Bootstrapping Extreme Value Estimators

Laurens de Haan*
Erasmus University Rotterdam and University of Lisbon

Chen Zhou†
Erasmus University Rotterdam and De Nederlandsche Bank

January 5, 2021

Abstract

This paper develops a bootstrap analogue of the well-known asymptotic expansion of the tail quantile process in extreme value theory. One application of this result is to estimate the variance of estimators of the extreme value index such as the Probability Weighted Moment (PWM) estimator. The context is the block maxima method or the peaks-over-threshold method.

Keywords: tail quantile process; block maxima; peak-over-threshold

* *E-mail addresses:* ldehaan@ese.eur.nl

† Postal address: Econometric Institute, Erasmus University Rotterdam, 3000DR Rotterdam, The Netherlands. Phone: +31(10)4082269.
E-mail addresses: zhou@ese.eur.nl

1 Introduction

Statistical inference based on extreme value theory uses only large observations in a sample. The large observations are selected by either the peaks-over-threshold (POT) method or by the block maxima method. Asymptotic theories for most estimators based on these two methods are established using the tail quantile process of the peaks-over-threshold or the quantile process of block maxima. However, the asymptotic variance obtained in such asymptotic theories, though mostly explicit, can be intricate. One example is the much used probability weighted moment (PWM) estimator in the block maxima method; see Ferreira and de Haan (2015). The direct motivation for the present research is to estimate this variance via the bootstrap. We derive a bootstrap version of the fundamental expansions for the (tail) quantile process in both the POT and block maxima methods. Consequently, for any statistical estimator in extreme value theory whose asymptotic property is established via the (tail) quantile process, the bootstrap mimics faithfully the original asymptotic behavior of the estimator.

Although the bootstrap is a widely used method for obtaining the distribution of a statistical estimator, it does not work automatically for estimators based on extreme value theory. A somewhat trivial example is that the bootstrapped full sample maxima lies below the actual full sample maxima. More formally, Bickel and Freedman (1981) shows the non-consistency of the bootstrap method, when using the sample maxima as an estimator of the right endpoint of a distribution. For the definition of the consistency of the bootstrap method, see Van der Vaart (1998, Section 23.2). Broadly speaking, proving consistency of the bootstrap in extreme value theory is a difficult issue because even if observations are drawn from a distribution that satisfies the extreme value conditions, the corresponding empirical distribution function does not satisfy the extreme value conditions. Consequently, it is not obvious that the bootstrap can be used to obtain the distribution of statistical estimators in extreme value theory. Therefore, our somewhat surprising result provides confidence for using the bootstrap in extreme value theory.

The first question regarding the bootstrap in extreme value theory is what to bootstrap. For the POT method, one can bootstrap just the selected peaks or one can bootstrap the original sample and reconstruct the peaks. We choose the latter for better expected performance because the former seems not to randomize sufficiently the observations. In the block maxima method, one can bootstrap just the block maxima of the original sample or one can bootstrap the original sample and reconstruct the block maxima. We again choose the latter for a similar reason.

For the POT method we wish to construct a bootstrap analogue of the following fundamental expansion of the tail quantile process (due to Drees (1998)). Let $\tilde{X}_1, \tilde{X}_2, \ldots$ be a sequence of i.i.d. random variables with a common distribution function F. Suppose that the sample maximum $\tilde{X}_{n:n}$, properly normalized, converges to one of the extreme value distributions $G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}$. Note that G_γ is a Fréchet distribution for γ positive, a negative Weibull distribution for γ negative and a Gumbel distribution for $\gamma = 0$. Denote $\tilde{X}_{1,n} \leq \cdots \leq \tilde{X}_{n,n}$ as the order statistics from a sample of n observations. For any intermediate sequence $k = k(n)$ satisfying $k \to \infty$, $k/n \to 0$ as $n \to \infty$, the process $\left\{\tilde{X}_{n-[ks],n}\right\}_{0\leq s\leq 1}$ is called the tail quantile process.

The asymptotic expansion of the tail quantile process relies on the second order condition of extreme value theory (cf. Appendix B, de Haan and Ferreira (2006)) as follows. Define $U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right)$ for $t > 0$ where $^{\leftarrow}$ indicates the left-continuous inverse function. Assume that there exist $\rho \leq 0$, a positive function a and an eventually positive or negative function A with $\lim_{t\to\infty} A(t) = 0$ such that for $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} =: \Psi_{\gamma,\rho}(x), \tag{1.1}$$

where

$$\Psi_{\gamma,\rho}(x) = \begin{cases} \frac{x^{\gamma+\rho}-1}{\gamma+\rho}, & \text{if } \rho < 0 \\ \frac{1}{\gamma} x^\gamma \log x, & \text{if } \rho = 0 \neq \gamma \\ \frac{1}{2}(\log x)^2, & \text{if } \rho = 0 = \gamma. \end{cases}$$

Here γ and ρ are called the extreme value index and the second order index respectively, while a and A are called the first and second order scale functions respectively. This second order condition implies the mentioned convergence of the sample maximum.

The following result is a version of Theorem 2.1 in Drees (1998); see also Corollary 2.4.5 in de Haan and Ferreira (2006).[1]

Proposition 1.1 *Assume the second order condition* (1.1) *holds. Let* $k = k(n) \to \infty$ *and* $\sqrt{k}A\left(\frac{n}{k}\right) = O(1)$ *as* $n \to \infty$*. There exist an appropriate version of the auxiliary functions* a *and* A*, denoted as* a_0 *and* A_0*, a sequence of independent Brownian motions* $W_1, W_2, \ldots$ *such that, for any* $\varepsilon > 0$*, as* $n \to \infty$*,*

$$\begin{aligned}&\sqrt{k}\left(\frac{\tilde{X}_{n-[kx],n} - b_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{x^{-\gamma}-1}{\gamma}\right)\\ =&x^{-\gamma-1}W_n(x) + \sqrt{k}A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(\frac{1}{x}\right) + x^{-\gamma-1/2-\varepsilon}o_P(1),\end{aligned} \tag{1.2}$$

holds uniformly for $x \in (0,2]$*, where*

$$b_0\left(\frac{n}{k}\right) := \begin{cases} U\left(\frac{n}{k}\right) & \text{if } \gamma \geq -1/2 \\ \tilde{X}_{n,n} + \frac{a_0\left(\frac{n}{k}\right)}{\gamma} + \frac{a_0\left(\frac{n}{k}\right)A_0\left(\frac{n}{k}\right)}{\gamma+\rho}1_{\rho<0} & \text{if } \gamma < -1/2. \end{cases}$$

The first goal of this paper is to establish a parallel result for the tail quantile process based on the bootstrapped observations. Take a bootstrap sample, that is, draw i.i.d. random variables $\tilde{X}_1^*, \tilde{X}_2^*, \ldots, \tilde{X}_n^*$ from the empirical distribution function F_n of $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$. After ordering the bootstrapped observations as $\tilde{X}_{1,n}^* \leq \cdots \leq \tilde{X}_{n,n}^*$, we consider the bootstrapped tail quantile process $\left\{\tilde{X}_{n-[ks],n}^*\right\}_{0\leq s\leq 1}$ for an intermediate sequence k and will prove a result analogue to (1.2); see Section 2. The bootstrap analogue that we develop has a similar structure except that in the expansion we have two independent Brownian motion terms, one due to the randomness of the original sample and the other one due to the bootstrap

[1] Notice that Corollary 2.4.5 in de Haan and Ferreira (2006) is subject to a technical error in the random shift function $b_0\left(\frac{n}{k}\right)$, which is fixed in this Proposition (personal communication with the authors).

randomness.

