of Theorem 3.1, we have

$$
\begin{aligned}
&\hat{\gamma}_{n,2}(k)-\hat{\gamma}_{n,3}(k) \\
= \; & (\tfrac{a(Y_{n,n-k})}{U(Y_{n,n-k})}-\gamma_+)(\tfrac{1}{1-\gamma_-}-\tfrac{1}{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}) \\
& +(\tfrac{a(Y_{n,n-k})}{U(Y_{n,n-k})}-\gamma_+)(P_n+d_1A(n/k)-\tfrac{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{4}Q_n \\
& -\tfrac{\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{4}d_2A(n/k)) \\
& +\tfrac{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}{12}\{[\tfrac{12\gamma_+}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}-\tfrac{24}{1-3\gamma_-}+\tfrac{6}{1-2\gamma_-}]P_n \\
& +[\tfrac{6(1-2\gamma_-)}{1-3\gamma_-}-\tfrac{3\gamma_+\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}+3]Q_n-(1-3\gamma_-)R_n \\
& +[\tfrac{6(1-2\gamma_-)}{1-3\gamma_-}d_2-\tfrac{24}{1-3\gamma_-}d_1+\tfrac{12\gamma_+}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}d_1-(1-3\gamma_-)d_3 \\
& +\tfrac{6}{1-2\gamma_-}d_1+3d_2-\tfrac{3\gamma_+\sqrt{(1-\gamma_-)(1-2\gamma_-)}}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)}d_2]A(n/k)\}+o_p(A(n/k)).
\end{aligned}
$$

The rest of proof is similar to the proof of Theorem 3.1. $\square$

Proof of Theorem 3.3. Let

$$E_n^{(1)} := \Big\{\omega: \text{ all of } |P_n|, |\frac{k}{n}Y_{n,n-k} - 1|, |d_n^{(1)} - d_1| \text{ and } \\ |D_n^{(1)} - D^{(1)}| \text{ are less than or equal to } k^{\delta_0 - 1/2}\Big\}$$

for some $\delta_0 \in (0, 1/2)$, where

$$D_n^{(1)} := \epsilon \frac{1}{k}\sum_{i=0}^{k-1}\left\{1 + (Y_{n,n-i}/Y_{n,n-k})^{\gamma_-} + 2(Y_{n,n-i}/Y_{n,n-k})^{\gamma_- + \rho + \epsilon}\right\}$$

and

$$D^{(1)} := \epsilon \int_1^\infty (1 + x^{\gamma_-} + x^{\gamma_- + \rho + \epsilon})x^{-2}dx.$$

Now take ϵ and t_0 as in the proof of Theorem 3.1. Then, provided $\frac{n}{k}(1 - k^{\delta_0 - 1/2}) \geq t_0$, we have $Y_{n,n-k} \geq t_0$ on $E_n^{(1)}$. Also, since A is regularly varying we have

$$|A(Y_{n,n-k}) - A(n/k)| < 2\epsilon A(n/k)$$

on $E_n^{(1)}$. Using these two facts and the inequalities in the beginning of the proof of Theorem 3.1, we find

$$\left|\frac{M_n^{(1)}(k)}{a(Y_{n,n-k})/U(Y_{n,n-k})} - \frac{1}{1 - \gamma_-} - P_n - A(n/k)d_1\right| < \epsilon A(n/k)$$

on the set $E_n^{(1)}$: so we have $o(A(n/k))$ instead of $o_p(A(n/k))$. Defining sets $E_n^{(2)}$ and $E_n^{(3)}$ related to the behavior of $M_n^{(2)}$ and $M_n^{(3)}$ we get similar inequalities for those.

Define

$$E_n := E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)}.$$

Using the mentioned inequalities and the fact that the conditions for the set E_n imply that P_n, Q_n, R_n and A are surely small, we can also replace $o_p(A(n/k))$ by $o(A(n/k))$ in the expansions given for $\hat{\gamma}_{n,2}(k)$ and $\hat{\gamma}_{n,3}(k)$ in the proof of Theorems 3.1 and 3.2 as long as we stay inside E_n.

Moreover the inequality $|\frac{k}{n}Y_{n,n-k} - 1| \leq k^{\delta_0 - 1/2}$ guarantees that we can replace

$$\frac{a(Y_{n,n-k})}{U(Y_{n,n-k})} - \gamma_+ \text{ by } \frac{a(n/k)}{U(n/k)} - \gamma_+$$

(cf. the limit relation for $-\gamma_+ + a(t)/U(t)$ in the proof of Theorem 3.1). Hence as in the proofs of Theorems 3.1 and 3.2 we find

$$\frac{\mathrm{E}\{\hat{\gamma}_{n,4}^2 \mathbf{1}(E_n)\}}{(\bar{V}^2(\gamma)/k + \bar{b}^2(\gamma,\rho)A_0^2(n/k))} \to 1.$$

Next we show that the contribution of the set E_n^c to the expectation may be neglected. For example by the definition of $\hat{\gamma}_{n,4}$

$$\mathrm{E}\{\hat{\gamma}_{n,4}^2 \mathbf{1}(|P_n| > k^{\delta_0 - 1/2})\} \le k^{2\delta-1}\,\mathrm{Pr}\{|P_n| > k^{\delta_0-1/2}\}$$

and by Bennett's inequality (cf. Petrov, 1975, Ch. III, §5) we can show

$$\mathrm{Pr}\{|P_n| > k^{\delta_0-1/2}\} \le k^{-\beta}$$

eventually for any $\beta > 0$. Hence

$$\frac{\mathrm{E}\{\hat{\gamma}_{n,4}^2 \mathbf{1}(|P_n| > k^{\delta_0-1/2})\}}{(\bar{V}^2(\gamma)/k + \bar{b}^2(\gamma,\rho)A_0^2(n/k))} \to 0, n \to \infty.$$

The reasoning in case any of the other conditions of the set E_n is violated is exactly the same (the inequality $\mathrm{Pr}\{|\frac{n}{k}Y_{n,n-k} - 1| > k^{\delta_0-1/2}\} \le k^{-\beta}$ can be obtained by translating the inequality for $\frac{k}{n}Y_{n,n-k}$ into one for its inverse $\frac{1}{k}\sum_{i=1}^n \mathbf{1}\{Y_i > \frac{n}{k}x\}$ and then applying Bennets inequality). This completes the proof of Theorem 3.3.

