A Location-invariant Non-positive Moment-type Estimator of the Extreme Value Index

Chuandi Liu[1] and Chengxiu Ling[1,2*]

[1]School of Mathematics and Statistics, Southwest University, 400715 Chongqing, China

[2]Department of Actuarial Science, University of Lausanne, Chamberonne 1015, Lausanne, Switzerland

January 2, 2018

Abstract: This paper investigates a class of location invariant non-positive moment-type estimators of extreme value index, which is highly flexible due to the tuning parameter involved. Its asymptotic expansions and its optimal sample fraction in terms of minimal asymptotic mean square error are derived. A small scale Monte Carlo simulation turns out that the new estimators, with a suitable choice of the tuning parameter driven by the data itself, perform well compared to the known ones. Finally, the proposed estimators with a bootstrap optimal sample fraction are applied to an environmental data set.

Keywords: location-invariant moment-type estimation; extreme value index; bootstrap methodology; extreme value statistics.

AMS 2000 subject classification: Primary 60G70; Secondary 65C05.

1 Introduction

Let $X_{1,n} \le \cdots \le X_{n,n}$ be the order statistics (o.s.) associated with a random sample $\underline{X}_n := (X_1, \ldots, X_n)$ with underlying distribution function F. Suppose that F belongs to the max-domain of attraction of a non-degenerate distribution function G, denoted by $F \in D(G_\gamma)$. Then G must be of the type of generalized extreme value distribution (cf. de Haan and Ferreira (2006))

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

where $(1+\gamma x)^{-1/\gamma} := e^{-x}$ if $\gamma = 0$. The parameter γ, the so-called extreme value index (EVI), is the primary parameter of extreme events. It is well-known that $F \in D(G_\gamma)$, $\gamma \in \mathbb{R} \Longleftrightarrow U \in GRV_\gamma$, where $U(t) = F^{\leftarrow}(1-1/t) =$

*Corresponding author. Email: lcx98@swu.edu.cn. The authors would like to thank Zuoxiang Peng and Enkelejd Hashorva for useful discussions.

$\inf\{y \in \mathbb{R} : F(y) \geq 1 - 1/t\}$ is the tail quantile function of $X \sim F$, and GRV_α stands for the class of generalized regularly varying functions at infinity with an index $\alpha \in \mathbb{R}$, that is, positive measurable functions g such that $\lim_{t\to\infty}(g(tx)-g(t))/a(t) = (x^\alpha - 1)/\alpha$, $x > 0$ for some auxiliary function a. For an unknown distribution $F \in D(G_\gamma)$ with $\gamma > 0$, the well-known Hill's estimators are based on the log-excesses over an o.s. $X_{n-k_0,n}$, which are given by

$$\hat{\gamma}_n^H(k_0) = \hat{\gamma}_n^H(k_0; \underline{X}_n) := \frac{1}{k_0}\sum_{i=1}^{k_0}\left(\ln X_{n-i+1,n} - \ln X_{n-k_0,n}\right),$$

where $k_0 = k_0(n)$ is an intermediate integer sequence, i.e., $\lim_{n\to\infty} k_0 = \lim_{n\to\infty} n/k_0 = \infty$. For $\gamma \in \mathbb{R}$, Dekkers et al. (1989) proposed a class of moment estimators as follows:

$$\hat{\gamma}_n^M(k_0) = \hat{\gamma}_n^M(k_0; \underline{X}_n) = \hat{\gamma}_n^H(k_0) + \widehat{\gamma}_n^{NM}(k_0)$$

where, with $\widehat{\gamma}_n^{(j)}(k_0) = \widehat{\gamma}_n^{(j)}(k_0; \underline{X}_n) = (1/k_0)\sum_{i=1}^{k_0}\left(\ln X_{n-i+1,n} - \ln X_{n-k_0,n}\right)^j$, $j = 1, 2$

$$\widehat{\gamma}_n^{NM}(k_0) = \widehat{\gamma}_n^{NM}(k_0; \underline{X}_n) = 1 - \frac{1}{2}\left(1 - \frac{(\widehat{\gamma}_n^{(1)}(k_0))^2}{\widehat{\gamma}_n^{(2)}(k_0)}\right)^{-1},$$

the so-called negative moment estimators since it is a consistent estimator for $\gamma < 0$. Further, Caeiro and Gomes (2010) studied the following alternative moment-type estimator for $\gamma < 0$:

$$\hat{\gamma}_n^{NM(\theta)}(k_0) = \hat{\gamma}_n^{NM(\theta)}(k_0; \underline{X}_n) = \widehat{\gamma}_n^{NM}(k_0) + \theta\hat{\gamma}_n^H(k_0), \quad \theta \in \mathbb{R}. \tag{1.1}$$

This class of estimators is highly flexible due to the tuning parameter θ. Note that the classes of estimators mentioned above are scale invariant but not location invariant, a property enjoyed by the EVI itself. Therefore, it is sensible to use the *peaks over random threshold* (PORT) methodology, introduced first by Fraga Alves (2001), and further studied by Ling et al. (2007, 2012) respectively for the Hill, moment and Weiss-Hill estimations of location invariant type. Typically, these estimators are based on a sample of excesses over a random threshold $X_{n-k,n}, k_0 \ll k \ll n$, that is, it is based on

$$\underline{X}_n^k := (X_{n,n} - X_{n-k,n}, \ldots, X_{n-k+1,n} - X_{n-k,n}), \tag{1.2}$$

corresponding to the PORT sample $\underline{X}_{n,q} := (X_{n,n} - X_{n_q,n}, \ldots, X_{n_q+1,n} - X_{n_q,n}), n_q = [nq]$ with $q = q_n := [(n-k)/n] \to 1$ as $n \to \infty$. Other results on PORT EVI-estimation for $q \in [0, 1)$ and heavy tail distributions, i.e., $F \in D(G_\gamma)$, $\gamma > 0$, can be found in Caeiro et al. (2016) for Pareto probability weighted moment estimations, Gomes and Henriques-Rodrigues (2016) for the mean-of-order-p estimations, and among others.

This paper aims to investigate the PORT-EVI estimation for the unknown distributions $F \in D(G_\gamma)$ with $\gamma \leq 0$. Typically, we are interested with the following location and scale invariant estimators:

$$\widehat{\gamma}_n^{NM(\theta)}(k_0, k) = \widehat{\gamma}_n^{NM(\theta)}(k_0; \underline{X}_n^k) = \widehat{\gamma}_n^{NM}(k_0, k) + \theta\widehat{\gamma}_n^H(k_0, k), \quad \theta \in \mathbb{R}, \tag{1.3}$$

which have the same functional form of the generalized negative moment estimators in (1.1) but with the original sample $\underline{X}_n$ replaced everywhere by the sample of excesses $\underline{X}_n^k$ in (1.2).