The same holds - *mutatis mutandis* - for the fundamental expansion of the quantile process of the block maxima. The original result in Ferreira and de Haan (2015) is as follows. Recall that $\tilde{X}_1, \tilde{X}_2, \ldots$ are a sequence of i.i.d. random variables with a common distribution function F. Define the block maxima of the sample $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ as

$$X_i = \max_{(i-1)m<j\leq im} \tilde{X}_j,$$

for $i = 1, 2, \ldots, k$, where $n = mk$. Let $X_{1,k} \leq X_{2,k} \leq \ldots \leq X_{k,k}$ be the order statistics of the block maxima $X_1, X_2, \ldots, X_k$.

Denote $V := \left(\frac{1}{-\log F}\right)^{\leftarrow}$. The asymptotic expansion of the quantile process of block maxima $\left\{X_{\lceil kx\rceil,k}\right\}_{x\in[1/(k+1),k/(k+1)]}$ relies on a different second order condition, this time regarding the V function: with replacing U by V in the second order condition (1.1), we assume that a similar limit relation holds with different auxiliary functions, a scale function $\tilde{a}$, a second order scale function $\tilde{A}$ and a different second order index $\tilde{\rho}$. The asymptotic expansion is given in the following Proposition; see Theorem 2.1 in Ferreira and de Haan (2015).

Proposition 1.2 *Assume the second order condition* (1.1) *holds for the V function. Let $m = m(n) \to \infty$ and $k = k(n) \to \infty$ in such a way that $\sqrt{k}\tilde{A}(m) \to \lambda \in \mathbb{R}$ as $n \to \infty$, where $\tilde{A}$ is the second order scale function for the V function. Then, there exist an appropriate version of the auxiliary functions $\tilde{a}$ and $\tilde{A}$, denoted as $\tilde{a}_0$ and $\tilde{A}_0$, and an appropriate sequence of Brownian bridges $\{B_k(\cdot)\}_{k=1}^{\infty}$ such that for any $\varepsilon > 0$, as $n \to \infty$,*

$$\sqrt{k}\left(\frac{X_{\lceil kx\rceil,k} - V(m)}{\tilde{a}_0(m)} - \frac{(-\log x)^{-\gamma} - 1}{\gamma}\right) = \frac{B_k(x)}{x(-\log x)^{1+\gamma}} + \sqrt{k}\tilde{A}_0(m)\Psi_{\gamma,\tilde{\rho}}\left(\frac{1}{-\log x}\right) + x^{-1/2-\varepsilon}(1-x)^{-1/2-\gamma-\tilde{\rho}-\varepsilon}o_P(1), \tag{1.3}$$

holds uniformly for all $x \in [1/(k+1), k/(k+1)]$.

Next, recall that the bootstrapped sample are i.i.d. random variables $\tilde{X}_1^*, \tilde{X}_2^*, \ldots, \tilde{X}_n^*$

drawn from the empirical distribution function F_n of $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$. Then, we construct the bootstrapped block maxima by

$$X_i^* = \max_{(i-1)m<j\leq im} \tilde{X}_j^* \quad \text{for } i = 1, \ldots, k$$

This setup shows that we bootstrap the original sample and reconstruct the bootstrapped block maxima from the bootstrapped sample instead of bootstrapping the block maxima from the original sample.

The second goal of the paper is to prove for the bootstrap block maxima a result similar to (1.3) for the original sample. This will be done in Section 3. The bootstrap analogue of this result that we develop has a similar structure except that in the expansion we have a Brownian bridge term due to the randomness of the bootstrap and a Brownian motion term due to the randomness of the original sample.

The bootstrap expansions that we prove lead easily (by integrating the various terms) to the asymptotic distribution of extreme value estimators. In Section 4, a few examples are given for applying our bootstrap expansions in the POT and the block maxima methods to estimate the distribution of estimators for the extreme value index. For the POT method, we use the PWM estimator (Hosking and Wallis (1987)) as an example. For the block maxima method, we use the PWM estimator in Hosking et al. (1985) as an example. We show that the sample variance of bootstrapped estimates can be a good estimator for the asymptotic variance of the original estimator.

The proof of our main results uses a simple representation of the bootstrap sample. For the POT method the representation for the bootstrap tail quantile process is as follows

Lemma 1.3 *Let F_n be the empirical distribution function of $\left\{\tilde{X}_j\right\}_{j=1}^n$. Let $Y_1^*, \cdots, Y_n^*$ be i.i.d. random variables following the standard Pareto distribution $1 - 1/x$, for $x > 1$, and* *independent of* $\left\{\tilde{X}_j\right\}_{j=1}^n$. *Then,*

$$\left\{\tilde{X}_j^*\right\}_{j=1}^n \stackrel{d}{=} \left\{F_n^{\leftarrow}(1-1/Y_j^*)\right\}_{j=1}^n , \tag{1.4}$$

$$\left\{\tilde{X}_{n-[ks],n}^*\right\}_{s\in(0,1]} \stackrel{d}{=} \left\{\tilde{X}_{n-[k\tilde{D}_n(s)],n}\right\}_{s\in(0,1]} , \tag{1.5}$$

where $\tilde{D}_n(s) = \frac{n}{kY_{n-[ks],n}^*}$.

Here $\left\{\tilde{X}_{n-[kx],n}\right\}$ is the tail quantile process of the original sample and $\tilde{D}_n(s) = \frac{n}{kY_{n-[ks],n}^*}$ with $\left\{Y_{n-[ks],n}^*\right\}$ the tail quantile process for the standard Pareto distribution. The two processes are independent. Then we combine the expansions of both processes. The combined expansion requires that $\tilde{D}_n(s)$ is in the correct range of the process $\left\{\tilde{X}_{n-[kx],n}\right\}$. The proof for the block maxima method is considerably more complicated. It is based on a similar representation for the bootstrap quantile process of the block maxima; see (3.4) below.

In Section 2 the fundamental expansion for the POT method is given along with an outline of the proof. The rest of the proof can be found in the Appendix A.1. Section 3 handles the block maxima method in a similar way with leaving detailed proofs to the Appendix A.2. Section 4 provides a few examples for applications. We show that the asymptotic expansions of the tail quantile process of the bootstrapped sample leads to a consistency result for the bootstrapped PWM estimator using the POT method. However, the asymptotic expansion of the quantile process of the bootstrapped sample maxima does not lead to a similar consistency result for the PWM estimator using the block maxima method. Section 5 provides simulation studies illustrating the theoretical findings.

2 The peaks-over-threshold method

Let $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ be a sequence of i.i.d. random variables with common distribution function F. Denote $U := (1/(1-F))^{\leftarrow}$ as the corresponding quantile function and $\tilde{X}_{1,n} \leq \tilde{X}_{2,n} \leq \cdots \leq \tilde{X}_{n,n}$ as the order statistics. Let $\tilde{X}_1^*, \tilde{X}_2^*, \ldots, \tilde{X}_n^*$ be an i.i.d. bootstrapped sample drawing from the empirical distribution function F_n of $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$.

We prove the following bootstrap analogue of the tail quantile process result in (1.2). A paper related to this result is Litvinova and Mervyn (2018).