□

Proof of Theorem 3.4. Given $\mathcal{X}_n := \{X_1, \cdots, X_n\}$, we have

$$M_{n_1}^{(1)*}(k_1) \stackrel{d}{=} \frac{1}{k_1}\sum_{i=1}^{k_1} \log U_n(Y_{n_1,n_1-i+1}) - \log U_n(Y_{n_1,n_1-k_1})$$

with $\{Y_{n_1,i}\}_{i=1}^{n_1}$ the order statistics from a distribution function $1-1/x \quad (x > 1)$ and independent of $\mathcal{X}_n$. By the same arguments as in the proof of Theorem 3.1 using Lemma 5.3 instead of Lemma 5.2 we get

$$\begin{aligned} &\frac{M_{n_1}^{(1)*}(k_1)}{a(Y_{n_1,n_1-k_1})/U(Y_{n_1,n_1-k_1})} \\ = \;& \frac{1}{1-\gamma_-} + P_{n_1} + \frac{A(n_1/k_1)}{(1-\gamma_-)(\bar{\rho}-\gamma_-)} + o_p(A(n_1/k_1)) + O_p\Big(\frac{\sqrt{n_1/k_1}\log n}{\sqrt{n}}\Big). \end{aligned}$$

Note that $\frac{\sqrt{n_1/k_1}\log n}{\sqrt{n}} = o(1/\sqrt{k_1})$, so that the last term can be absorbed into the second one. The expansion for $M_{n_1}^{(1)*}(k_1)$ is the same as for $M_{n_1}^{(1)}(k_1)$

given $\mathcal{X}_n$. Similarly for $M_{n_1}^{(2)*}(k_1)$ and $M_{n_1}^{(3)*}(k_1)$. So the result of Theorem 3.1 holds with k_0 replaced by

$$k_{0,1}^*(n_1) := \arg\inf_k \text{ as.} \operatorname{E}\{(\hat{\gamma}_{n_1,2}^*(k))^2|\mathcal{X}_n\}.$$

and n by n_1. A similar analogue holds to the results of Theorem 3.2. Finally in a way analoguous to what was done in Theorem 3.3 we can replace as. E by the non-asymptotic expectation. Hence the conclusion. $\square$

Proof of Corollary 3.1. Note that $\lim_{t\to\infty} t^{-\gamma}a(t)/U(t)$ is a positive constant in the case $\rho < \gamma < 0$ (see de Haan and Stadtmüller (1993)). Thus $A_0(t) \sim c_0 t^{\rho^*}$ which implies

$$s^-(1/t) \sim (-2c_0^2\rho^*)^{\frac{1}{1-2\rho^*}} t^{\frac{1}{1-2\rho^*}}.$$

The Corollary easily follows from Theorems 3.4 and 3.5. $\square$

Proof of Theorem 3.5. This follows by combining the results of Corollary 3.1 for $\bar{k}_{0,1}^*(n_1)$ and $\bar{k}_{0,1}^*(n_2)$. $\square$

Proof of Corollary 3.2. It easily follows from Theorems 3.1, 3.2 and 3.5. $\square$

Proof of Corollary 3.3. We can use the result of Corollary 3.2 and we only have to prove that $\hat{\rho}_n$ is a consistent estimator of ρ^*. By Theorem 3.4 the sequence $\bar{k}_{0,1}^*(n_1)$ is asymptotic to $c_1 n_1^{\frac{-2\rho^*}{1-2\rho^*}}$. Hence

$$\log \bar{k}_{0,1}^*(n_1)/\log n_1 \to \frac{-2\rho^*}{1-2\rho^*}$$

in probability. This gives the consistency. $\square$

Lemma 5.4. *If $F \in D_{dif}(G_\gamma)$, then (4.1) holds for $a_n = nU'(n)$ and $b_n = U(n)$ and for any $k \to \infty$, $k/n \to 0$ and $\theta \in (0,1]$, the stochastic process*

$$W_{n,k}(\theta) := \sqrt{k}\frac{X_{n,n-[k\theta]} - U(\frac{n}{k\theta})}{\frac{n}{k}U'(\frac{n}{k})}$$

converges (in the sense of convergence of all finite marginal distributions) to a Gaussian process $w(\theta)$ which has mean zero and covariance structure

$$Cov(w(\theta_1), w(\theta_2)) = \theta_1^{-\gamma}\theta_2^{-\gamma-1}, \quad 0 < \theta_1 \le \theta_2 \le 1.$$

Proof. See Theorem 2.3 of Cooil (1985). $\square$

Lemma 5.5. *If (4.3) holds and $n_1 = O(n^{1-\epsilon_0})$ for some $\epsilon_0 \in (0,1)$. Then for any $0 < \epsilon < 1$ there exists $t_0 > 0$ such that for all $t_0 \le t \le n_1(\log n_1)^2$ and $t_0 \le tx \le n_1(\log n_1)^2$*

$$\begin{array}{rl} & |\frac{\frac{U_n(tx)-U_n(t)}{a(t)}-\frac{x^\gamma-1}{\gamma}}{A(t)} - h_{\gamma,\rho}(x)| \\ \le & [\frac{\sqrt{tx}\log n}{n}+\epsilon]D(\gamma,\rho)x^{\gamma+\rho}e^{\epsilon|\log x|} \\ & +[\frac{\sqrt{t}\log n}{n}+\epsilon]D(\gamma,\rho) \\ & +\epsilon[1+x^\gamma+2x^{\gamma+\rho}e^{\epsilon|\log x|}] \\ & +\frac{D(\gamma,\rho)}{|A(t)|}\frac{\sqrt{t}\log n}{n}[\sqrt{x}+1] \qquad a.s. \end{array} \tag{5.4}$$

where $D(\gamma,\rho) > 0$ is a constant which only depends on γ and ρ.