The asymptotic expansions of the new proposed estimators in (1.3) are given in Theorem 2.2. We see that the tuning parameter θ affects partially its asymptotic biasness. This fact indicates that (1.3) with suitable choice of θ, may give a large variety of second-order asymptotically unbiased estimators of EVI, see e.g., Cai et al. (2012), Li and Peng (2009), Li et al. (2011), Gomes et al. (2013, 2016) for related discussions. Our second result, Theorem 2.4, gives the optimal sample fraction of (1.3) in the sense of minimal asymptotic mean square error.

Note that the asymptotic properties of the proposed PORT-EVI estimations depend on the unknown second-order parameter ρ, which restricts to some extent its application. Therefore, we carry out a small scale Monte Carlo simulation with the tuning parameter θ and the sample fraction k_0 chosen by a data-driven/bootstrap method (cf. Caeiro and Gomes (2010), Gomes et al. (2013)). We compare in Table 1 the finite sample behavior of the proposed estimator and the other location invariant ones for negative EVI, including the moment estimators in Ling et al. (2007), the Weiss-Hill estimators in Ling et al. (2012)), and the maximum likelihood and moment estimators in Hüsler et al. (2016). Finally, we give an application of our findings into a real-life data in environments.

We organize the paper as follows. In Section 2, we display main results. Section 3 is devoted to the application. The proofs are relegated to Section 4.

2 Main Results

Recall that $U(t) = \inf\{y \in \mathbb{R} : F(y) \geq 1 - 1/t\}$ the tail quantile function of $X \sim F$. In order to study the asymptotic distribution of the new PORT-EVI estimations given in (1.3), we need to strengthen the first-order condition $F \in D(G_\gamma)$ to be of second-order extended regular variation. Namely, suppose that there exist functions $a(\cdot)$, and $A(\cdot)$ with constant sign at infinity and $\lim_{t\to\infty} A(t) = 0$ such that (cf. de Haan and Ferreira (2006))

$$\lim_{t\to\infty} \frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right) =: H_{\gamma,\rho}(x), \quad x > 0,\ \gamma,\ \rho \leq 0. \tag{2.1}$$

The functions $a(\cdot)$ and $A(\cdot)$ are respectively referred to as the first-order and second-order auxiliary function of U. It is well-known that $A \in RV_\rho$, i.e., $\lim_{t\to\infty} A(tx)/A(t) = x^\rho$ for all $x > 0$.

The following lemma, similar to *B. 3.42* and *B. 3.43* in *Lemma B. 3.16* in de Haan and Ferreira (2006), p 398, is crucial to establish our main results.

Lemma 2.1. *Let $k = k(n), k_0 = k_0(k)$ be two intermediate sequences. If condition* (2.1) *is satisfied for $X \sim F$, then the following expansion holds locally uniformly for all $x > 0$*

$$\frac{\ln \frac{U((n/k_0)x)-U(n/k)}{U(n/k_0)-U(n/k)}}{\widetilde{a}(n/k, n/k_0)} = D_\gamma(x) + H_{\gamma,\rho}(x) A\left(\frac{n}{k_0}\right)(1+o(1)) - H_{\gamma,\gamma}(x)\widetilde{a}\left(\frac{n}{k}, \frac{n}{k_0}\right)(1+o(1)), \quad n\to\infty,$$

where $D_\gamma(x) = (x^\gamma - 1)/\gamma$ and $\widetilde{a}(n/k, n/k_0) = -\gamma(k/k_0)^\gamma$ for $\gamma < 0$, and $1/\ln(k/k_0)$ otherwise.

Hereafter, we write $\stackrel{d}{=}$ and $\stackrel{d}{\to}$ for equality in distribution and convergence in distribution, respectively. All the limits are taken as $n \to \infty$ unless otherwise stated. Further, denote with $Y \sim F_Y(y) = 1 - 1/y$, $y \geq 1$

$$\begin{cases} \mu_{j,\gamma} = \mathrm{E}\left\{(D_\gamma(Y))^j\right\}, \quad \nu_{j,\gamma,\rho} = \mathrm{E}\left\{j(D_\gamma(Y))^{j-1} H_{\gamma,\rho}(Y)\right\} \\ \widetilde{\nu}_{\gamma,\rho} = (1-\gamma)(1-2\gamma)\left(\frac{\nu_{2,\gamma,\rho}}{\mu_{2,\gamma}} - 2\frac{\nu_{1,\gamma,\rho}}{\mu_{1,\gamma}}\right), \quad \sigma_\gamma^2 = \frac{(1-\gamma)^2(1-2\gamma)(1-\gamma+6\gamma^2)}{(1-3\gamma)(1-4\gamma)}. \end{cases} \tag{2.2}$$

Theorem 2.2. *Suppose that condition* (2.1) *is satisfied for $X \sim F$. We have*

$$\widehat{\gamma}_n^{NM(\theta)}(k_0, k) \stackrel{d}{=} \gamma + \frac{Z_{k_0}}{\sqrt{k_0}}(1 + o_p(1)) + \widetilde{\nu}_{\gamma,\rho} A\left(\frac{n}{k_0}\right)(1 + o_p(1))$$
$$+ \left(\frac{\theta}{1-\gamma} - \widetilde{\nu}_{\gamma,\gamma}\right) \begin{cases} |\gamma|(k/k_0)^\gamma (1 + o_p(1)), & \gamma < 0, \\ 1/\ln(k/k_0)(1 + o_p(1)), & \gamma = 0, \end{cases} \tag{2.3}$$

where $Z_{k_0} \stackrel{d}{\to} N(0, \sigma_\gamma^2)$ and $\sigma_\gamma^2, \widetilde{\nu}_{\gamma,\rho}$ are given by (2.2).

Remark 2.3. *We see that the tuning parameter θ plays an important role in adjusting the bias of the estimators. For instance, we might choose θ to be exactly $(1-\gamma)(\widetilde{\nu}_{\gamma,\gamma} - \vartheta\widetilde{\nu}_{\gamma,\rho}/\gamma)$ leading to asymptotically unbiased estimators provided that the dominated term of the bias is $(k/k_0)^\gamma$ for $\gamma < 0$, that is, $\lim_{n\to\infty} A(n/k_0)(k_0/k)^\gamma = \vartheta \in \mathbb{R}$. Generally, if further $\lim_{n\to\infty} \sqrt{k_0}(k/k_0)^\gamma = \lambda \in \mathbb{R}$, then*

$$\sqrt{k_0}(\widehat{\gamma}_n^{NM(\theta)}(k_0, k) - \gamma) \stackrel{d}{\to} N(\mu^*, \sigma_\gamma^2) \quad \textit{with} \quad \mu^* = \lambda\left(\vartheta\widetilde{\nu}_{\gamma,\rho} - \gamma\left(\frac{\theta}{1-\gamma} - \widetilde{\nu}_{\gamma,\gamma}\right)\right).$$