Theorem 2.1 *Suppose the second order condition* (1.1) *holds. Let* $k = k(n) \to \infty$ *and* $\sqrt{k}A\left(\frac{n}{k}\right) = O(1)$ *as* $n \to \infty$. *There exists an appropriate version of the functions* a *and* A, *denoted as* a_0 *and* A_0, *two independent sequences of Brownian motions* $W_1, W_2, \ldots$ *and* $W_1^*, W_2^*, \ldots$ *such that as for any* $\varepsilon > 0$, *as* $n \to \infty$,

$$\begin{aligned}
&\sqrt{k}\left(\frac{\tilde{X}^*_{n-[ks],n} - b_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{s^{-\gamma}-1}{\gamma}\right) \\
&\quad = s^{-\gamma-1}W_n^*(s) + s^{-\gamma-1}W_n(s) + \sqrt{k}A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(\frac{1}{s}\right) + s^{-\gamma-1/2-\varepsilon}o_P(1),
\end{aligned} \tag{2.1}$$

holds uniformly for all $s \in [1/(k+1), 1]$, *where* $b_0\left(\frac{n}{k}\right)$ *is the same as in Proposition 1.1.*

Combining Theorem 2.1 with Proposition 1.1 yields the following result

Corollary 2.2 *Under the conditions in Theorem 2.1, as* $n \to \infty$,

$$\sqrt{k}\left(\frac{\tilde{X}^*_{n-[ks],n} - \tilde{X}_{n-[ks],n}}{a_0\left(\frac{n}{k}\right)}\right) = s^{-\gamma-1}W_n^*(s) + s^{-\gamma-1/2-\varepsilon}o_P(1),$$

holds uniformly for all $s \in [1/(k+1), 1]$.

Remark 2.1 *The Brownian motions* $\{W_n\}$ *stem from the randomness of the original sample whereas the Brownian motions* $\{W_n^*\}$ *stem from the randomness of the bootstrap procedure. The latter is independent of the original sample. Consequently, Corollary 2.2 implies a consistency result for the tail quantile process if* $\sqrt{k}A(n/k) \to 0$ *as* $n \to \infty$.

To prove the theorem, we need three auxiliary results. First, recall Proposition 1.1, which gives the expansion of the tail quantile process of the original observations $\left\{\tilde{X}_j\right\}_{j=1}^n$. Second, Lemma 1.3 relates the tail quantile process of the bootstrapped observations to the tail quantile process of the original observations in. Finally, Lemma 2.3 below guarantees that we can use the asymptotic expansion of the latter process to obtain the expansion of the former process.

We introduce the notation

$$A_n(s) \overset{P}{\asymp} B_n(s) \text{ uniformly for } s \in S, \text{ as } n \to \infty, \tag{2.2}$$

to indicate that both $A_n(s)/B_n(s) = O_P(1)$ and $B_n(s)/A_n(s) = O_P(1)$ hold, while the two $O_P(1)$ terms are uniformly bounded for all $s \in S$.

Lemma 2.3 *Assume that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. Then, as $n \to \infty$, uniformly for all $s \in [1/(k+1), 1]$,*

$$\tilde{D}_n(s) \overset{P}{\asymp} s \text{ and } \Pr(\tilde{D}_n(s) < 2) \to 1$$

The first step in the proof of Theorem 2.1 is the substitution of x in (1.2) with $\tilde{D}_n(s)$, which is made available by Lemma 2.3. We obtain that, under the conditions of the Theorem, as $n \to \infty$,

$$\sqrt{k}\left(\frac{\tilde{X}^*_{n-[ks],n} - b_0\left(\frac{n}{k}\right)}{a_0\left(\frac{n}{k}\right)} - \frac{(\tilde{D}_n(s))^{-\gamma} - 1}{\gamma}\right)$$
$$= (\tilde{D}_n(s))^{-\gamma-1} W_n\left(\tilde{D}_n(s)\right) + \sqrt{k} A_0\left(\frac{n}{k}\right) \Psi_{\gamma,\rho}\left((\tilde{D}_n(s))^{-1}\right) + (\tilde{D}_n(s))^{-\gamma-1/2-\varepsilon} o_P(1), \tag{2.3}$$

uniformly for $0 < s \leq 1$. For the proof of Theorem 2.1 it remains to expand the four terms involving $\tilde{D}_n(s)$.

The term $\frac{(\tilde{D}_n(s))^{-\gamma}-1}{\gamma}$ on the left hand side of (2.3) can be handled by taking $\xi = -\gamma$ in the following lemma.

Lemma 2.4 *Assume that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. There exists a sequence of Brownian motions $W_1^*, W_2^*, \ldots$ such that for any $\xi \in \mathbb{R}$, uniformly for all $s \in [1/(k+1), 1]$,* *as* $n \to \infty$,

$$\sqrt{k}\left(\frac{\left(\tilde{D}_n(s)\right)^{\xi} - s^{\xi}}{\xi}\right) = -s^{\xi-1}W_n^*(s) + s^{\xi-1/2-\varepsilon}o_P(1) = s^{\xi-1/2-\varepsilon}O_P(1). \tag{2.4}$$

The three terms on the right hand side of (2.3) are handled by the following lemmas.

Lemma 2.5 *Under the conditions in Theorem 2.1, uniformly for all* $s \in [1/(k+1), 1]$*, as* $n \to \infty$,

$$\left(\tilde{D}_n(s)\right)^{-\gamma-1} W_n\left(\tilde{D}_n(s)\right) = s^{-\gamma-1}W_n(s) + s^{-\gamma-1/2-\varepsilon}o_p(1).$$

Lemma 2.6 *Under the conditions in Theorem 2.1, uniformly for all* $s \in [1/(k+1), 1]$*, as* $n \to \infty$,

$$\sqrt{k}A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(\frac{1}{\tilde{D}_n(s)}\right) = \sqrt{k}A_0\left(\frac{n}{k}\right)\Psi_{\gamma,\rho}\left(\frac{1}{s}\right) + s^{-\gamma-1/2-\varepsilon}o_p(1).$$

Lemma 2.7 *Under the conditions in Theorem 2.1,, niformly for all* $s \in [1/(k+1), 1]$*, as* $n \to \infty$,

$$(\tilde{D}_n(s))^{-\gamma-1/2-\varepsilon}o_P(1) = s^{-\gamma-1/2-\varepsilon}o_p(1).$$

Then the theorem follows. Proof of all lemmas are left to Appendix A.1.

3 The block maxima method

In this section we present a bootstrap analogue of (1.3).

Theorem 3.1 *Suppose the second order condition* (1.1) *holds. Let* $n = mk$ *with* $k = k(n)$ *satisfying* $k, m \to \infty$, $k = O(n^l)$ *for some* $0 < l < 2/3$ *and* $k^{1/2}\exp\{(\log\log k)^2\}A(m) = O(1)$*, as* $n \to \infty$*. Then for all* $0 < \lambda < 1/2$*, there exist an appropriate version of the functions* a *and* A*, denoted as* a_0 *and* A_0*, a sequence of Brownian bridges* $B_1^*, B_2^*, \ldots$ *and,* *independently, a sequence of Brownian motions $W_1, W_2, \ldots$ such that as $n \to \infty$, uniformly for all $s \in [1/(k+1), k/(k+1)]$,*

$$\sqrt{k}\left(\frac{X^*_{\lceil ks\rceil,k} - \tilde{b}_0(m)}{a_0(m)} - \frac{(-\log s)^{-\gamma} - 1}{\gamma}\right) = \frac{B^*_k(s)}{s(-\log s)^{1+\gamma}} + \frac{W_n(-\log s)}{(-\log s)^{1+\gamma}} + s^{-1/2-\lambda}\frac{(1-s)^{1/2-\lambda}}{(-\log s)^{1+\gamma}} o_p(1), \qquad (3.1)$$

where

$$\tilde{b}_0(m) := \begin{cases} U(m) & \text{if } \gamma \geq -1/2, \\ \tilde{X}_{n,n} + \frac{a_0(m)}{\gamma} & \text{if } \gamma < -1/2, \end{cases}$$