Proof. Similar to the proof of Lemma 5.3. □

Proof of Theorem 4.1. By Lemma 5.4 we have

$$\begin{array}{rl} & \sqrt{k}(\hat\gamma_{n,\theta}(k)-\gamma) \\ = & \sqrt{k}(\frac{1}{-\log\theta}\log\frac{X_{n,n-[k\theta^2]}-X_{n,n-[k\theta]}}{X_{n,n-[k\theta]}-X_{n,n-k}}-\gamma) \\ = & \frac{\sqrt{k}}{-\log\theta}\log(1+\theta^\gamma\frac{X_{n,n-[k\theta^2]}-X_{n,n-[k\theta]}}{X_{n,n-[k\theta]}-X_{n,n-k}}-1) \\ \stackrel{d}{=} & \frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[k\theta^2]}-X_{n,n-[k\theta]}-\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}(1+o_p(1)) \\ = & [\frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[k\theta]}-U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})(X_{n,n-[k\theta]}-U(\frac{n}{k\theta}))+\theta^{-\gamma}(X_{n,n-k}-U(\frac{n}{k}))}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})} \\ & +\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}](1+o_p(1)) \\ & (\text{ note } \frac{X_{n,n-[k\theta]}-X_{n,n-k}}{\frac{n}{k}U'(\frac{n}{k})} \stackrel{p}{\to} \frac{\theta^{-\gamma}-1}{\gamma}) \\ \stackrel{d}{=} & [\frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[k\theta^2]}-U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})(X_{n,n-[k\theta]}-U(\frac{n}{k\theta}))+\theta^{-\gamma}(X_{n,n-k}-U(\frac{n}{k}))}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}} \\ & +\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}](1+o_p(1)) \\ \stackrel{d}{=} & \frac{1}{-\log\theta}\frac{1}{\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(w(\theta^2)-(1+\theta^{-\gamma})w(\theta)+\theta^{-\gamma}w(1))+o_p(1) \\ & +\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(1+o_p(1)), \end{array}$$

thus the asymptotic variance of $\sqrt{k}(\hat\gamma_{n,\theta}(k)-\gamma)$ equals

$$\frac{\gamma^2(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{(\log\theta)^2(\theta^{-\gamma}-1)^2}$$

and the asymptotic bias of $\sqrt{k}(\hat\gamma_{n,\theta}(k)-\gamma)$ equals

$$\sqrt{k}A(\frac{n}{k})\frac{\theta^{-\rho}}{-\log\theta}\frac{\gamma}{\theta^{-\gamma}-1}\frac{1-\theta^\rho}{\rho}\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho}.$$

By $A(t) = ct^{-\rho}$ we get in a way similar to the proof of Theorem 3.1

$$k_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}})^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}\} \to 1.$$

$\square$

Proof of Theorem 4.2. By Lemma 5.4 we have

$$\begin{array}{rl}
& \sqrt{k}(\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta}(k\theta^2)) \\
= & \sqrt{k}(\hat{\gamma}_{n,\theta}(k) - \gamma) - \sqrt{k}(\hat{\gamma}_{n,\theta}(k\theta^2) - \gamma) \\
\stackrel{d}{=} & [\frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[m\theta^2]}-U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})(X_{n,n-[k\theta]}-U(\frac{n}{k\theta}))+\theta^{-\gamma}(X_{n,n-k}-U(\frac{n}{k}))}{\theta^{-\gamma}(X_{n,-[k\theta]}-X_{n,n-k})} \\
& +\frac{\sqrt{k}}{\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})} \\
& -\frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[k\theta^4]}-U(\frac{n}{k\theta^4})-(1+\theta^{-\gamma})(X_{n,n-[k\theta^3]}-U(\frac{n}{k\theta^3}))+\theta^{-\gamma}(X_{n,n-[k\theta^2]}-U(\frac{n}{k\theta^2}))}{\theta^{-\gamma}(X_{n,n-[k\theta^3]}-X_{n,n-[k\theta^2]})} \\
& -\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^3})-(1+\theta^{-\gamma})U(\frac{n}{k\theta^2})+\theta^{-\gamma}U(\frac{n}{k\theta^2})}{\theta^{-\gamma}(X_{n,n-[k\theta^3]}-X_{n,n-[k\theta^2]})}](1+o_p(1)) \\
& (\text{ note } \frac{X_{n,n-[k\theta^3]}-X_{n,n-[k\theta^2]}}{\frac{n}{k}U'(\frac{n}{k})} \stackrel{p}{\to} \theta^{-2\gamma}\frac{\theta^{-\gamma}-1}{\gamma}) \\
\stackrel{d}{=} & \frac{1}{-\log\theta}\frac{1}{\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(w(\theta^2)-(1+\theta^{-\gamma})w(\theta)+\theta^{-\gamma}w(1))+o_p(1) \\
& +\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(1+o_p(1)) \\
& -\frac{1}{-\log\theta}\frac{1}{\theta^{-3\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(w(\theta^4)-(1+\theta^{-\gamma})w(\theta^3)+\theta^{-\gamma}w(\theta^2))+o_p(1) \\
& -\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^4})-(1+\theta^{-\gamma})U(\frac{n}{k\theta^3})+\theta^{-\gamma}U(\frac{n}{k\theta^2})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-3\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(1+o_p(1)),
\end{array}$$