Note that the optimal choice of the tuning parameter is not in general available since it depends on the unknown second-order parameter ρ. We consider next the case that $A(t) \sim ct^\rho, c \neq 0, \rho < 0$ and $\gamma < 0$, and establish the optimal sample fraction $k_0 = k_0^{(opt)}$ in the sense of minimal asymptotic mean squared error (AMSE) of $\widehat{\gamma}_n^{NM(\theta)}(k_0, k)$, i.e.,

$$k_0^{(opt)} = \arg\min_{k_0}\left(\frac{\sigma_\gamma^2}{k_0} + \left(\widetilde{\nu}_{\gamma,\rho} A\left(\frac{n}{k_0}\right) + \gamma\left(\widetilde{\nu}_{\gamma,\gamma} - \frac{\theta}{1-\gamma}\right)\left(\frac{k}{k_0}\right)^\gamma\right)^2\right), \quad \gamma < 0.$$

For simplicity of notation, denote with $\widetilde{c} = \gamma\left(\widetilde{\nu}_{\gamma,\gamma} - \theta/1-\gamma\right)$

$$k_0^{(1)} = \left(\frac{\sigma_\gamma^2}{-2\gamma^3\widetilde{c}^2}\right)^{\frac{2\gamma}{2\gamma-1}} k^{\frac{2\gamma}{1-2\gamma}} \quad \text{and} \quad k_0^{(2)} = \left(\frac{\sigma_\gamma^2}{-2\rho c^2 \widetilde{\nu}_{\gamma,\rho}^2}\right)^{\frac{1}{1-2\rho}} n^{\frac{2\rho}{2\rho-1}}. \tag{2.4}$$

Theorem 2.4. *If condition* (2.1) *is satisfied with $A(t) \sim ct^\rho, c \neq 0, \rho < 0$ and $\gamma < 0$, then the optimal sample fraction $k_0^{(opt)}$ is given as follows.*

(a) For $\gamma \geq \rho$, $k_0^{(opt)} \sim k_0^{(1)}$;

(b) For $\gamma < \rho$, we have

(i) If $k \ll n^{\rho(1-2\gamma)/(\gamma(1-2\rho))}$, then $k_0^{(opt)} \sim k_0^{(1)}$;

(ii) If $k \gg n^{\rho(1-2\gamma)/(\gamma(1-2\rho))}$, then $k_0^{(opt)} \sim k_0^{(2)}$;

(iii) If $k \sim Dn^{\rho(1-2\gamma)/(\gamma(1-2\rho))}$ with some $D>0$, then $k_0^{(opt)} \sim D_1 n^{2\rho(2\rho-1)}$ with $D_1 = D_1(\gamma, \rho, D)$ the solution of

$$2\left(\widetilde{c}\left(\frac{D}{D_1}\right)^\gamma + c\widetilde{\nu}_{\gamma,\rho} D_1^\rho\right)\left(\rho c D_1^{1+\rho} - \gamma\widetilde{c} D_1\left(\frac{D}{D_1}\right)^\gamma\right) = \sigma_\gamma^2.$$

Corollary 2.5. *Under the same notation and conditions as in Theorem 2.4. If $\gamma \geq \rho$, or $\gamma < \rho$ and $k \ll n^{\rho(1-2\gamma)/(\gamma(1-2\rho))}$, then the asymptotic bias of $\widehat{\gamma}_n^{NM(\theta)}(k_0,k)$ is $\sigma_\gamma/\sqrt{-2\gamma k_0^{(1)}}$ given that the optimal sample fraction $k_0 \sim k_0^{(1)}$. Moreover, setting $\widetilde{\gamma}_n^{NM(\theta)}(k_0,k) = \widehat{\gamma}_n^{NM(\theta)}(k_0,k) - \sigma_{\widehat{\gamma}_n^{NM(\theta)}(k_0,k)}\Big/\sqrt{-2\widehat{\gamma}_n^{NM(\theta)}(k_0,k)k_0}$ with $k_0 \sim k_0^{(1)}$, we have*

$$\sqrt{k_0}(\widetilde{\gamma}_n^{NM(\theta)}(k_0,k) - \gamma) \ \xrightarrow{d} \ N(0,\sigma_\gamma^2).$$

Remark 2.6. *Note that the precise optimal choices of θ and k_0, given by Remark 2.3 and Theorem 2.4 above, depend on the unknown second order index ρ involved, which results in certain restrictions of its applications.*

3 Applications

In this section, we conduct a small-scale Monte Carlo simulation and a real-life environmental data-set application of the proposed estimators given by (1.3). We consider the proposed PORT-EVI estimations $\widehat{\gamma}_n^{NM(\theta)}(k_0,k)$ with (θ,k_0) replaced by a data-driven/bootstrap estimation $(\widehat{\theta},\widehat{k}_0^*)$ subsequently, with similar methods as those by Gomes et al. (2013) and Draisma et al. (1999). First, set

$$\widehat{\theta} = \widehat{\theta}(k_1,k_2) = \arg\min_\theta \sum_{k_1 \leq i \leq k_2} T_{k,n}^2(i,\theta), \quad k_1 = [k^{0.02}] + 2, \ k_2 = [k^{0.98}], \tag{3.1}$$

where $T_{k,n}(i,\theta) = \widehat{\gamma}_n^{NM(\theta)}([i/2],k) - \widehat{\gamma}_n^{NM(\theta)}(i,k)$. Second, we compute the bootstrap sample fraction $\widehat{k}_0^*$ as follows.

Step 1: Generate B independent bootstrap samples $(x_1^*,\ldots,x_{n_2}^*)$ and $(x_1^*,\ldots,x_{n_2}^*,\ldots,x_{n_1}^*)$ from the excess sample $x_{n-i+1,n} - x_{n-k,n}, i = 1,\ldots,k$ with $n_1 = o(k)$ and $n_2 = [n_1^2/k] + 1$;

Step 2: Compute the lth bootstrap observed value $t_{n_j,n}^*(i,\widehat{\theta},l), i = 2,\ldots,n_j - 1, j = 1,2$ of $T_{n_j,n}(i,\widehat{\theta})$ based on the lth bootstrap sample for $l = 1,\ldots,B$;

Step 3: Obtain with $\mathrm{MSE}^*(i,n_j) = B^{-1}\sum_{l=1}^B (t_{n_j,n}^*(i,\widehat{\theta},l))^2, \ i = 2,\cdots,n_j - 1, j = 1,2$

$$\widehat{k}_T^*(n_j) = \arg\min_i \mathrm{MSE}^*(i,n_j), \ j = 1,2 \quad \text{and} \quad \widehat{\rho}^* = \ln \widehat{k}_T^*(n_1)\big/\big(2\ln(\widehat{k}_T^*(n_1)/n_1)\big);$$