Remark 3.1 *The Brownian motions $\{W_n\}$ stem from the randomness of the original sample whereas the Brownian bridges $\{B^*_k\}$ stem from the randomness of the bootstrap procedure.*

Remark 3.2 *Compared to the expansion for the quantile process based on the original block maxima in* (1.3)*, the shift function $\tilde{b}_0(m)$ in Theorem 3.1 differs in two aspects. For $\gamma \geq -1/2$, the shift depends on the function U, not the function V. For $\gamma < -1/2$, the shift function is random.*

Remark 3.3 *Notice that the second order condition* (1.1) *differs from the original second order conditions assumed for the block maxima approach as in Ferreira and de Haan (2015): the condition* (1.1) *is more suitable for the POT method. Nevertheless, the two types of second order conditions imply each other if $-1 < \rho \leq 0$, see Drees et al. (2003). Therefore, for $-1 < \rho \leq 0$, the conditions required are comparable. If $\rho < -1$, the second order condition* (1.1) *implies the alternative second order condition in Ferreira and de Haan (2015), with an alternative second order scale function A^* that is regularly varying with index -1. In this case the requirement that as $n \to \infty$, $k = O(n^l)$ for $0 < l < 2/3$ is only slightly more restrictive than $\sqrt{k}A^*(n/k) = O(1)$, required in the original block maxima result in Ferreira and de Haan (2015).*

The general idea behind the proof of the main theorem is similar to the proof of Theorem

2.1, but the steps taken are more complicated. First, we establish an extended version for the asymptotic expansion of the tail quantile process of the original observations (see Proposition 3.2). Second, we relate the quantile process of the bootstrapped block maxima $\{X_i^*\}_{i=1}^k$ to the tail quantile process of the original observations $\left\{\tilde{X}_j\right\}_{j=1}^n$ (see Lemma 3.3 and Lemma 3.4 below). Finally, we can use the asymptotic expansion of the latter process to obtain the expansion of the former process. We first present the three auxiliary results, and then show the steps toward proving the main theorem at the end of this Section.

Proposition 3.2 *Assume that the second order condition* (1.1) *holds. Assume that an intermediate sequence $k := k(n)$ satisfies that $k \to \infty$, $k^{1/2} \exp\{(\log\log k)^2\} A(m) = O(1)$ and $k = O(n/(\log n)^3)$ as $n \to \infty$. Then there exist a version of the functions a and A, denoted as a_0 and A_0, and a sequence of Brownian motions W_n such that for all $0 < \varepsilon < 1/2$, as $n \to \infty$,*

$$\sqrt{k}\left(\frac{\tilde{X}_{n-[kx],n} - \tilde{b}_0(m)}{a_0(m)} - \frac{x^{-\gamma}-1}{\gamma}\right) = x^{-\gamma-1}W_n(x) + x^{-\gamma-1/2-\varepsilon}o_P(1), \tag{3.2}$$

holds uniformly for $x \in (0, (\log k)^2]$, where $\tilde{b}_0(m)$ is defined as in Theorem 3.1.

Notice that this proposition is a generalization of Proposition 1.1 in the sense that the range of x is extended. The extension is useful in the proof of Theorem 3.1.

Recall that F_n is the empirical distribution function of $\left\{\tilde{X}_j\right\}_{j=1}^n$ and $\tilde{X}_{1,n} \le \tilde{X}_{2,n} \le \ldots \le \tilde{X}_{n,n}$ are the order statistics of the original observations $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$. Denote $\Phi(x) := \exp(-1/x)$ for $x > 0$, the distribution function of the standard Fréchet distribution. We have the following representation result.

Lemma 3.3 *Let $Z_1^*, \cdots, Z_k^*$ be i.i.d. random variables with distribution function Φ, inde-*

pendent of $\left\{\tilde{X}_j\right\}_{j=1}^n$. *Then*

$$\{X_i^*\}_{i=1}^k \overset{d}{=} \{F_n^{\leftarrow}(\Phi(mZ_i^*)\}_{i=1}^k\,, \tag{3.3}$$

$$\left\{X_{\lceil ks\rceil,k}^*\right\}_{s\in[\frac{1}{k+1},\frac{k}{k+1}]} \overset{d}{=} \left\{\tilde{X}_{n-[kD_n(s)],n}\right\}_{s\in[\frac{1}{k+1},\frac{k}{k+1}]}\,, \tag{3.4}$$

where $D_n(s) = m\left(1-\Phi\left(mZ_{\lceil ks\rceil,k}^*\right)\right)$.

The representation in (3.4) suggests that the expansion for the process $\left\{X_{\lceil ks\rceil,k}^*\right\}_{s\in[1/(k+1),k/(k+1)]}$ can be obtained by substituting x in (3.2) with $D_n(s)$ for $s \in [1/(k+1), k/(k+1)]$. The following lemma guarantees that such a substitution is allowed. Recall the notation $\overset{P}{\asymp}$ defined in (2.2).

Lemma 3.4 *Assume that* $m/\log k \to \infty$ *as* $n \to \infty$*. Then, as* $n \to \infty$*, uniformly for all* $s \in [1/(k+1), k/(k+1)]$,

$$D_n(s) \overset{P}{\asymp} \frac{1}{Z_{\lceil ks\rceil,k}^*} \overset{P}{\asymp} -\log s.$$

For the proof of Theorem 3.1 we need some auxiliary results. Proposition 3.2, Lemma 3.3 and Lemma 3.4 imply that under the conditions of Theorem 3.1, as $n\to\infty$,

$$\sqrt{k}\left(\frac{X_{\lceil ks\rceil,k}^* - \tilde{b}_0(m)}{a_0(m)} - \frac{(D_n(s))^{-\gamma}-1}{\gamma}\right)$$
$$= (D_n(s))^{-\gamma-1}W_n(D_n(s)) + (D_n(s))^{-\gamma-1/2-\varepsilon}o_P(1), \tag{3.5}$$

uniformly for all $s \in [1/(k+1), k/(k+1)]$. Next, we approximate the various terms in (3.5) under the same conditions as in Theorem 3.1.

The term $\frac{(D_n(s))^{-\gamma}-1}{\gamma}$ on the left hand side of (3.5) can be handled by taking $\xi = -\gamma$ in the following lemma.

Lemma 3.5 *Under the conditions in Theorem 3.1, for any* $0 < \lambda < 1/2$*, there exists a sequence of Brownian bridges* $B_1^*, B_2^*, \ldots$ *such that as* $k \to \infty$*, uniformly for all* $s \in [1/(k+$

1), k/(k+1)],

$$\sqrt{k}\left(\frac{(D_n(s))^{\xi}-1}{\xi}-\frac{(-\log s)^{\xi}-1}{\xi}\right)$$
$$=-\frac{B_k^*(s)}{s(-\log s)^{1-\xi}}+s^{-1/2-\lambda}\frac{(1-s)^{1/2-\lambda}}{(-\log s)^{1-\xi}}o_P(1), \tag{3.6}$$

for any $\xi\in\mathbb{R}$.

Consequently, for any $\tau>0$*, as* $n\to\infty$*, uniformly for all* $s\in[1/(k+1),k/(k+1)]$,

$$\frac{(D_n(s))^{\xi}-(-\log s)^{\xi}}{\xi}=k^{-1/2}s^{-1/2}(1-s)^{1/2}(-\log s)^{-1+\xi}(\log k)^{\tau}O_P(1). \tag{3.7}$$

Next, the two terms on the right hand side of (3.5) are handled by the following lemmas.