thus the asymptotic variance of $\sqrt{k}(\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta}(k\theta^2))$ equals

$$\frac{\gamma^2(1+\theta^{-2\gamma-1})(\theta^{-1}-1)(1+\theta^{-2})}{(\log\theta)^2(\theta^{-\gamma}-1)^2}$$

and the asymptotic bias of $\sqrt{k}(\hat{\gamma}_{n,\theta}(k) - \hat{\gamma}_{n,\theta}(k\theta^2))$ equals

$$\sqrt{k}A(\frac{n}{k})\frac{\theta^{-\rho}}{-\log\theta}\frac{1-\theta^\rho}{\rho}\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho}\frac{\gamma}{\theta^{-\gamma}-1}(1-\theta^{-2\rho}).$$

By $A(t) = ct^{-\rho}$ we get in a way similar to the proof of Theorem 3.1

$$\bar{k}_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})(1+\theta^{-2})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}\} \to 1.$$

$\square$

Proof of Theorem 4.3. Similar to the proof Theorem 3.3. □

Proof of Theorem 4.4. Similar to the proof of Theorem 3.4 by using Lemma 5.5 instead of Lemma 5.3. □

Proof of Theorem 4.5. Similar to the proof of Corollary 3.3. □

6 Simulations

The bootstrap procedure was tested on various distribution functions in a small simulation study: 200 samples of size 10.000 are generated from each distribution function. To each sample the bootstrap method was applied, with $\epsilon = 0.05$, that is $n_1 = 708$ and $n_2 = 502$, and 200 bootstrap samples. The distributions are Cauchy, generalized Pareto distribution (GPD) with $\gamma = 1/4$ and $\gamma = -1/4$, generalized extreme value distribution (GEV) with $\gamma = -1/4$ and $\gamma = -3/2$ and finally the distribution with $U(t) = H_{\gamma,\rho}(t)$,equation (3.2) with $\gamma = -1/4$, $\rho = -1/10$ and $\rho = -1$ (to have a distribution that allows a free choice of ρ).

Figure 1 illustrates the results for the Cauchy distribution: The bottom graph shows the observed and theoretical mean squared error of the γ estimate as a function of k: the solid line represents the observed mean squared error (i.e. the sample mean of the estimated $(\hat{\gamma}_{n,2}(k) - \gamma)^2$ of the individual simulated samples), and the dashed line the theoretical value calculated as $V^2(\gamma) + (A(n/k)\ b(\gamma,\rho))^2$. The vertical line indicates the sample mean of the k_0 estimates. The two components of the MSE, bias and variance are illustrated in the top and middle graphs. Table 1 summarizes the simulation results. For each parameter the table reports

- the theoretical value,
- statistics of the bootstrap estimates (sample mean, standard deviation and MSE of the estimates produced by the bootstrap procedure in the individual samples)
- observed optimal k, and the sample mean and MSE of the γ estimate at this k.

The general conclusion is that the bootstrap procedure gives reasonable estimates for the sample fraction to be used. It is reasonable in terms of the MSE of the γ estimate: for all but the last distribution the MSE of $\hat{\gamma}_{n,2}(\hat{k_0})$ is of the order of the MSE of the estimate at the observed optimal k.

The second order parameter ρ is only estimated correctly for the Cauchy distribution. The difficulty of estimating ρ has also been reported by others (see [3]) and is subject for further study.

For three of the distributions no theoretical values for k_0 has been given. These distributions have $\rho = \gamma$, a situation excluded in theorems (3.1) and (3.2). In this situation one cannot decide which part of the bias,

$$a(n/k)/U(n/k)\,\frac{1}{1-\gamma} \text{ or } A(n/k)\,\big(-2(1-\gamma)^2\,(1-2\gamma)\,d_1 + \frac{(1-\gamma)^2}{2}\,d_2\big)$$

is dominant. Most standard distributions with negative γ turn out to fall in this category.

Clearly more work needs to be done: first of all the performance of the ρ estimator and the $\gamma = \rho$ situation need clarification. And of course the effects of the number of bootstrap replications and the size of the bootstrap samples have to be studied.

7 Application

In the Neptune project (see de Haan and de Ronde (1997) for a review) we studied the joint distribution of extremes of wave-height, wave-period and sea levels. The project aimed at estimating failure probabilities of sea walls, based on the joint distribution of the extremes of the variables. A small dataset of 828 measurements covering 10 years at the Eierland station in the North sea was available. As is clear from figure 2, the wave height data series does not behave very nicely, but it was what was available to us. The difficulty in selecting a number of order statistics, makes the series a nice candidate for the bootstrap procedure (at the time we decided that only 27 order statistics should be used for estimation resulting in $\hat{\gamma}_{n,2} = 0$).

We applied the bootstrap method in the following way: as in the simulation experiment we used 200 bootstrap samples, resulting in an estimate of the optimal k. In order to evaluate and improve the precision of this estimate, the procedure was repeated with again 200 bootstrap samples, and the estimates averaged, until the average had an estimated standard error less then 2.

In figure 2 both the optimal $\hat{k}_0 = 259$, estimated with $\epsilon = 0.1$ and the corresponding $\hat{\gamma}_{n,2}(259) = 0.06$ are indicated. This value is reasonably in line with $\gamma = 0$, the value used in the Neptune project.