Step 4: Estimate k_0 and γ respectively by

$$\widehat{k}_0^* = \left[\frac{(1-2^{\widehat{\rho}^*})^{2/(1-2\widehat{\rho}^*)}(\widehat{k}_T^*(n_1))^2}{\widehat{k}_T^*(n_2)}\right] + 1 \quad \text{and} \quad \widehat{\gamma}^* = \widehat{\gamma}_n^{NM(\widehat{\theta})}(\widehat{k}_0^*,k).$$

We shall compare the following estimations of EVI at its bootstrap sample fraction $\widehat{k}_0^*$ with $B = 250, k = [n^p], n_1 = [k^p], p \in (0,1)$: our proposed bootstrap estimation (GNM for short), location invariant negative moment estimation (NM), moment estimation (M), corresponding to our estimations (1.3) with $\theta = 0,1$, Weiss-Hill estimation (WH) given by Ling et al. (2012)

$$\widehat{\gamma}_n^{WH}(k_0,k) = \widehat{\gamma}_n^{H}(k_0,k) + \frac{1}{\ln 2}\ln\frac{X_{n,n} - X_{n-m+1,n}}{X_{n,n} - X_{n-2m+1,n}}, \quad m = [(k_0+1)/2],$$

and the maximum likelihood and moment estimations (AML), investigated recently by Hüsler et al. (2016), the solution of the following equations of (γ, ϑ)

$$\begin{cases} \frac{1}{k_0}\sum_{i=1}^{k_0} \ln\left(1+\vartheta(X_{n-i+1,n}-X_{n-k_0,n})\right)=\gamma \\ \frac{1}{k_0}\sum_{i=1}^{k_0} \left(1+\vartheta(X_{n-i+1,n}-X_{n-k_0,n})\right)^{-1/\gamma}=1/2. \end{cases}$$

First, we generate a sample of size $n = 2000$ from the generalized extreme distributed (GEV) parent $X \sim G_\gamma((x-\mu)/\sigma) = \exp\left\{-(1+\gamma(x-\mu)/\sigma)^{-1/\gamma}\right\}$ with $(\gamma, \mu, \sigma) = (-0.5, -1, 1)$. Figure 1 shows that the $\widehat{\theta}(k_1, i)$ is stable for moderate i's since the $T_{k,n}(i, \theta)$ is very close to zero, and the data-driven choice of θ given by (3.1), is equal to 3.19, which makes the estimator $\widehat{\gamma}_n^{NM(\theta)}(k_0, k)$ with $\theta = 3.19$ possessing rather wider steady region than those with $\theta = 0, 1$ and 5. Second, we generate $N = 500$ random samples of size $n = 3000, 3500, 4000, 4500$ from GEV(γ, μ, σ)

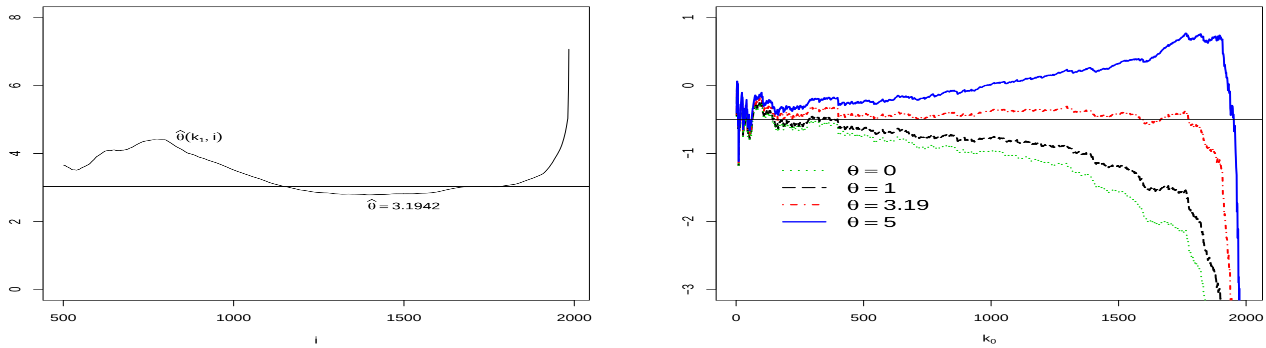

Figure 1: Graph of $(i, \widehat{\theta}(k_1, i))$ (left). Estimations $\widehat{\gamma}_n^{NM(\theta)}(k_0, k)$ for given k (right). Data is from the generalized extreme distribution $G_\gamma((x-\mu)/\sigma)$ with $(\gamma, \mu, \sigma) = (-0.5, -1, 1)$.

with $(\gamma, \mu, \sigma) = (-0.5, -1, 1)$. For ith sample, we calculate the data-driven tuning parameter $\widehat{\theta}^{(i)}$ for the GNM at the random threshold $X_{n-k,n}$ with $k = [n^p]$, and the bootstrap sample fraction $\widehat{k}_0^{i*}$, and the estimators of EVI at $\widehat{k}_0^{i*}$, denoted by $\widehat{\gamma}_n^{i*}, i = 1, \ldots, N$. We compare these estimators by their bias, the mean squared error (MSE), the coefficients of variation (CV) of the bootstrap sample fraction k_0 and the tuning parameter θ. Specifically,

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N} \widehat{\gamma}_n^{i*} - \gamma, \quad \text{MSE} = \frac{1}{N}\sum_{i=1}^{N} (\widehat{\gamma}_n^{i*} - \gamma)^2$$

and

$$CV_{k_0} = \frac{sd(k_0)}{\overline{k}_0} \quad \text{with} \quad \overline{k}_0 = \frac{1}{N}\sum_{i=1}^{N} \widehat{k}_0^{i*}, \quad sd(k_0) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{k}_0^{i*} - \overline{k}_0)^2}$$

$$CV_\theta = \frac{sd(\theta)}{\overline{\theta}} \quad \text{with} \quad \overline{\theta} = \frac{1}{N}\sum_{i=1}^{N} \widehat{\theta}^{(i)}, \quad sd(\theta) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{\theta}^{(i)} - \overline{\theta})^2}.$$

From Table 1, we conclude that the data-driven choice of θ by (3.1) is rather stable since the average θ is around 1.35, and the simulated CV$_\theta$ is around 6.95%. Further, with the chosen $\hat{\theta}$ and stable bootstrap sample fraction $\hat{k}_0^*$, our estimators (GNM) have the smallest bias and rather stable sample paths (recall the small value of CV$_{k_0}$ means stable bootstrap sample fraction). Finally, the MSE of our estimators is not always the smallest one, which might be caused by the simulated θ.