Lemma 3.6 *Under the conditions in Theorem 3.1,as* $k\to\infty$,

$$(D_n(s))^{-\gamma-1}W_n(D_n(s))=(-\log s)^{-\gamma-1}W_n(-\log s)+s^{-1/2-\lambda}\frac{(1-s)^{1/2-\lambda}}{(-\log s)^{1+\gamma}}o_P(1),$$

holds uniformly for all $s\in[1/(k+1),k/(k+1)]$.

Lemma 3.7 *Choose* ε *such that* $\varepsilon<\lambda$*. Then, under the conditions in Theorem 3.1, uniformly for all* $s\in[1/(k+1),k/(k+1)]$*, as* $k\to\infty$,

$$(D_n(s))^{-\gamma-1/2-\varepsilon}o_P(1)=s^{-1/2-\lambda}\frac{(1-s)^{1/2-\lambda}}{(-\log s)^{1+\gamma}}o_P(1).$$

Then the theorem follows. Proofs of Proposition 3.2 and all lemmas are left to Appendix A.2.

4 Applications

In this section, we apply our main results, Theorem 2.1 for the POT method and Theorem 3.1 for the block maxima method to obtain the asymptotic behavior of the bootstrapped version of a few estimators in extreme value theory. As in the case of the original estimator, the asymptotic behavior of the bootstrapped estimator follows directly from the expansion of the bootstrapped (tail) quantile process (Theorem 2.1 and 3.1). For the POT method, we use the PWM estimator (Hosking and Wallis (1987)) as an example and show that the bootstrap is *consistent* for the PWM estimator. Here consistency refers to the consistency of the bootstrap defined in Bickel and Freedman (1981). For the block maxima method, we use the PWM estimator in Hosking et al. (1985) as an example. We show that the bootstrap is *not consistent*. Nevertheless, the sample variance of bootstrapped estimates may still be a good estimator for the asymptotic variance of the original PWM estimator. The POT and block maxima methods are handled in two separate subsections.

4.1 The POT method

We start with the PWM estimator for the extreme value index using the POT method; see Hosking and Wallis (1987). Let $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ be a sequence of i.i.d. random variables with common distribution function F satisfying the second order condition (1.1). Denote $\tilde{X}_{1,n} \leq \cdots \leq \tilde{X}_{n,n}$ as the order statistics from a sample of n observations. Define the PWM estimator in the POT method as

$$\hat{\gamma}_{POT} := \frac{I_1 - 4I_2}{I_1 - 2I_2},$$

where the probability weighted moments I_q are given by

$$I_q = \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k}\right)^{q-1} (\tilde{X}_{n-i+1,n} - \tilde{X}_{n-k,n}),$$

for $q = 1, 2$. Here $k := k(n)$ is an intermediate sequence such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$.

The asymptotic behavior of the PWM estimator using the POT method is as follows; see e.g. equation (3.4) in Cai et al. (2013) with $q = 2$, $r = 1$.

Proposition 4.1 *Assume that $\gamma < 1/2$ and $\sqrt{k}A(n/k) = O(1)$ as $n \to \infty$. With the same Brownian motions W_n defined in Proposition 1.1, we have that as $n \to \infty$,*

$$\sqrt{k}(\hat{\gamma}_{POT} - \gamma) = \tilde{L}(W_n) + \tilde{b}_n(\gamma, \rho) + o_P(1),$$

where the random part is

$$\tilde{L}(W_n) = (2-\gamma)(1-\gamma)\left(2(2-\gamma)\int_0^1 s(s^{-\gamma-1}W_n(s) - W_n(1))ds \right.$$
$$\left. -(1-\gamma)\int_0^1 (s^{-\gamma-1}W_n(s) - W_n(1))ds\right),$$

and the asymptotic bias term is $\tilde{b}_n(\gamma, \rho) = \sqrt{k}A_0(n/k)\frac{(2-\gamma)(1-\gamma)}{(2-\gamma-\rho)(1-\gamma-\rho)}$, where A_0 is the same as in Proposition 1.1.

The proof of this proposition is based on the asymptotic expansion of the tail quantile process in Proposition 1.1. From this proposition, one may calculate the asymptotic variance of the PWM estimator as follows,

$$\mathrm{Var}(\tilde{L}(W_n)) = \frac{(1-\gamma)(2-\gamma)^2(1-\gamma+2\gamma^2)}{(1-2\gamma)(3-2\gamma)},$$

see Theorem 3.6.1 in de Haan and Ferreira (2006). Practically, one may use the estimated γ to obtain a consistent estimate of the asymptotic variance. Therefore, using the bootstrap is not a necessary step for obtaining the variance of the estimator. Nevertheless, we show that using bootstrap can achieve the consistency of the estimator.

We consider the bootstrap sample $\tilde{X}_1^*, \tilde{X}_2^*, \ldots, \tilde{X}_n^*$ and construct the PWM estimator based on the bootstrap sample, denoted as $\hat{\gamma}^*_{POT}$. Notice that the expansion in Theorem 2.1 has a structure very similar to the one in Proposition 1.1, except an extra term. Similar to the corresponding proof in Cai et al. (2013), we obtain the following result.

Proposition 4.2 *Assume that $\gamma < 1/2$ and $\sqrt{k}A(n/k) = O(1)$ as $n \to \infty$. With the same Brownian motions W_n and W_n^* defined in Theorem 2.1, we have that as $n \to \infty$,*

$$\sqrt{k}(\hat{\gamma}^*_{POT} - \gamma) = \tilde{L}(W_n + W_n^*) + \tilde{b}_n(\gamma, \rho) + o_P(1),$$

where $\tilde{L}(\cdot)$ is the same operator as in Proposition 4.1 and $\tilde{b}_n(\gamma, \rho)$ is the same bias term therein. Here the Brownian motions W_n are the same sequence as that in Proposition 4.1, and they are independent of the Brownian motions W_n^*.

Note that the operator $\tilde{L}(\cdot)$ is a linear operator, i.e. $\tilde{L}(W_n + W_n^*) = \tilde{L}(W_n) + \tilde{L}(W_n^*)$. Therefore, by comparing $\hat{\gamma}^*_{POT}$ and $\hat{\gamma}_{POT}$, we obtain that as $n \to \infty$,

$$\sqrt{k}(\hat{\gamma}^*_{POT} - \hat{\gamma}_{POT}) = \tilde{L}(W_n^*) + o_P(1). \tag{4.1}$$

The relation (4.1) serves as an important step in proving the consistency of the bootstrap procedure as in the following Theorem. The proof is postponed to Appendix A.3.

Theorem 4.3 *Assume that $\gamma < 1/2$ and $\sqrt{k}A(n/k) = o(1)$ as $n \to \infty$. The PWM estimator using the POT method is consistent: as $n \to \infty$,*

$$\sup_{x \in \mathbb{R}} \left| \Pr\left(\sqrt{k}(\hat{\gamma}^*_{POT} - \hat{\gamma}_{POT}) \le x \mid \tilde{X}_1, \ldots, \tilde{X}_n\right) - \Pr\left(\sqrt{k}(\hat{\gamma}_{POT} - \gamma) \le x\right) \right| \to 0,$$

where $\tilde{L}(\cdot)$ is the same operator as in Proposition 4.1 and $\tilde{b}_n(\gamma, \rho)$ is the same bias term therein. Here the Brownian motion W_n is the same sequence as that in Proposition 4.1, and is independent of the Brownian motion W_n^*.