The results for different values of ϵ, determining the size of the bootstrap sample are shown in table 2. The optimal k is not very sensitive to ϵ, but it decreases with ϵ only when $\epsilon > 0.2$, but those values correspond to a very small second bootstrap size.

A Second order conditions

The second order relations in the Sections 3 and 4 are different. The reason for this stems from the expansion of the logarithms. Let us try to proceed from one to the other. The domain of attraction condition is

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma} \quad \text{for all } x > 0. \tag{A.1}$$

It follows, if $U(\infty) > 0$, that

$$\lim_{t\to\infty} \frac{\log U(tx) - \log U(t)}{a(t)/U(t)} = \frac{x^{\gamma_-} - 1}{\gamma_-}. \tag{A.2}$$

So far there are no complications. The natural second order condition related to (A.1) is

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = H(x) \tag{A.3}$$

for some function A (positive or negative) with $A(t) \to 0$ $(t \to \infty)$. Now we try to work towards a second order condition for $\log U$. Starting from (A.3)

$$\begin{aligned} & \log U(tx) - \log U(t) \\ = \; & \log(1 + (\tfrac{U(tx)}{U(t)} - 1)) \\ = \; & \tfrac{U(tx)}{U(t)} - 1 - \tfrac{1}{2}(\tfrac{U(tx)}{U(t)} - 1)^2 + \cdots \end{aligned}$$

So that (let us take $\gamma < 0$ for example)

$$\begin{aligned} & \tfrac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \tfrac{x^\gamma - 1}{\gamma} \\ = \; & (\tfrac{U(tx)-U(t)}{a(t)} - \tfrac{x^\gamma - 1}{\gamma}) - \tfrac{U(t)}{2a(t)}(\tfrac{U(tx)}{U(t)} - 1)^2 + \cdots \end{aligned}$$

Now in some cases the first term is dominant (the "nice" situation), but in other cases the second term is dominant. And sometimes there is no relative limit. The various cases are dealt with in the next Theorem and the Remark.

Theorem A. Assume $U(\infty) > 0$ and there exist functions $a(t) > 0$ and $A(t) \to 0$ such that

$$\frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} \to H_{\gamma,\rho}(x)$$

where

$$H_{\gamma,\rho}(x) = \frac{1}{\rho}[\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^{\gamma}-1}{\gamma}] \quad (\rho \leq 0).$$

Suppose that $\gamma \neq \rho$. Then

$$\lim_{t\to\infty} \frac{\frac{a(t)}{U(t)} - \gamma_+}{A(t)} = c \in [-\infty, \infty]$$

where

$$c = \begin{cases} 0 & \text{if } \gamma < \rho \\ \frac{\gamma}{\gamma+\rho} & \text{if } \gamma > -\rho \\ \frac{\gamma}{\gamma+\rho} & \text{if } 0 < \gamma < -\rho \text{ and } \lim_{t\to\infty}(U(t) - a(t)/\gamma) = 0 \\ \pm\infty & \text{if } \rho < \gamma \leq 0 \\ \pm\infty & \text{if } 0 < \gamma < -\rho \text{ and } \lim_{t\to\infty}(U(t) - a(t)/\gamma) \neq 0 \\ \pm\infty & \text{if } \gamma = -\rho. \end{cases}$$

Furthermore

$$\frac{\frac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{\tilde{A}(t)} \to H_{\gamma_-,\rho'}(x)$$

where

$$\tilde{A}(t) = \begin{cases} A(t) & \text{if } c = 0 \\ \gamma_+ - \frac{a(t)}{U(t)} & \text{if } c = \pm\infty \\ \rho A(t)/(\gamma+\rho) & \text{if } c = \gamma/(\gamma+\rho), \end{cases}$$

$\tilde{A}(t) \in RV_{\rho'}$,

$$\rho' = \begin{cases} -\gamma & \text{if } 0 < \gamma \leq -\rho \\ \gamma & \text{if } \rho < \gamma \leq 0 \\ \rho & \text{if } \gamma < \rho \text{ or } \gamma > -\rho. \end{cases}$$

Remark A. Hence $\rho' = 0$ if $\gamma = 0$.

Proof. Suppose that $\gamma \neq 0$. Then from the proof of Lemma 5.2 we have

$$\frac{U(tx) - a(tx)/\gamma - (U(t) - a(t)/\gamma)}{a(t)A(t)} \to -\frac{1}{\gamma}\frac{x^{\gamma+\rho}-1}{\gamma+\rho}.$$

If $\gamma + \rho > 0$, then

$$\frac{U(t) - a(t)/\gamma}{a(t)A(t)} \to -\frac{1}{\gamma(\gamma+\rho)}.$$

Hence

$$\frac{a(t)/U(t)-\gamma}{A(t)} = \frac{a(t)\gamma}{U(t)}\frac{a(t)/\gamma-U(t)}{a(t)A(t)} \to \gamma/(\gamma+\rho).$$

If $\gamma+\rho=0$, i.e., $\gamma=-\rho>0$, then

$$\frac{U(t)-a(t)/\gamma}{a(t)A(t)} \to \pm\infty.$$

Hence

$$\frac{a(t)/U(t)-\gamma}{A(t)} = \frac{a(t)\gamma}{U(t)}\frac{a(t)/\gamma-U(t)}{a(t)A(t)} \to \pm\infty.$$

If $\gamma+\rho<0$, then

$$\begin{cases} U(t)-a(t)/\gamma \to c_0 \in (-\infty,0)\cup(0,\infty) \\ \frac{U(t)-a(t)/\gamma-c_0}{a(t)A(t)} \to -\frac{1}{\gamma(\gamma+\rho)}. \end{cases}$$