$n=3000$	NM	M	GNM	WH	AML	$n=3500$	NM	M	GNM	WH	AML
Bias$\times$100	-23.62	-8.96	**3.82**	11.73	10.22	Bias$\times$100	-22.02	-12.33	**2.18**	10.33	11.02
MSE$\times$100	7.37	1.89	2.61	1.98	6.22	MSE$\times$100	6.37	3.96	3.16	3.18	1.61
CV$_{k_0}$ $\times$ 100	45.06	35.08	43.81	52.42	42.82	CV$_{k_0}$ $\times$ 100	46.93	42.35	38.22	47.86	43.94
CV$_\theta$ $\times$ 100	–	–	6.94	–	–	CV$_\theta$ $\times$ 100	–	–	5.58	–	–
$\overline{\theta}$	–	–	1.36	–	–	$\overline{\theta}$	–	–	1.13	–	–
$n=4000$	NM	M	GNM	WH	AML	$n=4500$	NM	M	GNM	WH	AML
Bias$\times$100	-19.67	-7.27	**4.34**	10.43	-10.08	Bias$\times$100	-19.81	-9.90	**3.88**	9.94	-9.00
MSE$\times$100	5.79	0.96	1.33	1.61	1.78	MSE$\times$100	5.18	3.95	1.84	1.49	3.10
CV$_{k_0}$$\times$100	54.63	32.36	42.94	46.25	54.38	CV$_{k_0}$$\times$100	45.63	40.59	36.88	46.83	50.57
CV$_\theta$ $\times$ 100	–	–	8.73	–	–	CV$_\theta$ $\times$ 100	–	–	6.56	–	–
$\overline{\theta}$	–	–	1.31	–	–	$\overline{\theta}$	–	–	1.45	–	–

Table 1: Comparisons of the proposed location invariant estimator GNM with NM, M, WH and AML at bootstrap sample fraction $\widehat{k}_0^*$ by the Bias and MSE of EVI, and the CV of k_0 and θ.

Finally, we consider an environmental data set which records the daily average wind speed in Duplin airport during the period 1961–1978. From Figure 2, we see that our estimator $\widehat{\gamma}_n^{NM(1.3078)}(k_0, k)$ with $k=[n^{0.99}]$ possesses rather more stable sample paths than those for the NM and M estimators. It turns out that the wind data possesses a Weibull tail. Further, in order to compare the seasonal difference of the wind data, we carry out the bootstrap algorithm for the sample fraction k_0 with $k=[n^{0.99}], B=250, n_1=[k^{0.99}]$ for the data collected in Spring, Summer, Autumn, Winter, respectively. Table 2 gives the estimations of θ, k_0 and γ involved in our proposed estimations, and the bootstrap 95% confidence interval (CI) of γ given below (recall σ_γ in (2.2) and Remark 2.3)

$$(\widehat{\gamma}_L^*, \widehat{\gamma}_U^*) = \widehat{\gamma}^* \pm 1.96\sigma_{\widehat{\gamma}^*}/\sqrt{\widehat{k}_0^*}.$$

We see that the wind data for Spring and Autumn has lighter tails than those for Winter and Summer.

Period	$\hat{\theta}$	$\hat{k}_0^*$	$\hat{\gamma}^*$	$(\hat{\gamma}_L^*, \hat{\gamma}_U^*)$	Period	$\hat{\theta}$	$\hat{k}_0^*$	$\hat{\gamma}^*$	$(\hat{\gamma}_L^*, \hat{\gamma}_U^*)$
Spring	1.2060	595	-0.1902	$(-0.2704, -0.1101)$	Summer	1.5181	1070	0.0012	$(-0.0588, 0.0612)$
Autumn	1.3207	969	-0.0993	$(-0.1598, -0.0387)$	Winter	1.3301	570	-0.0307	$(-0.1108, 0.0495)$

Table 2: Estimations of θ, k_0, γ and 95% CI estimation of γ, denoted by $(\widehat{\gamma}_L^*, \widehat{\gamma}_U^*)$. Data is the daily wind speed during 1961–1978 in Dublin airport.

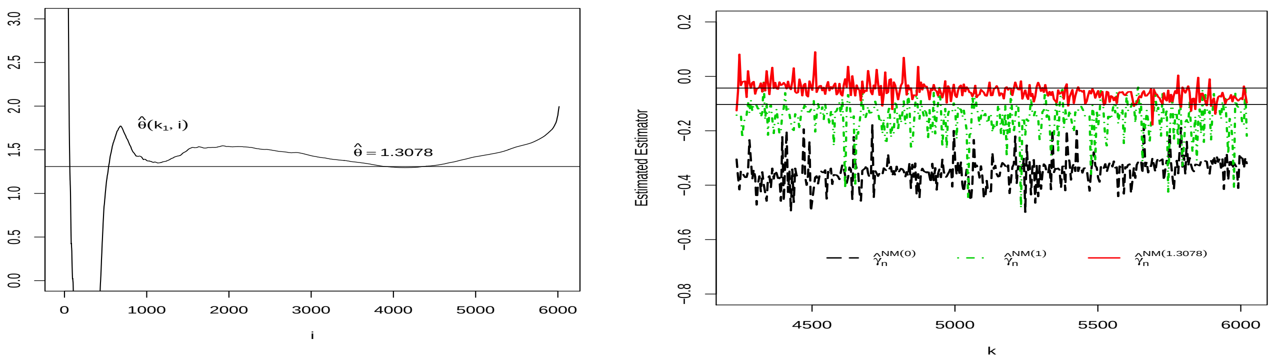

Figure 2: Graph of $(i, \widehat{\theta}(k_1, i))$ (left). Sample paths of $\widehat{\gamma}_n^{NM(\theta)}(\widehat{k}_0^*, k)$ with different θ for the daily average wind speed collected in 1963–1971 (right). The estimated 95% CI is indicated by the horizontal lines (right).

4 Proofs

Proof of Lemma 2.1 Clearly, it follows from (2.1) that

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma} =: D_\gamma(x)$$

holds locally uniformly for all $x > 0$. Therefore,

$$\widetilde{a}(n/k, n/k_0) := \frac{a(n/k_0)}{U(n/k_0) - U(n/k)} = \begin{cases} -\gamma \left(\frac{k_0}{k}\right)^{-\gamma}(1 + o(1)), & \gamma < 0, \\ -\frac{1}{\ln(k_0/k)}(1 + o(1)), & \gamma = 0. \end{cases} \tag{4.1}$$

Using further the Taylor's expansion $\ln(1 + x) = x - x^2/2 + O(x^3),\ x \to 0$, we have

$$\begin{aligned} \ln\left(\frac{U(n/k_0 x) - U(n/k)}{U(n/k_0) - U(n/k)}\right) &= \ln\left(1 + \frac{a(n/k_0)}{U(n/k_0) - U(n/k)}\big(D_\gamma(x) + A(n/k_0)H_{\gamma,\rho}(x)(1 + o(1))\big)\right) \\ &= \widetilde{a}(n/k, n/k_0)\big(D_\gamma(x) + A(n/k_0)H_{\gamma,\rho}(x)(1 + o(1))\big) \\ &\quad - \frac{1}{2}\left(\widetilde{a}(n/k, n/k_0)\right)^2 \left(D_\gamma(x) + O\big(A(n/k_0)\big)\right)^2 (1 + o(1)), \quad n \to \infty. \end{aligned}$$