Theorem 4.3 motivates the bootstrap procedure for obtaining the asymptotic variance of $\hat{\gamma}_{POT}$: by bootstrapping d times and obtain estimators $\hat{\gamma}^*_l$, $l = 1, 2, \ldots, d$, the sample variance of these bootstrapped estimates

$$s_{boot}^2 := \frac{1}{d-1} \sum_{l=1}^{d} \left(\hat{\gamma}_l^* - \frac{1}{d} \sum_{l=1}^{d} \hat{\gamma}_l^* \right)^2 ,$$

approximates the asymptotic variance of the original estimator $\hat{\gamma}_{POT}$.

We remark that although Theorem 4.3 requires the condition $\sqrt{k}A(n/k) = o(1)$ as $n \to \infty$ which assumes away the bias in the original estimator. This is only necessary for obtaining the consistency result. Nevertheless the aforementioned procedure for obtaining the asymptotic variance is also valid if the bias is present: i.e. $\sqrt{k}A(n/k) = O(1)$ as $n \to \infty$. In this case, the equation (4.1) is still valid. Therefore, the sample variance of $\{\hat{\gamma}_l^*\}$, is still a good approximate of $\mathrm{Var}(\tilde{L}(W_n))/k$.

An alternative way to approximate the asymptotic variance involves the original estimator $\hat{\gamma}_{POT}$. Theorem 4.3 implies that the following statistic

$$s_{alternative}^2 = \frac{1}{d} \sum_{l=1}^{d} \left(\hat{\gamma}_l^* - \hat{\gamma}_{POT} \right)^2 ,$$

is also an approximation of the asymptotic variance of the original estimator.

For any other estimator for the extreme value index using the POT method, as long as its asymptotic behavior can be established using the tail quantile process in Proposition 1.1 and the asymptotic limit of the estimator depends on a linear operator of the limit of the tail quantile process, then similar results as in Proposition 4.5 and Theorem 4.3 can be established. In other words, the bootstrap procedure is *consistent* if the asymptotic bias is zero. Further, the sample variance of the bootstrapped estimates can approximate the variance of the original estimator. Examples of such estimators are the Pickands' estimator (Pickands (1975)), the maximum likelihood estimator (Smith (1987)) and the negative Hill estimator (Falk (1995)).

4.2 The block maxima method

Analogously to the POT method, we investigate the bootstrapped PWM estimator for the extreme value index using the block maxima method, which differs from the PWM estimator using the POT method. The estimator was introduced in Hosking et al. (1985) with its asymptotic normality proved in Ferreira and de Haan (2015).

Let $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ be a sequence of i.i.d. random variables with common distribution function F satisfying the second order condition (1.1). Define the block maxima of the sample as $X_i = \max_{(i-1)m<j\leq im} \tilde{X}_j$, for $i = 1, 2, \ldots, k$, where $n = mk$. Let $X_{1,k} \leq X_{2,k} \leq \ldots \leq X_{k,k}$ be the order statistics of the block maxima $X_1, X_2, \ldots, X_k$. The PWM estimator using the block maxima method, denoted as $\hat{\gamma}_{BM}$, is the solution of the equation

$$\frac{3^{\hat{\gamma}_{BM}} - 1}{2^{\hat{\gamma}_{BM}} - 1} = \frac{3\beta_2 - \beta_0}{2\beta_1 - \beta_0},$$

where the probability weighted moments β_q are given by

$$\beta_q = \frac{1}{k} \sum_{i=1}^{k} \frac{(i-1)\cdots(i-q)}{(k-1)\cdots(k-q)} X_{i,k},$$

for $q = 0, 1, 2$. Here $k := k(n)$ is an intermediate sequence such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$.

Theorem 2.3 in Ferreira and de Haan (2015) shows the asymptotic behavior of the PWM estimator using the block maxima method. We cite this result in a simpler form where the sequence $k = k(n)$ is chosen such that no asymptotic bias appears.

Proposition 4.4 (A simpler version of Theorem 2.3 in Ferreira and de Haan (2015)) *Assume the conditions in Theorem 3.1. Further assume $\gamma < 1/2$. There exists a series of standard Brownian bridge $B_1, B_2, \cdots$ such that as $n \to \infty$,*

$$\sqrt{k}(\hat{\gamma}_{BM} - \gamma) = L\left(\frac{B_k(s)}{s(-\log s)^{1+\gamma}}\right) + o_P(1),$$

where $L(\cdot)$ is a linear operator on the sample path of a stochastic process.

Here we omit the detailed formula for L because we only need its linearity property. Notice that the operator L, though being explicit, is very complicated. This makes the asymptotic variance of the PWM estimator using the block maxima method almost intractable. This is a real situation where the bootstrap can help.

We consider the bootstrap sample $\tilde{X}_1^*, \tilde{X}_2^*, \ldots, \tilde{X}_n^*$, reconstruct the block maxima and the PWM estimator based on the bootstrap sample, denoted as $\hat{\gamma}_{BM}^*$. Based on the asymptotic expansion of the bootstrapped tail quantile process in Theorem 3.1, we get the following result.

Proposition 4.5 *Assume the conditions in Theorem 3.1. Further assume $\gamma < 1/2$. With the same Brownian motions W_n and Brownian Bridges B_n^* defined therein, we have that as $n \to \infty$,*

$$\sqrt{k}(\hat{\gamma}_{BM}^* - \gamma) = L\left(\frac{B_k^*(s)}{s(-\log s)^{1+\gamma}} + \frac{W_n(-\log s)}{(-\log s)^{1+\gamma}}\right) + o_P(1),$$

where $L(\cdot)$ is the same operator as in Proposition 4.4. Here the Brownian bridge B_k^* is independent of the Brownian bridge B_k.

We note that a consistency result analog to Theorem 4.3 cannot be established for the bootstrapped PWM estimator. In other words, the bootstrapped PWM estimator using the block maxima method is not *consistent*. Since $L(\cdot)$ is a linear operator, by comparing the expansion of $\hat{\gamma}_{BM}^*$ with $\hat{\gamma}_{BM}$, we get that as $n \to \infty$,

$$\sqrt{k}(\hat{\gamma}_{BM}^* - \hat{\gamma}_M) = L\left(\frac{B_k^*(s)}{s(-\log s)^{1+\gamma}}\right) + L\left(\frac{W_n(-\log s)}{(-\log s)^{1+\gamma}}\right) - L\left(\frac{B_k(s)}{s(-\log s)^{1+\gamma}}\right) + o_P(1).$$

Different from (4.1), the additional term $L\left(\frac{W_n(-\log s)}{(-\log s)^{1+\gamma}}, \gamma\right) - L\left(\frac{B_k(s)}{s(-\log s)^{1+\gamma}}\right)$ is a common bias in the limit distribution of $\sqrt{k}(\hat{\gamma}_{BM}^* - \hat{\gamma}_M)$ when considering different bootstrapped estimators. Therefore, the conditional distribution of the bootstrapped estimator does not converge to the same distribution as the original estimator.

Nevertheless, since the bias term is common for all bootstrapped estimates, we can use the bootstrap procedure to approximate the asymptotic variance by considering the sample variance of the bootstrapped estimates. Repeat the bootstrap procedure for d times and get obtain estimators $\hat{\gamma}_l^*$, $l = 1, 2, \ldots, d$. When considering $\hat{\gamma}_l^* - \frac{1}{d}\sum_{l=1}^{d} \hat{\gamma}_l^*$, the common term $L\left(\frac{W_n(-\log s)}{(-\log s)^{1+\gamma}}, \gamma\right) - L\left(\frac{B_k(s)}{s(-\log s)^{1+\gamma}}\right)$ cancels out. Consequently, the sample variance of $\{\hat{\gamma}_l^*\}$,

$$s_{boot}^2 := \frac{1}{d-1}\sum_{l=1}^{d}\left(\hat{\gamma}_l^* - \frac{1}{d}\sum_{l=1}^{d}\hat{\gamma}_l^*\right)^2,$$

can be used as an approximation of the variance of $\frac{1}{\sqrt{k}} L\left(\frac{B_k^*(s)}{s(-\log s)^{1+\gamma}}, \gamma\right)$, which equals to the asymptotic variance of the original estimator $\hat{\gamma}_{BM}$.