For $\gamma+\rho<0$ and $\gamma>0$,i.e., $0<\gamma<-\rho$, we have

$$\begin{array}{ll} & \frac{a(t)/U(t)-\gamma}{A(t)} \\ = & \frac{a(t)\gamma}{U(t)}\left(\frac{a(t)/\gamma-U(t)+c_0}{a(t)A(t)} - \frac{c_0}{a(t)A(t)}\right) \\ \to & \begin{cases} \pm\infty & \text{if } c_0 \neq 0 \\ \gamma/(\gamma+\rho) & \text{if } c_0 = 0 \end{cases} \end{array}.$$

For $\gamma+\rho<0$ and $\gamma<0$, i.e., $\gamma<0$, we $|a(t)/(U(t)A(t))| \in RV_{\gamma-\rho}$. Hence

$$a(t)/(U(t)A(t)) \to \begin{cases} \pm\infty & \text{if } \gamma-\rho>0 \quad \& \quad \gamma<0 \\ 0 & \text{if } \gamma-\rho<0 \quad \& \quad \gamma<0. \end{cases}$$

Suppose that $\gamma=0$ and $\rho<0$. Then from the proof of Lemma 5.2 $a(t)\to c_1 \in (-\infty,0)\cup(0,\infty)$. Hence

$$a(t)/(U(t)A(t)) \sim c_1/(U(\infty)A(t)) \to \pm\infty.$$

We have now proved the first part of the theorem.

Note that $a(t)/U(t)\to\gamma_+$. For $\gamma\le 0$, we have

$$\begin{aligned} \log\frac{U(tx)}{U(t)} &= \log\{1+\tfrac{a(t)}{U(t)}[\tfrac{x^{\gamma_-}-1}{\gamma_-}+A(t)H_{\gamma,\rho}(x)+o(A(t))]\} \\ &= \tfrac{a(t)}{U(t)}[\tfrac{x^{\gamma_-}-1}{\gamma_-}+A(t)H_{\gamma,\rho}(x)+o(A(t))] \\ &\quad -\tfrac{1}{2}(\tfrac{a(t)}{U(t)})^2[\tfrac{x^{\gamma_-}-1}{\gamma_-}+A(t)H_{\gamma,\rho}(x)+o(A(t))]^2+o((\tfrac{a(t)}{U(t)})^2), \end{aligned}$$

i.e.,

$$
\begin{aligned}
& \tfrac{\log U(tx)-\log U(t)}{a(t)/U(t)} - \tfrac{x^{\gamma_-}-1}{\gamma_-} \\
= \; & A(t)H_{\gamma,\rho}(x) + o(A(t)) \\
& -\tfrac{a(t)}{2U(t)}[(\tfrac{x^{\gamma_-}-1}{\gamma_-})^2 + 2\tfrac{x^{\gamma_-}-1}{\gamma_-}A(t)H_{\gamma,\rho}(x) + o(A(t))] + o(\tfrac{a(t)}{U(t)}).
\end{aligned}
$$

For $\gamma > 0$, we have

$$
\begin{aligned}
& x^{-\gamma}\tfrac{U(tx)}{U(t)} \\
= \; & x^{-\gamma} + \tfrac{a(t)}{U(t)}\tfrac{1-x^{-\gamma}}{\gamma} + x^{-\gamma}\tfrac{a(t)}{U(t)}[A(t)H_{\gamma,\rho}(x) + o(A(t))] \\
= \; & 1 + (x^{-\gamma}-1)(1-\tfrac{a(t)}{\gamma U(t)}) + x^{-\gamma}\tfrac{a(t)}{U(t)}[A(t)H_{\gamma,\rho}(x) + o(A(t))],
\end{aligned}
$$

i.e.,

$$
\begin{aligned}
& \tfrac{\log U(tx)-\log U(t)}{a(t)/U(t)} - \log x \\
= \; & -(\log x + \tfrac{x^{-\gamma}-1}{\gamma})\tfrac{U(t)}{a(t)}(a(t)/U(t)-\gamma) \\
& +x^{-\gamma}A(t)H_{\gamma,\rho}(x) + o(A(t)) + o(a(t)/U(t)-\gamma).
\end{aligned}
$$

So the second part of the theorem follows easily. □

Remark B. It is not true that a second order condition for U always implies a second order condition for $\log U$: Let $\gamma = \rho$ and define

$$U'(t) = t^{\gamma-1}\exp\{\int_1^t s^{\gamma-1}(2+\sin(\log\log s))\,ds\}.$$

From the representation (2.5) of de Haan and Resnick (1996) we find

$$\frac{\frac{U(tx)-U(t)}{tU'(t)} - \frac{x^{\gamma}-1}{\gamma}}{t^{\gamma}[2+\sin(\log\log t)]} \to \int_1^x u^{\gamma-1}\frac{u^{-\gamma}-1}{-\gamma}\,du.$$

Hence

$$
\begin{aligned}
& a(t)/(U(t)A(t)) \\
= \; & \tfrac{tU'(t)}{U(t)t^{\gamma}[2+\sin(\log\log t)]} \\
\sim \; & \tfrac{\exp\{\int_1^{\infty} s^{\gamma-1}[2+\sin(\log\log s)]\,ds}{U(\infty)[2+\sin(\log\log t)]}
\end{aligned}
$$

which does not have a limit.

References

[1] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*. Cambridge University Press, New York, 1987.

[2] B. Cooil. Limiting multivariate distributions of intermediate order statistics. *Ann. Probab.*, 13:469–477, 1985.

[3] J. Danielsson, L. de Haan, L. Peng, and C.G. de Vries. Using a bootstrap method to choose the sample fraction in the tail index estimation. 1996. Submitted.