Consequently, the desired result follows by noting that $H_{\gamma,\gamma}(x) = \big((x^\gamma - 1)/\gamma\big)^2/2$ and the fact that condition (2.1) holds locally uniformly for all $x > 0$. □

Proof of Theorem 2.2 Let $Y_{1,n} \le Y_{2,n} \le \cdots \le Y_{n,n}$ denote the increasing order statistics of sample $Y_1, \cdots, Y_n$ from the parent $Y \sim F_Y(y) = 1 - 1/y, y \ge 1$. Since $\{X_{n-i+1,n}\}_{i=1}^n \stackrel{d}{=} \{U(Y_{n-i+1,n})\}_{i=1}^n$, we have

$$\widehat{\gamma}_n^{NM(\theta)}(k_0, k) \stackrel{d}{=} \theta\widehat{\gamma}_n^{(1)}(k_0, k) + \frac{\widehat{\gamma}_n^{(2)}(k_0, k) - 2(\widehat{\gamma}_n^{(1)}(k_0, k))^2}{2(\widehat{\gamma}_n^{(2)}(k_0, k) - (\widehat{\gamma}_n^{(1)}(k_0, k))^2)} \tag{4.2}$$

with

$$\widehat{\gamma}_n^{(j)}(k_0, k) = \frac{1}{k_0}\sum_{i=1}^{k_0}\left(\ln\frac{U(Y_{n-i+1,n}) - U(Y_{n-k,n})}{U(Y_{n-k_0,n}) - U(Y_{n-k,n})}\right)^j, \quad j > 0.$$

Recalling that $\nu_{j,\gamma,\rho}$ and $\mu_{j,\gamma}$ are given in (2.2), we verify the following asymptotic expansions

$$\frac{\widehat{\gamma}_n^{(j)}(k_0, k)}{(\widetilde{a}(n/k, n/k_0))^j} \stackrel{d}{=} \mu_{j,\gamma} + \frac{\sigma_{j,\gamma}}{\sqrt{k_0}}Z_{k_0}^{(j)}(1 + o_p(1)) + \nu_{j,\gamma,\rho}A\left(\frac{n}{k_0}\right)(1 + o_p(1)) - \nu_{j,\gamma,\gamma}\widetilde{a}\left(\frac{n}{k}, \frac{n}{k_0}\right)(1 + o_p(1)), \tag{4.3}$$

where, with $\mathbf{C}$ the correlation matrix with correlation entry $(\mu_{j_1+j_2,\gamma}-\mu_{j_1,\gamma}\mu_{j_2,\gamma})/\sigma_{j_1,\gamma}\sigma_{j_2,\gamma}, \sigma^2_{j,\gamma}=\mu_{2j,\gamma}-\mu^2_{j,\gamma}$

$$(Z^{(j_1)}_{k_0}, Z^{(j_2)}_{k_0})^\top \overset{d}{\to} N_2\left(\mathbf{0},\mathbf{C}\right), \quad j_1, j_2>0.$$

Indeed, it follows from de Haan and Ferreira (2006) that $\{Y_{n-i+1,n}/Y_{n-k_0,n}\}_{i=1}^{k_0} \overset{d}{=} \{Y_{k_0-i+1,k_0}\}_{i=1}^{k_0}$ and

$$\sqrt{k_0}\Big((k_0/n)Y_{n-k_0,n}-1\Big) \overset{d}{\to} N(0,1), \quad \sqrt{k}\Big((k/n)Y_{n-k,n}-1\Big) \overset{d}{\to} N(0,1).$$

Thus,

$$\frac{\widetilde{a}(Y_{n-k,n},Y_{n-k_0,n})}{\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)}=1+o_p(k_0^{-1/2}), \quad \frac{A\left(Y_{n-k_0,n}\right)}{A\left(n/k_0\right)}=1+o_p(1).$$

Consequently, it follows from Lemma 2.1 that

$$\begin{aligned}
&\frac{\widehat{\gamma}^{(j)}_n(k_0,k)}{(\widetilde{a}(n/k,n/k_0))^j}=\frac{\widehat{\gamma}^{(j)}_n(k_0,k)}{(\widetilde{a}(Y_{n-k,n},Y_{n-k_0,n}))^j}\left(\frac{\widetilde{a}(Y_{n-k,n},Y_{n-k_0,n})}{\widetilde{a}(n/k,n/k_0)}\right)^j\\
&=\frac{1}{k_0}\sum_{i=1}^{k_0}\Big(D_\gamma(Y_{k_0-i+1,k_0})+H_{\gamma,\rho}(Y_{k_0-i+1,k_0})A(Y_{n-k_0,n})(1+o_p(1))\\
&\quad-H_{\gamma,\gamma}(Y_{k_0-i+1,k_0})\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1))\Big)^j\left(1+o_p\left(\frac{1}{\sqrt{k_0}}\right)\right)\\
&=\frac{1}{k_0}\sum_{i=1}^{k_0}\left(D_\gamma(Y_i)\right)^j+\frac{1}{k_0}\sum_{i=1}^{k_0}j\left(D_\gamma(Y_i)\right)^{j-1}H_{\gamma,\rho}(Y_i)A\left(\frac{n}{k_0}\right)(1+o_p(1))\\
&\quad-\frac{1}{k_0}\sum_{i=1}^{k_0}j\left(D_\gamma(Y_i)\right)^{j-1}H_{\gamma,\gamma}(Y_i)\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1))+o_p\left(\frac{1}{\sqrt{k_0}}\right)\\
&=\mathrm{E}\left\{\left(D_\gamma(Y)\right)^j\right\}+\frac{\sigma_{j,\gamma}}{\sqrt{k_0}}\frac{\sqrt{k_0}}{\sigma_{j,\gamma}}\left(\frac{1}{k_0}\sum_{i=1}^{k_0}(D_\gamma(Y_i))^j-\mathrm{E}\left\{\left(D_\gamma(Y)\right)^j\right\}\right)+o_p\left(\frac{1}{\sqrt{k_0}}\right)\\
&\quad+\mathrm{E}\left\{j\big(D_\gamma(Y)\big)^{j-1}H_{\gamma,\rho}(Y)\right\}A\left(\frac{n}{k_0}\right)(1+o_p(1))-\mathrm{E}\left\{\left(j\big(D_\gamma(Y)\big)^{j-1}H_{\gamma,\gamma}(Y)\right)\right\}\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1))\\
&=\mu_{j,\gamma}+\frac{\sigma_{j,\gamma}}{\sqrt{k_0}}Z^{(j)}_{k_0}(1+o_p(1))+\nu_{j,\gamma,\rho}A\left(\frac{n}{k_0}\right)(1+o_p(1))-\nu_{j,\gamma,\gamma}\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1)), \quad j>0,
\end{aligned}$$

which, together with the Cramér–Wold device and the Liapounov's theorem (cf. Chung (1974), p 200), implies (4.3).