Compared to the POT method, when using the BM method, we cannot use the alternative approach involving the original estimator $\hat{\gamma}_{BM}$ to approximate the asymptotic variance, i.e. the variation of the bootstrapped estimates around the original estimator cannot be used as an estimator for the variance of the original estimator. Intuitively, since the block maxima estimator depends on the order of the observations, the original estimator also possesses randomness due to the random order.

Similar to the POT method, for any estimator of the extreme value index using the block maxima method, if its asymptotic property can be established using the quantile process in Proposition 3.2 such that the asymptotic limit of the estimator depends on a linear operator of the limit of the quantile process, then the sample variance of bootstrapped estimates can be a good estimator for the asymptotic variance of the original estimator. An example of such an estimator is the maximum likelihood estimator under the block maxima as analyzed in Dombry and Ferreira (2019).

5 Simulations

In this section, we perform various simulation studies to illustrate the theoretical results obtained in Section 4. Throughout the simulation study, we consider the PWM estimators. Unless otherwise specified, we consider observations drawn from the Pareto distribution $F(x) = 1 - x^{-1/\gamma}$, with an extreme value index $\gamma > 0$.

Firstly, we compare the consistency result when using the POT method (Theorem 4.3) with the inconsistency result when using the block maxima method. To validate the asymptotic results, we simulate samples with a very large sample size $n = 10000$ and a true extreme value index $\gamma = 0.2$. We choose $k = 1000$ in the estimation. For each method, we first simulate $m = 100$ samples and estimate γ using the PWM estimator for each sample, as $\hat{\gamma}_i$, $1 \leq i \leq m$, and then plot the histogram of the estimation errors $\hat{\gamma}_i - \gamma$. Next, we simulate a single sample and also estimate γ by the PWM estimator as $\hat{\gamma}$. Based on the single sample, we bootstrap $b = 100$ times to obtain b bootstrap estimates $\hat{\gamma}_i^*$, $1 \leq i \leq b$. We plot the histogram of the bootstrap errors $\hat{\gamma}_i^* - \hat{\gamma}$ for this sample. In Figure 1, panels (a) and (b) show the two histograms when using the POT and the block maxima methods respectively.

We observe that for the POT method, the two histograms are comparable. In particular, the range of the errors are similar. By contrast, for the block maxima method, the two histograms differ: in terms of the range, the bootstrapped errors are shifted towards the left for this specific simulation. In addition, we perform the Kolmogorov-Smirnov (KS) test to test whether the two types of errors share the same distribution. For the POT method, we obtain a KS statistic at 0.11 with p-value 0.581. For the block maxima method, we obtain a KS statistic at 0.47, with p-value virtually zero. Hence, we conclude that the bootstrap errors have a different distribution than the estimation errors only when using the block maxima method.

We further explore the tests between the distributions of the two types of errors, when using different k values. For $k = 500, 510, \ldots, 2000$, we repeat the same procedure as above and plot the obtained p-values in Figure 2.

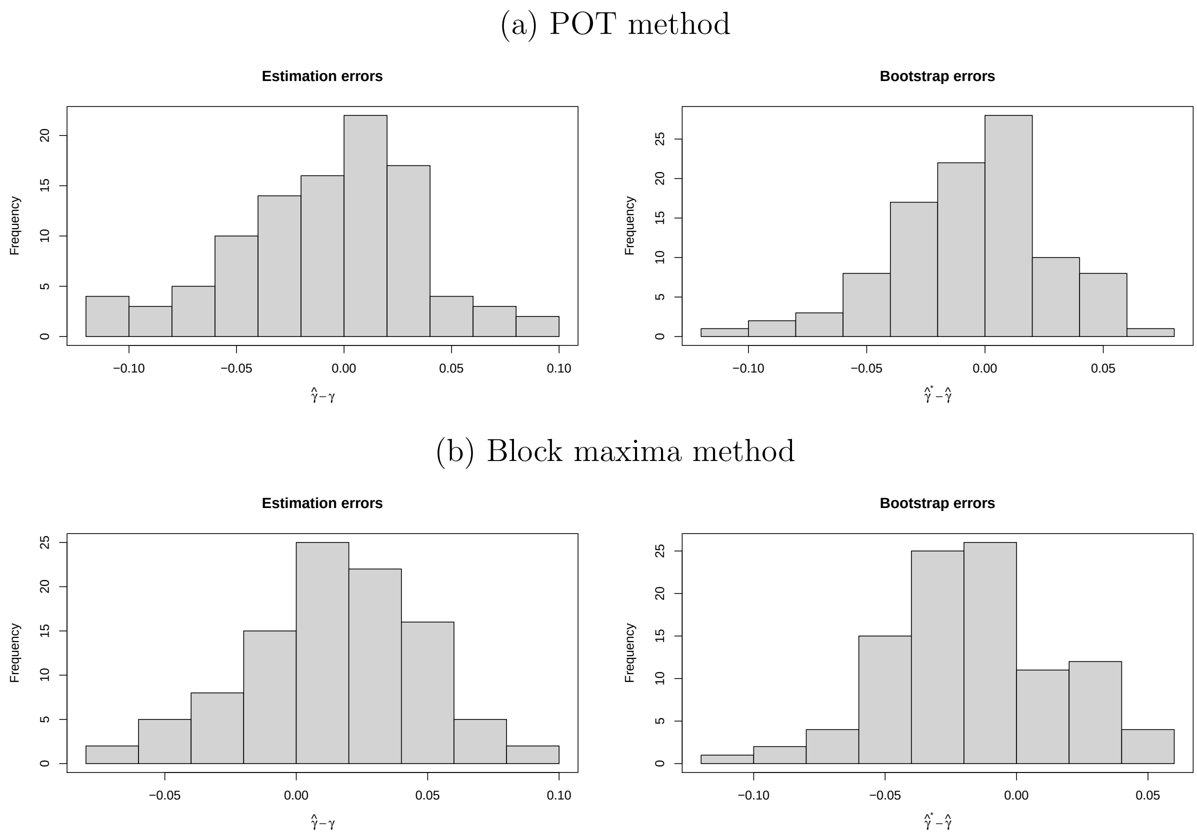


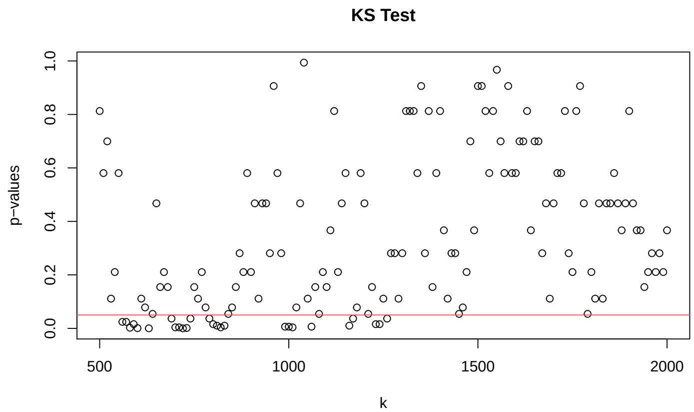
Figure 1: Histograms for estimation errors and bootstrap errors

For the POT method, we observe that for most levels of k, we do not reject that the two types of errors have a different distribution. However, for the block maxima method, for all k above 800, the null hypothesis of having the same distribution is rejected under 5% significance level. This result agrees with the theoretical results regarding consistency.