[4] L. de Haan. On regular variation and its application to the weak convergence of sample extremes. Technical Report CWI Tract 32, CWI, Amsterdam, 1970.

[5] L. de Haan and L. Peng. Rates of convergence for bivariate extremes. *Journal of Multivariate Analysis*, 61(2):195–230, 1997.

[6] L. de Haan and S.I. Resnick. Second order regular variation and rates of convergence in extreme value theory. *Annals of Probability*, 24:97–124, 1996.

[7] L. de Haan and U. Stadtmüller. Generalized regular variation of second order. *J. Austral. Math. Soc. (Series A)*, 61:381–395, 1996.

[8] Laurens de Haan and John de Ronde. Sea and wind: Multivariate extremes at work. *Extremes*, 1(1):7–45, 1998.

[9] A.L.M. Dekkers and L. de Haan. Optimal choice of sample fraction in extreme value estimation. *Journal of Multivariate Analysis*, 47(2):173–195, 1993.

[10] A.L.M. Dekkers, J.H.J. Einmahl, and L. de Haan. A moment estimator for the index of an extreme value distribution. *Ann. Statis.*, 17(4):1833–1855, 1989.

[11] H. Drees. On smooth statistical tail functionals. *Scandinavian Journal of Statistics,*, 25(1):187–210, 1998.

[12] H. Drees and E. Kaufmann. Selecting the optimal sample fraction in univariate extreme value estimation. *Stoch. Proc. Appl.*, 75:149–172, 1998.

[13] W. Feller. *An Introduction to Probability Theory and its Applications*, volume II. Wiley, New York, 1966.

[14] J. Geluk and L. de Haan. Regular variation, extensions and tauberian theorems. Technical Report CWI Tract 40, CWI, Amsterdam, 1987.

[15] J. Pickands III. Statistical inference using extreme order statistics. *Ann. Statis.*, 3:119–131, 1975.

[16] J. Pickands III. The continuous and differentiable domains of attraction of the extreme value distributions. *Ann. Probab.*, 14:996–1004, 1986.

[17] E. Omey and E. Willekens. π-variation with remainder. *J. London Math. Soc.*, 37:105–118, 1988.

[18] T.T. Pereira. Second order behaviour of domains of attraction and the bias of generalized pickands' estimator. In J. Lechner J. Galambos and E. Simin, editors, *Extreme Value Theory and Applications III, Proc. Gaithersburg Conference (NIST special publ)*, 1993.

[19] V. Petrov. *Sums of independent random variables.* Springer, Berlin, 1975.

[20] G. Shorack and J. Wellner. *Empirical Processes with Applications to Statistics.* John Wiley & Sons, 1986.

Table 1: Simulation study. Statistics of the bootstrap estimates compared to theoretical values and observed (empirical) estimates of k_0, ρ and γ (see section 6).

		Bootstrap estimator			observed	
Parameter	theoretical	mean	stdev	MSE	mean	MSE
Cauchy						
k_0	1563.9	1354.3	667.62		1546	
ρ	-2	-1.9459	0.52107			
$\gamma(k_0)$	1	0.99828	0.091811	0.0084	1.0148	0.00166
GPD(1/4)						
k_0	641.97	1140.3	632.26		587	
ρ	-0.25	-1.7416	0.48564			
$\gamma(k_0)$	0.25	0.28981	0.062448	0.0055	0.28548	0.0032
GPD(-1/4)						
k_0		719.86	483.82		1403	
ρ	-0.25	-1.6816	0.40212			
$\gamma(k_0)$	-0.25	-0.21809	0.053223	0.0038	-0.21945	0.0018
$H_{-1/4,-1/10}$						
k_0	94.697	160.34	150.39		92	
ρ	-0.1	-0.95093	0.21389			
$\gamma(k_0)$	-0.25	-0.12098	0.18311	0.0500	-0.12362	0.0286
$H_{-1/4,-1}$						
k_0		325.54	215.94		347	
ρ	-0.25	-1.173	0.27012			
$\gamma(k_0)$	-0.25	-0.22321	0.1224	0.0156	-0.22616	0.00401
GEV(-1/4)						
k_0		746.38	441.34		1239	
ρ	-0.25	-1.7196	0.39267			
$\gamma(k_0)$	-0.25	-0.24751	0.052803	0.0028	-0.25281	0.0010
GEV(-3/2)						
k_0	957.98	3022.9	1012.2		1083	
ρ	-1	-3.1803	0.54967			
$\gamma(k_0)$	-1.5	-1.8992	0.44293	0.3546	-1.5899	0.01641

Table 2: Wave height data ($n = 828$). The effect of the bootstrap sample size, determined by ϵ, on the estimated ρ, optimal number of order statistics $\hat{k_0}$ and $\hat{\gamma}_{n,2}(k_0)$. See section 7.

ϵ	n_1	n_2	$\hat{\rho}$	$\hat{k}_0$	$\hat{\gamma}_{n,2}$	std.err
0.05	592	424	-5.3795	270	0.074693	0.061144
0.1	423	217	-4.9549	250.86	0.045954	0.063972
0.15	303	111	-4.5453	258.55	0.059485	0.063004
0.2	216	57	-4.4864	258.11	0.059485	0.058986
0.25	155	30	-3.9546	223.01	-0.0011046	0.070302
0.3	111	15	-3.2259	137.83	-0.22552	0.094392