Next, we show that (4.3) implies (2.3). Indeed, since it follows from $\mu_{j,\gamma}=\mathrm{E}\left\{(D_\gamma(Y))^j\right\}$ that

$$\mu_{1,\gamma}=\frac{1}{1-\gamma}, \quad \mu_{2,\gamma}=\frac{2}{(1-\gamma)(1-2\gamma)}, \quad \mu_{2,\gamma}-2\mu^2_{1,\gamma}=2\gamma\sigma^2_{1,\gamma}, \quad \gamma\le 0,$$

Setting below $c_{j,\gamma,\rho}=\nu_{2,\gamma,\rho}-2j\mu_{1,\gamma}\nu_{1,\gamma,\rho},\ j=1,2$, we have by Taylor's expansion $1/(1+x)=1-x(1+o(1)),\ x\to 0$

$$\begin{aligned}
&\frac{\widehat{\gamma}^{(2)}_n(k_0,k)-2(\widehat{\gamma}^{(1)}_n(k_0,k))^2}{2(\widehat{\gamma}^{(2)}_n(k_0,k)-(\widehat{\gamma}^{(1)}_n(k_0,k))^2)}\\
&=\frac{2\gamma\sigma^2_{1,\gamma}+\dfrac{\sigma_{2,\gamma}Z^{(2)}_{k_0}-4\mu_{1,\gamma}\sigma_{1,\gamma}Z^{(1)}_{k_0}}{\sqrt{k_0}}(1+o_p(1))+c_{2,\gamma,\rho}A\left(\dfrac{n}{k_0}\right)(1+o_p(1))-c_{2,\gamma,\gamma}\widetilde{a}\left(\dfrac{n}{k},\dfrac{n}{k_0}\right)(1+o_p(1))}{2\left(\sigma^2_{1,\gamma}+\dfrac{\sigma_{2,\gamma}Z^{(2)}_{k_0}-2\mu_{1,\gamma}\sigma_{1,\gamma}Z^{(1)}_{k_0}}{\sqrt{k_0}}(1+o_p(1))+c_{1,\gamma,\rho}A\left(\dfrac{n}{k_0}\right)(1+o_p(1))-c_{1,\gamma,\gamma}\widetilde{a}\left(\dfrac{n}{k},\dfrac{n}{k_0}\right)(1+o_p(1))\right)}\\
&=\left(\gamma+\frac{\sigma_{2,\gamma}Z^{(2)}_{k_0}-4\mu_{1,\gamma}\sigma_{1,\gamma}Z^{(1)}_{k_0}}{2\sigma^2_{1,\gamma}\sqrt{k_0}}(1+o_p(1))+\frac{c_{2,\gamma,\rho}}{2\sigma^2_{1,\gamma}}A\left(\frac{n}{k_0}\right)(1+o_p(1))-\frac{c_{2,\gamma,\gamma}}{2\sigma^2_{1,\gamma}}\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1))\right)\\
&\quad\times\left(1-\frac{\sigma_{2,\gamma}Z^{(2)}_{k_0}-2\mu_{1,\gamma}\sigma_{1,\gamma}Z^{(1)}_{k_0}}{\sigma^2_{1,\gamma}\sqrt{k_0}}(1+o_p(1))-\frac{c_{1,\gamma,\rho}}{\sigma^2_{1,\gamma}}A\left(\frac{n}{k_0}\right)(1+o_p(1))+\frac{c_{1,\gamma,\gamma}}{\sigma^2_{1,\gamma}}\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1))\right)
\end{aligned}$$

$$= \gamma + \frac{(1-2\gamma)(1-\gamma)}{\sqrt{k_0}}\left(\frac{\sigma_{2,\gamma}}{\mu_{2,\gamma}}Z^{(2)}_{k_0} - 2\frac{\sigma_{1,\gamma}}{\mu_{1,\gamma}}Z^{(1)}_{k_0}\right)(1+o_p(1)) + \tilde{\nu}_{\gamma,\rho}A\left(\frac{n}{k_0}\right)(1+o_p(1)) - \tilde{\nu}_{\gamma,\gamma}\tilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1))$$

$$=: \gamma + \frac{Z_{k_0}}{\sqrt{k_0}}(1+o_p(1)) + \tilde{\nu}_{\gamma,\rho}A\left(\frac{n}{k_0}\right)(1+o_p(1)) - \tilde{\nu}_{\gamma,\gamma}\tilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1)), \tag{4.4}$$

where, $\widetilde{\nu}_{\gamma,\rho} = (1-2\gamma)(1-\gamma)\left(\frac{\nu_{2,\gamma,\rho}}{\mu_{2,\gamma}} - 2\frac{\nu_{1,\gamma,\rho}}{\mu_{1,\gamma}}\right)$, and $Z_{k_0} \xrightarrow{d} N(0,\sigma_\gamma^2)$ follows by (4.3) with

$$\sigma_\gamma^2 = \lim_{n\to\infty}\mathrm{Var}\left\{(1-2\gamma)(1-\gamma)\left(\frac{\sigma_{2,\gamma}}{\mu_{2,\gamma}}Z^{(2)}_{k_0} - 2\frac{\sigma_{1,\gamma}}{\mu_{1,\gamma}}Z^{(1)}_{k_0}\right)\right\} = \frac{(1-\gamma)^2(1-2\gamma)(1-\gamma+6\gamma^2)}{(1-3\gamma)(1-4\gamma)}.$$

Moreover, it follows from (4.3) with $j=1$ that

$$\widehat{\gamma}_n^{(1)}(k_0,k) = \frac{1}{1-\gamma}\widetilde{a}\left(\frac{n}{k},\frac{n}{k_0}\right)(1+o_p(1)),$$

which together with (4.2) and (4.4) implies (2.3). We complete the proof of Theorem 2.2. □

Proof of Theorem 2.4 (a) Since $\lim_{n\to\infty} A(n/k_0)(k_0/k)^\gamma = 0$ for $\gamma \ge \rho$, we have with $\widetilde{c} = \gamma(\widetilde{\nu}_{\gamma,\gamma} - \theta/(1-\gamma))$

$$\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \frac{\sigma_\gamma^2}{k_0} + \widetilde{c}^2\left(\frac{k}{k_0}\right)^{2\gamma}.$$

If $k_0 \ll k^{2\gamma/(2\gamma-1)}$, then $1/\sqrt{k_0} \gg (k/k_0)^\gamma$, and thus $\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \sigma_\gamma^2/k_0$ is a decreasing function of k_0. If $k_0 \gg k^{2\gamma/(2\gamma-1)}$, then $1/\sqrt{k_0} \ll (k/k_0)^\gamma$, and thus $\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \widetilde{c}^2\,(k/k_0)^{2\gamma}$ is an increasing function of k_0. Therefore, we choose $k_0 = O(k^{2\gamma/(2\gamma-1)})$ in order to balance the bias and variance of $\widehat{\gamma}_n^{NM(\theta)}(k_0,k)$ and obtain $k_0^{(opt)} \sim k_0^{(1)}$ given by (2.4).