Finally, we show that the bootstrap estimate for the standard deviation of the PWM estimators are applicable when using either the POT or the block maxima method. To illustrate the usefulness of bootstrap, we consider typical sample size used in application $n = 2000$. In addition, we choose $k = 100, 105, \ldots, 500$.

Although the asymptotic standard deviations for both PWM estimators, using either the POT or the block maxima method, are explicitly given in their asymptotic theories, the calculation for that using the block maxima method is rather cumbersome. To avoid calculating the true standard deviation of an estimator, we conduct a pre-simulation using $m = 1000$

(a) POT method



(b) Block maxima method

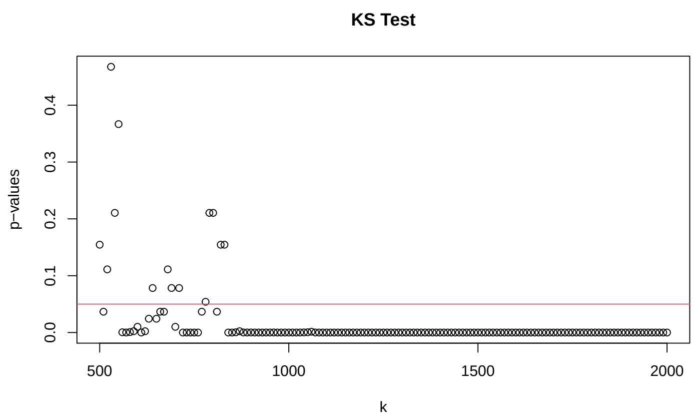


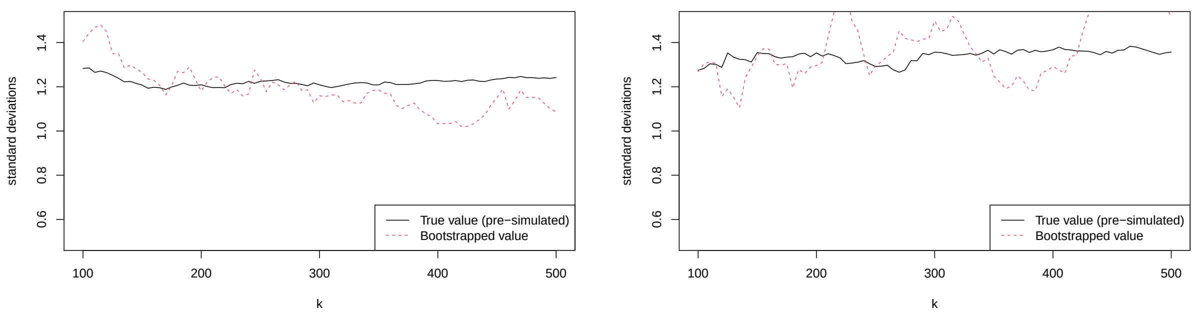
Figure 2: P-values for testing the same distributions between estimation and bootstrap errors

samples. The true standard deviation is then approximated by the sample standard deviation of across the m estimates. We present its value multiplied with $\sqrt{k}$ as indicated by the black solid lines. Since the asymptotic theory shows that the speed of convergence is $1/\sqrt{k}$ for both PWM estimators, the scaled true standard deviation is expected to remain at a horizontal level. Next, we simulate a single sample, and plot the bootstrap standard deviation for an estimator based on $b = 100$ bootstrapped samples, also scaled by $\sqrt{k}$, using the red dash lines.

Figure 3 shows the results for the POT and block maxima methods in panels (a) anb (b) respectively. Besides the Pareto distribution (left), we also consider the standard normal distribution (right) as the underlying data generating process. Notice that the standard normal distribution corresponds to $\gamma = 0$, with a second order index $\rho = 0$ for both the U and V functions.

All four figures show that the bootstrapped standard deviation is around the true value obtained via pre-simulations. Notice that we employed only $b = 100$ bootstrapped sample to maintain a very fast computation for the bootstrapped standard deviation. The accuracy is expected to be further improved when considering higher b.

(a) POT method



(b) Block maxima method

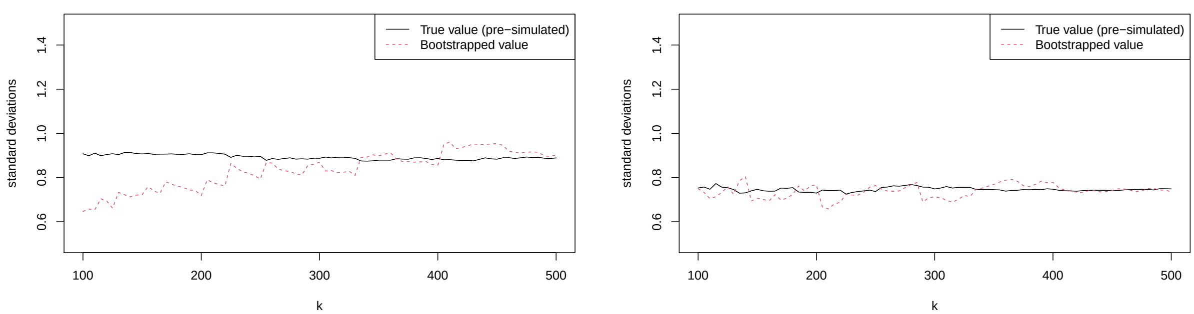


Figure 3: Bootstrapped standard deviations: the left (right) figure is based on the Pareto distribution with $\gamma = 0.2$ (the standard normal distribution with $\gamma = 0$)

Acknowledgement We thank the AE and two referees for their insightful comments. Laurens de Haan was financially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the projects UIDB/00006/2020 and PTDC/MAT-STA/28649/2017.

References

Bickel, P. J. and D. A. Freedman (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics 9*(6), 1196–1217.

Cai, J.-J., L. de Haan, and C. Zhou (2013). Bias correction in extreme value statistics with index around zero. *Extremes 16*(2), 173–201.

de Haan, L. and A. Ferreira (2006). *Extreme value theory: an introduction*. Springer.

Dombry, C. and Ferreira (2019). Maximum likelihood estimators based on the block maxima method. *Bernoulli 25*(3), 1690–1723.

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

Drees, H., L. de Haan, and D. Li (2003). On large deviation for extremes. *Statistics & probability letters 64*(1), 51–62.

Falk, M. (1995). Some best parameter estimates for distributions with finite endpoint. *Statistics: A Journal of Theoretical and Applied Statistics 27*(1-2), 115–125.

Ferreira, A. and L. de Haan (2015). On the block maxima method in extreme value theory: PWM estimators. *The Annals of Statistics 43*(1), 276–298.

Hosking, J. R. and J. R. Wallis (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics 29*(3), 339–349.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics 27*(3), 251–261.

Litvinova, S. and J. Mervyn (2018). Bootstrapping tail statistics: tail quantile process, hill estimator, and confidence intervals for high-quantiles of heavy tailed distributions. *Monash Business School Working Paper 12/18.*

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics 3*(1), 119–131.

Smith, R. (1987). Estimating tails of probability distributions. *Annals of Statistics 15*(3), 1174–1207.

Van der Vaart, A. W. (1998). *Asymptotic statistics.* Cambridge university press.