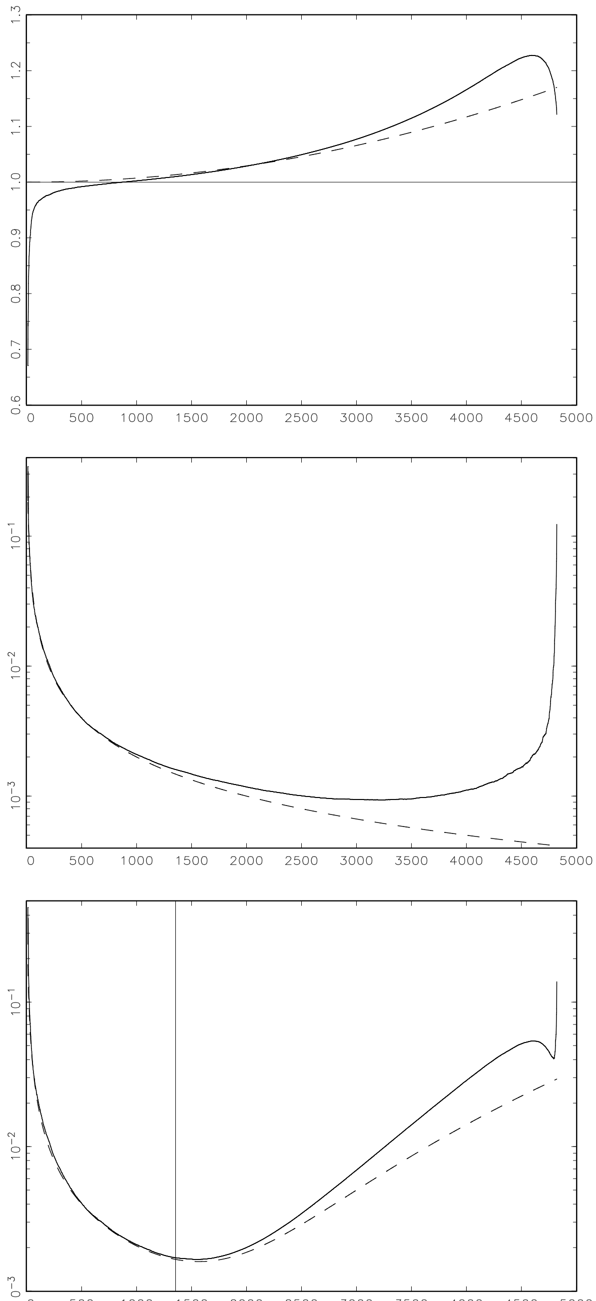

Figure 1: The estimate $\hat{\gamma}_{n,2}$ for the Cauchy distribution: from top to bottom the sample mean (i.e. the average of all simulations), the variance and the mean squared error against the number of order statistics. The solid lines represent the observed values and the dashed line the theoretical values. See section 6.

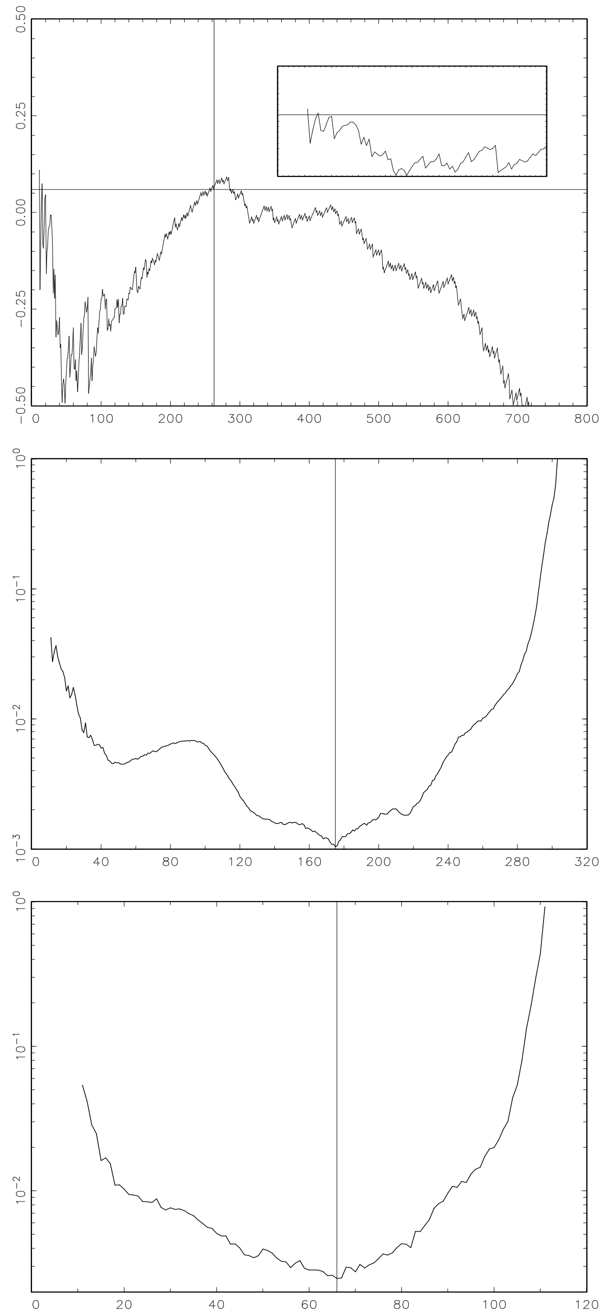

Figure 2: Wave height γ. The top graph shows the $\hat{\gamma}_{n,2}$ estimate as a function of the number of order statistics k. The optimal k_0 and the $\hat{\gamma}_{n,2}$ estimate are indicated by the vertical resp. horizontal lines. The inset enlarges the graph for the top 100 order statistics. The other graphs show the bootstrap estimates of $\mathrm{E}\{(\hat{\gamma}_{n,2}-\hat{\gamma}_{n,3})^2\}$ as a function of $k(n_1)$ resp. $k(n_2)$ (200 bootstrap repeats; $\epsilon = 0.1$). See section 7.