(b) For $\gamma < \rho$, we show the three cases that $k \ll, \gg n^{\rho(1-2\gamma)/(\gamma(1-2\rho))}$ and $k \sim Dn^{\rho(1-2\gamma)/(\gamma(1-2\rho))}$ for some $D>0$ subsequently. For case (i), we have

$$k^{\gamma/(\gamma-\rho)}n^{\rho/(\rho-\gamma)} \ll k^{\gamma/(\gamma-1/2)} \ll n^{\rho/(\rho-1/2)}.$$

Hence, if $k_0 \ll k^{\gamma/(\gamma-1/2)}$, then $1/\sqrt{k_0} \gg (k_0/k)^{-\gamma}, 1/\sqrt{k_0} \gg (k_0/n)^{-\rho}$, and thus $\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \sigma_\gamma^2/k_0$ is decreasing with respect to k_0. On the other hand, if $k_0 \gg k^{\gamma/(\gamma-1/2)}(\gg k^{\gamma/(\gamma-\rho)}n^{\rho/(\rho-\gamma)})$, then $1/\sqrt{k_0} \ll (k_0/k)^{-\gamma}, (k_0/n)^{-\rho} \ll (k_0/k)^{-\gamma}$, implying $\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \widetilde{c}^2(k_0/k)^{-2\gamma}$ which is increasing w.r.t. k_0. Therefore, the optimal $k_0^{(opt)} = O(k^{2\gamma/(2\gamma-1)})$ such that $\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \sigma_\gamma^2/k_0 + \widetilde{c}^2(k_0/k)^{-2\gamma}$ reaches its minimum. Consequently, we have $k_0^{(opt)} \sim k_0^{(1)}$.

For case (ii), note that

$$n^{\rho/(\rho-1/2)} \ll k^{\gamma/(\gamma-1/2)} \ll k^{\gamma/(\gamma-\rho)}n^{\rho/(\rho-\gamma)}.$$

Similar arguments as for case (i) give $k_0^{(opt)} = O(n^{2\rho/(2\rho-1)})$ such that $1/\sqrt{k_0} = O((n/k_0)^\rho) \gg ((k/k_0)^\gamma)$ and

$$\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = \sigma_\gamma^2/k_0 + c^2\widetilde{\nu_{\gamma,\rho}^2}(k_0/n)^{-2\rho}$$

reaches its minimum. Consequently, we have $k_0^{(opt)} \sim k_0^{(2)}$.

For case (iii), note that

$$k^{\gamma/(\gamma-1/2)} \sim D^{\gamma/(\gamma-1/2)}n^{\rho/(\rho-1/2)} \sim D^{\frac{\gamma(1/2-\rho}{(\gamma-1/2)(\gamma-\rho)}}k^{\gamma/(\gamma-\rho)}n^{\rho/(\rho-\gamma)}.$$

Similar arguments as for case (i) give $k_0 \sim D_1 n^{\rho/(\rho-1/2)}$ such that

$$\mathrm{MSE}_\infty(\widehat{\gamma}_n^{NM(\theta)}(k_0,k)) = n^{\rho/(\rho-1/2)} \left(\frac{\sigma_\gamma^2}{D_1} + \left(\widetilde{c} \left(\frac{D}{D_1} \right)^\gamma + c\widetilde{\nu}_{\gamma,\rho} D_1^\rho \right)^2 \right)$$

reaches its minimum. We complete the proof of Theorem 2.4. □

Acknowledgments The authors would like to thank the referees for their important suggestions which significantly improved this contribution. C. Liu and C. Ling acknowledge the National Natural Science Foundation of China grant (11604375) and the Natural Science Foundation Project of CQ (cstc2016jcyjA0036). C. Ling is also supported by the China Postdoctoral Science Foundation (2016M602624) and Chinese Government Scholarship (201708505031).

References

[1] Caeiro, F., Gomes, M.I. (2010) An asymtotically unbiased moment estimator of a negative extreme value index. *Discussiones Mathematicae: Probability and Statistics.* **30**: 5–19.

[2] Caeiro, F., Gomes, M.I., Rodrigues, L.H. (2016) A location-invariant probability weighted moment estimation of the extreme value index. *International Journal of Computer Mathematics.* **93**(4): 676–695.

[3] Cai, J., de Haan, L., Zhou, C. (2012) Bias correction in extreme value statistics with index around zero. *Extremes.* **16**: 173–201.

[4] Chung, K.L. (1974) *A course in Probability Theory,* 2nd ed. Academic, New York.

[5] Danielsson, J., de Haan, L., Peng, L., de Vries, C.G. (2001) Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis.* **76**: 226–248.

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

[7] de Haan, L., Ferreira, A. (2006) *Extreme Value Theory: An introduction.* Springer, New York.

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

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

[10] Fraga Alves, M.I. (2001) A location invariant Hill-type estimator. *Extremes.* **4**: 199–217.

[11] Gomes, M.I., Henriques-Rodrigues, L., Caeiro, F. (2013) *Advances in Theoretical and Applied Statistics*, Pages 143–153. Springer Berlin Heidelberg.

[12] Gomes, M.I., Henriques-Rodrigues, L. (2016) Competitive estimation of the extreme value index. *Statistics* & *Probability Letters.* **117**: 128–135.

[13] Hill, B. (1975) A simple general approach to inference about the tail of a distribution. *Annals of Statistics.* **3**: 1163–1174.

[14] Hüsler, J., Li, D., Raschke, M. (2016). Extreme value index estimator using maximum likelihood and moment estimation. *Communication in Statistics - Theory and Methods.* **45**(12): 3625–3636.

[15] Li, D., Peng, L. (2009) Does bias reduction with external estimator of second-order parameter work for endpoint? *Journal of Statistical Planning and Inference.* **139**: 1937–1952.

[16] Li, D., Peng, L., Xu, X. (2011) Bias reduction for endpoint estimation. *Extremes.* **14**: 393–412.

[17] Ling, C., Peng, Z., Nadarajah, S. (2007) A location invariant moment-type estimator II. *Theory of Probability and Mathematical Statistics.* **77**: 177–189.

[18] Ling, C., Peng, Z., Nadarajah, S. (2012) Location invariant Weiss-Hill estimator. *Extremes.* **15**: 197–230.